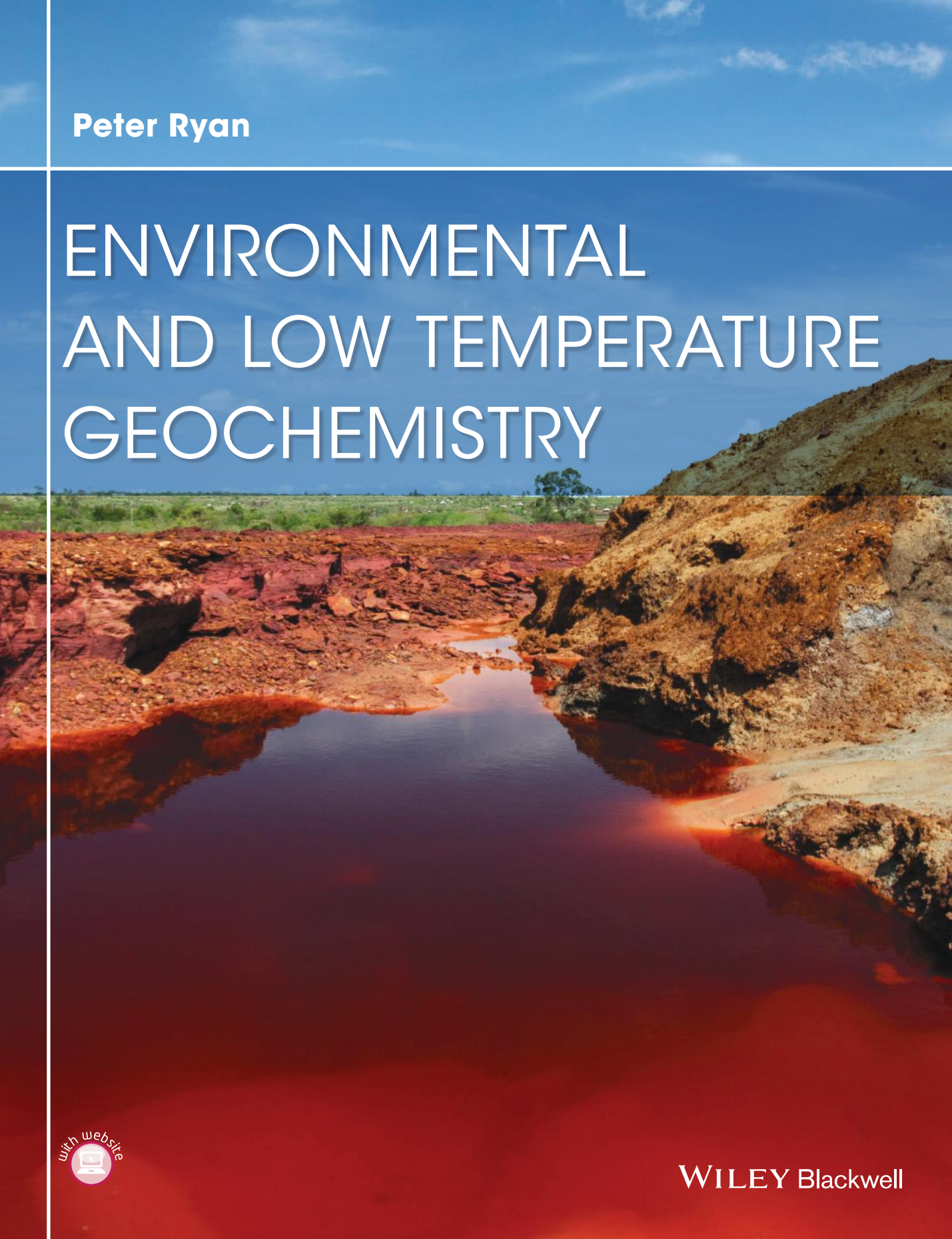




Peter Ryan

ENVIRONMENTAL AND LOW TEMPERATURE GEOCHEMISTRY



WILEY Blackwell

Environmental and Low Temperature Geochemistry

Environmental and Low Temperature Geochemistry

Peter Crowley Ryan

WILEY Blackwell

This edition first published 2014 © 2014 by Peter Crowley Ryan

Registered Office

John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Offices

9600 Garsington Road, Oxford, OX4 2DQ, UK

The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

The right of the author to be identified as the author of this work has been asserted in accordance with the UK Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty: While the publisher and author(s) have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. It is sold on the understanding that the publisher is not engaged in rendering professional services and neither the publisher nor the author shall be liable for damages arising herefrom. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Ryan, Peter, 1966– author.

Environmental and low temperature geochemistry / Peter Ryan.

pages cm

Includes bibliographical references and index.

ISBN 978-1-118-86735-8 (cloth) – ISBN 978-1-4051-8612-4 (pbk.) 1. Environmental geochemistry. I. Title.

QE516.4.R93 2014

551.9–dc23

2013050119

A catalogue record for this book is available from the British Library.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Cover image: Acid mine drainage. Large pool of red coloured water caused by acid mine drainage; the leeching of acidic minerals from earth exposed to water or air during mining. This is the site of a former copper mine that was closed in the 1960s. Photographed in Macalder, Kenya. Phil Hill/Science Photo Library.

Set in 10/13pt Photina by SPi Publisher Services, Pondicherry, India

Contents

Acknowledgements, xii

About the Companion Website, xiii

1 BACKGROUND AND BASIC CHEMICAL PRINCIPLES: ELEMENTS, IONS, BONDING, REACTIONS, 1

- 1.1 An overview of environmental geochemistry – history, scope, questions, approaches, challenges for the future, 1
- 1.2 The naturally occurring elements – origins and abundances, 2
- 1.3 Atoms and isotopes: a brief review, 6
- 1.4 Measuring concentrations, 8
 - 1.4.1 Mass-based concentrations, 8
 - 1.4.2 Molar concentrations, 9
 - 1.4.3 Concentrations of gases, 10
 - 1.4.4 Notes on precision and accuracy, significant figures and scientific notation, 10
- 1.5 Periodic table, 11
- 1.6 Ions, molecules, valence, bonding, chemical reactions, 14
 - 1.6.1 Ionic bond strength, 14
 - 1.6.2 Covalent bonds, 16
 - 1.6.3 Electronegativity, 17
 - 1.6.4 Metallic bonds, hydrogen bonds and van der Waals forces, 18
- 1.7 Acid–base equilibria, pH, K values, 19
- 1.8 Fundamentals of redox chemistry and chemical reactions, 21

- 1.9 Chemical reactions, 23
- 1.10 Equilibrium, thermodynamics and driving forces for reactions: systems, gibbs energies, enthalpy and heat capacity, entropy, volume, 23
 - 1.10.1 Systems, species, phases and components, 24
 - 1.10.2 First law of thermodynamics, 26
 - 1.10.3 Second law of thermodynamics, 27
 - 1.10.4 Enthalpy, 27
 - 1.10.5 Heat capacity, 29
 - 1.10.6 Gibbs free energy, 30
 - 1.10.7 Gibbs free energy and the equilibrium constant, 31
- 1.11 Kinetics and reaction rates: distance from equilibrium, activation energy, metastability, 33
 - 1.11.1 Reaction rate, reaction order, 34
 - 1.11.2 Temperature and the Arrhenius equation, 36
- Review questions, 37
- References, 37

2 SURFICIAL AND ENVIRONMENTAL MINERALOGY, 39

- 2.1 Introduction to minerals and unit cells, 40
- 2.2 Ion coordination, Pauling's rules and ionic substitution, 42
 - 2.2.1 Coordination and radius ratio, 42

- 2.2.2 Bond-strength considerations, 45
 - 2.2.3 Pauling's and Goldschmidt's rules of ionic solids, 45
- 2.3 Silicates, 48
 - 2.3.1 Nesosilicates, 49
 - 2.3.2 Inosilicates, 50
 - 2.3.3 Phyllosilicates, 52
 - 2.3.4 Tectosilicates, 58
- 2.4 Clay minerals (T–O minerals, T–O–T minerals, interstratified clays), 58
 - 2.4.1 Smectite, 59
 - 2.4.2 Smectites with tetrahedrally derived layer charge, 60
 - 2.4.3 Smectites with octahedrally derived layer charge, 60
 - 2.4.4 Vermiculite, 62
 - 2.4.5 Illite, 62
 - 2.4.6 Chlorite and Berthierine, 63
 - 2.4.7 Kaolin (kaolinite and halloysite), 63
 - 2.4.8 Interstratified clay minerals, 64
 - 2.4.9 Trace metals and metalloids in clay minerals, 64
- 2.5 Crystal chemistry of adsorption and cation exchange, 64
 - 2.5.1 Cation exchange, 66
 - 2.5.2 Double-layer complexes, 68
- 2.6 Low-temperature non-silicate minerals: carbonates, oxides and hydroxides, sulfides, sulfates, salts, 70
 - 2.6.1 Carbonates, 70
 - 2.6.2 Oxides and hydroxides, 71
 - 2.6.3 Sulfides and sulfates, 72
 - 2.6.4 Halide and nitrate salts, 74
- 2.7 Mineral growth and dissolution, 74
- 2.8 Biomineralization, 78
- Review questions, 79
- References, 80

3 ORGANIC COMPOUNDS IN THE ENVIRONMENT, 82

- 3.1 Introduction to organic chemistry: chains and rings, single, double, and triple bonds, functional groups, classes of organic compounds, organic nomenclature, 82
 - 3.1.1 Definition of organic compounds, 82
 - 3.1.2 Hybridization of carbon atoms in organic compounds, 83
 - 3.1.3 Alkanes, 84
 - 3.1.4 Alkenes, 86
 - 3.1.5 Functional groups, 86
 - 3.1.6 Aromatic hydrocarbons and related compounds, 88
 - 3.1.7 Nitrogen, phosphorus and sulfur in organic compounds, 92
 - 3.1.8 Pharmaceutical compounds, 93
- 3.2 Natural organic compounds at the earth surface, 94
 - 3.2.1 Fossil fuels, 95
- 3.3 Fate and transport of organic pollutants, controls on bioavailability, behavior of DNAPLs and LNAPLs, biodegradation, remediation schemes, 96
 - 3.3.1 Solid–liquid–gas phase considerations, 96
 - 3.3.2 Solubility considerations, 97
 - 3.3.3 Interactions of organic compounds and organisms, 98
 - 3.3.4 Adsorption of organic compounds, 99
 - 3.3.5 Non-aqueous phase liquids (NAPLs) in the environment, 103
 - 3.3.6 Biodegradation, 104

- 3.3.7 Remediation, 105
 - 3.4 Summary, 106
 - Questions, 106
 - References, 106
- 4 AQUEOUS SYSTEMS – CONTROLS ON WATER CHEMISTRY, 108**
 - 4.1 Introduction to the geochemistry of natural waters, 108
 - 4.1.1 Geochemistry and the hydrologic cycle, 108
 - 4.2 The structure of water – geometry, polarity and consequences, 113
 - 4.3 Dissolved versus particulate: examples of solutions and suspensions, 114
 - 4.3.1 Dissolved vs. particulate vs. colloidal, 115
 - 4.4 Speciation: simple ions, polyatomic ions and aqueous complexes, 116
 - 4.5 Controls on the solubility of inorganic elements and ions, 117
 - 4.5.1 The ratio of ionic charge: ionic radius and its effect on solubility, 119
 - 4.5.2 Reduction–oxidation reactions, 120
 - 4.5.3 Half-cell reactions, 120
 - 4.5.4 Redox reactions in the environment, 122
 - 4.5.5 pH and acid–base consideration, 123
 - 4.5.6 Ligands and elemental mobility, 124
 - 4.6 Ion activities, ionic strength, TDS, 125
 - 4.6.1 Ion activity product, 126
 - 4.6.2 Ionic strength, 126
 - 4.6.3 Total dissolved solids, 126
 - 4.7 Solubility products, saturation, 127
 - 4.8 Co-precipitation, 128
 - 4.9 Behavior of selected elements in aqueous systems, 129
 - 4.9.1 Examples of heavy metals and metalloids, 129
 - 4.9.2 Eh–pH diagrams, 132
 - 4.9.3 Silicon in solutions, 136
 - 4.10 Effect of adsorption and ion exchange on water chemistry, 137
 - 4.10.1 Ionic potential, hydration radius and adsorption, 138
 - 4.10.2 Law of mass action and adsorption, 138
 - 4.10.3 Adsorption edges, 140
 - 4.10.4 Adsorption isotherms, 142
 - 4.11 Other graphical representations of aqueous systems: piper and stiff diagrams, 143
 - 4.12 Summary, 146
 - Questions, 147
 - References, 147
- 5 CARBONATE GEOCHEMISTRY AND THE CARBON CYCLE, 149**
 - 5.1 Carbonate geochemistry: inorganic carbon in the atmosphere and hydrosphere, 149
 - 5.1.1 Atmospheric CO₂, carbonate species and the pH of rain, 150
 - 5.1.2 Speciation in the carbonate system as a function of pH, 151
 - 5.1.3 Alkalinity, 152
 - 5.1.4 Carbonate solubility and saturation, 155
 - 5.1.5 The effect of CO₂ partial pressure on stability of carbonate minerals, 157
 - 5.1.6 The effect of mineral composition on stability of carbonate minerals, 157
 - 5.2 The carbon cycle, 158
 - 5.2.1 Oxidation states of carbon, 158
 - 5.2.2 Global-scale reservoirs and fluxes of carbon, 159
 - 5.2.3 Fixation of carbon into the crust, 161
 - 5.2.4 Rates of flux to and from the crust, 164

- 5.2.5 The ocean reservoir, 166
- 5.2.6 Fixation of C into oceans, 166
- 5.2.7 Long-term viability of oceans as C sink, 168
- 5.2.8 The atmospheric reservoir, 171
- 5.2.9 Sequestration, 174

Questions, 175

References, 176

6 BIOGEOCHEMICAL CYCLES – N, P, S, 177

- 6.1 The nitrogen cycle, 180
 - 6.1.1 Nitrogen valence, nitrogen species, 181
 - 6.1.2 Processes operating within the nitrogen cycle, 182
 - 6.1.3 Global scale reservoirs and fluxes of nitrogen, 184
 - 6.1.4 Human perturbation of the nitrogen cycle and resulting environmental impacts, 186
- 6.2 The phosphorus cycle, 190
 - 6.2.1 P cycling in soils, 191
 - 6.2.2 The global phosphorus cycle, 193
 - 6.2.3 Phosphorus and eutrophication, 194
- 6.3 Comparison of N and P, 195
- 6.4 The sulfur cycle, 196
 - 6.4.1 Sulfur valence, sulfur species, 196
 - 6.4.2 The global S cycle, 197
 - 6.4.3 The marine S cycle, 198
 - 6.4.4 Soils and biota, 200
 - 6.4.5 Atmosphere, 200
 - 6.4.6 River flux, 201
- 6.5 Integrating the C, N, P and S cycles, 202

Questions, 203

References, 203

7 THE GLOBAL ATMOSPHERE: COMPOSITION, EVOLUTION AND ANTHROPOGENIC CHANGE, 206

- 7.1 Atmospheric structure, circulation and composition, 206
 - 7.1.1 Structure and layering of the atmosphere, 207
 - 7.1.2 Geological record of atmospheric composition, 208
 - 7.1.3 Climate proxies, 209
 - 7.1.4 Orbital control on C, 210
 - 7.1.5 Composition of the current atmosphere, 213
 - 7.1.6 Air circulation, 214
- 7.2 Evaporation, distillation, CO₂ dissolution and the composition of natural precipitation, 218
- 7.3 The electromagnetic spectrum, greenhouse gases and climate, 219
 - 7.3.1 Electromagnetic spectrum, 219
 - 7.3.2 Re-radiation from earth surface, 219
 - 7.3.3 Greenhouse effect and heat trapping, 222
- 7.4 Greenhouse gases: structures, sources, sinks and effects on climate, 223
 - 7.4.1 Molecular structures and vibrations of greenhouse gases, 223
 - 7.4.2 Greenhouse gases, radiative forcing, GWPs, 224
 - 7.4.3 Global warming, 227

Questions, 228

References, 228

8 URBAN AND REGIONAL AIR POLLUTION, 230

- 8.1 Oxygen and its impact on atmospheric chemistry, 231
- 8.2 Free radicals, 232
- 8.3 Sulfur dioxide, 234

- 8.4 Nitrogen oxides, 237
- 8.5 Carbon monoxide, 238
- 8.6 Particulate matter, 240
- 8.7 Lead (Pb), 242
- 8.8 Hydrocarbons and air quality: tropospheric ozone and photochemical smog, 242
- 8.9 Stratospheric ozone chemistry, 245
- 8.10 Sulfur and nitrogen gases and acid deposition, 249
- 8.11 Trace elements in atmospheric deposition: organochlorine pesticides, mercury and other trace elements, 252
 - 8.11.1 Pesticides in air, 252
 - 8.11.2 Hg in air, 253
 - 8.11.3 As, Cd and Ni, 254
- Questions, 255
- References, 256

9 CHEMICAL WEATHERING AND SOILS, 258

- 9.1 Primary minerals, mineral instability, chemical weathering mechanisms and reactions, soil-forming factors, and products of chemical weathering, 258
 - 9.1.1 Goldich stability sequence, 259
 - 9.1.2 Weathering rates, 260
 - 9.1.3 Chemical weathering, 261
 - 9.1.4 Consequences of chemical weathering: dissolved species and secondary minerals, 264
 - 9.1.5 Geochemical quantification of elemental mobility in soil, 265
 - 9.1.6 Quantifying chemical weathering: CIA, 267
 - 9.1.7 Soil profile, 268
 - 9.1.8 Soil-forming factors, 268
 - 9.1.9 Soil classification – soil orders and geochemical controls, 273
- 9.2 Secondary minerals, controls on their formation, and mineral stability diagrams, 275

- 9.2.1 Factors controlling soil mineralogy, 275
- 9.2.2 Mineral stability diagrams, 276
- 9.3 Soils and the geochemistry of paleoclimate analysis, 281
- 9.4 Effects of acid deposition on soils and aquatic ecosystems, 282
 - 9.4.1 Increased solubility of Al in acidic soil solution, 283
 - 9.4.2 Displacement of adsorbed nutrient cations, 284
 - 9.4.3 Leaching of base cations enhanced by increased NO_3 and SO_4 , 285
 - 9.4.4 Decrease of soil buffering capacity and base saturation, 286
 - 9.4.5 Acid deposition and heavy metals, 287
- 9.5 Soils and plant nutrients, 287
- 9.6 Saline and sodic soils, 289
- 9.7 Toxic metals and metalloids, 290
- 9.8 Organic soil pollutants and remediation (fuels, insecticides, solvents), 294
- Questions, 295
- References, 296

10 STABLE ISOTOPE GEOCHEMISTRY, 299

- 10.1 Stable isotopes – mass differences and the concept of fractionation, 299
- 10.2 Delta (δ) notation, 302
- 10.3 Fractionation: vibrational frequencies, temperature dependence, 304
 - 10.3.1 Stable isotopes and chemical bond strength, 305
 - 10.3.2 Temperature-dependent stable-isotope fractionation, 305

10.3.3	Equilibrium and non-equilibrium isotope fractionation, 307
10.4	$\delta^{18}\text{O}$ and δD , 309
10.4.1	Paleotemperature analysis using oxygen and hydrogen isotopes, 314
10.4.2	Oxygen and hydrogen isotopes as tracers in the hydrologic cycle, 314
10.4.3	Application of oxygen and hydrogen isotopes to paleosol climate records, 316
10.5	$\delta^{15}\text{N}$, 316
10.6	$\delta^{13}\text{C}$, 318
10.6.1	Carbon isotope analysis of paleoenvironment, 320
10.6.2	Carbon isotopes in hydrology and chemical weathering, 321
10.7	$\delta^{34}\text{S}$, 321
10.7.1	Fraction of sulfur isotopes, 322
10.8	Non-traditional stable isotopes, 324
10.8.1	$\delta^{65/63}\text{Cu}$, 325
10.8.2	$\delta^{56/54}\text{Fe}$, 326
10.8.3	$\delta^{202/198}\text{Hg}$, 326
10.8.4	$\delta^{26}\text{Mg}$ and $\delta^{44/42}\text{Ca}$, 328
10.8.5	$\delta^{37/35}\text{Cl}$, 330
10.9	Summary, 330
	Questions, 331
	References, 331

11 RADIOACTIVE AND RADIOGENIC ISOTOPE GEOCHEMISTRY, 335

11.1	Radioactive decay, 335
11.1.1	Decay mechanisms and products, 336
11.1.2	Half-lives, decay rates and decay constants, 337
11.2	Radionuclides as tracers in environmental geochemistry, 341

11.2.1	$^{206}\text{Pb}/^{207}\text{Pb}$, 341
11.2.2	$^{87}\text{Sr}/^{86}\text{Sr}$, 342
11.3	Radionuclides as environmental contaminants, 342
11.3.1	Controls on U, Th and their decay products, 342
11.3.2	Refined uranium ores and associated nuclear wastes, 346
11.3.3	Geological disposal of high-level radioactive wastes, 349
11.4	Geochronology, 350
11.4.1	^{14}C , cosmogenic radionuclides and earth-surface dating techniques, 350
11.4.2	Common radioactive decay methods of dating sediments and minerals, 359
11.4.3	$^{234}\text{U}/^{238}\text{U}$ and ^{234}U disequilibrium, 364
	Questions, 367
	References, 367

APPENDIX I CASE STUDY ON THE RELATIONSHIP OF VOLATILE ORGANIC COMPOUNDS (VOCs), MICROBIAL ACTIVITY, REDOX REACTIONS, REMEDIATION AND ARSENIC MOBILITY IN GROUNDWATER, 371

I.1	Site information, contaminant delineation, 371
I.2	Remediation efforts, 372
I.3	Sources of PCE and As, 374
I.4	Mobilization of arsenic, 374
	References, 377

APPENDIX II INSTRUMENTAL ANALYSIS, 378

II.1	Analysis of minerals and crystal chemistry, 378

II.1.1	Electron microscopy (SEM, TEM and many other acronyms), 378
II.1.2	X-ray diffraction, 379
II.1.3	FTIR, 382
II.1.4	Elements in solution by ICP-AES, ICP-MS, AAS, 384
II.1.5	XRF, 385
II.1.6	X-ray absorption spectroscopy (XAS) techniques (EXAFS, XANES), 385

II.1.7	Isotopic analysis: mass spectrometry, 387
II.1.8	Chromatography, 389
	References, 389

**APPENDIX III TABLE OF
THERMODYNAMIC DATA OF SELECTED
SPECIES AT 1 ATM AND 25 °C, 390**

	Index, 394

Acknowledgements

This text would not have come to fruition without the support of many, first and foremost Margaret Crowley Ryan for her encouragement and suggestions. Our daughters Emma and Sophie helped me to think through certain aspects of the text, and I also would like to thank my parents, Quinn and Jane Ryan, for their support from day one. My graduate advisors Bob Reynolds, Gray Thompson, Don Winston and Page Chamberlain challenged and encouraged me in my early years as a researcher and writer, and Jeff Olson encouraged me to take on this project in its earliest stages. Ian Francis, Kelvin Matthews, Delia Sanford and Alison Woodhouse at Wiley-Blackwell have been supportive and patient throughout.

Inspiration for different sections of this text come from many excellent geochemistry and environmental chemistry texts, particularly those of Berner and Berner, Drever, Faure, Langmuir, Moore and Reynolds, Spiro and Stigliani and Walther. They are referenced at

many points in the text and the well-worn copies on my shelf attest to their influence on my understanding of geochemistry. I am fortunate to be a member of a great cohort of colleagues at Middlebury College, both in the Geology Department and in the Environmental Studies Program. Students in GEOL 323 served as guinea pigs for early versions and made insightful comments — they as well as my colleagues are too numerous to mention but their support has been vital. I am grateful to Amy Larson Rhodes for thorough editing of some sections. Any time I got bogged down with writer's block, exile by the stones turned up to 11 pushed me forward, so thanks to Keith and Mick et al. I would also like to cite my long-running collaboration with Jon Kim of the Vermont Geological Survey as well as support from Javier Huertas, Rafa Bellver and colleagues at the Instituto Andaluz de Ciencias de la Tierra in Granada, Spain.

About the Companion Website

This book is accompanied by a companion website:
www.wiley.com/go/ryan/geochemistry

The website includes:

- Powerpoints of all figures from the book for downloading
- PDFs of tables from the book

Background and Basic Chemical Principles: Elements, Ions, Bonding, Reactions

1.1 AN OVERVIEW OF ENVIRONMENTAL GEOCHEMISTRY – HISTORY, SCOPE, QUESTIONS, APPROACHES, CHALLENGES FOR THE FUTURE

“The best way to have a good idea is to have a lot of ideas.” (Linus Pauling)

“All my life through, the new sights of nature made me rejoice like a child.” (Marie Curie)

Environmental geochemistry encompasses research at the intersection of geology, environmental studies, chemistry and biology, and at its most basic level, is designed to answer questions about the behavior of natural and anthropogenic substances at or near the surface of Earth. The scope includes topics as diverse as trace metal pollution, soil formation, acid rain and sequestration of atmospheric carbon, and most problems in environmental geochemistry require understanding of the relationships among aqueous solutions, geological processes, minerals, organic compounds, gases, thermodynamics, kinetics,

and microbial influences, to name a few. A good example is the fate and transport of lead in the environment. In some areas much of the lead in the Earth surface environment is (was) derived from combustion of leaded gasoline, because even after it had been banned, Pb tends to persist in soil. The original distribution of Pb was controlled at least in part by atmospheric processes ranging from advection to condensation and precipitation.

The fate and transport of Pb deposited on the land surface is controlled by the interactions and relationships among lead atoms, solids compounds (e.g. inorganic minerals or organic matter), potential for uptake into plants or other organisms, and the composition of water in soils, lakes or streams (including dissolved gases like O₂ and CO₂). In cases where Pb falls on soils bearing the carbonate anion (CO₃²⁻), the formation of lead carbonate (PbCO₃) can result in sequestration of lead in a solid state where it is largely unavailable for uptake by organisms. If the PbCO₃ is thermodynamically stable, the lead can remain sequestered (i.e. stored), but changes in chemical regime can destabilize carbonates. For example,

acidic precipitation that lowers the pH of soil can cause dissolution of PbCO_3 , but how much of the carbonate will dissolve? How rapidly? Much like the melting of ice at 10°C , geochemical processes are kinetically controlled (some more than others), so even in cases where phases exist out of equilibrium with their surroundings, we must know something about rate laws in order to predict how fast reactions (e.g. dissolution of PbCO_3) will occur.

If Pb is dissolved into an aqueous form, additional questions of fate and transport must be addressed – will the Pb remain in solution, thus facilitating its uptake by plants? Or will it be carried in solution into a nearby surface water body, where it could be consumed by a fish or amphibian? Or will other soil solids play a role in its fate? Will it become adsorbed to the surface of a silicate clay or organic matter, transported downstream until it ultimately desorbs in lake sediments? We also need to consider the possibility that the PbCO_3 does not dissolve, but rather is physically eroded into a stream or lake, where it might dissolve or remain a solid, possibly becoming consumed by a bottom feeder, from which point it could biomagnify up the food chain.

Environmental geochemistry has its origins in ground-breaking advances in chemistry and geology ushered in by the scientific breakthroughs of the late 19th and early 20th centuries, particularly advances in instrumental analysis. The Norwegian Victor M. Goldschmidt is considered by many to be the founder of geochemistry, a reputation earned by his pioneering studies of mineral structures and compositions by X-ray diffraction and optical spectrograph studies. These studies led Goldschmidt to recognize the importance and prevalence of isomorphous substitution in crystals, a process where ions of similar radii and charges can substitute for each other in crystal lattices.

Goldschmidt's peer, the Russian Vladimir I. Vernadsky, had come to realize that minerals form as the result of chemical reactions, and furthermore, that reactions at the Earth surface

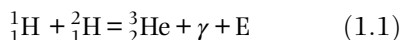
are strongly mediated by biological processes. The application of geochemistry applied to environmental analysis mainly arose in the 1960s and 1970s with growing concern about contamination of water, air and soil. The early 1960s saw publication of Rachel Carson's *Silent Spring*, and early research on acid rain at Hubbard Brook in New England (F.H. Bormann, G.E. Likens, N.M. Johnson, and R.S. Pierce) emphasized the interdisciplinary thinking required for problems that spanned atmospheric, hydrologic, soil, biotic and geologic realms.

Current research in environmental geochemistry encompasses problems ranging from micrometer-scale (e.g. interactions between minerals and bacteria or X-ray absorption analysis of trace metal speciation), local-scale (e.g. acid mine drainage, leaking fuel tanks, groundwater composition, behavior of minerals in nuclear waste repositories) to regional (acid rain, mercury deposition, dating of glacier retreat and advance) and global (climate change, ocean chemistry, ozone depletion). Modern environmental geochemistry employs analytical approaches ranging from field mapping and spatial analysis to spectrometry and diffraction, geochemistry of radioactive and stable isotopes, and analysis of organic compounds and toxic trace metals. While the explosion of activity in this field makes it impossible to present all developments and to acknowledge the research of all investigators, numerous published articles will be cited and highlighted in throughout the text, and a case study that integrates many concepts is presented in Appendix I.

1.2 THE NATURALLY OCCURRING ELEMENTS – ORIGINS AND ABUNDANCES

The chemical elements that comprise Earth were mainly produced during the Big Bang approximately 12–15 billion years ago. The early universe was extremely hot (billions of degrees)

and for the first few seconds was comprised only of matter in its most basic form, quarks. After about 15 s, the atomic building blocks known as neutrons, protons, electrons, positrons, photons and neutrinos began to form from quarks, and within moments after the Big Bang, the first actual atoms formed. Protons combined with neutrons and electrons to form hydrogen (${}^1_1\text{H}$) and its isotope deuterium (${}^2_1\text{H}$, or D), which rapidly began to form helium (He) through fusion, a process in which the nuclei of smaller atoms are joined to create larger, heavier atoms: $2\text{H} \rightarrow \text{He} + \text{energy}$, or to be more precise:



where γ is the symbol for gamma radiation emitted during nuclear fusion. (note: basic principles of atomic theory are presented in Section 1.3).

Small amounts of lithium were probably also produced in the first few minutes or hours after the Big Bang, also by fusion (in a simple sense represented as $\text{H} + \text{He} = \text{Li}$). From a graph of the abundances of elements in the solar system (Fig. 1.1), it is clear that H and He are the most abundant, and that, for the most part, element abundance decreases exponentially with increasing atomic number up to atomic number 50, and beyond that elements are quite rare.

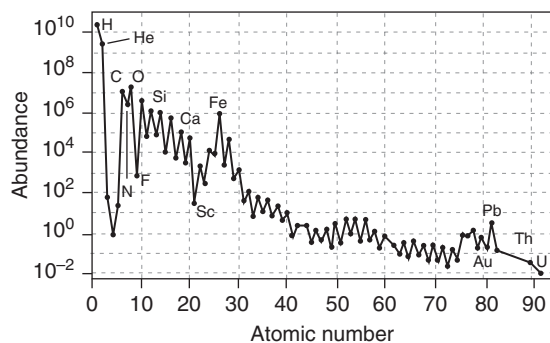
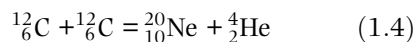
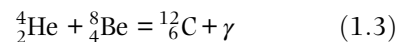
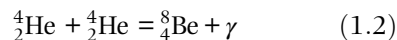


Fig. 1.1 Abundance of elements in the solar system normalized to $\text{Si} = 10^6$ on a logarithmic y -axis – this is a standard means of normalizing and plotting values for this type of data set.

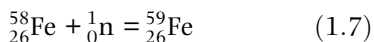
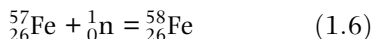
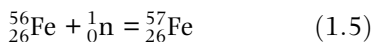
It is also clear that some elements (e.g. Li, Be, B, Sc) appear to be anomalously uncommon as compared to their neighbors, whereas others seem to be present in anomalously high concentrations (e.g. Fe, Ni, Pb). It is also interesting to note the sawtooth pattern produced by alternation of relatively abundant even-numbered elements as compared to neighboring odd-numbered elements (a phenomenon described by the Oddo–Harkins rule). How do these observed trends relate to the processes that formed the elements? The answer lies in basic principles of nuclear fusion.

Elements larger than He were formed by fusion in stars in the first few million to hundreds of millions of years after the Big Bang by a process generally referred to as **stellar nucleosynthesis**. Gravitational forces had produced a contracting, spinning disc-like mass of primordial H and He known as the solar nebula that contained the energy necessary to form heavier elements by fusion as follows:

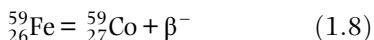


This process skipped over ${}^3_3\text{Li}$ and ${}^5_5\text{B}$ and the ${}^8_4\text{Be}$ that formed was very unstable and was either rapidly transformed to ${}^{12}_6\text{C}$ before it decayed (see middle reaction above) or was destroyed by radioactive decay.

Fusion was able to form elements up to iron (${}^{56}_{26}\text{Fe}$), but beyond Fe no heat is produced during fusion – that is, the process is no longer exothermic when fusing nuclei heavier than Fe. In fact, the iron nucleus is so stable that fusion reactions involving iron actually consume energy (it is an endothermic process), so without the heat needed to fuel fusion reactions, another process had to take over to form the heavier elements. This process is known as neutron capture and can be represented like this:



The Fe^{59} atom is unstable and undergoes spontaneous radioactive decay by beta emission as follows:



In this case, the negatively charged beta particle (sometimes written as ${}^0_{-1}e$ or e^-) effectively represents the transformation of a neutron (${}^1_0\text{n}$) to a proton (${}^1_1\text{p}$). Neutron capture combined with radioactive decay then formed progressively heavier elements up to the heaviest naturally occurring element, uranium. Some elements such as ${}^{56}\text{Ni}$ and ${}^{56}\text{Co}$ are unstable and undergo radioactive decay to form stable ${}^{56}\text{Fe}$, helping to explain the relative abundance of Fe as compared to elements with similar atomic number.

Neutron capture (equations 1.5–1.7) takes place via two main mechanisms. The **r-process** (“r” is for rapid) takes place in core-collapse supernovae, where there is a high flux of neutrons and extremely high temperatures (e.g. $> 10^9$ K), and nucleosynthesis involves a rapid series of neutron capture reactions starting (typically) with ${}^{56}\text{Fe}$ – the r-process explains the origin of $\sim 50\%$ of atoms heavier than Fe. The other main means by which heavy elements can be produced is known as the **s-process** (“s” is for secondary), in which nucleosynthesis occurs by means of slow neutron capture. The difference is that s-process neutron-capture nucleosynthesis occurs in asymptotic giant branch (AGB) stars, which have lower temperatures (e.g. 10^3 – 10^4 K) than supernovae, and thus the s-process requires pre-existing (hence the “secondary” nature) heavy isotopes that can function as seed nuclei.

Getting back to fusion, it is clear that progressive fusion reactions involving atoms with even

numbers of protons will lead to the sawtooth pattern in Fig. 1.1, but there is also another contributing factor to this pattern. The **Oddo–Harkins rule** states that atoms with an even number of protons in their nuclei are more stable than their odd-numbered counterparts. This is because, during nucleosynthesis, nuclei with an unpaired proton were more likely to capture an additional proton, producing a more stable proton arrangement in the nucleus of atoms with even numbers of protons. For additional information on nucleosynthesis and the origin of the elements, the reader is referred to the accessible and more-detailed presentation in Gunter Faure’s text *Principles and Applications of Geochemistry*.

As the universe continued to cool, galaxies and solar systems began to form. The solar nebula that was to form our solar system cooled and began to solidify into small masses known as chondrites and eventually larger masses known as planetesimals (on the order of tens of km in diameter). Those bodies closest to the early sun were more enriched in heavier elements (especially Si, Al, Mg, Fe, Ca, Na, K), in part because centrifugal forces effectively flung lighter elements (H, He) preferentially to the farther reaches of the solar system (other important influences include temperature, pressure, redox conditions and nebular density, but these factors that will not be covered here). The end result is that the inner planets are terrestrial and rocky (Mercury through Mars) and enriched in heavier elements, whereas the outer planets are gaseous (Jupiter and beyond) and enriched in lighter elements (think of the possibility of methane oceans or methane ice on Jupiter’s moon Titan).

Planets ultimately formed when gravitational forces caused accretion of planetesimals. The accretion of what was to become Earth produced heat that left the proto-planet in a molten or semi-molten state and allowed relatively dense Fe and nickel (Ni) to sink to the **core of the Earth**, whereas relatively light silicon (Si), aluminum (Al), magnesium (Mg), calcium (Ca), sodium (Na) and

potassium (K) floated to the top to form **Earth's crust**, leaving the Fe-Mg-Ni-Cr-Si **mantle** in between core and crust. While this is a broad generalization, the result is a differentiated Earth (Table 1.1), one where average continental crust (~ between 25–60 km thick) has a felsic composition much like that of granite, whereas oceanic crust (~ between 5 and 10 km thick) is compositionally mafic (or basic) and comprised mainly of basalt, rock that is less silica-rich and relatively enriched in Fe, Mg and Ca relative to continental crust. The mantle is comprised of rock types like

peridotite and dunite and has an ultramafic (or ultrabasic) composition.

One example of how compositional differentiation of the Earth is important in an environmental sense is related to soils formed by weathering of rocks derived from the mantle – these ultramafic soils tend to be depleted in plant nutrients such as Ca and K, and enriched in the trace metals Ni and Cr (and in some places, arsenic). Finally, the primordial atmosphere of Earth was comprised of CO₂, H₂O and N gases derived from volcanic eruptions. Oxygen in its form as O₂ gas is a relatively

Table 1.1 Chemical differentiation of Earth with major elements and selected trace elements. Major elements (SiO₂ through P₂O₅) are presented in units of wt% oxides and the trace elements are presented in concentrations of parts per million (ppm, or mg/kg). Data sources are as follows: granite is the United States Geological Survey granite standard “G-2”; basalt and ultramafic data are averages from Turekian and Wedepohl (1961) and Vinogradov (1962); sandstone and carbonate rock data are unpublished analyses of early Paleozoic sedimentary rocks from northwestern Vermont performed by the author; and shale is the North American Shale Composite (Gromet et al., 1984). Trace element values with asterisks are averages from Turekian and Wedepohl (1961) and Vinogradov (1962). Additional resource include the text *The Continental Crust: Its Composition and Evolution* by Taylor and McLennan (1985) and the chapter by Rudnick and Gao (2003) in *Treatise on Geochemistry*.

	Granite	Basalt	Ultramafic	Sandstone	Shale	Carbonate
SiO ₂	69.4	49.3	42.1	71.3	64.8	8.34
TiO ₂	0.48	1.86	0.05	0.70	0.80	0.12
Al ₂ O ₃	15.4	15.3	2.25	10.6	17.0	1.52
Fe ₂ O ₃ *	2.66	12.0	13.7	5.03	5.70	1.07
MnO	0.03	0.22	0.20	0.10	0.25	0.07
MgO	0.75	7.39	38.2	2.52	2.83	21.2
CaO	1.96	9.9	2.22	3.03	3.51	66.7
Na ₂ O	4.08	2.47	0.66	1.56	1.13	0.21
K ₂ O	4.48	0.98	0.02	4.61	3.97	0.55
P ₂ O ₅	0.14	0.25	0.04	0.16	0.15	0.07
SUM	99.4	99.7	99.4	99.7	100.1	99.8
As	0.25	2.2	0.8	1.0	28	1.6
Ba	1880	315	0.7	595	636	99
Cr	10*	185	1800	20	125	10
Co	4.6	47	175	2.3	20*	3.6
Cu	11	94	15	2.5	45*	1.3
Ni	10*	150	2000	10.1	58	3.3
Pb	30	7	0.5	8.3	20*	2.8
Th	13*	3.5	0.005	4.4	12.3	0.54
U	3*	0.75	0.002	1.3	2.7	1.6
Zn	86	120	40	50	95*	13

recent addition to Earth's atmosphere, having begun to accumulate slowly and in a step-wise manner in the atmosphere after the appearance of photosynthetic algae 3 billion years ago.

1.3 ATOMS AND ISOTOPES: A BRIEF REVIEW

A schematic sketch of a carbon atom, consisting of a central positively charged nucleus surrounded by a negatively charged "cloud" of electrons, is presented in Fig. 1.2. All atoms consist of a nucleus that contains positively charged **protons** and neutral **neutrons**, subatomic particles with a mass of 1 atomic mass unit (1 amu, or 1 Dalton [Da]).

The actual mass of a proton is 1.6726×10^{-24} g, so it is more convenient to say that the mass of a proton is one amu or Da. The +1 charge on a proton = 1.602×10^{-19} coulomb. The number of protons in the nucleus is what distinguishes atoms of one element from another – hydrogen has 1 proton, helium has 2, carbon has 6, uranium has 92 protons, and so on. If there is a nucleus with some other number of protons than

92, it is not uranium. The number of protons is commonly referred to as the **atomic number** (Z).

The mass of a neutron is effectively the same as the mass of a proton (1 amu; 1.6749×10^{-24} g), and sum of protons and neutrons is the **atomic mass** or **mass number** or **atomic weight** – all are essentially synonymous terms, (see back inside cover for atomic weights chart). Virtually all of the mass of an atom is contained in its nucleus. The carbon atom in Fig. 1.2 contains a nucleus with 6 protons and 6 neutrons, a configuration represented in this text and in many other places as $^{12}_6\text{C}$. So, while the number of neutrons is not explicitly given in this notation, it is implied, and in the case of the most common form of uranium, $^{238}_{92}\text{U}$, there must be 146 neutrons in the nucleus along with 92 protons to produce the atomic mass of 238.

The carbon atom in Fig. 1.2 contains 6 protons and likely also contains 6 neutrons, but it could contain 7 or 8 neutrons. Differences in the number of neutrons (6, 7 or 8) give rise to the three **isotopes** of carbon. Carbon-12, or ^{12}C , is the most abundant isotope of carbon and it contains 6 neutrons; ^{13}C contains 7 neutrons and it is much rarer than ^{12}C (Chapter 10). ^{12}C and ^{13}C are both stable

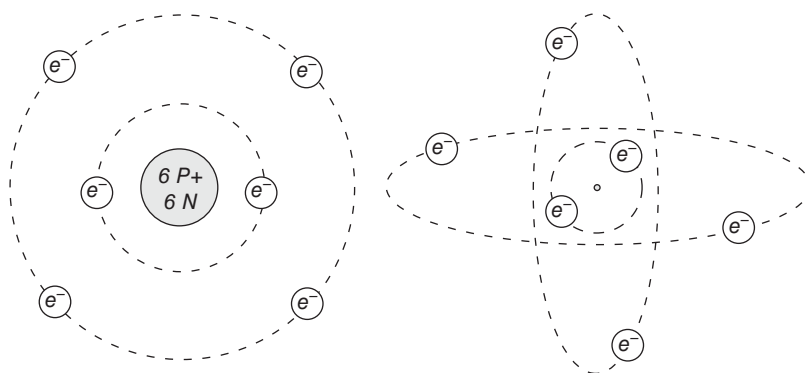


Fig. 1.2 Schematic sketches of Bohr models of a carbon atom (^{12}C) showing, on the left, 6 protons and 6 neutrons in the nucleus with 6 electrons in 2 separate orbitals, and on the right, an atom that attempts to show the actual size of the nucleus compared to the electron cloud, yet even here the nucleus is ~ 100 times larger than an actual nucleus. The example on the right also introduces the idea of separate orbitals in the outer electron "shell".

isotopes of carbon and the ratios of these two isotopes in plants, rocks, waters and sediments have proven very useful in environmental analysis. Carbon also has a **radioactive isotope**, ^{14}C , atoms of which are produced in the atmosphere in the presence of cosmic rays and ultimately undergo radioactive decay to nitrogen. This topic and many others in the field of isotope geochemistry are covered in Chapters 10 and 11.

Balancing the positive charge of the protons are **electrons** located in specified positions, or energy levels, outside of the nucleus. The charge on an electron is exactly opposite that of a proton – it is -1 , or -1.602×10^{-19} coulomb, and neutrally charged atoms contain equal numbers of protons and electrons. Furthermore, while the “electron cloud” surrounding the nucleus is of a virtually negligible mass, it occupies a volume that is orders of magnitude larger than the volume of the nucleus (atoms are mostly open space).

For the purpose of this text and virtually all environmental geochemistry, the main concern to investigators is the outermost shell of electrons (the **valence shell**), because that is the part of the atom most intimately involved in bonding and ionization. Neutral atoms with atomic number (Z) < 20 contain somewhere between 1 and 7 electrons in their valence (outermost) shell; the noble gases Ne and Ar contain 8 valence-shell electrons. The most stable valence shell electron configuration for elements with $Z < 20$ is one that consists of 8 electrons, so all atoms with $Z < 20$ seek to produce ions or form bonds that result in eight valence electrons. This is done either by losing, gaining or sharing electrons (Section 1.6). In elements with $Z > 20$, the presence of d and f orbitals makes the **octet** rule inapplicable; nonetheless, heavier elements (Fe^{+2} and Fe^{+3} ; Ni^{+2} ; U^{+4} and U^{+6}) lose electrons in predictable manners to form more-stable configurations, and the forms of these ions are commonly known and enable prediction of their behaviors.

One convenient way to depict valence electrons is with **Lewis electron dot diagrams** (Fig. 1.3).

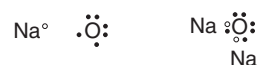


Fig. 1.3 Lewis electron dot diagrams showing valence electrons of sodium and oxygen in their ground states (left) and in the compound Na_2O (right), where each sodium has lost an electron to oxygen, resulting in two Na^+ and one O^{2-} .

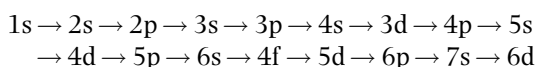
Sodium, with one valence electron, and oxygen, with six, can be depicted as such:

Sodium satisfies the octet rule by losing an electron, becoming Na^{+1} . The neutral oxygen atom with six valence electrons can most easily satisfy the octet rule by gaining two electrons, producing O^{2-} , the most common form of oxygen in nature.

The **Aufbau principle** describes how electron orbitals are populated in a systematic manner. In brief, electrons occupy orbitals of fixed energy levels and tend to occupy the lowest energy levels possible to create a stable atom. Orbitals are filled in a relatively predictable, sequential manner, starting with the lowest quantum number ($n = 1$), for which there is only one orbital (the s -orbital) which can be occupied by two electrons (of opposite spin).

For a neutral hydrogen atom with only one electron the symbol is $1s^1$. For neutral He with 2 electrons, the notation is $1s^2$. Quantum number 2 ($n = 2$) contains both s - and p -orbitals, where the s -orbital again can be occupied by one electron pair and the p -orbital can host up to 3 electron pairs. For neutral lithium with 3 electrons, the notation is $1s^2 2s^1$. For neutral oxygen, the notation is $1s^2 2s^2 2p^4$. (to satisfy the octet rule, oxygen gains two electrons, producing O^{2-} with an electron configuration like that of neon: $1s^2 2s^2 2p^6$). The third quantum level ($n = 3$) contains s -, p - and d -orbitals, where the s - and p -orbitals can contain 1 and 3 electron pairs, respectively, and the d -orbital can host 5 electron pairs ($10e^-$ s). However, the $3d$ orbital exists at a higher energy level than the $4s$ orbital, so the $4s$ orbital is filled before the $3d$.

The sequence in which these orbitals are filled is as follows:



Two examples of larger atoms are calcium and iron. The notation for calcium is:

$1s^2 2s^2 2p^6 3s^2 3p^6 4s^2$. Note that the second and third quantum levels contain eight electrons, but that the fourth contains only two. What must calcium do to satisfy the octet rule when forming an ion? It loses two electrons, taking on an electron configuration like that of argon ($1s^2 2s^2 2p^6 3s^2 3p^6$) and becoming Ca^{+2} . The notation for iron is:

$1s^2 2s^2 2p^6 3s^2 3p^6 3d^6 4s^2$ (26 electrons balance the 26 protons to produce a neutral Fe atoms). At this point the octet rule cannot be used to predict common oxidation states – iron tends to occur as Fe^{+2} or Fe^{+3} depending on redox conditions (more to come on redox later in this chapter).

1.4 MEASURING CONCENTRATIONS

Concentrations of elements or compounds are measured in a few common ways. Only in rare cases do scientists directly measure individual atoms/compounds, for even microscopic crystals of minerals at the micrometer scale such as hydroxides and silicate clays typically contain millions or billions of atoms.

1.4.1 Mass-based concentrations

One common means of measuring concentrations is by units such as milligrams per kilogram (mg/kg) for solids and milligrams per liter (mg/L) for liquids, or sometimes in units of micrograms ($\mu\text{g}/\text{kg}$ or $\mu\text{g}/\text{L}$) or nanograms (ng/kg or ng/L) for trace elements. Note that $\text{mg}/\text{L} = \mu\text{g}/\text{mL}$. Weight percent is a common mass-based

approach for expressing element concentration in cases where elements are in high concentrations (e.g. Table 1.1), a good example being Si, which comprises approximately 28% by weight (or mass) of the continental crust (or, as wt% oxide, approximately 59% of the crust – this conversion is presented below).

A few other useful facts to know about units of concentrations are:

For solids:

mg/kg is also known as parts per million (ppm), where mg is a milligram (10^{-3}g)

$\mu\text{g}/\text{kg}$ is also known as parts per billion (ppb), where μg is a microgram (10^{-6}g)

ng/kg is also known as parts per trillion (ppt), where ng is a nanogram (10^{-9}g)

For liquids:

mg/L is also known as parts per million (ppm)

$\mu\text{g}/\text{L}$ is also known as parts per billion (ppb)

ng/L is also known as parts per trillion (ppt)

Weight percent oxide is a common means of expressing major (abundant) elements in soils and rocks. Note in Table 1.1 that Si, Al, Fe and the other major elements are presented in units of SiO_2 , Al_2O_3 , Fe_2O_3 and so on. This is done partly by convention (or habit) and also because the major elements tend to occur in silicate minerals bonded to oxygen. The conversion factors for wt% element to wt% oxide for the common oxides are presented here in Table 1.2.

In the case of converting Ca to CaO, the conversion is determined as the molar mass of CaO divided by the molar mass of Ca, i.e. $56.08 \div 40.08 = 1.399$.

In the case of Al_2O_3 , the conversion factor is determined as the molar mass of Al_2O_3 divided by the mass of the equivalent amount of Al in the oxide, i.e. $\text{Al}_2\text{O}_3 \div (2 * \text{Al}) = 101.957 \div (2 * 26.98) = 1.889$. For Na to Na_2O , the conversion is the molar mass of Na_2O divided by the molar mass of $2 * \text{Na} = 61.98 \div (2 * 22.99) = 1.348$.

For elements that can occur in more than one oxidation state (e.g. Fe^{+2} , Fe^{+3}), values may be

Table 1.2 Conversion factors for wt% element to wt% oxide.

Element	C.F.	Oxide
Al	1.889	Al ₂ O ₃
Ca	1.399	CaO
Fe	1.286	FeO
Fe	1.430	Fe ₂ O ₃
K	1.205	K ₂ O
Mg	1.658	MgO
Mn	1.291	MnO
Na	1.348	Na ₂ O
P	2.291	P ₂ O ₅
Si	2.139	SiO ₂
Ti	1.668	TiO ₂

presented as either one of the oxidation states (i.e. either as FeO or Fe₂O₃), or as a combination of the two if the relative abundances of Fe⁺² and Fe⁺³ are known. In many cases where iron oxidation state is not known, all iron is presented in terms of Fe₂O₃.

However, mass-based measurements like µg/L or wt% are not always the best way to express concentrations. Consider for example groundwater with 98.7 µg/L nickel (⁵⁹Ni) and 98.7 µg/L of uranium (²³⁸U); i.e. both elements are present in equal concentrations of 98.7 ppb (evaporating the liter of water would leave 98.7 µg each of Ni and U). However, given that U atoms (238 g/mol) are ~4 times heavier than Ni (59 g/mol), there must be more Ni atoms – in fact, there are approximately four times as many Ni atoms as there are U atoms.

1.4.2 Molar concentrations

Quantifying concentrations on a molar basis has its roots in the work of Italian scientist Amadeo Avogadro, who in 1811 realized that equal volumes of gases at identical pressures and temperatures contain equal numbers of atoms (or molecules in the case of gases like N₂ and O₂),

even though their atomic masses differed. The term **mole** (abbreviated mol) describes the number of atoms of a given element required to form a mass equal to the atomic mass of the substance, in grams. For C, this mass is 12.011 grams. For U, this mass is 238.03 g, and so on. For all elements, the number of atoms required to form the atomic mass in grams is 6.0221×10^{23} atoms, a value known as **Avogadro's number**. 238.03 g of uranium (1 mol of U) contains 6.0221×10^{23} atoms; 4.002 g of helium (1 mol of He) contains 6.0221×10^{23} atoms.

The mole is a very useful concept in chemistry – most equations are expressed in terms of moles of reactants and products. In order to express Ni and U concentrations in terms of moles per liter (mol/L), or for trace elements like these, µmol/L, the mass concentration must be multiplied by the inverse of the molar mass (and a conversion for g to µg) as follows:

$$\text{For Ni: } 98.7 \mu\text{g/L} * 1 \text{ mol} / 58.693 \text{ g} * 1 \text{ g} / 10^6 \mu\text{g} \\ = 1.68 \times 10^{-6} \text{ mol/L}$$

$$\text{For U: } 98.7 \mu\text{g/L} * 1 \text{ mol} / 238.03 \text{ g} * 1 \text{ g} / 10^6 \mu\text{g} \\ = 0.415 \times 10^{-6} \text{ mol/L}$$

It is often helpful to express units in easy to communicate terms, so in this case mol/L would probably be converted to micromoles per liter (µmol/L), by multiplying mol/L by 10⁶ µmol/mol: $1.68 \times 10^{-6} \text{ mol/L} * 10^6 \mu\text{mol/mol} = 1.68 \mu\text{mol/L}$ of U in the groundwater

$$0.415 \times 10^{-6} \text{ mol/L} * 10^6 \mu\text{mol/mol} = 0.415 \mu\text{mol/L} \\ \text{of Ni in the groundwater}$$

In a solid (e.g. sediment or rock) with 1 mol/kg each of Fe (55.85 g/mol) and Al (26.98 g/mol): (1) there is an equal number of atoms of Fe and Al in the soil; and (2) Fe comprises a greater mass of the soil than Al (whether expressed as wt%, g/kg or mg/kg). Given that mass units are a common way of expressing concentration, it may be necessary to convert from molar units to mass units. A few algebraic calculations allow conversion from mol/kg to 3 common units, g/kg, mg/kg and wt%.

Given that 1 mol of Fe = 55.85 g:
 $55.85 \text{ g/mol} \times 1 \text{ mol/kg} = 55.85 \text{ g/kg Fe}$
 ... **or** $55.85 \text{ g/kg} \times 1000 \text{ mg/g} = 55850 \text{ mg/kg Fe}$
 ... **or** $55.85 \text{ g/kg} \times 0.1 = 5.855\% \text{ Fe (by weight)}$
 1 mol of Al = 27 g; $27 \text{ g/mol} \times 1 \text{ mol/kg} = 27 \text{ g/kg Al}$
 ... **or** $26.98 \text{ g/kg} \times 1000 \text{ mg/g} = 26980 \text{ mg/kg Al}$
 ... **or** $26.98 \text{ g/kg} \times 0.1 = 2.698\% \text{ Al (by weight)}$

1.4.3 Concentrations of gases

Atmospheric gas concentrations are typically expressed as the proportion of the total volume accounted for by a given gas. For example, the current atmospheric concentration of CO_2 is ~ 400 ppmv, indicating that 400 out of every one million molecules of gas in Earth's atmosphere is CO_2 . At the onset of the industrial revolution atmospheric CO_2 was 280 ppmv. Less-abundant gases are often expressed in terms of ppbv or pptv (parts per billion or trillion, volumetrically), and the major components of the atmosphere like the fixed gases N_2 , O_2 and Ar, are expressed in terms of percent (by vol): $\text{N}_2 = 78.1\%$, $\text{O}_2 = 20.9\%$ and $\text{Ar} = 0.9\%$ (the amounts vary depending on the amount of H_2O vapor in the air, which can range from 0 to 4% by volume). Expressed in this way, CO_2 comprises approximately 0.0400% percent of the atmosphere (vol%), but clearly units of ppmv are more useful for a gas like CO_2 . Gases also dissolve in liquids, and units of concentration in these cases are commonly mg/L or mmol/L.

1.4.4 Notes on precision and accuracy, significant figures and scientific notation

A few important topics related to data analysis and presentation of results are encompassed by the concepts of precision and accuracy. Simply

stated, **accuracy** describes how closely a measured value agrees with the actual value. The accuracy of chemical analyses can be tested by analyzing standards of known concentration. Consider a certified standard solution that contains 250 mg/L of aluminum (Al), and five analyses of this standard on your spectrometer produces results of 237, 271, 244, 262, and 240 mg/L. The mean of those five values is 251 mg/L – the average value is very close to the certified value of 250 mg/L. One way to express the accuracy of this test is as a percent difference from the certified value:

$$[(251 \text{ mg/L} - 250 \text{ mg/L}) \div 250 \text{ mg/L}] \times 100 = 0.4\%$$

However, the five results are somewhat lacking in **precision**, which is basically a measure of the reproducibility of results – how closely do measured results agree with each other? It is conceivable to produce results with a high degree of precision that are lacking in accuracy. For example, after recalibrating the spectrometer and re-analyzing the Al standard, values now are 277, 281, 274, 278 and 275 mg/L. The mean value of 277 mg/L is farther from the certified value of 250 mg/L (the difference from the certified value is 10.8%), but the results are definitely more precise.

One way to calculate uncertainty is to determine the standard deviation of the data. The formula for standard deviation and details on its appropriate use can be found in a statistics textbook, but using Microsoft Excel's standard deviation formula produces σ values as follows:

For the values 237, 271, 244, 262 and 240 mg/L,
 $\sigma = 14.9$

For the values 277, 281, 274, 278 and 275 mg/L,
 $\sigma = 2.74$

Which are better, analyses with greater accuracy but less precision (251 + 14.9 mg/L), or greater precision but less accuracy (277 + 2.74 mg/L)? Precise values are easier to correct because there is less uncertainty than if you have to deal with

accurate values plagued by low precision. This emphasizes one of the reasons why it is important to run standards when making analytical measurements. Of course, the ideal situation is to use a well-calibrated instrument for which accuracy and precision are both high, but regardless, it is imperative that researchers seek to quantify both parameters when making measurements.

In addition to quantifying uncertainty, it is very important to present numerical results in a manner that relates to the sensitivity of the measurement, or the degree of confidence associated with that measurement, while also trying to avoid propagation of error. Every measurement has some limited number of **significant digits** (or **significant figures**). Measurements of pH made using litmus paper can only be reported to 1 significant figure (e.g. pH = 6), whereas measurements made with well-calibrated probes may be reported to 3 significant figures (e.g. pH = 5.87).

As a rule, calculations should be carried out using all figures with each of their representative significant figures, and then the final result should be rounded at the end of the calculation. Using an example where the average concentration of uranium in groundwater in an aquifer is 14.3 µg/L, the aquifer volume is $1.241 \times 10^8 \text{ m}^3$ and average porosity (and thus % water in the saturated zone) is 18%, the resulting mass of U in the aquifer is:

$$14.3 \mu\text{g/L} * 1.241 \times 10^8 \text{ m}^3 * 1000 \text{ L/m}^3 * 0.018 \\ = 3.2 \times 10^{11} \mu\text{g U}$$

The final result is limited by the two significant figures in 18% (note that conversions involving liters to cubic meters, cm^3 to mL, etc. do not limit sig figs).

Also important to understand is the significance of zeros. Consider the following values: 7200, 700043, 0.0436, 0.043600. How many significant figures are reported for each, and why?

- 7200 is assumed to have 2 significant figures because any zero at the end of a number *and before a decimal point* is assumed to not be significant. How many significant figures does 847 000 possess? Three. 7200.0 contains 5 significant figures.
- 700 043 has 6 sig figs. Any zeros within a number are significant.
- 0.0436 has 3 sig figs because any zeros after a decimal point and before the first non-zero digit are not significant.
- 0.043600 has 5 significant figures because zeros that follow a non-zero digit after a decimal point are considered significant (as in 7200.0 example above).

Some of these examples serve to illustrate why scientists tend to express values in terms of **scientific notation**. Using scientific notation, the values above become:

$$7.2 \times 10^3 \text{ (2 sig figs)}$$

$$7.2000 \times 10^3 \text{ (if we needed to report 7200 to 5 sig figs, this is how it would look)}$$

$$7.00043 \times 10^5 \text{ (6 sig figs)}$$

$$4.36 \times 10^{-2} \text{ (3 sig figs)}$$

$$4.3600 \times 10^{-2} \text{ (5 sig figs)}$$

As a closing thought, one has to wonder about the precision of the road sign in southwestern Ecuador shown in Fig. 1.4. Being literal about significant figures, the distance to Engunga is precisely indicated as being between 14995 and 15005 m. (note: it is a very small town, so the precision may be justified).

1.5 PERIODIC TABLE

The Periodic Table of the Elements (front inside book cover) is one of the most useful tools to researchers and students of chemistry, geochemistry and biochemistry. First fully developed by the Russian chemist Dmitri Mendeleev in 1869 and refined ever since, it lists elements in order of atomic mass (from left to right in each row, and also from top to bottom in each column) and



Fig. 1.4 Road sign in southwestern Ecuador.

by similarities in chemical properties. Row 3 of the periodic table begins with sodium (atomic number $Z = 11$) and progresses to the right with increasing Z all the way up to Ar ($Z = 18$).

The atomic mass for each element represents a weighted average of the mass of the isotopes of that element, normalized to ^{12}C – observe that the atomic mass of sulfur (S, with $Z = 16$) is 32.065, reflecting weighted average of ^{32}S and heavier isotopes. The S atomic mass of 32.065 is derived from the abundances of the four stable S isotopes as is demonstrated in the following mass-balance calculation (data from Faure, 1986):

$$^{32}\text{S} = 95.02\% \quad ^{33}\text{S} = 0.75\% \quad ^{34}\text{S} = 4.21\% \\ ^{36}\text{S} = 0.02\%$$

$$^{32}\text{S}: 0.9502 * 31.97 = 30.377894$$

$$^{33}\text{S}: 0.0075 * 32.97 = 0.247275$$

$$^{34}\text{S}: 0.0421 * 33.97 = 1.430137$$

$$^{36}\text{S}: 0.0002 * 35.97 = \underline{0.007194}$$

$$\text{Sum} = 32.062500$$

Atomic masses are sometimes regarded as dimensionless numbers, but are also expressed in terms of amu (Daltons) and also in terms of gram molecular weights (i.e. the gram molecular weight, or mass of one gram of sulfur, is 32.065 g). All atomic masses are normalized to the mass of ^{12}C , and this is why the mass of ^{32}S is expressed as 31.97 g/mol.

Columns (or groups) of elements generally contain elements with similar valence electron

configurations, that is, the outermost shell of electrons tends to behave in a similar manner for elements of a given row. Good examples are column 1, the alkali metals, all of which lose one electron when they form chemical bonds, resulting in a series of +1 charged ions including Na^+ and K^+ ; column 2, the alkaline-earth metals that form divalent cations (e.g. Ca^{+2}); and column 8, the noble gases, including elements like argon and neon with complete valence shells (as a result, they do not form chemical bonds in nature). Elements in column 7, the halogens, tend to gain one electron when forming chemical bonds, resulting in halide anions with a -1 charge (e.g. Cl^-).

It may not come as a surprise that alkali metals tend to form bonds with halides, such as:



The point to understand here is that the periodic table presents information in a systematic way that can help to predict the behavior of elements in environmental systems. Like Na^+ , K^+ also can form bonds with chloride (Cl^-) to form a different salt, KCl (a substitute for NaCl in reduced-sodium diets). In fact, K^+ and Na^+ substitute for one another in many minerals. The periodic table also implies that arsenic (As) might substitute for phosphorous (P) in mineral structures, which it does.

In 2003, the environmental geochemist Bruce Railsback from the University of Georgia

developed an innovative new periodic table for geologists known as *An Earth Scientist's Periodic Table of the Elements and Their Ions*, one where elements are organized according to their occurrences in geological environments (<http://www.gly.uga.edu/railsback/PT.html>). This new formulation is designed to predict how elements and ions behave in the environment. Unlike the conventional periodic table originally envisioned by Mendeleev, the Earth scientist's periodic table organizes elements by charge, so it shows many elements multiple times because many elements have numerous valence states. It also contains abundant information on abundance of different

elements in soils, seawater, mantle vs. crust, ionic radii in crystals, and much more. For example, Fig. 1.5 shows the relative solubilities of oxides of various ions, and from this inset it is apparent that Al and Ti form insoluble oxides (corundum, Al_2O_3 ; rutile, TiO_2) but that Na, K, N and S do not (they are more likely to occur as soluble ions such as Na^+ , K^+ , NO_3^- and SO_4^{2-}).

The Earth scientist's periodic table also shows the numerous species of nitrogen that exist in natural systems (e.g. valence states and common molecules of N), predicted behaviors, attributes and affinities of the abundant and trace cations in natural systems, and more (Railsback, 2003).

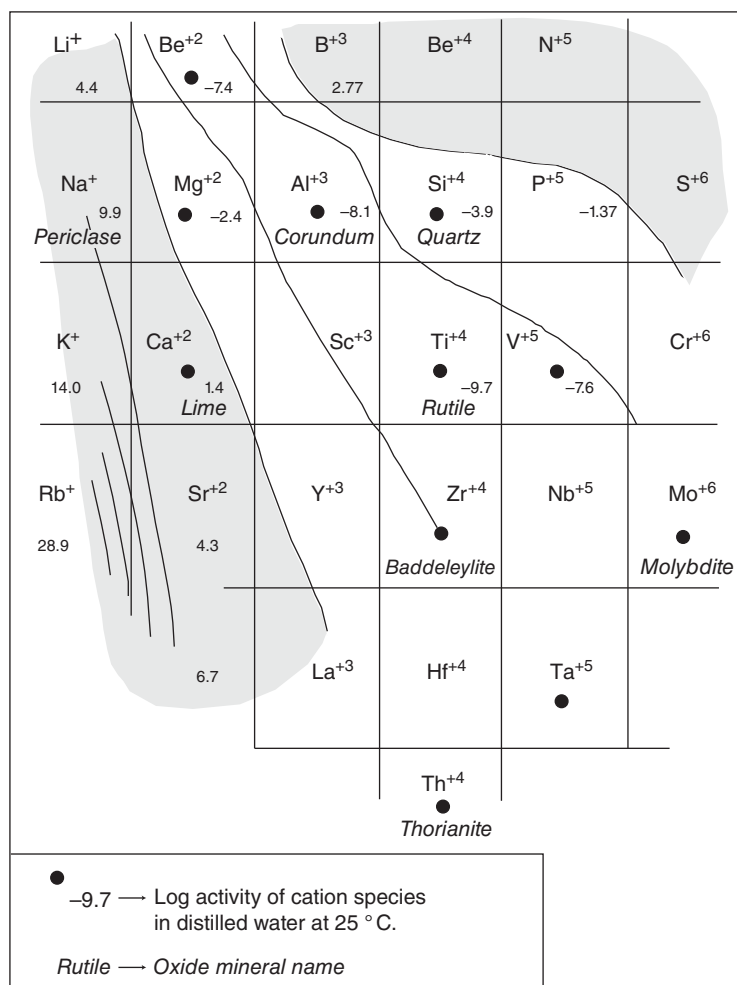


Fig. 1.5 Solubilities of ions as a function of ionic charge, from Inset 4 of *An Earth Scientist's Periodic Table of the Elements and Their Ions* (Railsback, 2003). Oxides of the elements are used as reference (e.g. lime, periclase, etc.). Note that low-charge base cations/alkali metals and cations of the alkaline Earth metals are relatively soluble, as are high-charge cations such as S^{+6} and N^{+5} , which form polyatomic anions (SO_4^{2-} , NO_3^-). Cations with +3 and +4 charges (e.g. Al^{+3} , Ti^{+4} , Zr^{+4}) are insoluble in most surficial environments, with activities in H_2O at 25 °C of $\sim 10^{-8}$ to 10^{-10} . (Railsback 2003.)

1.6 IONS, MOLECULES, VALENCE, BONDING, CHEMICAL REACTIONS

It is relatively rare to find elements in their native, neutral state in nature. A few examples include the diatomic gases hydrogen (H_2), oxygen (O_2) and nitrogen (N_2), solids like gold (Au) and graphite and diamond (both forms of C), and the noble gases (He , Ne , Ar , etc.). Most elements occur as ions, molecules or compounds in liquids (petroleum, alcohols, H_2O and the dissolved species it contains), solids (minerals, proteins, humus), and gases (CO_2 , CH_4 , NO_2). **Ions** are charged atoms, atoms that have either gained or lost one or more electrons from their neutral state. Examples of ions include the cation Na^{+1} and the anion S^{-2} . Most of the elements form cations, because metals generally lose electrons to form cations, and most of the elements are metals. A few common metal cations are Mg^{+2} , Al^{+3} , $\text{Cr}^{+3,+6}$, $\text{Fe}^{+2,+3}$ – note that some elements have more than one oxidation state. Generally only those elements in the upper right of the periodic table form anions (e.g. O^{-2} , N^{-3} , F^{-1} , etc.). **Oxidation state** is a term for the charge on an ion: “the oxidation state of iron in swamps is generally +2 (divalent), whereas in streams iron tends to occur in the +3 (trivalent) oxidation state”.

Molecules are formed when two or more atoms are joined by a chemical bond, and **compounds** are a specific type of molecule formed when two or more atoms of *different* elements are joined by chemical bonds. The gases H_2 , O_2 and N_2 are comprised of molecules but they are not compounds (they are diatomic gases). Examples of compounds include NaCl , H_2O and C_8H_{18} (octane). Most substances in nature, other than some gases and a few metals, are compounds.

There are two main types of **chemical bonds** responsible for forming molecules and compounds, ionic bonds and covalent bonds. These two bond types represent polar ends of the bonding spectrum, but it is useful to consider each type separately, and then examine intermediate cases.

We will also examine metallic bonds and Van de Waals bonds, but first will begin with ionic bonds.

Ionic bonding occurs between atoms of elements with very different valence electron configurations. Coulomb forces describe interactions between charged atoms or molecules (i.e. ions), and when the charges are opposite (i.e. involving cations and anions), the result is attraction (conversely, the Coulombic interaction between like-charged particles, e.g. 2 cations, causes repulsion).

The classic example of Coulombic attraction is the bond between an alkali metal cation (e.g. Na^{+}) and a halogen anion (e.g. Cl^{-}). The cation and anion are electrostatically attracted to each other and the result is formation of an ionic bond producing a solid, in this case cubic vitreous crystals of halite (NaCl).

Loss of an electron by sodium and its incorporation into the valence shell of chlorine can be viewed in the sense of a Lewis electron-dot diagram (Fig. 1.6):

The way this type of bond commonly forms in the natural world is when sodium and chloride ions are dissolved in water and ultimately become attracted to each other when their concentrations become sufficiently high to allow formation of solid crystals. (More details about the controls on aqueous processes are covered in detail in Chapters 4, 5, and 9).

1.6.1 Ionic bond strength

The strength of an ionic bond is largely controlled by the charges on the ions and by the ionic radii of the ions involved in the bonding – this is a concept that should be intuitively apparent because ions with higher (opposite) charges will be attracted

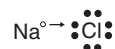


Fig. 1.6 Schematic representation of formation of an ionic bond by transfer of the Na valence electron to Cl .

more, and the closer the spacing of the ions, generally the stronger the bond.

The attraction of two ions can be quantified by Coulomb's Law:

$$F_c = k * (q_1 * q_2) / \epsilon * r \quad (1.10)$$

where k is a constant (described below), q_1 and q_2 are the values of ionic charges on the ions, ϵ is the dielectric constant and r is the distance between the nuclei of the two ions joined by the bond.

When F_c is negative the ions are attracted; the negative sign indicates that the system (the two ions) have shifted to a lower (more stable) energy state than is the case when the ions are separated. Given that all other terms (k , ϵ , r) are positive, negative F_c results from two ions with opposite charge, i.e. a cation and an anion.

The constant k is expressed as:

$$k = 1 / (4\pi\epsilon_0) \quad (1.11)$$

where ϵ_0 , the permittivity constant (sometimes known as P_0 or D), equals $8.854 \times 10^{-12} \text{ C}^2 \text{ J}^{-1} \text{ m}^{-1}$ and results in a value for the constant $k = 8.998 \times 10^9 \text{ J} \cdot \text{m} / \text{C}^2$, where $\text{J} = \text{Joules}$ and $\text{C} = 1.602 \times 10^{-19} \text{ Coulombs}$, a measure of charge.

The dielectric constant ϵ expresses the effect of the ambient environment on the strength of the bond, the best example being the difference in ionic bond strength in dry air as compared to water. At 20°C and 1 atm, the *dry air* value for $\epsilon = 1.0$, whereas the value for ϵ in *water* at $20^\circ\text{C} = 88$. Therefore, all other factors being equal, an ionic bond is 88 times weaker in water than in dry air.

For Na^+ and Cl^- ions, q^1 and q^2 are 1 and -1 . The ionic radius of Na^+ in a crystal of NaCl is $1.16 \text{ \AA}$ and the ionic radius for Cl^- is $1.67 \text{ \AA}$ (where an angstrom is 10^{-10} m). The reason that the Cl^- anion is larger than the Na^+ cation is that Cl^- contains an additional shell of electrons. (In general, anions have larger radii than cations because anions gain electrons when they form ions.)

The bond distance between nuclei will be:

$$\begin{aligned} &1.16 \times 10^{-10} \text{ m} + 1.67 \text{ \AA} \times 10^{-10} \text{ m} \\ &= 2.83 \times 10^{-10} \text{ m} \end{aligned}$$

so the attractive energy of the NaCl bond in air ($\epsilon = 1.0$) is:

$$\begin{aligned} F_c &= 8.998 \times 10^9 \text{ J} \cdot \text{m} / \text{C}^2 * (-1.602 \times 10^{-19}) \\ &\quad * (+1.602 \times 10^{-19}) \div 2.83 \times 10^{-10} \text{ m} \\ &= -8.15 \times 10^{-19} \text{ J} \end{aligned}$$

Note that units of charge in the numerator (C^2 , Coulombs squared) cancel with C^2 units from the constant, and also that m (meters, from r) in the denominator cancel with m in the numerator of k , leaving us with units of J (joules), which express energy lost (if $\text{J} < 0$) or gained (if $\text{J} > 0$), where negative values indicate transition to a more stable state.

This calculation applies to a single Na-Cl atom pair. If we wished to compute this for a mole we would multiply the result by Avogadro's number (one mole of NaCl would contain 6.022×10^{23} atoms each of Na and Cl):

$$\begin{aligned} &-8.15 \times 10^{-19} \text{ J/atom} * 6.022 \times 10^{23} \text{ atoms/mol} \\ &= -4.91 \times 10^5 \text{ J/mol} = -491 \text{ kJ/mol} \end{aligned}$$

If we compare this to the bond between K^+ (ionic radius = $1.52 \text{ \AA}$) and Cl^- , a bond that results in formation of KCl (sylvite or potassium chloride), we find that:

$$\begin{aligned} F_c &= 8.998 \times 10^9 \text{ J} \cdot \text{m} / \text{C}^2 * (-1.602 \times 10^{-19}) \\ &\quad * (+1.602 \times 10^{-19}) \div 3.19 \times 10^{-10} \\ m &= -7.23 \times 10^{-19} \text{ J} \end{aligned}$$

and

$$\begin{aligned} &-7.23 \times 10^{-19} \text{ J/atom} * 6.022 \times 10^{23} \text{ atoms/mol} \\ &= -4.35 \times 10^5 \text{ J/mol} = -435 \text{ kJ/mol} \end{aligned}$$

Considering the attraction of cation and anion, NaCl forms a stronger bond than KCl because the smaller ionic radius of Na allows the Na and Cl nuclei to be held closer.

In the case of magnesium chloride (MgCl_2), the ionic charge on Mg is +2, and its ionic radius in MgCl_2 is 0.86 Å. So in this case,

$$F_c = 8.998 \times 10^9 \text{ J} \cdot \text{m} / \text{C}^2 * (-1.602 \times 10^{-19}) \\ * (+2 * 1.602 \times 10^{-19}) \div 2.53 \times 10^{-10} \text{ m} \\ = -1.82 \times 10^{-18} \text{ J}$$

and

$$-1.82 \times 10^{-18} \text{ J/atom} * 6.022 \times 10^{23} \text{ atoms/mol} \\ = 1.10 \times 10^6 \text{ J/mol} = -1100 \text{ kJ/mol}$$

Based on Coulombic attraction, the higher charge and smaller radius of Mg^{+2} compared to K^+ and Na^+ make the MgCl_2 lattice energy approximately twice that of KCl and NaCl.

For ionic solids in water ($\epsilon = 88$ in the denominator), the lattice energies of NaCl, KCl and MgCl_2 are 5.58 kJ/mol, 4.95 kJ/mol and 12.5 kJ/mol, respectively, which is a quantitative way of saying that ionic bonds are approximately two orders of magnitude weaker in water than in dry air.

It is important to emphasize that the treatment presented above only considers the attraction between two ions and does not consider other factors associated with the strength of ionic bonds, two of which include (1) the effect of other ions and the geometry of the lattice structure (this can be assessed using the Madelung constant, which is not covered here), and (2) temperature; for example, the melting points for NaCl, KCl and MgCl_2 , respectively, are 801, 770 and 714 °C, differences which are not explained by Coulomb's Law.

1.6.2 Covalent bonds

Covalent bonds involve overlap of electron orbitals, i.e. they are bonds that form as a result of sharing of electrons between atoms. Unlike ionic bonding, where atoms have very different attractions to valence electrons, atoms that form covalent

bonds have similar valence electron configurations, and because their valence electrons are attracted to their nuclei with similar strength, it is impossible for one atom to lose its valence electron(s) to another atom to form a bond.

One classic example of covalent bonding involves the attraction involved with formation of the diatomic gases H_2 , N_2 , O_2 . When two oxygen atoms combine to form O_2 , there is clearly no difference in the valence electron configuration or in the attraction of each oxygen atom for its electrons. The atoms are effectively equal in structure, so the solution to the octet rule for O_2 involves reorganization of electrons and overlapping to produce valence shells that each contain 8 electrons (viewed as an electron dot diagram in Fig. 1.7).

Both O atoms have satisfied the octet rule the same way, by overlapping electron orbitals to produce a stable configuration. So, while each O atom has six valence electrons, two of them are shared with the adjacent O atom by orbital overlap to produce a strong chemical bond. In this case the bond between the oxygen atoms is a **double bond** because it involves 2 pairs of electrons, as contrasted to a typical **single bond** that involves only one pair of electrons (like H-O bonds in water molecules).

Covalent bonds also occur between elements with similar electronegativities (Fig. 1.8), including H and O in H_2O , C and O in CO_2 , and Si and O in SiO_2 . Chemical bonds that are a mix of ionic and covalent (e.g. SiO_2) are termed polar covalent bonds.

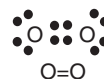


Fig. 1.7 Lewis electron dot diagram of an oxygen molecule (O_2) shown using two different notations. Note double bond between oxygen atoms consisting of two electron pairs depicted by the double dashed lines in the lower example.

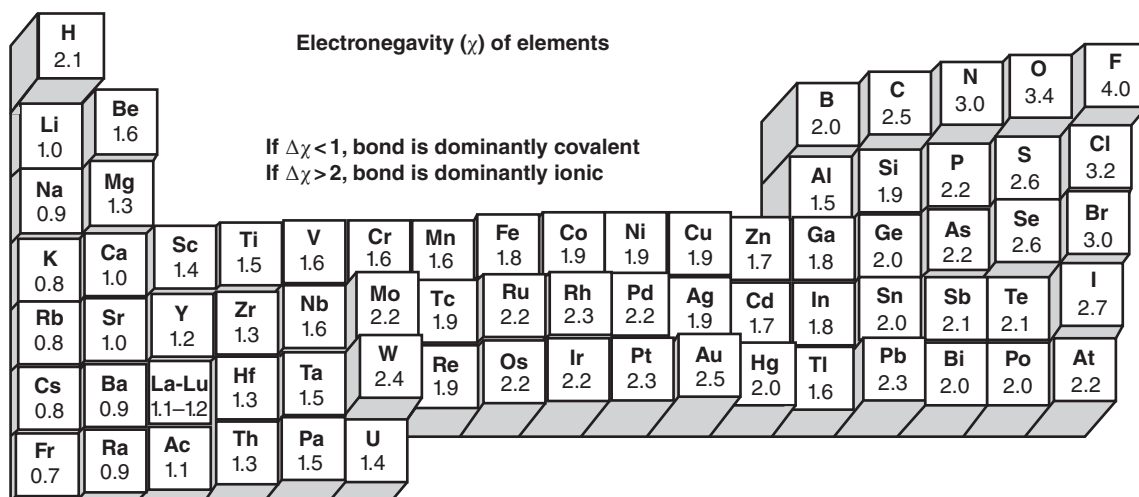


Fig. 1.8 Electronegativity of elements based on Linus Pauling's early 20th-century research.

1.6.3 Electronegativity

Electronegativity is a measure of the attraction of a nucleus for its valence electrons. Na and K have low electronegativities because they readily lose their valence electron to satisfy the octet rule, resulting in +1-charged cations. Conversely, fluorine and chlorine have strong attractions to the electrons in their valence shell and furthermore have the ability to pull valence electrons away from nearby atoms to form -1-charged anions and satisfy the octet rule – these elements have high electronegativity values. Electronegativity is important because it helps to predict the type of chemical bond that will form between elements – elements with large differences in electronegativity form dominantly ionic bonds, whereas elements with low differences in electronegativity form covalent or polar covalent bonds.

Figure 1.8 presents electronegativities of the elements and Fig. 1.9 shows the % ionic character of some selected chemical bonds expressed by difference in electronegativity. Nearly all bonds involving atoms of two elements (i.e. not including diatomic gases like H₂ or O₂ native elements such as Au or Ag) involve some aspect of electron transfer (ionic character) and orbital overlap

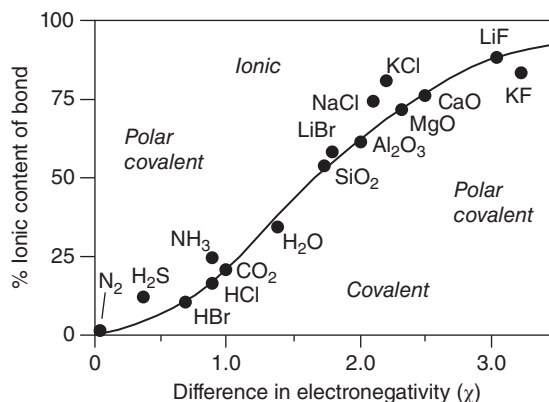


Fig. 1.9 Plot of relationship between difference in electronegativity of two atoms in a bond and the percent ionic character of the bond. Note that no bond is 100% ionic; however, diatomic gases (e.g. N₂, Cl₂, O₂) and the C-C bonds in diamond are 100% covalent. Based on early and middle 20th-century work of Linus Pauling and more recent research by Lu et al. (2006).

(covalent character) – the greater the difference in electronegativity, the greater the ionic character (and usually, the greater the tendency to dissociate and dissolve in water).

The electronegativity difference ($\Delta\chi$) for NaCl is 3.0 (Na) – 0.9 (Cl) = 2.1, resulting in a predominantly ionically bonded compound. For a

purely covalent molecule like N_2 , O_2 , Cl_2 or diamond (pure C), $\Delta\chi$ is zero. Methane (CH_4) and hydrogen sulfide (H_2S) are strongly covalent ($\Delta\chi = 0.4$ and 0.5 , respectively). SiO_2 is a mix of covalent and ionic (Fig. 1.9) and the term **polar covalent bond** (or **polar bond**) is used to describe this type.

CaCO_3 (calcium carbonate, the mineral calcite) contains bonds between Ca–O and C–O. The Ca–O bonds are predominantly ($\sim 70\%$) ionic ($\Delta\chi = 3.5 - 1.0 = 2.5$), whereas the C–O bond is predominantly ($\sim 80\%$) covalent ($\Delta\chi = 3.5 - 2.5 = 1.0$). Understanding this concept is important when predicting solubilities of minerals; for example, in water, calcite dissolves to produce Ca^{+2} ions and the polyatomic CO_3^{-2} anion in solution (see Chapter 5 for details of carbonate geochemistry). The covalent character preserves the C–O bond and as a result the carbonate anion (CO_3^{-2}) is a common constituent of natural waters.

Consider also the examples of NaCl (very ionic) and SiO_2 (which contains polar covalent bonds). Ionically bonded compounds like halite are highly soluble in water, but quartz is very insoluble, which helps to explain the common occurrence of white quartz sand beaches – if the covalent bonds in quartz were readily destroyed by water we would have no sandy beaches. In fact, we probably wouldn't have mountains or canyons either, for the strong bonds in rock-forming minerals are primarily polar covalent. In the absence of water, ionic bonds and covalent bonds are generally both strong, and appreciably stronger than the other types of chemical bonds in nature such as metallic bonds and dipole bonds (e.g. hydrogen bonds, Van der Waals bonds).

1.6.4 Metallic bonds, hydrogen bonds and van der Waals forces

Metallic bonds occur among metals such as Cr, Cu, Fe, Ni and Zn in solid compounds, but unlike ionic or covalent bonds, once valence electrons are released by a metal atom, they are not fixed

with a specific atom, but rather migrate through the crystal structure. This type of bond occurs in sulfide minerals such as pyrite, and also in native elements such as Cu, gold (Au) and silver (Ag). These bonds are weaker than covalent and ionic bonds and are part of the reason why metal-bearing sulfide minerals are relatively unstable at the Earth surface.

Dipole bonds, also known as **van der Waals bonds** (and sometimes called van der Waals forces), exist between electrically neutral molecules or compounds with some unequal distribution of charge. A great example of a dipole bond exists between water molecules. Water is a dipolar compound, with a positively charged pole and a negatively charged pole (remember that the molecule as a whole is neutral). Fig. 1.10 presents a schematic sketch of three adjacent water molecules attracted by dipolar bonds – in this case, the bond occurs between hydrogen atoms at the positively charged pole (δ^+) of water molecules and the negatively charged pole (δ^-) produced by valence electrons from the highly electronegative oxygen atom in the adjacent water molecule (Fig. 1.10).

This specific type of dipolar bond is known as a **hydrogen bond**, and although shown between adjacent H_2O molecules in this case, hydrogen bonds can also occur between H and O (or N) within a molecule. Note that the dots surrounding the O atoms represents electrons – oxygen satisfies the octet rule by sharing an electron with an electron from each of the H atoms, and the

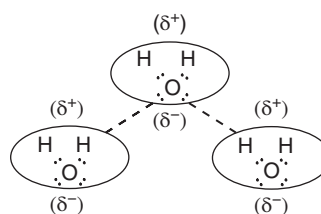


Fig. 1.10 Diagram of three water molecules showing polarity of the water molecule (δ^+ , δ^-) as well as hydrogen bonds (dashed lines) between H and O in adjacent molecules.

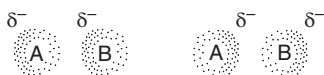


Fig. 1.11 Van der Waals bonds between adjacent atoms. A and B represent nuclei of adjacent atoms and dots are electrons. In the example on the left, electron clouds are shifted to the left, causing attraction of the atoms. In the example on the right, electron clouds are shifted to the right, also enhancing attraction of the atoms. If this type of alternating motion is synchronized, the atoms will be weakly attracted.

H–O bonds within the H_2O molecules (within the ovals) are single bonds (dominantly covalent) involving one pair of electrons each.

Water has a permanent dipole, but **van der Waals** bonds also exist between electrically neutral molecules or compounds where the electrostatic attractions are temporary, commonly existing as transitory states involving alternating positive and negative charge distributions that minimizes repulsion. The example in Fig. 1.11 represents two different transitory states of adjacent atoms where the dots are a schematic representation of an electron cloud. The delta symbol (δ^-) represents the side of the atom or molecule with the greater concentration of negative charge (Fig. 1.11):

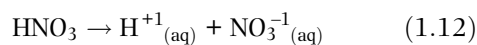
While only schematic, this diagram indicates how alternating electron distributions can lead to attraction between adjacent atoms or nonpolar compounds. Both A and B are transitory states, and in fact, both are unstable states – as soon as the electrons take on one configuration they are driven in the opposite direction by repulsive forces, causing the other configuration. This then drives the electron cloud back towards configuration A, and this alternation of transitory states facilitates atomic or molecular attraction. Although these bonds are far weaker than covalent, ionic or metallic bonds, they are important to understand because they control the melting and boiling points of many compounds, particularly nonpolar organic compounds including pesticides, fuels and solvents.

1.7 ACID-BASE EQUILIBRIA, pH, K VALUES

Chemical reactions in natural systems commonly occur in the presence of water, and understanding acid–base chemistry is crucial to understanding a large proportion of issues in geochemistry, including the solubility of minerals and trace metals, chemical weathering, the decomposition of organic matter, speciation of chemical elements and reactions in the atmosphere.

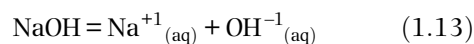
Following the convention established by the Swedish chemist Svante Arrhenius, an **acid** is a compound that releases hydrogen ions (H^{+1}) in solution (i.e. when dissolved in water). A few examples of classic acids are the inorganic acids hydrochloric acid (HCl), sulfuric acid (H_2SO_4) and nitric acid (HNO_3), and organic acids such as formic acid (HCOOH , also known as methanoic acid, it is what ants sting with) and acetic acid (HCH_2OOH , also known as ethanoic acid, or more commonly as vinegar).

Using nitric acid as an example, acids behave in solution as follows:



In water, HNO_3 dissociates to release hydrogen ions (and an equal amount of nitrate anions) into solution. The subscript (aq) is used here to indicate that the ions are dissolved in water (i.e. are aqueous). The hydrogen ion actually does not exist alone in aqueous solutions, but rather bonds with water molecules and exists as the hydronium ion (H_3O^{+}), but it is commonly represented in chemical equations as H^{+} .

The Arrhenius definition of a **base** is a compound that, when dissolved in water, releases hydroxyl anions (OH^{-1}) in solution. A few classic examples of bases are sodium hydroxide (NaOH), calcium hydroxide ($\text{Ca}[\text{OH}]_2$) and ammonium hydroxide (NH_4OH). In solutions, bases behave as follows:



Strong acids and bases dissociate completely, or nearly so. A strong base like NaOH may almost completely dissolve in solution, producing a high concentration of hydroxyl anions and resulting in a very basic solution. A weak base like NH_4OH (ammonium hydroxide) is much less soluble, so in solution it produces a much lower concentration of OH^- . Similarly, the mineral acids listed above are strong acids, while the organic acids (formic and acetic acid) are relatively weak and produce much lower concentrations of H^+ in solution.

The pH scale is the conventional means of expressing the acidity or alkalinity of a solution. It is defined by the concentration¹ of H^+ ions in solution according to the equation:

$$\text{pH} = -\log[\text{H}^+] \quad (1.14)$$

where $[\text{H}^+]$ is the concentration of hydrogen ions in solution. In a highly acidic solution with 10^{-2} mol/L of H^+ , the pH = 2. In an alkaline solution with 10^{-11} mol/L of H^+ , the pH = 11. In aqueous solutions, the product of the concentrations of H^+ and OH^- is 10^{-14} ($[\text{H}^+] * [\text{OH}^-] = 10^{-14}$). What this translates to is that if the concentration of $\text{H}^+ = 10^{-4}$, the concentration of $\text{OH}^- = 10^{-10}$. In this case, pH = 4 and qualitatively, it makes sense that the solution is acidic because there is far more H^+ than OH^- . In an alkaline solution with $[\text{OH}^-] = 10^{-2}$, the concentration of $[\text{H}^+] = 10^{-12}$ and the pH = 12.

Consider now the dissociation of two acids, one a strong acid (nitric acid, HNO_3) and one a relatively weak acid (carbonic acid, H_2CO_3).



¹ Note: this section discusses aqueous species (e.g. H^+ , NO_3^-) in terms of concentration (e.g. $[\text{NO}_3^-]$), yet often species in solution are represented by activity (e.g. $a \text{NO}_3^-$), a term that takes into account effects of other ions in solution on the effective concentration (often activities are less than actual concentrations). In this introductory section and in many other resources, concentrations are used; in Chapter 4, the concept of activity is introduced.

The dissociation of nitric acid can be represented according to the **Law of Mass Action** as follows:

$$K_{\text{aHNO}_3} = [\text{H}^+] * [\text{NO}_3^-] / [\text{HNO}_3] = 2.4 \times 10^1 \quad (1.17)$$

The concentrations of H^+ and NO_3^- are equal when HNO_3 dissolves and if we assume a HNO_3 concentration of 1, $[\text{H}^+] = [\text{NO}_3^-] = \sqrt{24} = 4.9$.

The same treatment for carbonic acid produces:

$$K_{\text{aH}_2\text{CO}_3} = [\text{H}^+] * [\text{HCO}_3^-] / [\text{H}_2\text{CO}_3] = 10^{-6.37} \quad (1.18)$$

If, as was done with HNO_3 , we assume an H_2CO_3 concentration of 1, $[\text{H}^+] = [\text{HCO}_3^-] = \sqrt{10^{-6.37}} = 6.5 \times 10^{-4}$. In other words, the concentration of H^+ produced by nitric acid in water is approximately ten thousand (10^4) times greater than H^+ produced by an equivalent amount of carbonic acid in water.

K_{a} values for some relatively common acids are presented in Table 1.3.

Two additional definitions of acids and bases include Brønsted and Lewis classifications.

A **Brønsted acid** is a substance that can donate a proton (i.e. H^+) to another substance, and a **Brønsted base** is a substance that can accept a proton from another substance. The following chemical reaction illustrates this relationship:



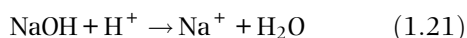
where HCl (the Brønsted acid) donates a proton that is accepted by the Brønsted base NH_3 (ammonia). The result is formation of a chloride anion and ammonium, where Cl^- can be termed a Brønsted base (it can accept a proton) and NH_4^+ is considered a Brønsted acid. Most **minerals** can be viewed as Brønsted bases because they consume H^+ when they undergo chemical weathering in soils. A good example is the weathering of gibbsite, $\text{Al}(\text{OH})_3$, in acidic soils:

Table 1.3 Selected acids and their K_a values, pK_a values and conjugate bases.

Acid				Conjugate Base	
Name	Formula	K_a	pK_a	Formula	Name
Hydroiodic acid	HI	3.2×10^9	9.5	I^-	Iodide
Hydrobromic acid	HBr	1.0×10^9	9.0	Br^-	Bromide
Hydrochloric acid	HCl	1.3×10^6	5.1	Cl^-	Chloride
Sulfuric acid	H_2SO_4	1.0×10^3	3.0	HSO_4^-	Hydrogen sulfate anion
Nitric acid	HNO_3	2.4×10^1	1.4	NO_3^-	Nitrate
Hydrogen sulfate ion	HSO_4^-	1.0×10^{-2}	-2.0	SO_4^{2-}	Sulfate anion
Sulfurous acid	H_2SO_3	1.3×10^{-2}	-1.9	HSO_3^-	Hydrogen sulfite anion
Phosphoric acid	H_3PO_4	7.1×10^{-3}	-2.2	$H_2PO_4^-$	Dihydrogen phosphate anion
Dihydrogen phosphate ion	$H_2PO_4^-$	6.3×10^{-8}	-7.2	HPO_4^{2-}	Hydrogen phosphate anion
Hydrogen phosphate ion	HPO_4^{2-}	4.2×10^{-13}		PO_4^{3-}	Phosphate anion
Nitrous acid	HNO_2	7.2×10^{-4}	-3.1	NO_2^-	Nitrite anion
Hydrofluoric acid	HF	6.8×10^{-4}	-3.2	F^-	Fluoride
Methanoic (formic) acid	HCOOH	1.8×10^{-4}	-3.7	$HCOO^-$	Methanoate (formate) anion
Benzoic acid	C_6H_5COOH	6.3×10^{-5}	-4.2	$C_6H_5COO^-$	Benzoate anion
Ethanoic acid	CH_3COOH	1.8×10^{-5}	-4.7	CH_3COO^-	Ethanoate (acetate) anion
Carbonic acid	H_2CO_3	4.4×10^{-7}	-6.4	HCO_3^-	Bicarbonate anion
Hydrogen carbonate ion	HCO_3^-	4.7×10^{-11}	-10.3	CO_3^{2-}	Carbonate anion
Hydrogen sulfide	H_2S	1.1×10^{-7}	-7.0	HS^-	Hydrogen sulfide anion
Ammonium ion	NH_4^+	5.8×10^{-10}	-9.2	NH_3	Ammonia



Consider also that NaOH, described as an Arrhenius base above because it yields dissolved OH^- in solution, is also considered a base by the Brønsted definition:



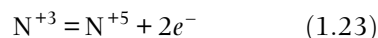
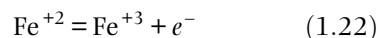
Lewis acids are substances that can accept electron pairs when forming bonds (H^+ is a good example), whereas **Lewis bases** are electron-pair donors (OH^- is a good example of this type of substance).

1.8 FUNDAMENTALS OF REDOX CHEMISTRY AND CHEMICAL REACTIONS

Reduction–oxidation (redox) chemistry refers to processes that take place when atoms gain or lose electrons, and often involve reactions where oxygen is transferred. Electron transfer facilitates exchange

of energy that is crucial to processes across the chemical spectrum, from aquifer and soil dynamics to photosynthesis and degradation of toxic organic compounds. In nature, redox reactions often involve changes to the oxidation state of elements like carbon, nitrogen, oxygen, sulfur, manganese and iron that can exist in different oxidation states (e.g. carbon exists in many oxidation states, including C^{-4} , C^0 , C^{+2} , C^{+4}), where the change from one oxidation state to another involves gain or loss of electrons.

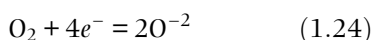
Oxidation refers to the loss of electrons by an atom. Two common examples are the oxidation of iron from its ferrous state (Fe^{+2}) to its ferric state (Fe^{+3}) by loss of one electron, and of nitrogen from N^{+3} to N^{+5} by loss of two electrons. These two reactions can be represented as follows:



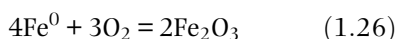
Where do those liberated electrons go? They probably were pulled away from the oxidized atom

because a neighboring atom had a greater attraction for those valence electrons. The oxidation of one atom cannot occur without a corresponding change to another atom. This change is **reduction** and it takes place when an atom gains electrons.

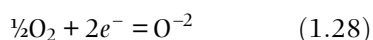
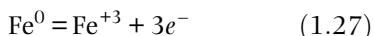
Two common examples are the reduction of oxygen gas (where oxygen is O^0) to oxygen anions (O^{-2}), and of C^{+4} (e.g. the C in CO_2) to molecular carbon, C^0 (e.g. the C in some organic matter).



While it is useful to examine individual examples of reduction or oxidation, loss of electrons from one atom results in gain of electrons for another atom. A simple example that you can probably visualize is the oxidation of iron metal (Fe^0) to iron oxide (Fe_2O_3) where Fe occurs in its trivalent or ferric state (Fe^{+3}).

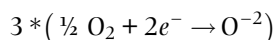
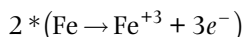


It helps to examine individual reduction and oxidation pairs in order to see where and how the exchange of electrons takes place:



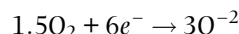
All elements in their pure state, like the Fe atom and O atom (represented as $\frac{1}{2}O_2$) shown above, have an oxidation state of zero.

The oxidation of iron metal (Fe^0) by oxygen gas (O_2) involves the loss of three electrons from a neutral iron atom and gain of two electrons by a neutral oxygen atom. Clearly, one oxygen atom cannot cause the oxidation of one iron atom from Fe^0 to Fe^{+3} , and this brings up the need to balance redox reactions, as follows:

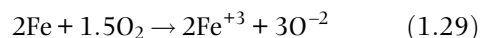


This results in a balanced pair of reactions (with respect to electrons) where two iron atoms lose

six electrons and three oxygen atoms gain 6 electrons:



and the paired redox reaction can be expressed as:



The electrons are excluded from this redox reaction because there are $6e^-$ on both sides (products and reactants) and thus cancel each other out, yet **electron flux is implied by the change in oxidation states of Fe and O atoms**. Terminologically, iron is the reducing agent that donates electrons to oxygen, causing oxygen to become reduced (and iron to be oxidized); oxygen is the oxidizing agent that pulls electrons from iron, which results in reduction of oxygen and oxidation of iron. In reality, the oxidation of iron ends up producing iron oxide, shown here as the mineral hematite:



Hematite consists of Fe in its most oxidized state (Fe^{+3} , ferric iron) and oxygen in the form that it takes in virtually all compounds, O^{-2} .

The terms “oxidizing” and “oxidized” are effectively synonymous. An *oxidized* soil will likely contain abundant available O_2 as well as minerals that are stable in an oxidized (or oxidizing) environment, e.g. iron oxides. This soil will be *oxidizing* because if an organic compound or sulfide were to be transported into the soil, an oxidation reaction would likely lead to destruction of the *reduced* substance. Similarly, *reduced* and *reducing* environments are effectively synonymous.

Redox chemistry comprises some of the most important reactions in the realm of geochemistry and biochemistry. Microbial activity often plays an important role in redox chemistry because electron transfer is an energy source – a classic example is the microbially mediated decomposition of fuels and solvents in soils, where the oxidation of organic carbon provides energy to the microbe and results in the transformation of leaked fuel into H_2O and CO_2 . More-detailed information on redox reactions is presented in Chapter 4 on aqueous geochemistry.

1.9 CHEMICAL REACTIONS

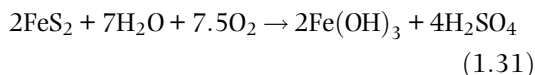
Reactions among elements and compounds have been presented in a few ways in this introduction thus far, including reference to chemical bonding, the formation of elements and redox chemistry. Chemical equations are algebraic expressions that represent quantities and charges of constituents involved in chemical reactions, and in some ways are the language of geochemistry, or at least are one of the languages. Accordingly, the following section will present a few fundamental concepts about chemical reactions, what they represent, how to balance them and how to interpret them.

First, a few general rules. Chemical reactions commonly take place in the presence of water, but if water is not produced or is not consumed by the reaction, it is not listed in the reaction. Given that the 1st Law of Thermodynamics states that matter can neither be created nor destroyed, but rather can only change forms, chemical reactions should not give the illusion that the 1st law is being violated – what this means is that all chemical reactions must be balanced. If there are four oxygen *atoms* expressed on the reactants side of the equation, then there also must be four oxygen *atoms* on the products side. Similarly, if the reaction is of the redox variety, the charges (sum of + and –) should be equal on both sides. If the reaction involves nuclear fusion or fission, then energy and mass should be equal on both sides of the equation.

1.10 EQUILIBRIUM, THERMODYNAMICS AND DRIVING FORCES FOR REACTIONS: SYSTEMS, GIBBS ENERGIES, ENTHALPY AND HEAT CAPACITY, ENTROPY, VOLUME

This section begins with the example of pyrite (FeS_2 or iron disulfide), a mineral that forms in O_2 -poor environments that include deep crustal levels and anoxic surface environments like swampy muds (technically iron sulfide that forms at Earth surface temperatures is a poorly ordered form of $\sim\text{FeS}$, e.g. mackinawite). Iron, in the Fe^{+2} state, and sulfur, in a combination of S^{-2} and S^0 states (average = S^{-1}), both occur in chemically reduced forms in the mineral pyrite – these oxidation states are stable in reducing/anoxic/anaerobic/ O_2 -poor environments.

Given the conditions under which pyrite forms, it is possible to predict its fate in oxidizing (O_2 -rich) environments, common examples being soils exposed to the oxygen-rich, water vapor-bearing atmosphere, or an O_2 -rich bubbling stream, where the stable forms of iron and sulfur are Fe^{+3} and S^{+6} . Under oxidizing conditions like these pyrite undergoes chemical oxidation, producing iron hydroxide with its characteristic rusty orange stains typical of many rock outcrops and soils. The reaction of pyrite, water and oxygen to produce iron hydroxide, sulfuric acid and free electrons (oxidation!) can be expressed as follows:



(In reality this reaction occurs in two or more steps, often in the presence of sulfur-oxidizing bacteria). Viewing this reaction in terms of the extent of iron oxidation would likely result in a characteristic pattern observed for many reactions in nature (Fig. 1.12).

In Fig. 1.12, the initial system (pyrite in contact with atmosphere) is out of equilibrium – both the

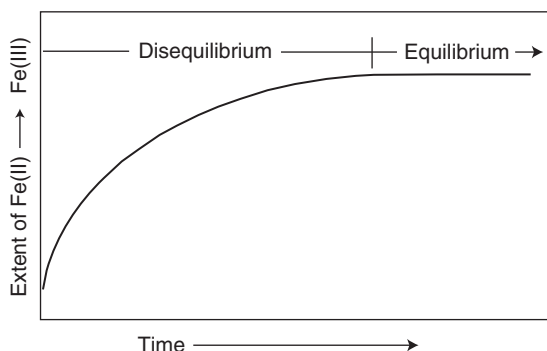


Fig. 1.12 Typical progression of a geochemical system towards equilibrium. A high degree of disequilibrium in the early stages causes rapid rates of change, but as the system approaches equilibrium, rates decrease logarithmically.

Fe and S are unstable in reduced forms when exposed to O_2 -rich air (e.g. by a landslide that exposes fresh un-oxidized rock) and the pyrite begins to react. This reaction is expressed as $Fe(II) \rightarrow Fe(III)$ in the graph, and note that initially the reaction occurs rapidly but as the reaction progresses the rate steadily decreases and ultimately ceases altogether. There are two probable explanations:

1 the system has reached **equilibrium**, a condition where net concentrations of products and reactants do not change. **Dynamic equilibrium** sometimes occurs in natural systems, when the rate of formation of products equals the rate of formation of reactants. If the $Fe(OH)_3$ and H_2SO_4 produced by pyrite oxidation are not leached out of the system, a dynamic equilibrium may be established where the rate of the forward reaction shown above is equal to the rate of the reverse reaction – i.e. FeS_2 , O_2 and H_2O are produced at the same rate as are $Fe(OH)_3$ and H_2SO_4 . At dynamic equilibrium, reactions are taking place, namely “products” are dissolving to produce “reactants” at the same rate that “reactants” produce “products”, but the net concentrations of products and reactants do not change with time. However, any change to a variable involved in a

dynamic equilibrium (e.g. concentration of reactant or product, volume, pressure, temperature) will cause a shift to counter the change. For example, removing a reactant will shift the reaction toward the direction of reactants – the rate will slow, or reactants will be produced at the expense of products. Or, if products are lost from the system, the reaction will continue to form products, which leads to the other possibility:

2 The reaction has run to completion. In some natural systems, where products are lost due to leaching (e.g. of H_2SO_4 in the case above) or degasification (e.g. of CO_2 with decomposition of hydrocarbons), a dynamic equilibrium cannot be established. If the pyrite reaction above stopped because all $Fe(II)$ had been consumed to produce $Fe(III)$ (i.e. if the reaction were to run to completion because products are being lost from the system), that system will not reach dynamic equilibrium. The natural environment differs from the laboratory in that reactants are often lost from soils, rocks and groundwaters, and dynamic equilibrium may not apply. In other cases, reactions in nature may not reach equilibrium because reaction rates are very slow – the reaction never proceeds past the early convex part of the disequilibrium curve. This often occurs in soils, where igneous minerals such as amphiboles and pyroxenes that are stable in high-temperature, low- O_2 environments persist in a state of disequilibrium because the rates at which they decompose in weathering environments are relatively slow.

1.10.1 Systems, species, phases and components

In spite of certain limitations, examining environmental systems through the lens of **equilibrium thermodynamics** can be very useful. It can help determine the direction in which chemical changes will take place (e.g. pyrite is unstable and will oxidize in contact with the atmosphere) and also to infer rates because the farther a system

is from equilibrium, the faster it will react to reach equilibrium. A common term used in thermodynamics is **system**, which is a somewhat arbitrary definition of the components we wish to consider. Depending on the question, a system might be an aquifer, or a pore within an aquifer, or the entire hydrologic cycle; it could be an entire granitic pluton, or it might be a micrometer-sized fluid inclusion in a quartz crystal. It really depends on the scale of study. If the question is climate change, the entire troposphere might be considered as the system, or a smaller system comprised only of a landfill might be considered if the main concern is a single source of carbon (e.g. CH_4).

Systems can be **open** (e.g. a stream, the atmosphere, a leaking landfill), where material is added or lost, or **closed**, where flow of material is restricted (e.g. tiny pores within impermeable fine-grained sediments). In some cases, systems are closed with respect to solids but open with respect to gases or heat, and in other cases systems can be defined as closed to physical and thermal flux, in which case they are **isolated**. Systems are comprised of components, phases and species.

In geochemistry, **species** are microscopic entities, commonly ions or gases such as Ca , CO_3^{2-} , SO_4^{2-} , CO_2 or H_2S , whereas **phases** are defined as physically separable parts of a system, typically minerals, liquids (e.g. H_2O) and distinct gases (dissolved CO_2 and O_2 in stream water, for example). Phases are comprised of species—for example, the **phase** calcite is comprised of the **species** Ca^{+2} and CO_3^{2-} , or of the **species** CaO and CO_2 (calcite can be defined in either way). While species are generally substances that can or do exist in nature, **components** do not necessarily exist in nature. Sometimes they are similar to species, but in other cases they may be mathematical expressions that can be useful in thermodynamic calculations, one example being KNa_{-1} , a mathematical operator used to indicate addition of K and subtraction of Na.

So, a soil with the mineral dolomite ($\text{CaMg}[\text{CO}_3]_2$), water, dissolved Ca^{+2} , dissolved Mg^{+2} , dissolved CO_3^{2-} , CO_2 gas and quartz could be defined as having four phases (dolomite, water, quartz, CO_2 gas) and 6 species or components (Ca^{+2} , Mg^{+2} , CO_3^{2-} , CO_2 , H_2O , and SiO_2).

Lastly, the **phase rule** (sometimes referred to as **Gibbs' phase rule**) relates components (C), phases (P) and degrees of freedom (F) according to this simple equation:

$$F = C - P + 2 \quad (1.32)$$

Degrees of freedom represent tangible changes to a system, typically temperature and pressure. In a system with 2 degrees of freedom, temperature and pressure can both change without producing a change in the state of the system. The classic example involves a simple system involving only one component, H_2O , which can exist in 3 phases (solid, liquid and vapor forms of H_2O), as displayed in the H_2O phase diagram (Fig. 1.13).

There is one point in the H_2O phase diagram where all three phases coexist – the triple point.

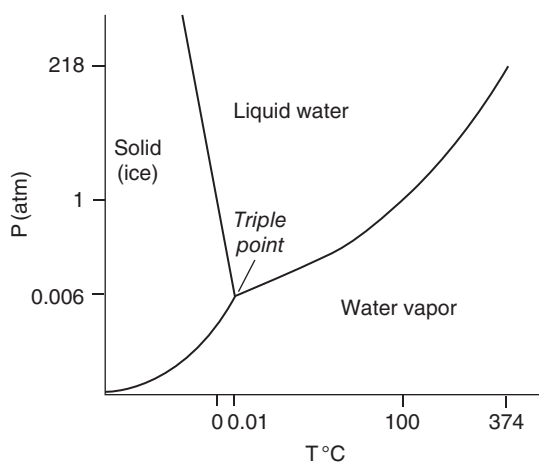


Fig. 1.13 Phase diagram for H_2O depicting one component (H_2O) and three phases (solid, liquid and gaseous water).

At the triple point, there is one component (H_2O ; $C = 1$) and three phases (solid, liquid and vapor; $P = 3$). The phase rule indicates that for this system,

$$F = 1 - 3 + 2 = 0$$

In other words, there are no degrees of freedom. Any change in either T or P will produce a change in the state of the system. Increasing T at constant P will cause ice to melt and liquid to vaporize. Increasing P at constant T will cause vapor to condense and ice to melt, producing a system with only one phase. Consider for a moment a system with only one phase, liquid water at 100°C and 200 atm of pressure. In this case, the phase rule indicates that (for a system where $C = 1$ and $P = 1$) there will be two degrees of freedom, and this is borne out in the diagram above – if either T or P changes, or even if both T and P simultaneously change, there will be no change to the state of the system. It will remain liquid water until the system either cools or decreases in pressure to the point where P – T reaches the phase boundary (line) between liquid and solid or between liquid and vapor.

1.10.2 First law of thermodynamics

The term thermodynamics implies an approach based on changes in **heat**, and the field did originate in studies of the transformation of heat energy to mechanical energy during the 1800 s. Energy can change forms, e.g. **potential energy** to heat (or vice versa), as in the combustion of organic matter, a process where **stored chemical potential energy** is converted to heat energy. Thermodynamics deals with transfers of energy, and one of the fundamental principles is the **First Law of Thermodynamics**, often known as the **Law of Conservation of Matter and Energy**, which can be stated as follows:

$$\Delta E = Q - W \quad (1.33)$$

Or, for small increments of change, this equation can be stated as:

$$dE = dq - dW \quad (1.34)$$

E is the **internal energy** of a system, Q or dQ represent **heat flux**, and W represents **work** done on the system. The first law essentially states that energy is neither created nor destroyed – matter and energy may change forms, new phases may be produced, gases may convert to solids, or solids may dissolve, but the net energy of a closed system never changes (or, for an open system, the net energy of the open system and its surroundings never changes). The kinetic energy of friction along a geological fault may produce heat, or the chemical potential energy of the bonds in hydrocarbons may produce heat during combustion, but the net change in energy is zero.

Consider an experiment you could perform while sitting in your overstuffed chair reading this text, the example of heat (Q) produced by the work (W) of rubbing hands together. The system can be defined as your body and its immediate surroundings. The energy you need to move your shoulders and arms to produce friction comes from the stored energy in the chemical bonds in the food you recently ate, and in a simplified approximation, this chemical energy is transformed to mechanical energy. The friction between your hands converts mechanical energy to heat energy, causing your hands to warm up (and raising the temperature of the system). This heat will soon be lost to the surroundings (an issue we will soon address with the 2nd Law), but if we consider the net energy of the system (body and immediate surroundings), it has not increased or decreased, but rather merely has changed forms. Of course, the heat will eventually escape from your surroundings and ultimately you will need to eat more food to provide the chemical energy needed to reinitiate the cycle

(conversion of chemical energy in food to heat energy is an irreversible reaction in our bodies). In terms of the equation of the 1st Law, the work (W) done will have a positive sign, as will the heat term Q (heat is gained by the system, so the term Q will have a positive sign, because in geochemistry, the flow of heat is considered positive for any system that gains heat). So, in a semi-quantitative sense, W and Q will both be positive and the values will be equal, and $\Delta E = Q - W = 0$, reflecting the fact that there is no change in the net energy of our body and its surroundings (at least in the short term).

Work can be defined as follows:

$$dw = P \cdot dV \quad (1.35)$$

where P is external **pressure** on a system and dV is the change in **volume** of the system. A typical way to consider work relative to volume is to quantify the work done by expansion during the change of state from liquid water to water vapor. At constant P , a positive dV term (expansion increases V) will produce a positive dw , indicating that work has been done on the surroundings by the system. If you substitute the PdV term for W into the equation of the first law presented above, you will arrive at this equation for the first law, expressed in terms of change in **internal energy**:

$$dE = dq - PdV \quad (1.36)$$

1.10.3 Second law of thermodynamics

The **Second Law of Thermodynamics** deals with **entropy**, a measure of the degree of **disorder** within a system. Any system tends toward a state of increasing randomness unless energy is added to the system to increase order. Increasing entropy is a **spontaneous** process and energy must be added to produce order. In geochemistry,

an example of entropy as a spontaneous process is chemical weathering of a granite, for example, where minerals with ordered crystal lattice structures (a low entropy state) are decomposed into soluble aqueous species such as Na^+ , K^+ , and $\text{Si}(\text{OH})_4$ that are then scattered across the globe. Only with addition of energy (e.g. the internal heat of the Earth) can some order be restored (i.e. entropy decreased). An everyday example of entropy is this: the natural state of a kitchen or living room will progress toward a state of disorder (greater entropy) characterized by dirty dishes, potato chip bags, music scores and old newspapers scattered about unless we expend energy to restore order.

Entropy (S) can be represented by the following equation, where q is heat and T is temperature and the process is reversible:

$$dS = dq/dT \quad (1.37)$$

Entropy can be quantified in terms of change in heat content per change in temperature, and units of S are joules/deg (non-SI units are cal/deg; 1 cal = 4.18 J). Consider the entropy of liquid water and water vapor at standard temperature (25 °C) and pressure (1 atmosphere) – standard T and P are considered **standard state conditions** and are commonly used to make calculations immensely simpler. S for H_2O_l is 69.9 J/mol/K, and for H_2O_g it is 188.7 J/mol/K (Appendix III). Intuitively, the value of S is greater for the gaseous, more disordered state of H_2O than for the liquid state.

1.10.4 Enthalpy

Another important consideration in thermodynamics is **enthalpy** (H), the **heat content** of a system. In some cases enthalpy is expressed in units of calories (cal) or kilocalories (kcal), which makes sense because calories are a common

measure of heat in everyday life; however, the SI units are joules (J) (or kilojoules, kJ).

Enthalpy is typically expressed as follows:

$$H = E + PV \quad (1.38)$$

or

$$dH = dE + PdV \quad (1.39)$$

But often the most important consideration related to enthalpy is the change in H ($\Delta H = H_2 - H_1$) during a reversible reaction, where the two values of H represent enthalpies associated with different states of matter, like liquid water and water vapor, or with elements in different bonding arrangements, e.g. Fe and O_2 vs. Fe_2O_3 . ΔH is an important parameter in geochemistry because it expresses heat absorbed or released during changes of state (e.g. evaporation) or during chemical reactions that produce minerals, ions or molecules. If heat is gained during a reaction (i.e. ΔH is positive), the reaction is endothermic; conversely, if heat is lost (ΔH is negative), the reaction is exothermic.

Combustion of organic matter is clearly exothermic – we burn firewood and hydrocarbons to produce heat. Boiling of water, the transformation of andalusite to sillimanite during prograde metamorphism, and the maturation of petroleum in sedimentary basins are all endothermic processes – they absorb heat. The stored heat can then later be released; for example, stored heat in petroleum is released during the exothermic reaction known as combustion. It is worth pointing out the difference between two terms that represent heat, Q and H . Q represents flow of heat or heat transfer, for example from hot Hawaiian lava into cool ocean water; H represents heat stored within a system, such as the stored heat in petroleum, water vapor or sillimanite.

Given that reactions either consume or produce heat, we can calculate the difference in enthalpy between the reactants and products and determine the amount of heat produced or

consumed by the reaction. This is important because it is one of the ways that thermodynamics can help to predict the behavior of environmental systems such as soils and waters; typically, spontaneous processes produce (i.e. release) heat, i.e. spontaneous processes usually are exothermic (caveat: while this is generally true it is not always the case – for example, when some salts dissolve in water the solution gets colder. The process absorbs heat from the water yet it is a spontaneous process because the increase in entropy is more important than the positive ΔH).

Enthalpies of reactions (ΔH°_R) are determined by summing the **enthalpies of formation (ΔH°_f)** of all reactants (standard state conditions) and subtracting this term from the sum of ΔH°_f values of all products (standard state conditions):

$$\Delta H^\circ_R = \sum n_x * H^\circ_{fx}(\text{products}) - \sum n_x * H^\circ_{fx}(\text{reactants}) \quad (1.40)$$

Enthalpies of formation are available from various sources and selected examples are presented in **Appendix III**. By convention, $H^\circ_f = 0$ for elements in their pure state (e.g. Fe, Si, Na) and for gases such as H_2 , N_2 and O_2 . In the ΔH°_R reaction above, the H°_f for each reactant or product (represented by the variable x) is multiplied by the number of moles (n) expressed in the reaction.

We can examine the reaction of Fe and O_2 to form hematite (Fe_2O_3) by the chemical reaction $2 Fe + 1.5 O_2 = Fe_2O_3$. Values of H°_f (in kJ/mol):

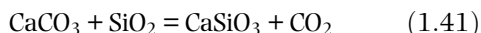
$$\begin{aligned} H^\circ_f(Fe) &= 0 \\ H^\circ_f(O_2) &= 0 \\ H^\circ_f(Fe_2O_3) &= -824.2 \end{aligned}$$

It is crucial to remember to multiply H°_f values by molar abundances presented in the chemical reaction.

$$\begin{aligned} \Delta H^\circ_R &= (1 * -824.3) - (2 * 0 + 1.5 * 0) \\ &= -824.3 \text{ kJ/mol.} \end{aligned}$$

The negative ΔH°_R for this reaction implies that it is spontaneous at standard temperature and pressure, and that iron will oxidize to form iron oxide.

Under certain conditions, calcite (CaCO_3) and quartz (SiO_2) react to form wollastonite (CaSiO_3) plus CO_2 according to this reaction:



Is this a spontaneous process at the Earth's surface? (i.e. 1 atm and 25 °C). Values of H°_f (in kJ/mol) are (Appendix III):

$$\begin{aligned} H^\circ_f(\text{CaCO}_3) &= -1207.4 \\ H^\circ_f(\text{SiO}_2) &= -910.7 \\ H^\circ_f(\text{CaSiO}_3) &= -1630 \\ H^\circ_f(\text{CO}_2) &= -393.5 \end{aligned}$$

The enthalpy of the reaction can then be determined as follows:

$$\begin{aligned} \Delta H^\circ_R &= [(1 \text{ mol} * -1630 \text{ kJ/mol}) \\ &\quad + (1 \text{ mol} * -393.5 \text{ kJ/mol})] \\ &\quad - [(1 \text{ mol} * -1207.4 \text{ kJ/mol}) \\ &\quad + (1 \text{ mol} * -910.7 \text{ kJ/mol})] = +94.6 \text{ kJ} \end{aligned}$$

This reaction requires addition of 94.6 kJ of heat (per mol of each reactant given the stoichiometry of the reaction), indicating that it is not spontaneous, and evidence from geology confirms this conclusion because calcite and quartz do not react to form wollastonite and carbon dioxide until systems reach medium-grade metamorphic temperatures and pressures (~500 °C, > 1 kb).

1.10.5 Heat capacity

Heat capacity is the amount of heat required to raise the temperature of a given amount (1 mol or 1 g) of a substance by 1 °C. It is defined as the ratio

of heat added to resulting temperature change, where the greater amount of heat required to result in temperature increase corresponds to higher heat capacity.

$$C = dq/dT \quad (1.42)$$

C = heat capacity and dq and dT are changes in heat and temperature. Rearranging emphasizes the point that, for a given amount of heat added (dq), the magnitude of temperature increase will be lower if the heat capacity is high.

$$dT = dq/C \quad (1.43)$$

At constant volume, heat capacity can be expressed as:

$$C_v = (\delta q/\delta T)_v \quad (1.44)$$

Given equation 1.36, and realizing that $PdV = 0$ at constant volume:

$$C_v = (\delta E/\delta T)_v \quad (1.45)$$

(note: here, the symbol δ indicates that the derivation is performed with a restriction, in this case constant volume. The subscript v indicates that volume of the system is constant).

At constant pressure:

$$C_p = (\delta q/\delta T)_p \quad (1.46)$$

and equation 1.36 can be rearranged slightly to give:

$$dq = dE + PdV \quad (1.47)$$

Now, substituting the right side of equation 1.47 into the numerator of the right side of equation 1.46:

$$C_p = (\delta E/\delta T)_p + P(\delta V/\delta T)_p \quad (1.48)$$

And given equation 1.39 (effectively, $\delta H = \delta E + P\delta V$),

$$C_p = (\delta H / \delta T)_p \quad (1.49)$$

Examples of heat capacities of substances (at 25 °C) include liquid water, which has a relatively high value ($C = 4.19 \text{ J/g}\cdot\text{K}$) relative to dry rock, which has values in the range of 0.7 to 0.9 J/g·K, or steel, which has an even lower value of 0.47 J/g·K. The paucity of water in deserts means that rock is the dominant control on temperature change and this results in drastic temperature swings, both diurnally and seasonally; by comparison, in moister regions such as tropical forests or temperate coastal regions, temperature extremes are minimized by the high heat capacity of water.

Heat capacity is an **extensive property**, i.e. it is dependent on the amount of the substance in question – the greater the amount of substance, the greater the amount of heat needed to be added to achieve the same change in temperature. That said, values of heat capacity are typically normalized to a mass of 1 gram, meaning that effectively heat capacity is an **intensive property**, i.e. when normalized (divided by mass), it reflects a characteristic of a given substance independent of amount (i.e. mass).

1.10.6 Gibbs free energy

The **Gibbs Free Energy** (G) of a system accounts for changes in both enthalpy and entropy during reactions.

$$\Delta G^\circ_R = \Delta H^\circ_R - T\Delta S^\circ_R \quad (1.50)$$

Any reaction that produces a decrease in the Gibbs free energy is spontaneous – any reaction for which ΔG°_R is negative is spontaneous. Systems tend towards lower-energy states in the absence of new addition of energy, so a decrease in G will produce a more **stable** system. In a schematic way (Fig. 1.14), stability of a physical

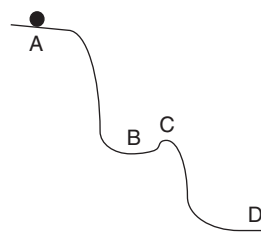


Fig. 1.14 Schematic diagram illustrating relative stability conditions, from unstable at points A and C, to metastable at point B and stable at point D. Note that the metastable point B can be reached from points A or C, and that to shift from metastable (B) to stable (D), some energy must be introduced. If not, the metastable condition may persist for a long time.

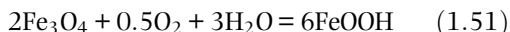
system can be used to illustrate this point. The ball at point A is unstable with respect to location, and to decrease this instability (high energy state), it will roll down to point B, and if it can overcome the slight energy barrier at point C, it will eventually achieve its most stable configuration by rolling down to point D.

Point B might be referred to as a **metastable** condition, one that is not the most stable configuration (that would be D), but one that may play a strong role in system behavior.

An example of metastability in the geochemical realm is the occurrence of the mineral halloysite, a disordered form of kaolin that can form in tropical soils when highly unstable igneous minerals such as olivine and pyroxene rapidly dissolve (over hundreds to thousands of years). Rapid dissolution may lead to formation of metastable halloysite rather than kaolinite due to kinetic factors; once formed, halloysite may require hundreds of thousands to millions of years to finally transform into thermodynamically stable kaolinite.

Determining Gibbs free energies for geochemical systems allows prediction of their behavior, much like the simple analysis of the ball between points A and D. For example, if magnetite (Fe_3O_4), water and oxygen gas exist in a soil at standard conditions, we might ask, is the assemblage stable or

of out of equilibrium? Is magnetite more or less stable in this soil than an iron hydroxide such as goethite? The possibility that the system will react to form goethite (FeOOH) to achieve a lower-energy state could be assessed according to this reaction:



$\Delta G^\circ_{\text{R}}$ for the reaction is computed according to this equation:

$$\Delta G^\circ_{\text{R}} = \sum n_x * G^\circ_{\text{fx}}(\text{products}) - \sum n_x * G^\circ_{\text{fx}}(\text{reactants}) \quad (1.52)$$

Values for G°_{f} (in kJ/mol) are (Appendix III):

$$\begin{aligned} G^\circ_{\text{f}}(\text{Fe}_3\text{O}_4) &= -1015.5 \\ G^\circ_{\text{f}}(\text{O}_2) &= 0 \\ G^\circ_{\text{f}}(\text{H}_2\text{O}) &= -237.2 \\ G^\circ_{\text{f}}(\text{FeOOH}) &= -488.6 \end{aligned}$$

These values produce the following equation:

$$\begin{aligned} \Delta G^\circ_{\text{R}} &= [6 \text{ mol} * -488.6 \text{ kJ/mol}] \\ &\quad - [(2 \text{ mol} * -1015.5 \text{ kJ/mol}) \\ &\quad + (0.5 \text{ mol} * 0 \text{ kJ/mol}) \\ &\quad + (3 \text{ mol} * -237.2 \text{ kJ/mol})] \\ &= -189.0 \text{ kJ} \end{aligned}$$

What this negative Gibbs free energy value demonstrates is that magnetite is unstable in the presence of water and oxygen at the Earth's surface and will react to produce goethite, a common soil mineral. What this reaction does not tell us is how fast this reaction will take place. Is there an energy barrier? Does the system need to overcome an activation energy like that at point C in the ball example above? Does diffusion of O_2 or H_2O to magnetite surfaces limit reaction rate? In reality, magnetite can persist in soils in a metastable state for at least thousands of years, and that fact is not evident from this $\Delta G^\circ_{\text{R}}$ calculation. Section 1.11 on kinetics addresses some questions related to rates of processes.

1.10.7 Gibbs free energy and the equilibrium constant

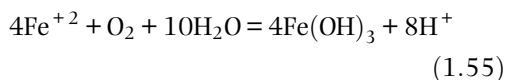
The Gibbs Free Energy of a reaction can also be expressed in terms of the equilibrium constant (K_{eq}), which is much like the K_{a} discussed earlier in this chapter. Consider a general reaction at equilibrium where:



A and B, and C and D are reactants and products, respectively, and the lowercase a , b , c and d represent molar fractions of reactant or product. In this general case:

$$K_{\text{eq}} = [\text{C}]^c [\text{D}]^d / [\text{A}]^a [\text{B}]^b \quad (1.54)$$

$[\text{C}]^c$ represents the concentration of reactant C raised to the c power, and so on. Products are in the numerator and reactants in the denominator. Or, for a more tangible example, we can write the K_{eq} for oxidation of ferrous iron (Fe^{+2}). The chemical reaction is:



and:

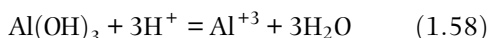
$$K_{\text{eq}} := [\text{Fe}(\text{OH})_3]^4 * [\text{H}^+]^8 / [\text{Fe}^{+2}]^4 * [\text{O}_2]^8 * [\text{H}_2\text{O}]^{10} \quad (1.56)$$

However, by convention, the concentration of water is given the value of 1, and pure solid phases like ferrihydrite ($\text{Fe}(\text{OH})_3$) are similarly assigned values of 1, so in this case the K_{eq} simplifies to:

$$K_{\text{eq}} := [\text{H}^+]^8 / [\text{Fe}^{+2}]^4 * [\text{O}_2]^8 \quad (1.57)$$

which means that the main controls on the reaction are the concentrations of H^+ and Fe^{+2} and the availability and reactivity (fugacity) of O_2 .

A high value for the K_{eq} indicates that the reaction will proceed in the direction of products. In other words, for reactions that favor products, the numerator will be greater than the denominator, producing a high K_{eq} . In cases where reactants are more favored (i.e. reactants are predicted to occur in higher concentrations than products), K_{eq} values are small. For example, the dissolution of gibbsite in the presence of acid, represented by the following reaction,



has a $K_{\text{eq}} = 10^{8.11}$ at 25 °C and 1 bar, which is often expressed in log form as $\log K_{\text{eq}} = 8.11$ (Langmuir, 1997). Expressed in K_{eq} form:

$$K_{\text{eq}} = [\text{Al}^{+3}]/[\text{H}^+]^3 = 10^{8.11} \quad (1.59)$$

it is clear that the system strongly favors dissolution of $\text{Al}(\text{OH})_3$ and formation of dissolved Al^{+3} . (Of course, pH of the solution will strongly control the probability of this reaction). Conversely, the dissolution of fluorite ($\text{CaF}_2 = \text{Ca}^{+2} + 2\text{F}^-$) has a $\log K_{\text{eq}} = -10.6$, and this low value indicates that fluorite is very insoluble in water. The product of the concentrations of $[\text{Ca}^{+2}]$ and $[\text{F}^-]^2$ is extremely low ($K_{\text{eq}} = 10^{-10.6}$), indicating that the reactant side of the equation is favored.

$\Delta G^\circ_{\text{R}}$ for this reaction or any reaction can be expressed as:

$$\Delta G^\circ_{\text{R}} = -RT \ln K_{\text{eq}} \quad (1.60)$$

This is the Van't Hoff equation, where R (the gas constant) is equal to 8.314 J/cal/mol K (or 8.314×10^{-3} kJ/mol K) and T is temperature (in K), which produces units of J/mol (or kJ/mol) for $\Delta G^\circ_{\text{R}}$.

In most cases, Gibbs free energy data are determined for systems at 25 °C (298.15 K), so for systems at 25 °C (reasonably representative of surface environments), $\Delta G^\circ_{\text{R}}$ can be further simplified and converted to $\log_{10}$ as:

$$\Delta G^\circ_{\text{R}} = -5.708 * \log(K_{\text{eq}}) \quad (1.61)$$

Typically enthalpy and entropy changes are not strongly temperature-dependent, so given equation 1.54, the Gibbs Free Energy, ΔG , should vary in a linear manner with temperature. Rearranging equation 1.50 slightly and presenting the variables in a generic sense gives us:

$$\Delta G/T = \Delta H/T - \Delta S \quad (1.62)$$

By then substituting the right side of equation 1.60 for ΔG :

$$RT \ln K/T = -\Delta H/T + \Delta S \quad (1.63)$$

Then dividing and rearranging:

$$\ln K = -\Delta H/RT + \Delta S/R \quad (1.64)$$

Viewed in the following manner, equation 1.64 is an equation of a line: $\ln K = y$, $-\Delta H/R$ is m (slope), $1/T$ is x and $\Delta S/R$ is b (constant). Two characteristic types of plots (Fig. 1.15) result when different values for T are plugged in to equation 1.64.

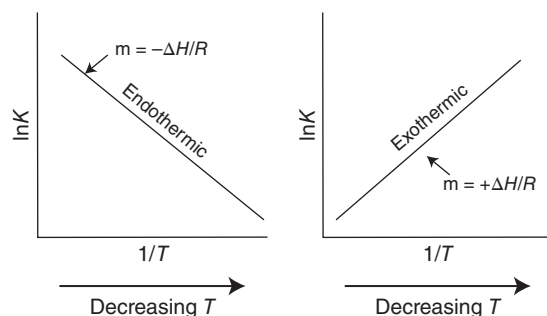


Fig. 1.15 Arrhenius plots showing relationship between equilibrium constant (K) and temperature for endothermic and exothermic reactions. For endothermic reactions, increasing temperature increases K , i.e. adding heat increases reaction rate; for exothermic reactions, the opposite is true.

Taking the derivative of equation 1.64 with respect to temperature gives a different formulation of the van't Hoff equation:

$$d(\ln K)/dT = \Delta H/RT^2 \quad (1.65)$$

Note that the term $\Delta S/R$ disappears because both terms are constants (assuming S does not change with temperature).

Integrating equation 1.65 between temperatures T_1 and T_2 produces equation 1.66, another expression of the van't Hoff equation, one which makes it possible, given a K value (e.g. K_1) at standard temperature (e.g. T_1), to determine an unknown value of K (e.g. K_2) at a different temperature (i.e. T_2).

$$\ln(K_2/K_1) = -\Delta H/R(1/T_2 - 1/T_1) \quad (1.66)$$

For calculations using this equation, R is the gas constant (8.314 J/K·mol) and temperature is in Kelvin. It is straightforward and can be useful (note: using this equation K is unitless, as it also is when expressed in equations 1.60 through 1.64).

1.11 KINETICS AND REACTION RATES: DISTANCE FROM EQUILIBRIUM, ACTIVATION ENERGY, METASTABILITY

Thermodynamic information is often insufficient to fully understand behaviors or states of ions and compounds in nature, especially in low-temperature systems like soils, waters and the atmosphere. Soils, for example, almost always contain minerals that are thermodynamically *unstable* at the Earth's surface, such as the olivine, augite and plagioclase that form at $T > 800^\circ\text{C}$ in cooling basalt lava, as well as minerals that are thermodynamically *stable* at low temperatures (25°C , for example), such as calcite, kaolinite

and hematite. While the latter three can come to equilibrium with surface waters, olivine, augite and plagioclase do not. Thermodynamics tells us that. What thermodynamics does not quantify is rate. Consider magnetite in tropical soils; thermodynamics (e.g. ΔG) indicates that it will transform to goethite, but it does not indicate how fast.

Questions about rates of processes are addressed by the study of **kinetics**. Reactions that occur slowly, are not reversible or do not take place in a system at equilibrium for whatever reason (e.g. soils) are often best understood using kinetic approaches. The same can be said for **heterogeneous reactions**, those that involve various states of matter such as solid minerals and amorphous solids mixed with liquid and gas phases, conditions that tend to occur in soils, sediments, streams, aquifers and the atmosphere.

Rates of geochemical reactions depend on numerous factors, including temperature, pressure, redox conditions, pH, mineral composition, abundance of organic matter, pore water composition, diffusion rates, bond types, biotic factors, system composition, and more. In many cases, geochemical rates have been determined in laboratory studies that require extrapolation to natural environments and are often prone to large uncertainties (in some cases, an order of magnitude or more). The multitude of potential biotic factors can be harder to quantify than the main inorganic controls, yet biotic effects can be very strong. For example, plants can alter the pH of soil by exuding H^+ from their roots to enhance dissolution of nutrient-bearing minerals and microbes may accelerate oxidation and reduction reactions. Thus, the complex array of variables influencing reaction rates can complicate kinetic analysis.

Often geochemical processes are controlled by a **rate-limiting step**, a step in a process that is much slower than others. Consider the **dissolution** of a mineral grain, a process that could involve five steps: (1) diffusion of reactant(s) toward the mineral surface (where a common example of a reactant is H^+); (2) sorption

of reactants onto the mineral surface; (3) formation of a bond between the reactant and the part of the mineral grain under attack (perhaps the O atom bonded to a Ca ion in a feldspar); (4) desorption of the newly formed complex between reactant and mineral component (e.g. OH^-), and finally (5) transport of the newly formed product away from the mineral surface by diffusion. If any of one of the steps 1 through 5 is slower than others, it will be the rate limiting step. In clay-rich sediments, diffusion can be a rate-limiting step, whereas in highly insoluble minerals, step 3 (and/or 4) can be the rate-limiting step. Kinetic limitations on **crystal growth** are similar, and can be envisioned by reversing steps 1 through five.

1.11.1 Reaction rate, reaction order

Reaction rates are in part controlled by the order of the reaction, where reaction order is defined as the dependence of reaction rate on concentrations or molar ratios of species involved in reactions. Reactions can be of zeroth, first, second, or third order (0^{th} , 1^{st} and 2^{nd} are by far most common in environmental geochemistry). Rates of 0^{th} -order reactions occur at rates independent of the concentration of the reactant(s); rates of 1^{st} -order reactions are controlled by the concentration of a reactant or product; rates of 2^{nd} -order reactions are controlled by concentrations of two reactants or products, or a reactant (or product) squared.

A **0^{th} -order reaction**, one where reactant A undergoes transformation to product P, can be represented by the simple chemical equation $\text{A} \rightarrow \text{P}$. The following equation describes change in concentration of A (represented as $[\text{A}]$) with time (dt):

$$d[\text{A}]/dt = -k \quad (1.67)$$

where k is a constant, probably determined by laboratory experiments or by studies of

natural systems, and the negative sign indicates decreasing abundance of A with time. In integrated form, this equation is:

$$[\text{A}] = [\text{A}]_0 - kt \quad (1.68)$$

$[\text{A}]_0$ represents the concentration of A at time zero (where examples of t_0 could be when sediments were deposited in a floodplain or when a gaseous compound was emitted into the atmosphere), and the negative sign for kt correlates to decreasing A with time. So the concentration of A at any time is merely determined by the amount of A at t_0 [$[\text{A}]_0$], how much time has passed (t), and the rate constant (k) that governs reaction of A into P. But the amount of A does not control the rate, and this distinction is important (when compared to 1^{st} -order reactions, e.g.). Furthermore, the concentration of P does not affect reaction rate. Plotted graphically (Fig. 1.16), it becomes clear that, for 0^{th} -order processes, $[\text{A}]$ decreases in a linear fashion with time, and rate does not vary with time.

Examples of 0^{th} -order reactions include the dissolution of many types of minerals in aqueous solutions, notably salts.

A **1^{st} -order reaction** is one where the rate depends on the concentration of a reactant (or product) raised to the first power – that is,

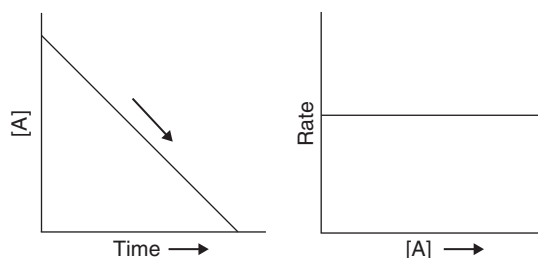


Fig. 1.16 Two graphical representations of zeroth (0^{th} -order) chemical reactions, where $[\text{A}]$ represents concentration of a chemical species (e.g. ion, mineral or compound). In the example on the left, decrease in concentration of A is linear, and on the right, the rate of change is constant.

1st-order reactions are proportional to the concentration of the reactant. A good example of a 1st-order reaction is radioactive decay, an example of which is provided below. The equation for a 1st-order reaction representing decreasing A with time is:

$$d[A]/dt = -kA \quad (1.69)$$

Literally, the rate of decrease of A depends on the rate constant (k) multiplied by the amount of A present at time t . In its integrated form, this equation is:

$$[A] = [A]_0 * e^{-kt} \quad (1.70)$$

Graphically, 1st-order reactions depict non-linear changes in concentration with time as well as rates that vary as a function of the concentration of reactants (Fig. 1.17) – note that the rate (expressed as, e.g., mol/year) varies but the rate constant (expressed as a percent or proportion, does not).

In the case of radioactive decay, the amount of daughter isotope (e.g. ²²²Rn produced by decay of ²²⁶Ra) produced per unit time decreases as the amount of parent isotope (e.g. ²²⁶Ra) decreases – as reactant is used up during the reaction, the *rate* of formation of products decreases (which is not to say that the amount of product decreases – it continues to increase, but at a progressively slower rate). In Fig. 1.17, [A] decreases exponentially (and the product increases parabolically, following the dashed line) while the *rate* of change is controlled by the amount of A (reactant) – less A equals lower rate. In the graph on the left, the solid line would represent amount of parent isotope (and the dashed line amount of daughter isotope) in a radioactive decay reaction.

Rates of **2nd-order reactions** are proportional to the concentration of a reactant squared (dependent on A where $2A \rightarrow B$), or in some cases to the product of the molar concentrations of two reactants (where $A + B \rightarrow C$).

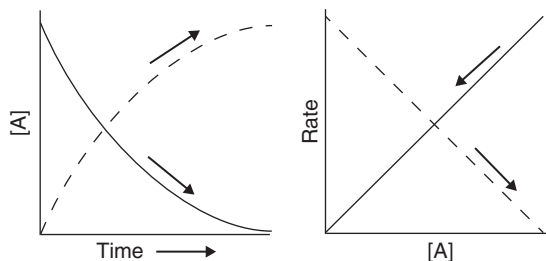


Fig. 1.17 Two graphical representations of first (1st)-order chemical reactions, where [A] represents concentration of a chemical species. On the left, change in concentration of A is logarithmic (solid line) or exponential (dashed line); on the right, for both cases (increasing or decreasing concentration), rate of change decreases in a linear fashion with time. Arrows in diagrams indicate direction of forward reaction (emphasizing that, on the right, the solid line represents decreasing [A] with time).

$$d[A]/dt = -kA^2 \quad \text{or} \quad d[A]/dt = -kAB \quad (1.71)$$

and integrated, the first equation can be expressed as:

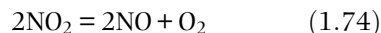
$$1/[A] = k*t + C \quad (1.72)$$

And given that $[A] = [A]_0$ when $t = 0$,

$$1/[A] - 1/[A]_0 = k*t \quad (1.73)$$

When depicted graphically (Fig. 1.18), 2nd-order reactions exhibit rapid initial changes in concentrations and progressively slower rates with time, and rate increases exponentially with higher concentrations of reactant(s).

An example of a 2nd order reaction is that of nitrogen dioxide to nitrogen monoxide plus oxygen, one that plays an important role in atmospheric chemistry:



We can summarize reaction rates with this equation:

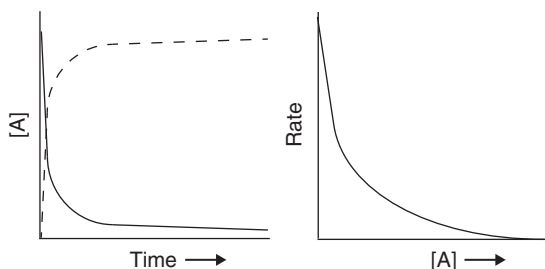


Fig. 1.18 Two graphical representations of second (2^{nd})-order chemical reactions, where $[A]$ represents concentration of a chemical species. In the example on the left, $[A]$ decreases exponentially, and on the right, we see that rate of change decreases exponentially with time.



If the reaction rate depends on the concentration of none of the components, then it is a 0^{th} -order reaction. If rate depends on the concentration of B, it is 1^{st} order, and if it depends on the concentration of A, it is 2^{nd} order. If it depends on the concentration of A, B and the product A_2B , it is 3^{rd} order (and 1^{st} order with respect to B and A_2B , and 2^{nd} order with respect to A), but this is a rare occurrence in geochemical systems.

1.11.2 Temperature and the Arrhenius equation

Temperature also plays an important role in reaction rate – higher temperatures foster higher reaction rates. Generally, reaction rates double for every increase of 10°C . The greater energy imparted by higher temperature tends to enhance interactions between reactants, thus enhancing the probability of productive interactions. Higher temperature can also overcome **activation energies** that can sometimes serve as barriers, like the one depicted schematically in Fig. 1.14. The Arrhenius equation relates temperature to the rate constant:

$$k = A * e^{-E_a/RT} \quad (1.76)$$

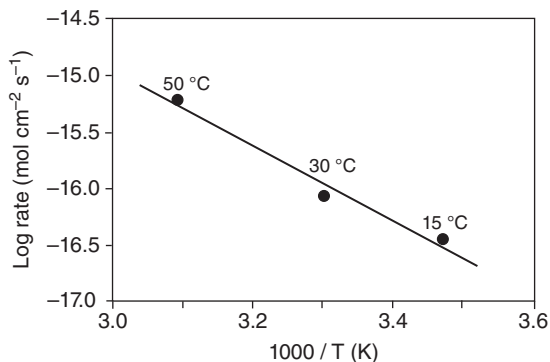


Fig. 1.19 Effect of temperature on chemical weathering rate of rhyolite as a function of temperature. (Yokoyama and Banfield 2002. Reproduced with permission of Elsevier.)

where E_a is activation energy, R is the gas constant, T is temperature (K), and A is the temperature-independent term known as the pre-exponential factor or the A factor – it serves to convert the product term into values appropriate for k of different reaction orders. In a qualitative sense, and because e is raised to the *negative* E_a/RT , increasing temperature will increase k , speeding up the reaction. Graphs of the effect of temperature on reaction rate are typically plotted as inverse of T ($1/T$, where T is in Kelvin) versus log of the rate constant k , a type of graph known as an Arrhenius plot. Figure 1.19 depicts the dissolution rate of powdered rhyolite (compositionally similar to granite; Table 1.1) as a function of temperature:

The rates shown in Fig. 1.19 increase more than ten-fold with a temperature increase of approximately 35°C (note log scale of y -axis).

The preceding introduction to equilibrium thermodynamics and kinetic constraints is meant to present some of the concepts, approaches and limitations contained within these two important fields. In the following chapters some of these approaches will be applied to predicting and understanding behaviors of environmental systems. For more detailed treatments of this topic, the reader is referred to excellent

and much more thorough treatments of these topics in geochemistry texts by James Drever (*The Geochemistry of Natural Waters*), Gunter Faure (*Principles of Geochemistry*) and Donald Langmuir (*Aqueous Environmental Geochemistry*). The *Kinetics of Geochemical Processes* issue of *Reviews in Mineralogy* edited by Lasaga and Kirkpatrick (1981) also presents various perspectives on this topic.

REVIEW QUESTIONS

1. Magnetite, hematite and goethite all occur in tropical soils. Rank them in order of thermodynamic stability in oxidized soil at 25 °C. Which is more likely to be released into soil water, a trace element (e.g. arsenic) substituted into the structure of hematite or goethite? Explain.
2. What is (are) the oxidation state(s) of manganese in each of the following minerals? MnOOH MnO₂ Mn₃O₄ MnO MnCO₃
3. Recalculate the chemical compositions of the igneous rocks given below in terms of the weight percent concentrations of the elements and rank them in terms of their abundance. Also calculate concentrations of each element in terms of mg/kg and mol/kg.

	#1	#2
SiO ₂	45.0	68.0
Al ₂ O ₃	3.5	15.0
FeO	8.0	3.2
Fe ₂ O ₃	0.0	1.3
MgO	39.0	1.7
CaO	3.2	3.4
Na ₂ O	0.32	3.1
K ₂ O	0.04	3.6
Cr ₂ O ₃	0.42	1.1×10^{-3}
NiO	0.25	9.2×10^{-4}

4. Why are elements with even atomic numbers more abundant than their neighbors with odd atomic numbers? Explain with appropriate nuclear reactions.

5. Why is Fe more abundant than its neighbors with similar atomic number? Describe with appropriate nuclear reactions.
6. Why is Pb more abundant than its neighbors with similar atomic number? Describe with appropriate nuclear reactions.
7. List the units of concentration that are commonly used in the following cases:
 - A. Acids in water.
 - B. Salts in water.
 - C. Metals in soil.
 - D. Metals in water.
 - E. SiO₂ in rock.
8.
 - A. Write the chemical reaction for the dissolution of barite.
 - B. Calculate the change in enthalpy associated with the dissolution of BaSO₄ (barite).
 - C. Based on your result, predict how the solubility of barite varies with temperature.
 - D. Calculate the Gibbs free energy for the dissolution of barite into ions.
 - E. Calculate the solubility of barite at 25 °C and 1 atm.
 - F. Given that Ra co-precipitates with Ba, predict the solubility of Ra in groundwaters (a) ~ lacking Ba (e.g. 1 ppb) and (b) rich in Ba (e.g. 100 ppb).
9. Which compound is more soluble in water, MgO or CaO? Explain.
10. Examine the kinetics of dissolution of a salt like NaCl by experimentation using an approach like that described by Velbel (2004).

REFERENCES

- Drever, J.I., 1997. *The Geochemistry of Natural Waters: Surface and Groundwater Environments*, 3rd edn. Simon and Schuster, Upper Saddle River, New Jersey, USA.
- Faure, G., 1986. *Principles of Isotope Geology*. John Wiley & Sons, Canada.

- Faure, G., 1998. *Principles and Applications of Geochemistry* (2nd edn). Prentice-Hall, Upper Saddle River, New Jersey, USA.
- Gromet, L.P., Dymek, R.F., Haskin, L.A., and Korotev, R.L., 1984. The "North American shale composite": Its compilation, major and trace element characteristics. *Geochimica et Cosmochimica Acta* 48: 2469–2482.
- Langmuir, D.L., 1997. *Aqueous Environmental Geochemistry*. Prentice Hall, Upper Saddle River, New Jersey, USA.
- Lasaga, A.C., Kirkpatrick, R.J. (eds.), 1981. *Kinetics of Geochemical Processes: Reviews in Mineralogy* (Vol. 8). Mineralogical Society of America, Washington, DC, USA, 398 pp.
- Lu, H., Dai, D., Yang, P., Li, L., 2006. Atomic orbitals in molecules: general electronegativity and improvement of Mulliken population analysis. *Physical Chemistry Chemical Physics*, 8: 340–346. DOI: 10.1039/B511516G
- Railsback, L.B., 2003. An Earth scientist's periodic table of the elements and their ions. *Geology* 31: 737–740. ** Versions in English and other languages are available at: <http://www.gly.uga.edu/railsback/PT.html>
- Rudnick, R.L. Gao, S., 2003. The Composition of the Continental Crust, in: *The Crust*, Rudnick, R.L. (Ed.), in: Holland, H.D., Turekian, K.K. (eds.), *Treatise on Geochemistry*, Vol. 3. Elsevier-Pergamon, Oxford, pp. 1–64.
- Taylor, S.R., McLennan, S.M. 1985. *The Continental Crust: Its Composition and Evolution*. Blackwell Scientific Publishers, Oxford, England, 312 pp.
- Turekian, K.K., Wedepohl, K.H., 1961. Distribution of the elements in some major units of the Earth's crust. *Geological Society of America Bulletin* 72: 175–192.
- Velbel, M.A. 2004. Laboratory and homework exercises in the geochemical kinetics of mineral-water reaction: rate law, Arrhenius activation energy, and the rate-determining step in the dissolution of halite. *Journal of Geoscience Education* 52: 52–59.
- Vinogradov, A.P., 1962. Average contents of chemical elements in the principal types of igneous rocks of the Earth's crust. *Geochemistry* 7: 641–664.
- Yokoyama, T., Banfield, J.F., 2002. Direct determinations of the rates of rhyolite dissolution and clay formation over 52,000 years and comparison with laboratory measurements. *Geochim. Cosmochim. Acta* 66: 2665–2681.

Surficial and Environmental Mineralogy

By virtue of their abundance in terms of mass and volume – and their reactivity in soils, sediments, rocks and aqueous systems – minerals and other solids play a vital role in determining partitioning and speciation of elements and compounds. Without considering the relationships between water and minerals it is often difficult to understand or predict the composition of water in aquifers, soils or streams, and in some cases, mineral structures or compositions make them hazardous to human health.

Minerals and the solutions associated with them are dynamic; for example:

- under certain conditions (including redox, pH, dissolved solids, microbial community) a given mineral may dissolve, releasing into solution components that can range from major ions to trace elements;
- under a different set of conditions, a given mineral will crystallize from solution, and in doing so will remove from solution major or trace components.

Some minerals, such as asbestiform amphiboles, are hazardous by mere virtue of their crystal

form, whereas others contain uranium and other radioactive elements that can affect the quality of soil, air and drinking water. Minerals are the primary source of most plant nutrients and often strongly control behavior of toxic substances at the Earth surface.

The surfaces of fine-grained minerals, notably clays and hydroxides (and solid organic matter known as *humus*), contain reactive sites that can adsorb and release ions – in order to understand whether a given contaminant or nutrient will remain in solution, or whether it will adsorb to a mineral surface, you must know the quantities and types of minerals present. Some minerals are very effective at ion exchange and play important roles in contaminant retention and nutrient cycling, while others are virtually non-reactive. So, whether the geochemical question relates to paleoclimate analysis or aquifer geochemistry, nutrient cycling or buffering of acid rain, fate and transport of organic contaminants or sediment diagenesis, minerals and related solids are bound to play an important role. Thus, this chapter first systematically presents information

on minerals in a “textbook” manner then highlights minerals that play important roles in surface and shallow crustal environments.

Various minerals have been referred to thus far in this text – quartz, hematite, magnetite, kaolinite, and calcite are a few examples – but now consider what defines a mineral (after Klein and Hurlbut, 1993):

- Minerals are **naturally occurring** solids formed in soils, volcanoes, sedimentary, igneous and metamorphic rocks, and other geological environments.
 - Minerals are **inorganic** – there are numerous naturally occurring organic solids (e.g. cellulose, organic salts such as sodium acetate), but they are not minerals. At the same time mineralogy can be strongly controlled by biotic factors.
 - Minerals possess ordered **crystalline** lattice structures (repeated arrangement of ordered atoms). One consequence of this is that obsidian (disordered volcanic glass) is not a mineral.
 - They also possess **definite (but not fixed) chemical compositions** – some minerals (e.g. quartz, SiO_2) have nearly fixed compositions, but many minerals vary within certain limits because ions with similar charge and radii (e.g. Na^+ for K^+ or As^{+3} for Fe^{+3}) substitute in crystalline lattice structures.
 - Minerals are **homogeneous solids** with characteristic physical properties – the ordered crystalline lattice expresses itself from the atomic scale to the macroscopic scale and results in characteristic physical properties such as density, hardness, color and crystal shape.
- Of course, for every rule there is likely to be an exception, and for minerals perhaps the best example is calcite that occurs in carbonate shells of marine organisms. Calcite often forms by inorganic processes (via evaporation of water in soils, or crystallization from over-saturated metamorphic fluids) and satisfies all of the criteria for “mineral” listed above, but clearly the biogenic form fails the “inorganically formed” criterion. However, because the calcite in seashells is identical

in composition and atomic structure to inorganic calcite, it is given a waiver of sorts and is considered a mineral. This also applies to minerals that begin to grow on the surfaces of microbial cells.

2.1 INTRODUCTION TO MINERALS AND UNIT CELLS

Minerals generally form when chemical bonds are created among ions dissolved in aqueous solutions (although they can also form by other processes, e.g. from magma or via solid-state transformations from other minerals). Aqueous solutions occur in low-temperature environments like soils and surface waters as well as mid- and high-temperature sedimentary, metamorphic, igneous and hydrothermal systems. In terms of thermodynamics, minerals crystallize to lower the internal energy of a system.

One tangible example of mineral growth (crystallization) occurs in soils and lakes of arid regions, where evaporation drives up concentrations of salt ions (e.g. sodium and chloride) to the point that they become oversaturated and combine to form mineral salts like halite (NaCl), for example. The chemical bonds that form among ions of Na^+ and Cl^- result in an ordered three-dimensional arrangement of atoms; at its

In the world of mineralogy and geochemistry, “low-temperature” environments generally refer to those at or near the Earth surface, whereas high-temperature environments include those deeper in the crust, i.e. igneous and metamorphic systems. Most geochemists consider tropical soils with mean annual temperature of 27 °C and alpine systems with mean annual temperature of 5 °C as low-temperature systems.

most basic, irreducible level, this arrangement produces an entity known as the **unit cell**, which can be defined as the smallest unit that can be repeated in three dimensions to produce the lattice structure of the mineral. The unit cell consists of a specific group of atoms bonded to each other in a specific ratio and geometric arrangement. Repeated in three dimensions, the unit cell produces a crystalline lattice that manifests itself in the form of a **crystal**, and typically the macroscopic crystal that you can hold in your hand possesses the same **form** and **symmetry** as that exhibited by the unit cell. Examples of the cubic halite unit cell are shown in Fig. 2.1, in two dimensions on the left and in three dimensions on the right:

In these diagrams, the Cl^- anion is larger than the Na^+ cation – Na^+ has one less electron shell than does Cl^- .

Minerals are electrically neutral – the sum of positive charges imparted by cations is balanced by the negative charges of anions. This is clearly illustrated in the 2D diagram, where the dashed line represents the boundaries of a unit cell that contains the equivalent of two full chloride anions and two full sodium cations (one complete Cl^- and four quarters of Cl^- atoms, and four

half Na^+ atoms). The composition of this unit cell is thus Na_2Cl_2 , with a net charge of zero. These same concepts are illustrated in three dimensions on the right of Fig. 2.1, where the halite unit cell possesses 6 one-half Cl atoms on each face of the cube, and 8 one-eighth Cl atoms on each corner:

$$(6 * \frac{1}{2}) + (8 * \frac{1}{8}) = 4 \text{ complete Cl atoms.}$$

The center of this unit cell contains one complete Na atom and 12 Na atoms on the edges of the unit cell (one-quarter of each of these 12 atoms rests inside the unit cell):

$$(12 * \frac{1}{4}) + 1 = 4 \text{ complete Na atoms.}$$

So the **stoichiometry** of this 3D unit cell is Na_4Cl_4 , but typically chemical compositions of minerals are reduced to the simplest integer form, so the conventional form of the chemical composition of halite is NaCl . In terms of relationships between unit cells and macroscopic crystals, halite unit cells are cubic, and so are halite crystals (examine table salt under a microscope and you will find small cubes) – the lattice structure exhibited in the unit cell is carried upward in scale to cubic crystals visible to the naked eye.

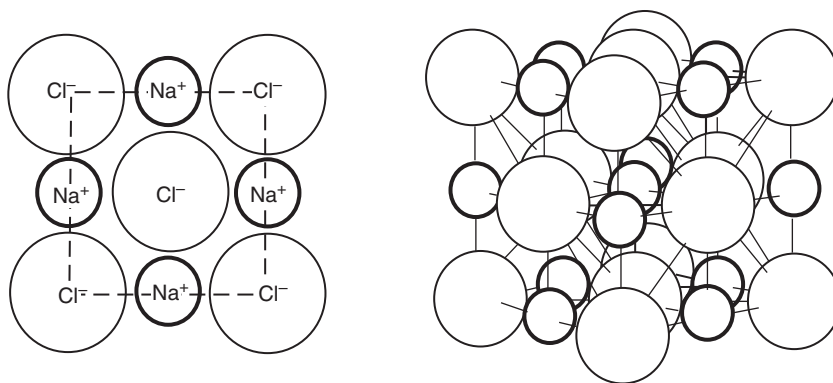


Fig. 2.1 Sketches of halite (NaCl), with a 2-dimensional example on the left and a 3-dimensional example on the right. Note the relative sizes of cations (Na^+ in this case) compared to anions (Cl^- in this case) as well as the ordered crystalline lattice structure.

Figure 2.2 depicts a schematic 2D representation of four adjacent halite unit cells.

This 2D representation contains the equivalent of 8 Na atoms and 8 Cl atoms. The charges balance and the stoichiometry are correct. Repeated billions of times in three dimensions,

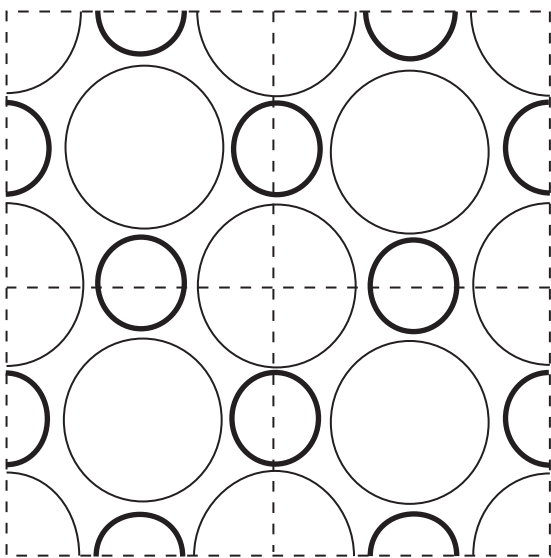


Fig. 2.2 Schematic two-dimensional sketch of a unit cell of halite. Small spheres (circles) are Na^+ and large ones are Cl^- .

the halite unit cell will produce a macroscopic crystal of halite.

2.2 ION COORDINATION, PAULING'S RULES AND IONIC SUBSTITUTION

Ions in crystals occur in various configurations depending on charges and radii of cations and anions – the most common anion by far is oxygen (O comprises ~47% of rocks and sediments of the Earth's crust; Si is second at 28%), although occasionally other anions such as Cl^- , F^- , S^{2-} occur in minerals.

2.2.1 Coordination and radius ratio

Figure 2.3 contains two examples of geometrical arrangements of cations coordinated with anions, one in trigonal coordination and the other in octahedral coordination.

On the left of Fig. 2.3, the cation occurs in **trigonal coordination**, a planar configuration with three oxygen atoms and a cation in the same plane (effectively, the plane of the page). The **coordination number** for the cation is 3. Set three

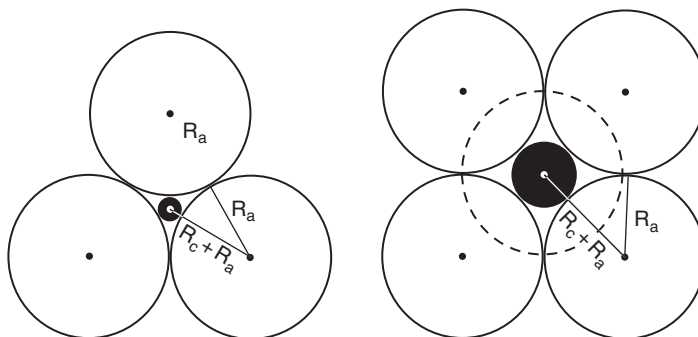


Fig. 2.3 Sketch of cations in three-fold (trigonal, on left) and six-fold (octahedral, on right) coordination. The cation in each case is represented by the small black circle with a white center representing the nucleus, and the anions (oxygen in most minerals) are the large white circles (not to scale). The dashed circle in the sketch on the right represents one anion above the plane of the four anions, as well as one anion below the plane of the four anions.

baseballs on a table as close as possible with a marble in the center and you have a trigonally coordinated marble. In the diagram on the right of Fig. 2.3, the cation occurs in **octahedral coordination**, an arrangement that consists of three planes of atoms, one with four O atoms that form a square around a central cation (in the plane of the page), and two planes that each contain one O atom, one above and one below the page – these O atoms fit in the dimple created by the 4 O atoms in the square. These basal and apical O atoms are represented in the diagram by the dashed circle. The **coordination number** for octahedral coordination is 6. You would need six baseballs and a larger marble to visualize octahedral coordination.

Geochemists and mineralogists commonly express relationships between anions and cations by the **radius ratio**, an approach devised by Linus Pauling in the 1920s and defined as the ratio of the radius of the cation to the radius of the anion. Table 2.1 presents crystal radii and ionic radii for common cations and anions, where an **angstrom (Å)** is equal to 10^{-10} m. Values are presented for both crystal radii and ionic radii. Crystal radii reflect the space occupied by an ion in a crystal lattice whereas ionic radii correspond to the inherent radius of the ion unaffected by distortions associated with bonding in a crystal. Both are used in geochemistry, and unfortunately it is often not clearly defined which is used.

Note that dimensions of unit cells are commonly expressed in Å, and sometimes in nanometers (where $1 \text{ nm} = 10^{-9} \text{ m}$). Pauling used X-ray diffraction of minerals to determine the radii of ions in crystals, where the radius is defined as the distance from the nucleus to the valence electron shell, and the atom is envisioned basically as a hard sphere. (Note in Table 2.1 that the radius of ions varies a little depending on the coordination number.)

In both diagrams in Fig. 2.3, R_a represents the ionic radius of the anion and R_c is the cation ionic radius, but because the cation is so small compared to the anion, conventional diagrams show

$R_c + R_a$. Regardless, in these diagrams R_c is identified by the white line segment that crosses the cation (the segment becomes black inside the nucleus for visibility).

For a cation in trigonal 3-fold coordination, the range of ionic radius ratios (R_c/R_a) is 0.155 to 0.225. If the ionic radius ratio is smaller than 0.155, the cation is too small, bond lengths between cation and anions are too great, and the bond is unstable; in a sense, the cation would rattle around inside the interstice created by the anions and anions would be too close to each other to be stable. Conversely, if the radius ratio is greater than 0.225, the cation is too large for trigonal coordination – it would distort the triangle to such an extent that the bonds would be unstable (the coordinating anions would be separated too far from each other). In this case, the cation would likely be better accommodated in 4-fold tetrahedral coordination.

Determining the radius ratio of Si (ionic radius = 0.34 Å) in 4-fold coordination relative to O (ionic radius = 1.30 Å) gives a value of 0.26, a perfect match for tetrahedral coordination (Fig. 2.4). The radius ratio of NaCl ($1.10 \text{ Å} / 1.72 \text{ Å}$) is 0.64, indicating that Na atoms should occur in 6-fold coordination with Cl. In Fig. 2.1 (3D diagram on right), one Cl atom is directly above the central Na atom, one is directly below, and four Cl atoms are in the same horizontal plane as the Na atom, so visually and by radius ratio we see that the coordination number (CN) for Na in halite is 6.

Figure 2.4 presents some common examples of cation sites and geometries in minerals. Not all possibilities are shown, and note that there are three common cubic geometries. The one presented is body-centered cubic, but other varieties include face-centered cubic (e.g. halite) and edge-centered (also known as simple) cubic. Furthermore, large cations can exist in sites with a coordination number of 12, a good example being the hexagonal close-packed potassium ions (K^+) that occur in interlayers of micas.

Table 2.1 Radii of selected major and trace element ions in minerals. For some elements with multiple oxidation states (e.g. Cr, Mn), selected oxidation states that commonly occur in minerals are given. Ions are arranged from lowest charge (O^{-2}) to highest charge, first for common rock-forming minerals (O through Si), followed by heavy metals plus As, then other useful ions (S through U). Note that coordination number (CN) corresponds to geometry, i.e. CN of 4 produces tetrahedral geometry, CN of 6 produces octahedral geometry, CN of 8 produces cubic or dodecahedral geometry, and CN = 12 results in close-packed (CP) geometry, as explained in the text. (Shannon 1976.)

Ion	CN	Geometry	Crystal radius (Å)	Ionic radius (Å)
O^{-2}	4	Tet	1.24	1.38
O^{-2}	6	Oct	1.26	1.40
K^{+1}	8	Cubic, dodec	1.65	1.51
K^{+1}	12	CP	1.78	1.64
Na^{+1}	6	Oct	1.16	1.02
Na^{+1}	8	Cubic, dodec	1.32	1.18
Mg^{+2}	6	Oct	0.86	0.72
Ca^{+2}	6	Oct	1.14	1.00
Ca^{+2}	8	Cubic	1.26	1.12
Fe^{+2}	6	Oct	0.92	0.78
Fe^{+3}	6	Oct	0.79	0.65
Al^{+3}	4	Tet	0.53	0.39
Al^{+3}	6	Oct	0.68	0.54
Si^{+4}	4	Tet	0.40	0.26
Mn^{+2}	6	Oct	0.97	0.83
Mn^{+4}	6	Oct	0.67	0.53
Co^{+2}	6	Oct	0.89	0.75
Ni^{+2}	6	Oct	0.83	0.69
Cu^{+2}	6	Oct	0.87	0.73
Zn^{+2}	6	Oct	0.88	0.74
As^{+3}	4	Tet	0.54	0.40
As^{+3}	6	Oct	0.72	0.58
As^{+5}	4	Tet	0.48	0.34
As^{+5}	6	Oct	0.60	0.46
Cr^{+3}	6	Oct	0.76	0.62
Cr^{+6}	4	Oct	0.40	0.26
Ti^{+4}	6	Oct	0.75	0.61
S^{-2}	6	Oct	1.7	1.84
Cl^{-1}	6	Oct	1.67	1.81
Rb^{+1}	8 to 12	Cubic, dodec, CP	1.75–1.86	1.61–1.72
Sr^{+2}	8 to 12	Cubic, dodec, CP	1.40–1.58	1.26–1.44
Ba^{+2}	8 to 12	Cubic, dodec, CP	1.56–1.75	1.42–1.61
Pb^{+2}	8 to 12	Cubic, dodec, CP	1.43–1.63	1.29–1.49
Ra^{+2}	8 to 12	Cubic, dodec, CP	1.62–1.84	1.48–1.70
Zr^{+4}	6	Oct	0.86	0.72
Th^{+4}	8	Cubic, dodec	1.19	1.05
U^{+4}	6 to 8	Oct, cubic	1.03–1.14	0.89–1.0
U^{+6}	6	Oct	0.87	0.73

Oct = octahedral; dodec = dodecahedral; Tet = tetrahedral; CP = closest-packed; CN = coordination number. Transition metals can occur in high-spin or low-spin orbital configurations. At the Earth surface and shallow crust, high-spin states occur and thus high-spin values are given where applicable (e.g. $Fe^{+2,+3}$, Co^{+2} , Mn^{+2}).

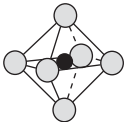
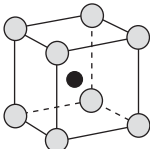
CN	Ion arrangement	Radius ratio		Example
2	Linear	<0.155		Carbon dioxide
3	Trigonal	0.155–0.225		C in carbonate
4	Tetrahedral	0.225–0.414		Si, Al in silicates (As, P)
6	Octahedral	0.414–0.732		Mg, Fe, Al, Ni, other heavy metals, As
8	Cubic e.g. body-centered	0.732–1.00		Large ions, e.g. Ca, K, Na (U, Th, Pb)

Fig. 2.4 Examples of geometric arrangement of cations and anions in minerals (or gases, e.g. CO₂), showing range of radius ratio (ratio of ionic radius of cation : anion) for each configuration.

2.2.2 Bond-strength considerations

Close packing of ions in crystals is important because bond strength, expressed as the energy of electrostatic attraction (F), is determined by the product of charges (absolute values) of the ions involved in the bond (e_1 , e_2) divided by the radius (r) squared, where the radius is the distance between nuclei of adjacent atoms in a chemical bond (Faure 1998):

$$F = (e_1 * e_2) / r^2 \quad (2.1)$$

From this equation it is clear that minimizing distance between nuclei of cations and anions enhances the strength of the bond, thus ions strive to coordinate themselves in close-packing arrangements. Chemical bonds also tend to be stronger when they involve ions of greater charge magnitude, so a bond between Na⁺ and Cl[−] would be

expected to be weaker than a Si⁺⁴–O^{−2} bond based on charge considerations alone. If we also consider radius of the Na–Cl bond (**2.8 Å**) compared to the radius of the Si–O bond (**1.6 Å**), we can arrive at a quantitative comparison of bond strengths:

$$F_{\text{Si-O}} = (4 * 2) / (1.6^2) = 3.1$$

$$F_{\text{Na-Cl}} = (1 * 1) / (2.8^2) = 0.13$$

Refer also to the example of ionic bond strength presented in Equations 1.10 and 1.11 in Chapter 1.

2.2.3 Pauling's and Goldschmidt's rules of ionic solids

The brilliant 20th-century mineralogist, chemist and two-time Nobel Prize winner Linus Pauling

devised a systematic approach to predict crystal chemistry and ionic arrangements in minerals (Pauling, 1960). The following presents what are commonly referred to as **Pauling's Rules for Ionic Crystals**:

- Rule 1 – the **Coordination Principle**: A polyhedron of coordinated anions (typically oxygen) forms around each cation, and the cation–anion distance and type of coordination polyhedron that results is determined by the **radius ratio** (section 2.2.1);

- Rule 2 – the **Principle of Local Charge Balance**: The most stable ionic structures are those for whom the sum of the strengths of the electrostatic bonds in a polyhedron equals the charge on the anion (typically, O^{2-}). In short, charges must balance. The first equation that governs this rule determines the value for S , the strength of an electrostatic bond,

$$S = e_{\text{cation}} / \text{CN} \quad (2.2a)$$

where e_{cation} is the charge on the cation, and CN is coordination number, and;

$$e_{\text{anion}} = \sum S_i \quad (2.2b)$$

where e_{anion} is the charge on the anion and S_i is the sum of the strengths of the electrostatic bonds. If $e_{\text{anion}} = \sum S_i$, then the crystal (polyhedron) is stable in terms of charge balance. As $\sum S_i$ values become progressively distant from 1, the crystal becomes progressively less stable, or it requires additional components to counter the charge imbalance (see calcite, below). If we return to the example of NaCl (Fig. 2.5), each Cl^{-1} ion is surrounded by six Na ions, so $S = 1/6$, and if we sum a bond strength (S) of $1/6$ over six Na–Cl bonds, we see that $\sum S_i = 1 = e_{\text{anion}}$.

The coordination of 6 Na about a Cl anion produces a stable configuration. Now if we consider C^{+4} in trigonal coordination with O^{2-} (Fig. 2.6), we observe that $S = 4/3$, and $S_i = 4/3 * 3 = 4$; $e_{\text{anion}} = -2$.

Clearly, trigonal coordination of CO_3 does not produce a charge-balanced ionic crystal. Three O^{2-} atoms combined with one C^{+4} atom produces $+4 - [3 * (-2)] = -2$; the -2 charge produces the CO_3^{2-} (carbonate) anion. In minerals, the charge imbalance is satisfied by inclusion of a divalent cation such as Ca^{+2} , Mg^{+2} or Fe^{+2} in the crystal structure to produce a carbonate mineral such as calcite (CaCO_3), magnesite (MgCO_3) or siderite (FeCO_3). Other examples of complex anions that

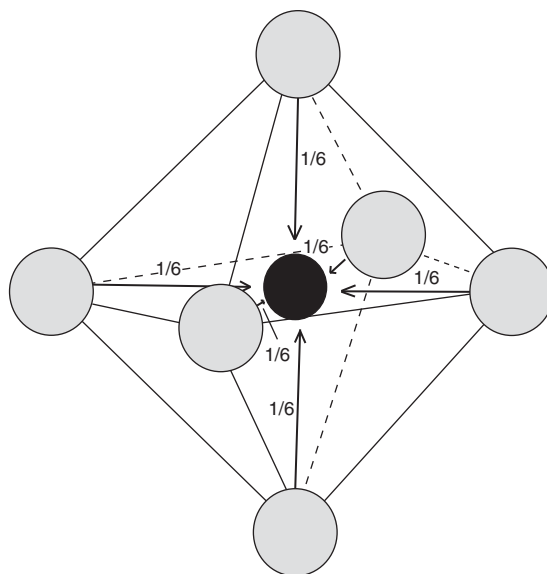


Fig. 2.5 Halite showing coordination of six Cl^{-} ions around a central Na^{+} ion.

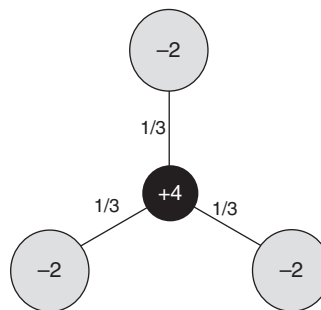


Fig. 2.6 Trigonal coordination of three oxygen atoms (O^{2-}) around a central carbon (C^{+4}) atom.

combine with cations to form minerals include silica (SiO_4^{-4}), phosphate (PO_4^{-3}) and sulfate (SO_4^{-2}).

• **Rule 3 – Sharing of Polyhedral Corners and Faces** reduces the stability of ionic crystals. This is so because sharing of faces and, to an even greater extent, edges, brings cations closer together. Cations strongly repel other cations, and the electrostatic repulsion between neighboring cations causes destabilization. In reality, sharing of polyhedra is common, but the most stable of these sharing arrangements are those where only one anion per polyhedron is shared with an adjacent polyhedron. The two examples in Fig. 2.7 illustrate this principle with tetrahedra:

The shorter cation–cation distance in the potential crystal on the right of Fig. 2.7 emphasizes the point that shared faces produce unstable arrangements due to cation repulsion.

• **Rule 4 – Valency and Sharing of Polyhedral Components:** Cations with high charge and small radius tend not to share polyhedra because of ensuing repulsion and because the high valency of the cation will limit the amount of charge available to the anion to distribute among other polyhedra (e.g. C^{+4} CO_3^{-2} , P^{+5} in PO_4^{-3} , and

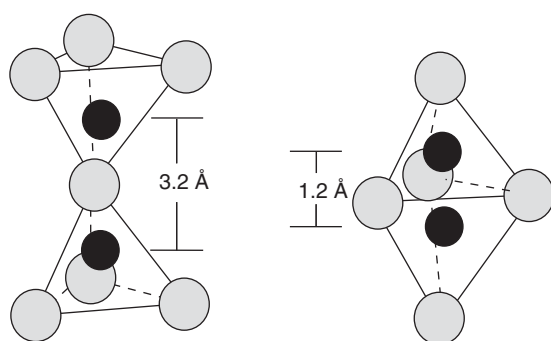


Fig. 2.7 Two potential arrangements of adjacent tetrahedra with cations (e.g. Si^{+4}) as black spheres and anions (e.g. O^{-2}) as gray spheres. Note that the arrangement on the left produces a greater distance between cations (3.2 Å) compared to the arrangement on the right (1.2 Å), so Pauling's Rule #3 indicates that the structure on the left is more stable.

S^{+6} in SO_4^{-2}). This rule is basically an extension of Rule 3.

• **Rule 5 – the Principle of Parsimony:** Although numerous different types of cation–anion arrangements, like those presented in Fig. 2.4 and Table 2.1, occur in crystalline solids, the number of essentially different kinds of constituents in a crystal tends to be small. For example, quartz and halite have only one structural arrangement each, and even complex minerals like the phyllosilicate clays possess only two to three cation–anion arrangements (commonly, tetrahedral, octahedral and 12-fold close-packed interlayer sites).

An additional implication of Pauling's Rules is that **ionic substitutions in crystals tend to occur when ions are of similar radii, ionic charge and electronegativity**. Na^+ and K^+ commonly substitute for each other in minerals; Mg^{+2} and Fe^{+2} commonly can occupy the same sites in minerals, as can the pairs Na^+ and Ca^{+2} , Fe^{+3} and Al^{+3} , Ni^{+2} and Mg^{+2} , Cr^{+3} and Fe^{+3} , and Al^{+3} and Si^{+4} . Note that in some cases, cations of slightly different charge can substitute for each other, provided that the charge imbalance is rectified elsewhere in the crystal. For example, in certain silicates, Ca^{+2} can substitute for Na^+ , resulting in a net gain of +1, provided that a paired substitution occurs, e.g. Al^{+3} for Si^{+4} or Mg^{+2} for Al^{+3} (to balance the gain of +1).

Rules of ionic substitution (also known as isomorphous substitution) in minerals were devised by Victor M Goldschmidt and are as follows:

- ionic radii must differ by $\leq 15\%$;
- ionic charge can differ by ± 1 if balanced by a paired substitution (Ca^{+2} for Na^+ , balanced by Al^{+3} for Si^{+4}), and rarely by ± 2 if balanced by paired substitution;
- ions with higher ionic potential (ionic charge : ionic radius) tend to occupy sites preferentially over those with lower ionic potential;
- the substituting ions must be able to form similar chemical bonds (they should have similar values of electronegativity).

One way that these rules can be applied to environmental geochemistry is to predict trace elements that might occur in common minerals. Consider iron oxides and hydroxides, minerals where ferric iron (Fe^{+3}) occurs in octahedral coordination. Using Pauling's and Goldschmidt's Rules, you should be able to predict that many trace elements can substitute for Fe^{+3} in minerals – trace elements with a charge within ± 1 of Fe^{+3} and radii within 15% of $0.65 \text{ \AA}$ include Ni^{+2} , Cu^{+2} , Co^{+2} , Zn^{+2} , As^{+3} , Cr^{+3} . Those that don't substitute for Fe^{+3} include Cr^{+6} (charge too high and radius too low) and Pb^{+2} (radius too large). A similar treatment could be applied to the potential for trace metals to substitute for Fe^{+2} in pyrite or for Ca^{+2} in calcite, and many more.

2.3 SILICATES

Silicates are by far the most abundant class of minerals in the crust and at the surface of the Earth. The two most abundant elements in the crust, including rocks, soils and sediments, are oxygen (46.6% by weight) and silicon (27.7% by weight), with Al, Fe, Ca, Mg, Na, K and Ti comprising nearly all of the remainder. Stated in terms of oxide weight percent, SiO_2 comprises approximately ~59% of the crust (Table 2.2) and it is for this reason, but not this reason alone, that silicate minerals are so abundant. The other reason is that silicon and oxygen combine to form one of the most stable and common bonds in nature. The polar covalent bond between Si^{+4} and O^{-2} forms the **silica tetrahedron**, which serves as the basis for all silicate minerals, whether they be high-temperature igneous minerals or low-temperature clay minerals that form in soils, sediments and sedimentary rocks.

The silica tetrahedron is comprised of a central Si atom surrounded by four O atoms; stoichiometrically, it is SiO_4^{-4} . The tetrahedral configuration maximizes distance between negatively charged

O^{-2} atoms while also minimizing distance between O^{-2} and Si^{+4} atoms (Fig. 2.8).

Atoms are not to scale in the drawing on the left to emphasize the tetrahedral arrangement of the atoms; in the diagram on the right, atoms are approximately to scale and emphasize (1) close packing of atoms, and (2) the fact that the small Si atom rests snugly in the interstice produced by the four O atoms. Four equal-sized spheres can help to visualize the arrangement of

Table 2.2 Composition of the Earth's crust presented as both wt% element (left) and wt% oxide (right). The difference between the sums listed and 100% reflects the abundance of all other elements.

Element (wt%)		Oxide (wt%)	
Oxygen	46.6	NA	
Silicon	27.7	SiO_2	59.3
Aluminum	8.1	Al_2O_3	15.3
Iron	5.0	Fe_2O_3	7.1
Calcium	3.6	CaO	5.0
Sodium	2.8	Na_2O	3.8
Potassium	2.6	K_2O	3.1
Magnesium	2.1	MgO	3.5
Titanium	0.60	TiO_2	1.0
Phosphorus	0.13	P_2O_5	0.30
Manganese	0.1	MnO	0.15
<i>SUM</i>	99.3	<i>SUM</i>	98.6

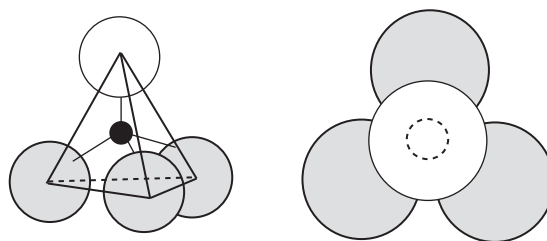


Fig. 2.8 Representations of the silica tetrahedron (SiO_4^{-4}). The three basal oxygen atoms are shaded gray, the apical oxygen is clear, and the silicon atom is black. The view of the example on the right is from above, looking down at the apical oxygen with three basal oxygens below (hidden Si atom is represented by dashed circle).

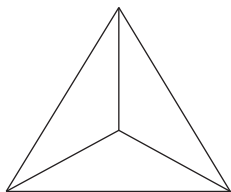


Fig. 2.9 Schematic representation of the silica tetrahedron.

the O atoms as well as the space occupied by the Si atom.

To simplify, the silica tetrahedron is often represented by the sketch in Fig. 2.9, where the four edges of the pyramid represent O atoms and the Si atom is implicitly in the center.

The following presentation of silicate minerals will use this symbol. This section on silicates will be presented systematically, beginning with the relatively simple nesosilicates, skipping a pair of groups (sorosilicates, cyclosilicates) and progressing to groups with progressively greater sharing of silica tetrahedra (inosilicates and phyllosilicates), eventually to the highly polymerized tectosilicates (where **polymerization** is extensive sharing of oxygen atoms among silica tetrahedra).

2.3.1 Nesosilicates

The simplest silicates in terms of structure and composition are the **nesosilicates** (also referred to as **orthosilicates**), a group that consists of isolated silica tetrahedra whose negative charge is balanced by cations located between tetrahedra. Unlike the other silicate groups, there is no polymerization of silica tetrahedra in nesosilicates – no oxygen atoms are shared between tetrahedra.

For the purpose of illustration, **olivine** is presented as an example of this group (Fig. 2.10). The olivine group is comprised of two end-members, forsterite (Mg_2SiO_4) and fayalite (Fe_2SiO_4). Note that the tetrahedra appear to exist as islands amongst divalent Mg (or Fe) cations.

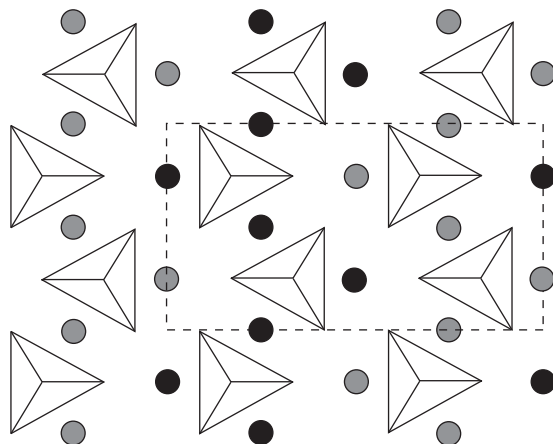


Fig. 2.10 Sketch of an olivine crystal showing isolated silica tetrahedra (triangles) as well as atoms of Mg or Fe represented by black and gray circles. Dotted line indicates parameters of an olivine unit cell.

The cations are indicated by black and gray circles, where the difference indicates something about the three-dimensional character of the olivine structure – in order to minimize cation–cation repulsion, the gray cations are located above the plane of the page while the black ones are below the plane of the page. One way to visualize an olivine unit cell is indicated by the dashed rectangle, within which are 4 complete silica tetrahedra, 4 complete Mg atoms, and 8 one-half Mg atoms, for a sum composition of $\text{Mg}_8(\text{SiO}_4)_4$, which is simplified to a formula unit of Mg_2SiO_4 .

Olivines are high-temperature minerals that commonly form in basaltic magmas and tend to decompose rapidly in weathering environments (e.g. soils). Examples of other nesosilicates include **zircon** (ZrSiO_4), **garnet** group minerals (which have the general formula $\text{X}_3\text{Y}_2(\text{SiO}_4)_3$, where X and Y are divalent and trivalent cations, respectively), and the Al_2SiO_5 polymorphs **andalusite**, **kyanite** and **sillimanite** (**polymorphs** have identical composition but different structures).

Examples of nesosilicates in low-temperature/environmental geochemistry include (1) rapid weathering of olivine in soils leads to formation

of clays and hydroxides, and (2) uranium and thorium can substitute for Zr in the zircon structure, which means that trace amounts of U or Th or their radioactive decay products could be released into soils or groundwater should the U or Th (or their decay products) be released from zircon crystals during chemical weathering or alpha recoil. The presence of U also makes zircon a valuable tool for dating rocks using the U–Pb radiometric dating method (Chapter 11).

2.3.2 Inosilicates

Linear chains of linked silica tetrahedra form the foundation of the **inosilicate** class of minerals, including: the **pyroxenes**, which consist of single chains of silica tetrahedra and **amphiboles**, which are composed of double chains of silica tetrahedra. In the single chains (Fig. 2.11), each silica tetrahedron shares two basal oxygen atoms, one with each adjacent tetrahedron, producing the fundamental unit SiO_3^{2-} – if two oxygen atoms are equally shared with adjacent tetrahedra, each

tetrahedron possesses two full O atoms and 2 one-half O atoms. The -2 charge is balanced by divalent cations such as Ca, Mg, Fe and in some cases, pairs of (Na^{+1} and Al^{+3}) or (Na^{+1} and Fe^{+3}).

Common pyroxenes include the clinopyroxenes augite ($[\text{Na,Ca,Mg,Fe}] \text{SiO}_3$), diopside ($\text{CaMgSi}_2\text{O}_6$) and jadeite ($\text{NaAlSi}_2\text{O}_6$) and the orthopyroxenes enstatite (MgSiO_3) and ferrosilite (FeSiO_3). Pyroxenes are common rock-forming minerals, comprising significant parts of numerous igneous and metamorphic rock types. A schematic sketch of a generalized pyroxene is presented in Fig. 2.11, where black spheres are cations and two single chains of tetrahedra are shown.

Note how cation–cation repulsion is minimized by positioning of cations relative to each other and to silica tetrahedra.

In **amphibole double chains**, one-half of the tetrahedra share two oxygen atoms (like those in pyroxenes), while the other half share 3 oxygen atoms (all but the apical oxygen). The result is linear chains of hexagonal rings that produce the $\text{Si}_4\text{O}_{11}^{6-}$ fundamental unit (often expressed as

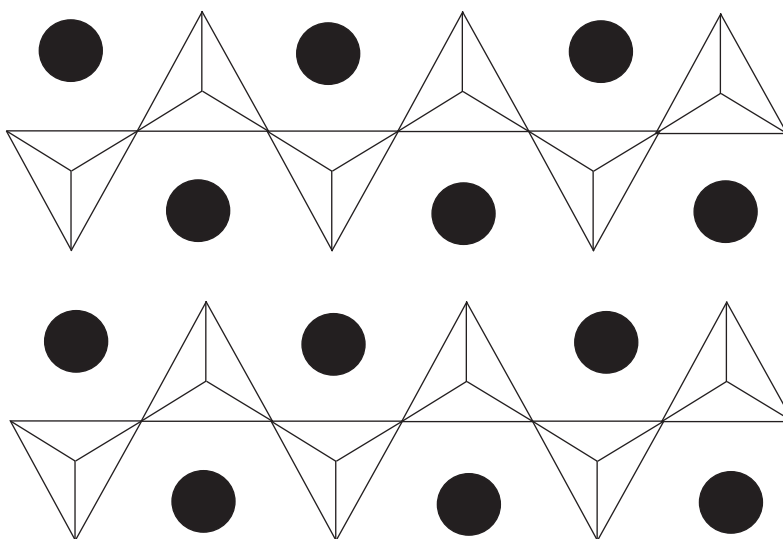


Fig. 2.11 Sketch of pyroxene group structure. Note linked silica tetrahedra that form chain structures, and occurrence of cations (e.g. Ca^{+2} , Mg^{+2} , Fe^{+2} , Na^{+}) in interstices.

Si_8O_{22} in many structural formulae). Shown in Fig. 2.12 is a schematic of an amphibole depicting a double-chain with relatively large cations within the hexagonal rings and smaller cations in sites outside of the double chain. What are not shown in this picture are the numerous different cation sites, positions of hydroxyl (OH^-) groups, and any real information about 3-dimensional structure.

Amphiboles are among the most complex minerals in terms of stoichiometry. Two relatively simple amphiboles, in terms of composition, are **tremolite**, $\text{Ca}_2\text{Mg}_5\text{Si}_8\text{O}_{22}(\text{OH})_2$, and **anthophyllite**, $(\text{Mg},\text{Fe})_7\text{Si}_8\text{O}_{22}(\text{OH})_2$. The most common

amphibole, hornblende, has the more complex formula $(\text{Na},\text{Ca})_{2-3}(\text{Mg},\text{Fe},\text{Al})_5(\text{Al},\text{Si})_8\text{O}_{22}(\text{OH})_2$, where the range from 2 to 3 cations in the Na–Ca site depends on the ratio of Na:Ca.

Some amphiboles, particularly **crocidolite** ($\text{Na}_2\text{Fe}^{2+}_3\text{Fe}^{3+}_2\text{Si}_8\text{O}_{22}(\text{OH})_2$) and to a lesser extent tremolite, amosite, anthophyllite and actinolite, are a human health hazard when inhaled as dust because (1) the shape of the fibers (Fig. 2.13) makes them difficult to cough out, (2) amphiboles are relatively stable in lungs and thus persist for long periods of time, and (3) factors 1 and 2 can result in tissue damage that leads

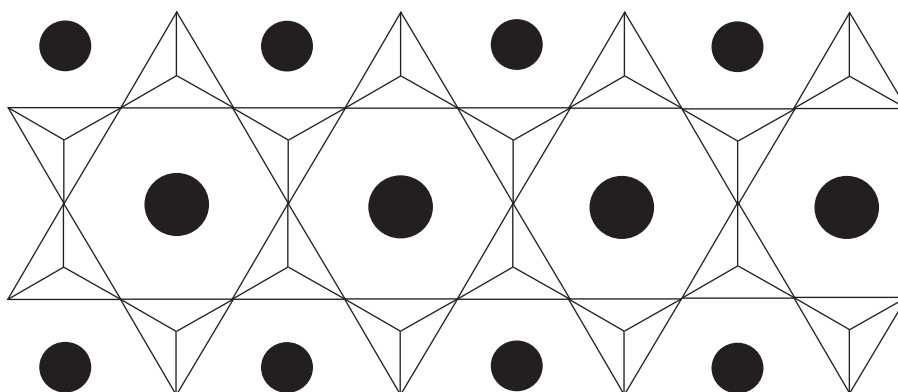


Fig. 2.12 Sketch of amphibole group structure. Note cross-linked chains of silica tetrahedra and occurrence of cations (e.g. Ca^{+2} , Mg^{+2} , Fe^{+2} , Na^+) in interstices. Larger sites mainly hold Na^+ or Ca^{+2} whereas smaller sites hold Al^{+3} or Mg^{+2} .

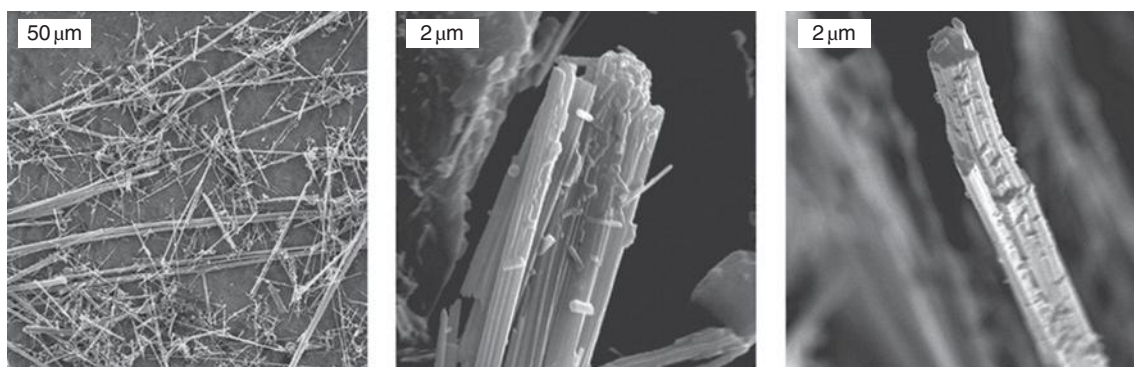


Fig. 2.13 Morphology of asbestiform amphibole. (US Geological Survey Microbeam Lab; <http://usgsprobe.cr.usgs.gov/picts2.html>).

to development of asbestosis, mesothelioma or lung cancer.

In reality, all silicate mineral dusts are hazardous to some extent, but the features mentioned above make *asbestiform amphiboles* particularly problematic.

2.3.3 Phyllosilicates

Phyllosilicates, as the name implies (the greek word for leaf is φύλλο [phyllo]), are sheet silicates formed by extensive 2-dimensional linkages of silica tetrahedra that extend virtually infinitely. All three basal oxygens are shared, leaving only the apical oxygen not involved in silica polymerization, thus producing a fundamental stoichiometric unit of $\text{SiO}_2^{-2.5}$, which is expressed in mineral formulas as $(\text{Si}_2\text{O}_5^{-2})$ or $(\text{Si}_4\text{O}_{10}^{-4})$. Al can substitute for Si in phyllosilicates in up to 50% of tetrahedral sites, although typically Al for Si substitution occurs in $\leq 25\%$ of total tetrahedral sites.

Phyllosilicates figure prominently in environmental geochemistry due to their ability to adsorb substances, especially cations, and their relatively high reactivity. Both attributes are due in part to the very fine-grain size (e.g. $< 2 \mu\text{m}$) of many phyllosilicates. Members of this mineral group are common in soils, sediments, hydrothermal systems and in sedimentary rocks, and via exchange and dissolution–precipitation reactions they influence the composition of surface water, groundwater, soils and sediments.

Figure 2.14 is a view looking down the crystallographic *c*-axis onto a sheet of silica tetrahedra. Three O atoms per tetrahedron are in the same plane as the page, and the fourth (the apical oxygen) sits above the plane of the page. By virtue of the fact that this tetrahedral sheet is not charge balanced, it must form chemical bonds to satisfy the residual negative charge. As a consequence, all phyllosilicates contain two main structural elements, tetrahedral sheets and octahedral sheets.

Tetrahedral sheets are comprised of silica tetrahedra (often with some Al substituted for Si) like the one depicted in Fig. 2.14, whereas **octahedral sheets** are comprised of cations (commonly, Al, Fe, and Mg, but many trace metals also fit – see Table 2.1) in six-fold coordination with a combination of oxygen atoms and OH^- (hydroxyl) anions; three anions form a triangle above the cation and three form a triangle below the cation, producing 6-fold octahedral geometry.

Octahedral configurations are shown in Fig. 2.15. The sketches attempt to show the same

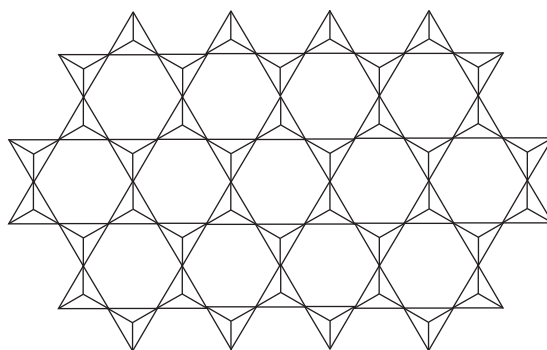


Fig. 2.14 Arrangement of silica tetrahedra in phyllosilicates, showing laterally extensive development of sheet-like structures. This diagram shows a top-down view onto a tetrahedral sheet of a phyllosilicate.

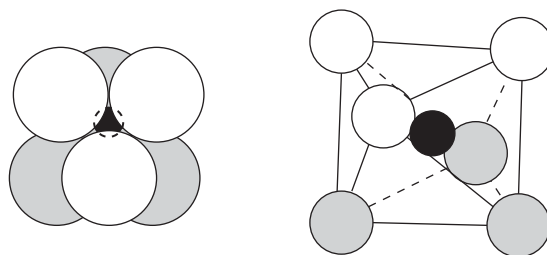
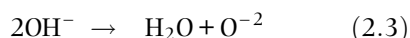


Fig. 2.15 Sketches of octahedral coordination of anions (typically O^{2-} or OH^-) around a cation (black). Anions are gray and white to allow the viewer to correlate the basal plane (gray spheres) to the upper plane (white spheres).

structure in two common ways to emphasize (1) the existence of two planes of trigonally coordinated anions (e.g. O^{2-}) sandwiching a cation (e.g. Si^{+4}), and (2) the way in which lines can be drawn between anions to emphasize the 8 sides from which the term octahedron is derived (diamonds are octahedral).

The minerals brucite [$Mg_3(OH)_6$] and gibbsite [$Al_2(OH)_6$] have structures that are very similar to phyllosilicate octahedral sheets, except for the fact that in these minerals all anions are OH^- , whereas in phyllosilicate octahedral sheets 1/3 to 2/3 of octahedral sheet anions are apical oxygen atoms from the tetrahedral layer and the rest are OH^- .

The presence of hydroxyl groups makes phyllosilicates (like amphiboles) **hydrous**, that is, they possess water in their structures in the form of hydroxyl (OH^-) groups. If heated to $600^\circ C$, the hydroxyl anions in hydrous minerals combine as follows to produce H_2O .



Their potential to dehydrate when heated is an important consideration in long-term nuclear waste disposal planning (i.e. will heat from radioactive decay alter impermeable clay barriers?), and also, water can be produced from dehydrating micas or chlorite (see below) during diagenesis and metamorphism, leading to regional fluid-flow events.

Figure 2.16 presents a view looking down the b -axis (edge-on), showing the relationship of tetrahedral and octahedral sheets in mica. In Fig. 2.14, the tetrahedral sheet shown is akin to looking down upon the upper tetrahedral sheet in Fig. 2.16. The hexagonal ring formed by linked tetrahedra is visible in both.

On the left of Fig. 2.16, Si (or Al) atoms in tetrahedral layers are small black circles, while octahedral cations are slightly larger black circles. O atoms are small white circles and OH^- anions are gray circles. Figure 2.16 depicts two **2:1 layers** on the left and two 2:1 layers on the right. The 2:1 layers are comprised of a pair of tetrahedral sheets sandwiched around an octahedral

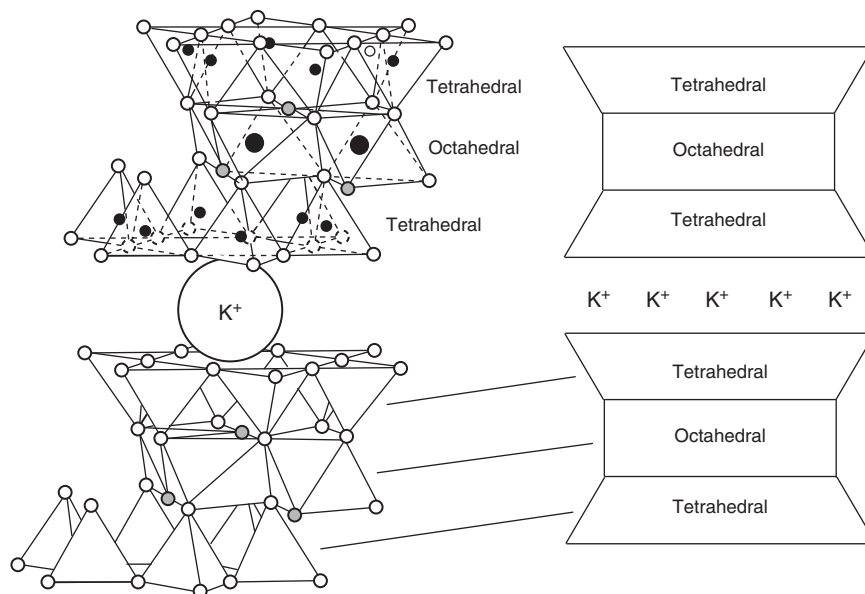


Fig. 2.16 Two representations of a mica, both showing tetrahedral and octahedral sheets as well as interlayer potassium. (Adapted from Moore and Reynolds 1997.)

sheet. The upper 2:1 layer on the left is sketched in a more transparent manner than the lower one in an attempt to show cation positions and cation–anion geometry. The 2:1 layers, in turn, are sandwiched around an **interlayer cation**, in this case K^+ . Clearly the atoms are not drawn to scale – for purposes of illustration, the K ion is drawn larger to show how it fits into the hexagonal holes of bounding tetrahedral sheets. The simplified sketches on the right were developed by Pauling to represent phyllosilicate structures.

Octahedral sheets are termed **dioctahedral** if they are comprised of two atoms of Al^{+3} (or, less commonly, two atoms of Fe^{+3}) per unit cell, and

trioctahedral if they are comprised of three atoms of Mg^{+2} (and/or Fe^{+2}) per unit cell.

The 2:1 Phyllosilicates

Pyrophyllite forms in Al-rich, low-grade metamorphic rocks and has the formula $Al_2Si_4O_{10}(OH)_2$. Its trioctahedral analog **talc** is $Mg_3Si_4O_{10}(OH)_2$. These two minerals will serve as a starting point for presenting different types of phyllosilicates. Pyrophyllite and talc can be represented schematically (Fig. 2.17), with cations per unit cell shown for tetrahedral and octahedral layers:

Positive charges ($2 * Al^{+3} + 4 * Si^{+4} = 22$; or $3 * Mg^{+2} + 4 * Si^{+4} = 22$) and the negative charges ($10 * O^{-2} + 2 * OH^{-1} = -22$) balance. Note that the interlayer is vacant, and the main force bonding the 2:1 layers is van Der Waals bonds. Figure 2.18 is pyrophyllite compared to mica (with interlayer K^+) – note difference in d (001), the spacing along the c-axis.

The micas can be developed conceptually by starting with a pyrophyllite unit cell and substituting one Al atom for one Si atom. The stoichiometry change is:

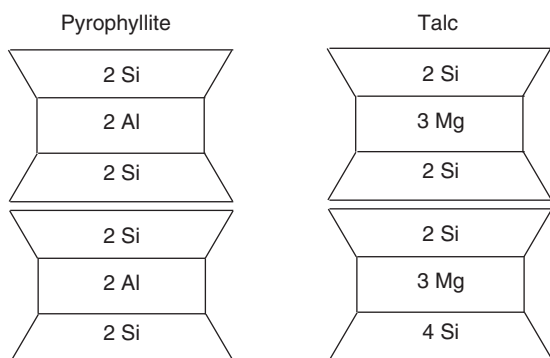
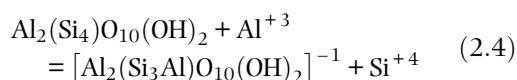


Fig. 2.17 Schematic of pyrophyllite (left) and talc (right). Note (a) substitution of 3 Mg (talc) for 2 Al (pyrophyllite) in octahedral sheet, and (b) absence of interlayer cations in both.

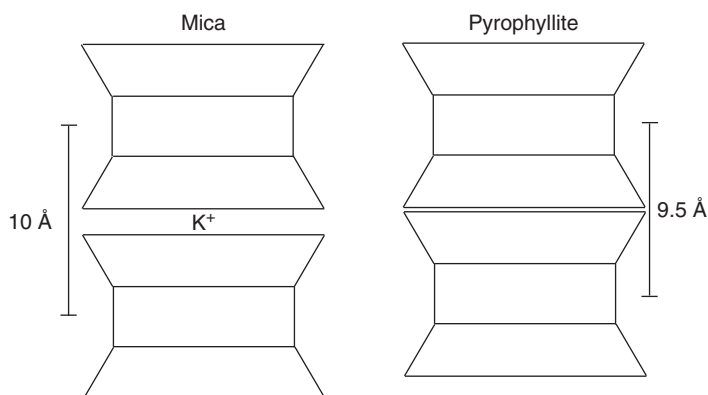


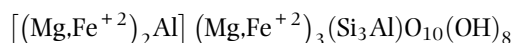
Fig. 2.18 Schematic sketches of mica and pyrophyllite. Note effect of presence/absence of interlayer potassium on the thickness of the unit cells in the c direction.

This produces a net charge of -1 of the 2:1 layer. No mineral can exist if it is not charge-balanced, so for phyllosilicates with a -1 charge on the 2:1 layer, cations must be introduced into the interlayer for charge balance. Phyllosilicates with layer charge of -1 and alkali cations (mainly K^+) in the interlayer are known as **micas**. The aluminous dioctahedral mica $KAl_2(Si_3Al)O_{10}(OH)_2$ is **muscovite**, a common mineral in granites and various metamorphic rock types. The trioctahedral mica **biotite** has a unit-cell composition $K(Mg,Fe)_3(Si_3Al)O_{10}(OH)_2$, with 3 Mg instead of 2 Al (note that real-world biotite has some substitution of Fe^{+2} for Mg^{+2} in the octahedral layer). By convention, octahedral and tetrahedral cations are presented separately in mineral formulas.

The **interlayer** K^+ ion in micas fits very snugly into the hexagonal rings formed by linked silica tetrahedral – when extended across the interlayer space, opposed hexagonal rings produce a 12-fold coordination site (i.e. from the hexagonal rings of 6 O atoms above and 6 O atoms below), and the electrostatic bond between -1 -charged 2:1 layers and $+1$ -charged interlayer cations holds 2:1 layers together. However, the interlayer bond is still the weakest in the mineral, so micas and other phyllosilicates tend to **cleave** along this plane, producing sheet-like crystals. In cases where the charge on the 2:1 layer is <1 , the result is a series of clay minerals including **smectite** (layer charge ranging from -0.3 to -0.5), **vermiculite** (layer charge ~ -0.5), and **illite** (layer charge -0.7 to 0.9) – these and other environmentally significant clays are covered separately in the following section.

Chlorites comprise a group of phyllosilicates with a 2:1 layer, but instead of an interlayer cation like micas, they contain an octahedral sheet in the interlayer. The magnitude of the negative charge on the 2:1 layer is equal to the positive charge on the octahedral interlayer, e.g. by substitution of Al^{+3} for Mg^{+2} in a

$(Mg,Fe^{+2})_3(OH)_6$ octahedral sheet, which might result in a unit formula of:



The first set of brackets $[(Mg,Fe^{+2})_2Al]$ is the octahedral interlayer, the second set $(Mg,Fe^{+2})_3$ is the octahedral sheet within the 2:1 layer, and the third set (Si_3Al) is the tetrahedral sheet. Chlorites can be visualized as shown in Fig. 2.19.

Chlorites form in Mg–Fe-rich sedimentary and soil environments. In some cases (e.g. highly alkaline or highly acidic soils), aluminous dioctahedral chlorites form. Fe–Mg-rich chlorites also form during low-grade metamorphism of basalts and other mafic (Mg–Fe-rich, K-poor) rocks.

The 1:1 Phyllosilicates

Phyllosilicates that contain repetitions of one tetrahedral sheet and one octahedral sheet (T–O) are known as 1:1 phyllosilicates (Fig. 2.19). The dioctahedral forms have the composition $Al_2Si_2O_5(OH)_4$ and are known as **kaolin group** minerals; trioctahedral 1:1 minerals form the **serpentine group**, with the general composition $Mg_3Si_2O_5(OH)_4$.

The most common form of kaolin is **kaolinite**, a very common soil mineral, especially in well-leached soils like those of warm temperate and tropical climates. The other main varieties of kaolin, **halloysite**, **dickite** and **nacrite**, differ from kaolinite in the way that adjacent $Al_2Si_2O_5(OH)_4$ 1:1 layers are stacked (or rotated) with respect to each other. Halloysite is a disordered form of kaolin that mainly occurs in soils and hydrothermal environments, whereas dickite and nacrite are higher-temperature forms of kaolin (e.g. $> 150^\circ C$) that form in hydrothermal or diagenetic environments.

Serpentines may crystallize in soils and sediments, one example being berthierine, which has the general formula $(Mg,Fe,Al)_{2.5-3}$

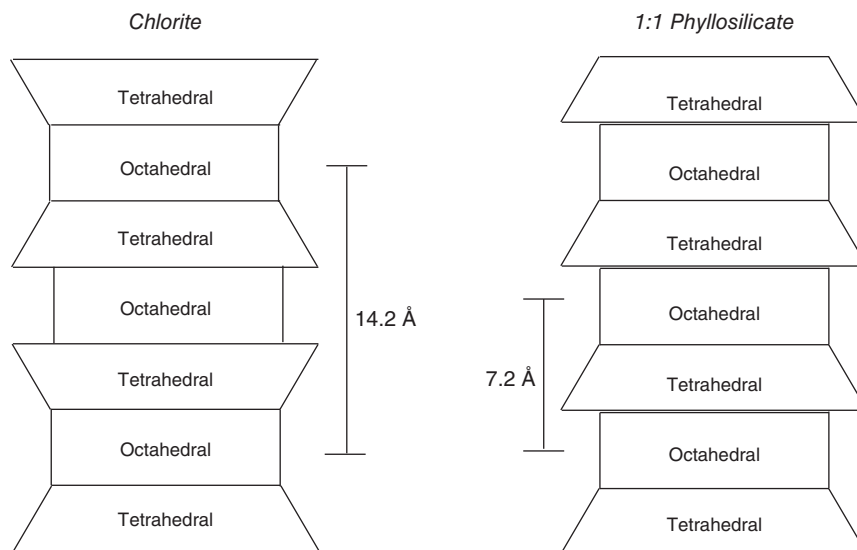


Fig. 2.19 Chlorite (left) and a 1:1 phyllosilicate (e.g. serpentine or kaolinite). Note that the interlayer in chlorite is an octahedral sheet, and that both chlorite and 1:1 phyllosilicates consist of alternating octahedral and tetrahedral sheets, the difference being the orientation of every other tetrahedral sheet.

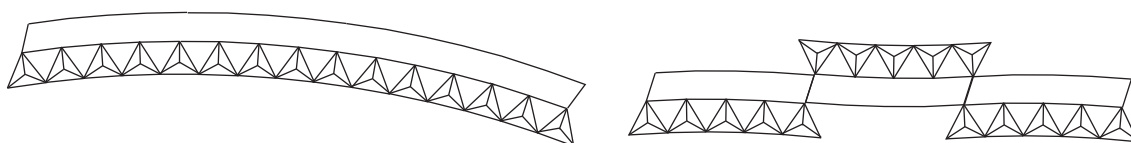


Fig. 2.20 The two ways that serpentines deal with the “mismatch problem”. On the left, the slightly larger octahedral sheet in chrysotile is draped over the curved tetrahedral sheet, ultimately forming rolled fibers. In antigorite (on the right), the mismatch is solved by short-period alternations (often of ~17 tetrahedra in length) of tetrahedral and octahedral sheets.

$[\text{SiAl}]_{2-2.5}\text{O}_5[\text{OH}]_4$), and they also occur in Mg-rich low-grade metamorphic rocks. Serpentine has three main polytypes, **antigorite**, **lizardite** and **chrysotile**, each with the approximate composition $\text{Mg}_3\text{Si}_2\text{O}_5(\text{OH})_4$, and each of which is differentiated from the others by different stacking arrangements.

Chrysotile occurs in an **asbestiform** morphology. The cause is tetrahedral–octahedral mismatch – in all serpentines the octahedral sheet is slightly larger than the tetrahedral sheet, and this mismatch is addressed by slight curvature of the two sheets in order form shorter hydrogen

bonds. If the apical oxygen atoms in adjacent tetrahedra rotate away from each other a few degrees they become better-suited to bond with the octahedra that have been stretched by the fact that they need to accommodate 3 Mg atoms. Figure 2.20 shows how chrysotile deals with the mismatch by forming rolled 1:1 layers, whereas antigorite contains periodic inversions of 1:1 layers known as a modulated structure that results in a wavy pattern.

The 1:1 layers in chrysotile are like rolled-up carpet, each layer of carpet being equivalent to a 1:1 layer, and this results in the fibrous or

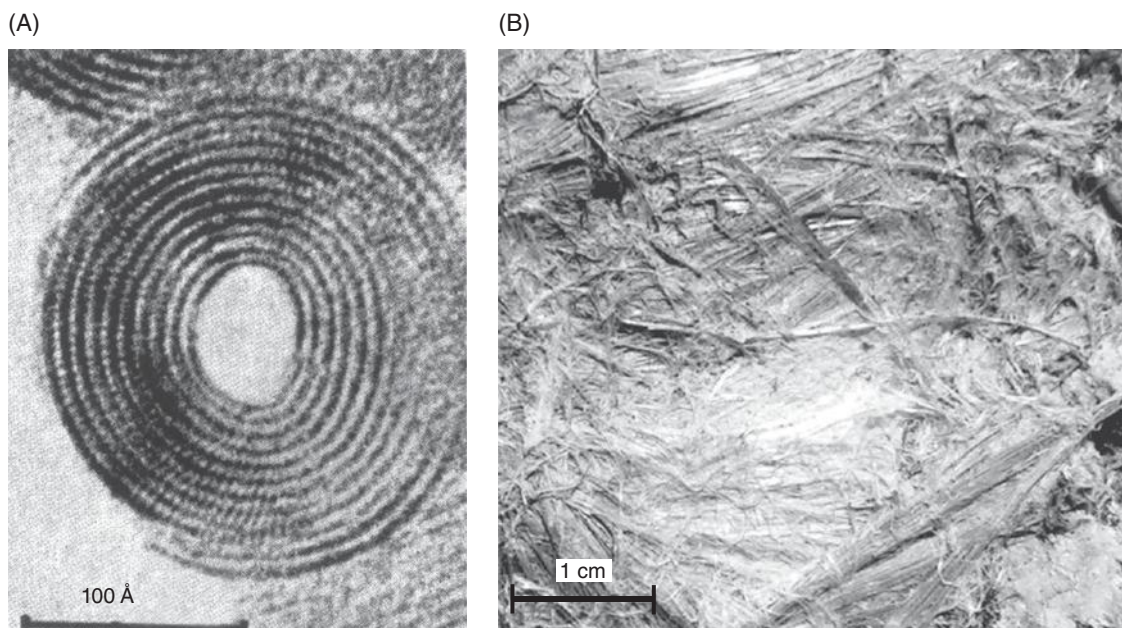


Fig. 2.21 Rolled tetrahedral and octahedral sheets viewed down long axis of chrysotile from a transmission electron microscope image (A) (Yada 1971, reproduced with permission from IUCR) and fibrous white chrysotile on a serpentine rock surface (B).

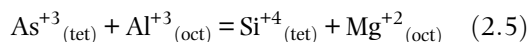
asbestiform character of chrysotile. The gently rolling waves of antigorite 1:1 layers produce a platy morphology that is typical of most phyllosilicates.

The two examples in Fig. 2.21 exhibit the characteristic morphology of chrysotile – on the left is a transmission electron microscope image, and on the right is a cm-scale hand specimen of chrysotile from the Belvidere Mine in northern Vermont, USA. The images show:

- Individual 1:1 layers in the TEM image, which are highlighted by schematic T–O layers near the top of the image.
- Distinctive white chrysotile fibers in the image on the right (note the 1 cm scale bar). The Belvidere Mine specimen also contains green antigorite (on the bottom right hand side) that exhibits platy morphology (parallel to the rock face) produced by the periodic alternations sketched above.

The fibrous shape of chrysotile means that it is a type of **asbestos**, and as with asbestiform amphiboles this raises concerns about lung disease. Due to its higher solubility in pulmonary fluids chrysotile may be less toxic than asbestiform *amphibole* but often amphibole asbestos (e.g. tremolite) is intermixed with chrysotile in ore deposits so the effects are hard to isolate.

Serpentine can incorporate arsenic to concentrations of hundreds of mg/kg in serpentinites (Hattori et al., 2005; Ryan et al., 2011). The crystal radii (Table 2.1) and charges of two arsenic ions, As^{+3} (CR = 0.54 Å) and As^{+5} (CR = 0.48 Å), are sufficiently similar to Si^{+4} (CR = 0.40 Å) to allow substitution (note that the crystal radius of Al^{+3} , a common tetrahedral ion, is 0.53 Å). Insertion of As^{+3} into a Si^{+4} tetrahedral site can be charge balanced by insertion of Al^{+3} into a Mg^{+2} octahedral site, producing a charge-balanced unit cell.



Substitutions of divalent and trivalent trace metals (particularly Cu^{+2} , Cr^{+3} , Ni^{+2} and Zn^{+2}) can occur in many phyllosilicates, particularly in the octahedral sheet of many ferromagnesian clay minerals. This topic is covered in more detail in Section 2.4.

2.3.4 Tectosilicates

Tectosilicates are the most highly polymerized silicates – all oxygen atoms are shared among adjacent tetrahedra. Tectosilicates, also known as framework silicates, include some of the most common minerals in the Earth's crust, soils and sediments, including **quartz** and the **feldspar group**, which together comprise about 75% of the Earth's crust. The tectosilicates also include less common groups such as the **feldspathoids** (which can be thought of as Si-poor feldspars) and **zeolites**. Zeolites, with cavities ($\sim 2\text{--}10 \text{ \AA}$) in their structures, have strong sorption capacity and are used extensively in industrial and municipal operations for removal of toxins.

The sharing of all four oxygen atoms in SiO_4^{4-} tetrahedra results in the charge-balanced fundamental compositional unit of SiO_2 and unlike the other silicate groups, results in a mineral that requires no other cations – this mineral is quartz. Altering quartz slightly helps to understand feldspars: starting with a hypothetical Si_4O_8 component and substituting an Al^{+3} ion for a Si^{+4} ion results in $\text{AlSi}_3\text{O}_8^{-1}$. The -1 charge imbalance can be satisfied by inclusion of Na^{+1} in the structure, producing the common igneous mineral **albite** ($\text{NaAlSi}_3\text{O}_8$), or K^{+1} to produce **potassium feldspars** (**microcline**, **orthoclase** and **sanidine**). In Fig. 2.22, some tetrahedra have edges (apical oxygens) facing toward the reader; the remaining ones (gray) have a face toward the reader. Note that there are two planes of

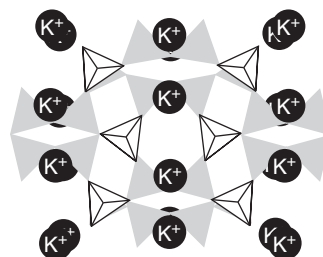


Fig. 2.22 Sketch of potassium feldspar showing potassium ions and silica tetrahedra. Gray triangles represent silica tetrahedral whose apices point away from the viewer. Note also that $\frac{1}{4}$ of tetrahedral sites are occupied by Al.

K atoms, and although Si and Al tetrahedra are not differentiated on the diagram, $\frac{1}{4}$ of the tetrahedral sites contain Al and the remainder contain Si.

Feldspars form in metamorphic, igneous and hydrothermal systems but they are common in soils and sediments where they often decompose into soil clays (Chapter 9).

2.4 CLAY MINERALS (T-O MINERALS, T-O-T MINERALS, INTERSTRATIFIED CLAYS)

The term “clay” can mean two things. One is a specific mineral group known as the **clay minerals**, a class of phyllosilicates that tend to occur as particles finer than $2 \mu\text{m}$, and it is this mineral group that comprises the focus of this section. The other designation of “clay” is the **clay-size fraction**, a size class commonly defined as <2 micrometers (μm), but also sometimes defined as <4 or $<5 \mu\text{m}$ (where a μm is 10^{-6} m). Particles in this size range tend to be highly reactive due to their fine grain size and high ratio of surface area : volume, but in theory the *clay-size fraction* could include any mineral, whether quartz, calcite, feldspar, smectite, amphibole, iron oxides or

olivine; however, given their tendency to occur as crystals $< 2 \mu\text{m}$, clay minerals often dominate the clay fraction.

Some clay minerals are finer-grained, more disordered versions of coarser, higher-temperature phyllosilicates (e.g. micas and chlorite), but most clay minerals (e.g. smectite, kaolinite, halloysite and palygorskite) have no real coarsely crystalline, metamorphic or igneous analogs. They tend to form at low temperatures (e.g. $< 100^\circ\text{C}$) and tend to occur in soils, sediments, sedimentary rocks and other surficial or near-surface environments. Clay minerals are structurally distinct layer silicates with (generally) platy morphology, and they are characterized by their extensive compositional diversity and the fact that they are the most abundant reactive phases in low-temperature systems like soils and sediments.

Clay minerals are environmentally significant because:

- 1 Their fine-grained nature makes them much more reactive than most minerals, and for this reason they play a strong role in adsorption (or release) of trace metals, organic compounds and other potential contaminants; they also are the main minerals responsible for uptake and release of plant nutrients and they have the potential to buffer changes in pH, e.g. in response to acid precipitation.
- 2 They are common products of chemical weathering and thus are major components of soils; in addition to their role described in point 1, their composition often reflects the environment in which the soil formed and therefore can be used in paleoclimate analysis.
- 3 Their fine-grained character means that they are physically mobile, so when considering point 1, physical transport of clays can result in transport of adsorbed contaminants or plant nutrients.
- 4 Clays like illite and smectite are common cements in sedimentary rocks and thus control rheology and fluid flow in sedimentary aquifers and deeper realms within sedimentary basins where hydrocarbons occur.

The following section on clay minerals will first present the 2:1 layer clays, focusing on smectites, a diverse group of expandable clays, as well as other 2:1 clays (illite, glauconite, vermiculite and chlorite), followed by discussion of the 1:1 clays (i.e. kaolins and serpentines) and the interstratified clays (e.g. illite–smectite, kaolinite–smectite). The information presented on phyllosilicates near the end of Section 2.3 will serve as the starting point for this treatment of the clay minerals.

2.4.1 Smectite

Smectites are a compositionally diverse group of clay minerals that are abundant in a wide array of natural systems, from soils in temperate and arid climates to sediments, shales, marine and terrestrial sediments and weathered volcanic ash. They are particularly important environmentally because of their high cation exchange capacity and their ability to expand in the presence of water (Fig. 2.23).

The structure and layer charge of smectites can be conceptually derived from the 2:1 phyllosilicates with zero layer charge, i.e. pyrophyllite or talc. Phyllosilicates derive **layer charge** from ionic substitutions in the tetrahedral sheet or the octahedral sheet, common examples being Al^{+3} for Si^{+4} in the tetrahedral sheet and Mg^{+2} for Al^{+3} in the octahedral sheet; both result in residual negative charge.

Smectites have layer charges of approximately -0.2 to -0.5 , in contrast to the -1 layer charge of micas. Also in contrast to the micas, the interlayer cations that balance the negative 2:1 layer charge in smectites are not fixed, and instead consist of hydrated, low-charge cations such as Na^+ , Ca^{+2} and Mg^{+2} , and less commonly, K^+ or various transition metals (e.g. Ni^{+2} , Zn^{+2} , Cu^{+2}). Organic compounds, some of them contaminants (e.g. pesticides) can also occupy smectite interlayers. The composition of the smectite interlayer typically

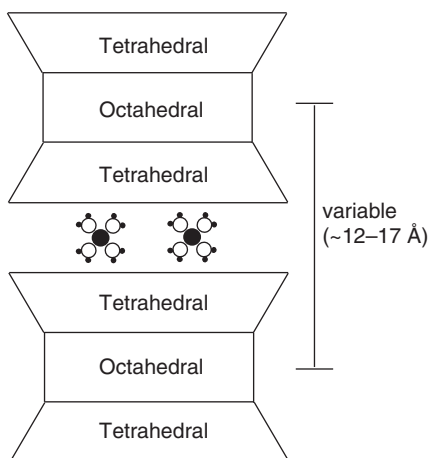
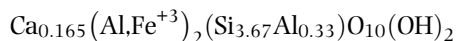


Fig. 2.23 Two smectite 2:1 layers with hydrated cations (black spheres) occupying the interlayer site. Layer charge on each of the 2:1 layers (T–O–T) is -0.2 to -0.5 and this negative charge attracts cations to the interlayer. Variability in c -axis spacing is due to ability to intercalate different amounts of interlayer water (where the water is bonded to cations).

reflects the composition of ambient pore water – if the water in the pores of smectite-rich sediment is rich in Na, so too will be the smectite interlayer. In a soil contaminated with soluble nickel, much of the nickel may be taken up in the smectite interlayer as exchangeable Ni^{+2} .

2.4.2 Smectites with tetrahedrally derived layer charge

Substituting 0.33 Al atoms into the *tetrahedral* sheet of pyrophyllite results in a 2:1 **beidellite** layer (Table 2.3) with -0.33 charge: $\text{Al}_2(\text{Si}_{3.67}\text{Al}_{0.33})\text{O}_{10}(\text{OH})_2$. The negative layer charge is shown balanced here by hydrated interlayer Ca:

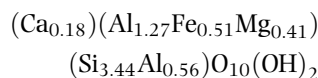


Note that this example contains some ferric iron in the interlayer. Some beidellites are very

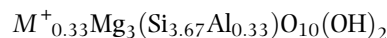
Table 2.3 Classification of smectites based on octahedral occupancy and origin of 2:1 layers charge (i.e. octahedral vs. tetrahedral sheets).

Dominant Origin of 2:1 Layer Charge		
	Tetrahedral Sheet	Octahedral Sheet
<i>Diocahedral Smectites</i>		
Al-rich	Montmorillonite	Beidellite
Fe-rich	N/A	Nontronite
<i>Triocahedral Smectites</i>		
Mg-rich	Hectorite (Mg + Li)	Saponite

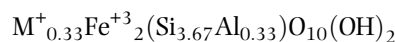
aluminous, with nearly no Fe or Mg substitution in the octahedral layer, yet other have extensive Fe for Al substitution – these ferruginous beidellites are common soil minerals, including this example from a Holocene tropical soil (Ryan and Huertas, 2009):



Two idealized compositions of other varieties of smectite with charge derived from the tetrahedral layer include **saponite**



and **nontronite**

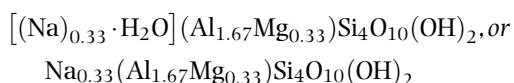


In both of these formulas, the interlayer cation is represented by the variable M^{+} .

2.4.3 Smectites with octahedrally derived layer charge

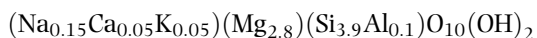
Substitution of 0.33 Mg atoms into the *octahedral* sheet of pyrophyllite results in $(\text{Al}_{1.67}\text{Mg}_{0.33})\text{Si}_4\text{O}_{10}(\text{OH})_2$, a 2:1 layer with a negative 0.33 charge balanced by hydrated cations, in this case Na^{+} .

Two common ways to represent the formula unit are:



Aluminous smectite with octahedrally derived charge is known as **montmorillonite**, a common smectite that forms by alteration of volcanic ash in marine environments – montmorillonite is very abundant in marine sediments and low-grade shales and soils eroded from these sources (i.e. many terrestrial soils and sediments).

Stevensite is a trioctahedral clay that derives its layer charge from vacancies in the octahedral sheet and, in some cases, from minor substitution of Al for Si in the tetrahedral sheet, a concept illustrated by these data adopted from Hay et al. (1995):



Extensive substitution results in smectites that can vary appreciably from the idealized compositions, with

1 numerous intermediaries between smectite end-members; and

2 smectites with layer charge derived from substitutions into both tetrahedral and octahedral layers.

Furthermore, the ionic radii of many trace metals differ by $\leq 15\%$ from those of the common trioctahedral cations ferrous iron and magnesium ($\text{Fe}^{+2} = 0.78 \text{ \AA}$, $\text{Mg}^{+2} = 0.72 \text{ \AA}$), so metals such as Co^{+2} ($0.79 \text{ \AA}$), Cr^{+3} ($0.76 \text{ \AA}$), Ni^{+2} ($0.69 \text{ \AA}$) and Zn^{+2} ($0.74 \text{ \AA}$) are known to occur in clay octahedral sheets (Table 2.1). In some cases they are trace components, but in other cases they are major components of octahedral sheets, a good example being Ni-rich stevensite formed by the weathering of Ni-rich ultramafic rock in Greece (Christidis and Mitsis, 2006) with a composition of



Shrink–Swell Capacity of Smectites

Depending on available moisture, **smectite interlayers** can contain one layer of hydrated cations (a $2.8 \text{ \AA}$ one-layer complex), or two (a $5.6 \text{ \AA}$ two-layer complex), or three water layers (a $8.4 \text{ \AA}$ three-layer complex). This swelling capacity is a distinctive characteristic of smectites, also sometimes referred to as **expandable** clays or **shrink–swell** clays. When the interlayer cation is sodium, the amount of interlayer water can increase by orders of magnitude (Na^+ has a very high hydration potential), which can have serious negative consequences for agriculture, engineered structures and land stability. In climates with alternating rainy seasons and dry seasons, smectite-rich soils expand and shrink seasonally, wreaking havoc on buildings and roads and causing problems for

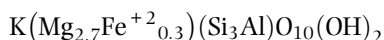
farmers due to alternating periods of impermeability (when saturated) and dessication cracking (dry periods) of farm fields.

However, the negative aspects of smectites are balanced by the fact that they are widely used in landfills and hazardous-waste repositories where water-borne contaminants must be contained. This is because hydrated smectites are virtually impenetrable to flow of water and their high potential to chemically adsorb trace metals and other contaminants means that they ultimately form a physical and chemical barrier to contaminant flow. This is true in both engineered systems and in natural systems, where, for example, clay-rich layers can form impermeable boundaries between shallow contaminated aquifers and deeper, confined aquifers.

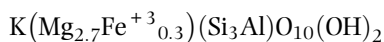
So, given that trace metals can occur in both octahedral sheets of trioctahedral clays and interlayers of smectites, clays play an important role in fate and transport of these elements.

2.4.4 Vermiculite

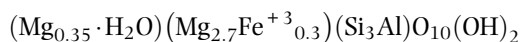
Vermiculite is structurally similar to smectite, but with higher layer charge (−0.6 to −0.9) and often much coarser particles. In fact, there is little distinction between high-charge smectites (those with −0.5 to −0.6 layer charge) and some vermiculites, at least in terms of layer charge and cation exchange capacity. Vermiculite often forms in soil by weathering of biotite, so we will begin by examining biotite and work our way towards vermiculite. An illustrative biotite composition (structural formula) is:



During chemical weathering ferrous iron typically oxidizes to ferric iron and this can have a destabilizing effect on many minerals. So, if we assume that all ferrous iron oxidizes, the result is:



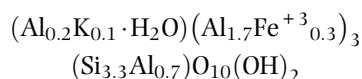
This produces a −0.7 charge on the 2:1 layer (it was −1.0 in the un-oxidized biotite), which is too low (i.e. insufficiently negative) to hold the potassium ion in the interlayer. The result is that potassium diffuses out of the interlayer (as a potential plant nutrient), and hydrated ions (commonly Mg but sometimes Al or other cations) migrate into the interlayer, producing this representative trioctahedral vermiculite:



The +0.7 charge on the interlayer ($\text{Mg}_{0.35}\text{H}_2\text{O}$) balances the 2:1 layer charge of −0.7.

Diocahedral vermiculite is derived from weathering of micas in acidic soils with high

concentrations of dissolved Al. The following formula is adopted from data presented in April et al. (1986) and, for purposes of illustration, does not include trace amounts (<0.07 atoms per formula unit) of octahedral Mg and Ti.

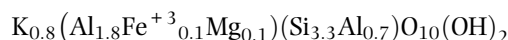


Note that (1) 2:1 layer charge is −0.7, in this case entirely derived from the tetrahedral sheet, and (2) not only is Al a major component of the octahedral sheet, but also the vermiculite interlayer, reflecting high Al concentrations in soil water.

Vermiculites possess very high cation exchange capacities and can play important roles in nutrient cycling.

2.4.5 Illite

Illite can be considered as a low-temperature analog muscovite. Muscovite contains a layer charge of −1 and is characterized by relatively large crystals (typically visible without a microscope), whereas illite possesses a layer charge of −0.7 to −0.9 and typically occurs in crystals only visible under a microscope. A typical illite composition is:



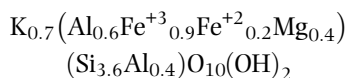
Note that the 2:1 layer charge is −0.8, with −0.7 derived from the tetrahedral sheet (Al for Si) and −0.1 derived from Mg for Al substitution in the octahedral sheet. K is fixed in the interlayer.

Illite commonly forms in sedimentary basins from precursor smectite (commonly montmorillonite), a reaction that is largely driven by temperature but also time (i.e. thermodynamics and kinetics both play a role in the smectite → illite reaction). Shallowly buried marine shales (depths <2 km, <75 °C) are dominated by smectite, whereas deeper shales are dominated by illite (~150–200 °C); intermediate shales contain a mixture of the two (more on this later in the

section on interstratified clays). If shales are further buried or overthrust so that they reach low-grade metamorphic temperatures of $\geq 250^\circ\text{C}$, illite transforms to muscovite, a point at which shales often transform to slates or phyllites. Because of the predictable transformation from smectite to illite to muscovite, this series can be used as a **geothermometer** to interpret burial and metamorphic histories of sedimentary rocks (factoring in compositional and kinetic controls on the smectite to illite transformation).

Illite is also a common soil mineral, most commonly derived from **detrital** or physical weathering processes but in some cases from **pedogenic** processes. It is stable in many weathering environments and can be inherited from parent material like granite and schist or via aeolian (wind) or fluvial (river-borne) processes. In other cases, where soils are potassium rich, illite can form in the soil environment by, for example, weathering of potassium feldspar.

The Fe-rich mica **glauconite** is a useful tool in paleoenvironmental analysis. It crystallizes on the seafloor at depths of ~ 250 m (near the edge of the continental shelf) within about 60° of the equator, so is useful in interpreting sedimentary environment, and because it is *syndepositional* (i.e. crystallized when sediments were being deposited), dating of glauconite by K–Ar and Rb–Sr radioactive decay methods can provide depositional ages. An average chemical composition obtained from chemical analyses of Thompson and Hower (1975) is:

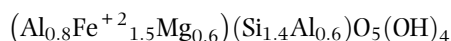


Note that glauconite derives its -0.7 layer charge from both the tetrahedral and octahedral sheets.

2.4.6 Chlorite and Berthierine

These two mineral types are grouped together here because of similarities in structure, composition and occurrence in soils, sediments and

sedimentary rocks. Chlorites are generally trioctahedral, as is berthierine, and they form in environments rich in Mg and Fe. The composition of berthierine is much like that of chlorite:



By doubling the composition of berthierine, you arrive at a reasonable chlorite composition (Section 2.3).

Now compare the structures of the two minerals, as shown in Fig. 2.19. The only appreciable difference is the orientation of every other tetrahedral sheet. In the case of berthierine, all tetrahedra are oriented in the same direction, producing a repeat distance along the *c*-axis of $7.2 \text{ \AA}$; in the case of chlorite, orientations of tetrahedra alternate, producing a repeat distance almost exactly double that of berthierine.

Like glauconite, berthierine (and its disordered analog odinite) is often a paleoenvironmental indicator in sedimentary environments, as it forms at or near the sediment/water interface (i.e. on the bottom of the water body, or within sediments, but at depths of only a few cm) in shallow marine lagoons, estuaries and river mouths located within 30° of the equator. These are regions where iron and silicate materials derived from intense tropical weathering are carried by rivers into shallow marine environments. With increasing temperature, berthierine transforms into chlorite, providing a geothermometer in Fe-rich sedimentary rocks (Hillier, 1994; Ryan and Reynolds, 1996).

2.4.7 Kaolin (kaolinite and halloysite)

Kaolinite and halloysite are the two polytypes of the 1:1 kaolin group that commonly occur in low-temperature environments like soils, sediments and sedimentary rocks. Both are nominally $\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$ although Fe^{+3} can substitute to a limited extent into octahedral sites (Petit and Decarreau, 1990).

Kaolinite and halloysite form in environments with high dissolved Al content and lower concentrations of Si and base cations (Na, K, Mg, Ca) relative to environments that foster formation of smectite, vermiculite, chlorite and illite. Tropical soils are often leached of base cations, resulting in Al-Si-rich solutions that foster formation of kaolin or halloysite.

One very important difference between the kaolin minerals and most other clays is that **kaolins have no layer charge** and thus have very low cation exchange capacities compared to other clay minerals, i.e. they have very limited attraction for many nutrients as well as trace metals and other cationic forms (refer to Section 2.5). It is for this reason that tropical soils tend to be nutrient poor (Chapter 9).

2.4.8 Interstratified clay minerals

Also known as mixed-layer clay minerals, interstratified clays are common in soils and sedimentary rocks, and the types, extent and changes in interstratification are controlled by important environmental factors such as composition, climate, temperature and time. The most common and widely studied interstratified clay series is illite/smectite (I/S). Low-temperature shales (e.g. 50 °C, 1 km burial) are dominated by smectite, and high-temperature shales (200 °C, 6 km burial) are dominated by illite. Between these two end-members exists a progressive series of interstratified clays ranging from smectite-rich I/S to intermediate I/S and illite-rich I/S (Fig. 2.24).

The nomenclature of interstratified clays is straightforward – the name of the component with the lower *d* spacing is listed first, followed by either a slash (used here) or dash, and then the name of the component with higher *d* spacing, hence names like illite/smectite and kaolinite/smectite, or abbreviated terms like I/S and K/S. Where proportions are known (e.g. 75% kaolinite layers and 25% smectite layers in kaolinite/

smectite), you will see terms like kaolinite (0.75)/smectite, abbreviated as K(0.75)/S.

In soils, interstratified kaolinite–smectite (progressing from smectite → kaolinite-smectite → kaolinite with leaching over time) and biotite–vermiculite (progressing from B → B/V → V with leaching over time) are common interstratified series.

2.4.9 Trace metals and metalloids in clay minerals

Trace metals such as Cr^{+3} , Ni^{+2} and Zn^{+2} (and less frequently Co^{+2} and Cu^{+2}) are known to occur in octahedral sheets of clay minerals, particularly smectites, serpentines, and chlorites (Newman and Brown, 1987). The literature contains numerous studies which document elevated concentrations of trace metals in clay structures; for example, Brindley and Maksimovic (1974) and Gaudin et al. (2004) describe Ni-rich layer silicates in soils and other low-temperature environments, and Oze et al. (2004) report Cr-rich chlorites in serpentine soils. Manceau et al. (2004) determined that, in a soil with 128 mg/kg Zn, >80% of the Zn is contained in interlayers and octahedral sheets of magnesian smectite, and Dähn et al. (2002) showed that Ni-rich smectite layers can crystallize on surfaces of pre-existing montmorillonite crystals in Ni-rich solutions – in these Ni-smectites, the nickel occurs as Ni^{+2} in octahedral sites.

2.5 CRYSTAL CHEMISTRY OF ADSORPTION AND CATION EXCHANGE

“Adsorption is probably the most important chemical process affecting the movement of contaminants in groundwater” (Drever, 1997). The composition of soil water and groundwater that feeds streams and lakes is similarly strongly controlled by adsorption–desorption reactions. Drever’s statement could also be extended to nutrient availability in natural and agricultural

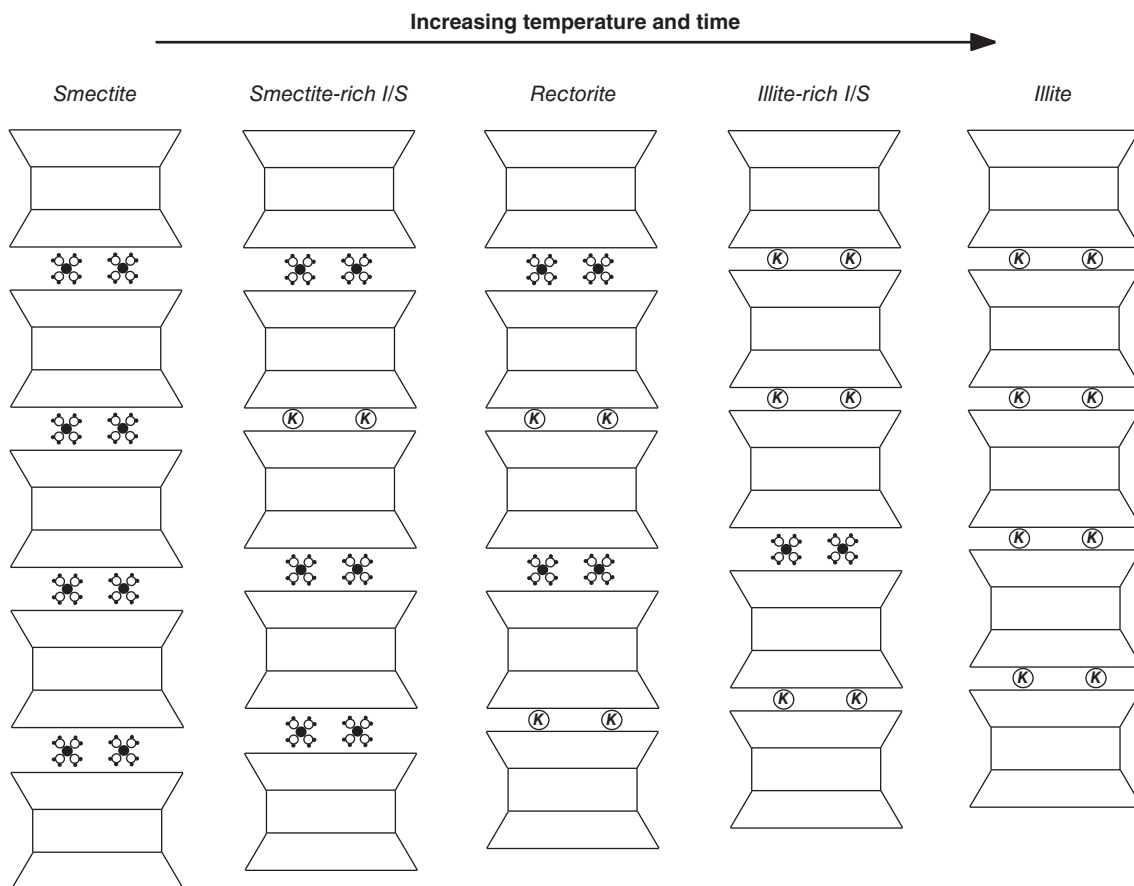


Fig. 2.24 The transition from smectite to illite via interstratified illite/smectite, including the ordered I/S phase rectorite, a mineral in which illite and smectite layers alternate in an ordered manner. In sedimentary basins, this transition typically takes place progressively with increasing burial over the temperature range of 50 °C to 200 °C.

systems as well as compositions of surface water. The adsorption and desorption of ions and other substances from surfaces of minerals and natural organic matter exerts a very strong control on cycling of plant nutrients by regulating nutrient uptake and leaching. An important consequence of adsorption may be that while an ion or molecule is very soluble in groundwater or other aqueous systems, if it is prone to adsorption onto surfaces of solid phases, its concentration in solution will be less than predicted when only considering dissolution–precipitation reactions (e.g. mineral stability).

It is important to realize that adsorption is a surface effect. Adsorbed ions or molecules can be released into solution without mineral dissolution, and adsorption–desorption reactions may only require slight changes in the concentrations of non-adsorbed ions or molecules in solution. Conversely, ions incorporated (or adsorbed) *into* mineral structures are generally only released into solution when the mineral dissolves. Adsorption–desorption reactions are **reversible** and take place at rates that are typically orders of magnitude more rapid than mineral dissolution–precipitation reactions. In fact, ion exchange reactions can occur

so rapidly that their rates are quite difficult to measure by conventional methods (McBride, 1994). To reiterate an important point, **adsorption** describes the attraction of substances (adsorbates) to the surface of a solid particle (adsorbent); **absorption**, on the other hand, is a process by which a substance enters into the internal structure of a particle or other absorbent. The term **sorption** includes both adsorption and absorption.

2.5.1 Cation exchange

Cation exchange is a particular type of adsorption characterized by attraction of cations to particle surfaces, and it is controlled by the following factors:

1 chemical attraction, which occurs when a dissolved phase forms a chemical bond with an atom (or atoms) exposed at the surface of the solid – an example might be a cation attracted to an exposed oxygen atom at crystal edge;

2 electrostatic attraction, where a dissolved phase is attracted to the surface of a charged particle – a good example is hydrated interlayer cations being attracted to the negatively charged 2:1 layer in smectite;

3 physical attraction, the weakest form of adsorption that is due primarily to van der Waals forces;

These processes are regulated by a combination of attributes related to the mineral itself, the substance that may be sorbed, and properties of the solution (e.g. soil water) in which the solid resides. The remainder of this section will consider attributes of minerals as they relate to surface exchange. Details of solution chemistry and its relationship to ion exchange and water chemistry are presented in Chapter 4 (Section 4.10).

Cation exchange is particularly pronounced in clay minerals and other fine-grained inorganic solids as well as natural organic matter. The solids are strong adsorbers and exchangers of cations, and hence have high **cation exchange capacity**. The focus here will be on clays because of their abundance, their (generally) high cation

exchange capacities and, hence, their importance in controlling aqueous chemistry. Attention will also be directed towards oxyhydroxides because of their pH-dependent surface charge that means in some cases (e.g. low-pH solutions) they possess **anion-exchange capacity**.

Cations in solution are attracted to clays and other fine-grained particles (colloids) by four main processes (the first two of which relate to the concept of surface area : volume ratio):

1 Charge on the particle: charged particles will electrostatically attract ions of the opposite charge – this is particularly true of smectite, vermiculite and other clay minerals with negative layer charges, each of which are strong cation exchangers.

2 Particle size: exposed atoms on mineral surfaces are potential adsorption sites, and the greater the amount of exposed surface area per volume of mineral, the greater the ability to adsorb. Fine-grained particles have **high ratios of exposed surface area to volume**, i.e. enormously high surface area relative to volume as compared to coarse minerals, and thus contain much higher proportions of potential exchangeable sites relative to the amount of mineral.

3 Interlayers or other exchange sites in minerals: interlayers in smectites and channels in zeolites are examples of additional adsorption sites. Comparing the available exchangeable sites in smectite compared to kaolinite (Fig. 2.25) emphasizes that each smectite unit cell contains an interlayer that can absorb cations, whereas kaolinite does not (nor do minerals like illite, chlorite and oxyhydroxides). This concept becomes even more apparent when considering that the *a*–*b* plane extends virtually infinitely in clays like smectite and vermiculite compared to their dimensions parallel to the *c*-axis; their interlayers are very extensive due their platy textures, meaning that smectites and vermiculites have extensive interlayer exchange sites.

4 Edge sites, dangling bonds, surface effects: these effects are derived from surface complexation, a process whereby hydrated ions attach to

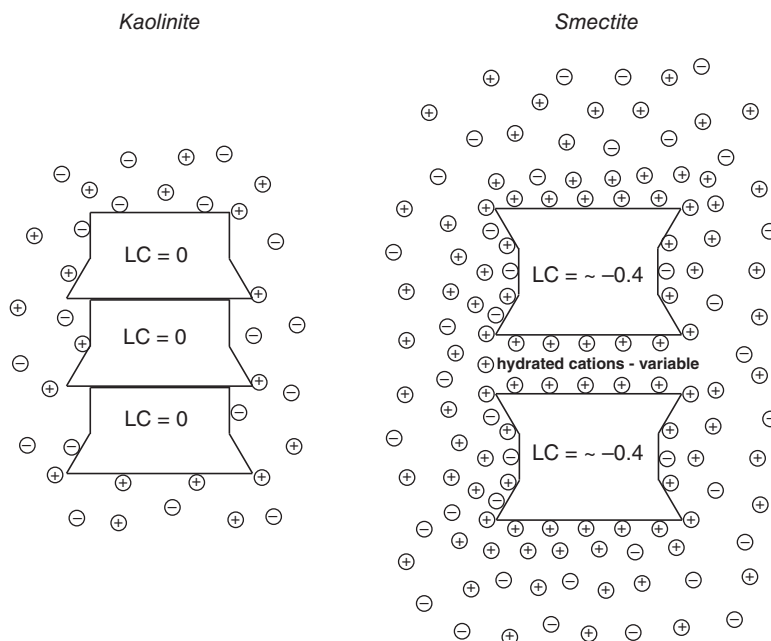


Fig. 2.25 Schematic representation of the differences in cation exchange between kaolinite and smectite. “LC” indicates layer charge, which is zero in kaolinite but is -0.2 to -0.5 per unit cell in smectite. CEC of smectites is also enhanced by expandable interlayer spaces and smaller crystal size (not shown here).

ions exposed at mineral surfaces. This occurs because edges of crystals commonly contain unsatisfied chemical bonds, a good example being a metal cation in octahedral coordination at the edge of a crystal that is bonded to fewer than 6 oxygens or hydroxyls. In acidic solutions, mineral surfaces tend to be dominated by exposed H^+ ions, making the H^+ -charged surface attractive to anions, whereas in alkaline solutions, surfaces are negatively charged and attract cations. These effects are most pronounced in oxides, hydroxides and silicate clays like kaolin and serpentine with little or no layer charge.

The pH at which the surface charge of a particle is zero in a solution in which only H^+ and OH^- sorb to the mineral surface is known as the **isoelectric point** (IEP); in natural solutions with a range of ions, the term is **point of zero charge** (PZC). At values less than the PZC or IEP, the surface charge is positive (dominated by H^+) and so the particle attracts anions. As pH increases and

H^+ (protons) are drawn away from the surface and into solution, the surface becomes negatively charged and acts as a cation exchanger. The values of PZCs and IEPs presented in Table 2.4 indicate the following:

- Smectites adsorb cations in all situations except for very acidic ($pH < 3$) environments.
- Kaolins possess weak anion-exchange capacity at $pH < 3$ to 4; at $pH > 4$, kaolins possess weak cation-exchange capacity.
- The isoelectric point of iron oxides and hydroxides occurs at $pH 7 \pm 2$, meaning that in acidic solutions they are anion exchangers, in alkaline solutions are cation exchangers, and in neutral solutions have little or no preference for cations or anions. These relationships can have consequences for cycling of nutrients such as the NO_3^- (nitrate) anion, anionic contaminants (e.g. certain pesticides and some forms of arsenic), trace metals and other aqueous species. Figure 2.26 depicts a **goethite** crystal in two

different solutions, one with a pH = 5 (i.e. below the isoelectric point) and the other at a pH = 9 (above the isoelectric point). Note the adsorption of protons at pH = 5 that causes **goethite** to adsorb anions.

Table 2.4 Isoelectric points and points of zero charge (PZC) of selected minerals (values = pH) and non-minerals*. Note that isoelectric points and PZC values are very similar, but unlike isoelectric points, which are only a function of protonation and deprotonation, PZCs are also controlled by solution chemistry (e.g. adsorption–desorption of metals and other ions). At pH values below the isoelectric point, the mineral is an anion exchanger; above this value, the mineral sorbs cations.

Mineral	Composition	IEP	PZC
Smectite	$\text{Ca}_{0.17}(\text{Al}, \text{Fe}^{+3})_2(\text{Si}_{3.66}\text{Al}_{0.34})\text{O}_{10}(\text{OH})_2$	1–3	2–3
Kaolin	$\text{Al}_2(\text{Si}_2\text{O}_5(\text{OH})_4$	4	2–4
Humus*	$\sim \text{C}_{20}\text{H}_{20}\text{O}_{15}\text{N}$	~ 2 –3	~ 2 –3
Allophane*	$\sim \text{Si}_{1.5}\text{Al}_2\text{O}_4(\text{OH})_4\text{H}_2\text{O}$	5–6	~ 5
Hematite	Fe_2O_3	6–8	4–7
Goethite	FeOOH	6–7	6–7
Gibbsite	$\text{Al}(\text{OH})_3$	~ 9	~ 10

* Note: all of the above substances are minerals with the exceptions of humus and allophane. Humus is organic matter and allophane lacks ordered internal structure.

It is important to note here that the most abundant minerals in soils that contribute to cation exchange (clay minerals such as smectite) have isoelectric points at pH < 3, meaning that at nearly all pH values in natural systems, the potential for cation exchange is far greater than the potential for anion exchange. Anions such as nitrate (NO_3^{-1}) and sulfate (SO_4^{-2}) tend to be poorly retained by soils except in highly evaporitic environments (where evaporation prevents leaching).

2.5.2 Double-layer complexes

Exchangeable ions occur on particle surfaces in a **double layer** consisting of *inner-sphere* and *outer-sphere* complexes. **Inner-sphere complexes**, also known as the fixed-layer or Stern layer, are typically held by electrostatic forces directly on the surface of the solid; for this reason, the inner-sphere complex is often entirely cations (assuming pH > PZC). **Outer-sphere complexes** (also known as diffuse or Gouy layers) are farther from the crystal surface. At the boundary between inner-sphere and outer-sphere complexes cations > anions, but the concentration of anions increases exponentially away from this boundary until anions and cations are balanced

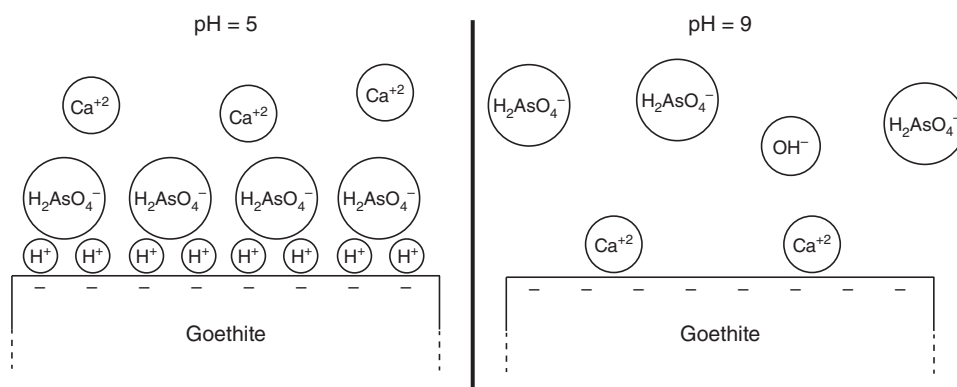


Fig. 2.26 Sketch of a goethite surface in solutions at pH = 5 (left) and pH = 9 (right), demonstrating the effect of IEP or PZC on ion-exchange capabilities of soil solids. Note the occurrence of anionic arsenate as inner sphere complex at pH = 5 (facilitated by H⁺) compared to the occurrence of Ca⁺² as an inner sphere complex at pH = 9. Negative signs near the goethite surface indicate a net negative charge at the mineral surface.

Table 2.5 Cation-exchange capacities of selected minerals (note that oxyhydroxides are anion exchangers at pH values approximately < 7). These values were determined at pH = 7; values vary as a function of pH and aqueous composition. (Sparks 1995; Drever 1997; Langmuir 1997.)

Mineral/phase	CEC (cmol _c /kg)	S _A (m ² /g)	Origin of CEC
Organic matter (humus)	100–500	300–1000	Sub-micrometer, localized charge
Vermiculite	100–200		High (–) layer charge, interlayer, often sub-micrometer
Smectites	80–150	500–800	(–) layer charge, sub-micrometer, interlayer
Illite	10–40	50–100	(–) layer charge, often sub-micrometer
Halloysite	10–40	20–100	Sub-micrometer, minor layer charge, interlayer
Kaolinite	1–10	10–40	Sub-micrometer, possibly minor layer charge in disordered kaolinite
Fe and Al oxyhydroxides	1–10	50–300	Dangling bonds, sub-micrometer CEC at pH < 7, AEC at pH > 7
Quartz, feldspars	1–2	<1	Weak van der Waals forces

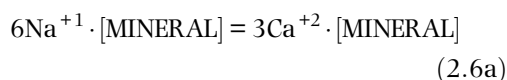
(charge-wise) – this is the outer limit of the exchange complex, and these concepts are illustrated in Fig. 2.25.

Cation-exchange capacity (Table 2.5) is expressed as molar units of exchangeable cation per mass of solid (where the solid is a given mineral or other type of solid such as organic matter). Units are centimoles of charge per kg (cmol_c/kg), although sometimes CEC is expressed in units of millimoles of charge per kg (mmol/kg). The outdated unit milli-equivalents per 100 g (meq/100 g) produces values that are identical to cmol_c/kg, but cmol_c/kg is preferred because it is consistent with modern terminology.

To illustrate the concept of cmol_c/kg, let us consider three different exchangeable cations, Na⁺, Ca⁺² and Al⁺³ and the quantity of charge contained in one mole of each:

- One mole of Na⁺ (22.98 g) contains one mole of positive charge, and one centimole (cmol) of Na⁺ (0.2298 g) contains one cmol of positive charge.
- One mole of Ca⁺² (40.08 g) contains two moles of positive charge, so one cmol of Ca⁺² (0.4008 g) contains two cmol of positive charge.
- One mole of Al⁺³ (26.98 g) contains three moles of positive charge, so 1 cmol Al⁺³ (0.2698 g) contains 3 cmol_c.

Now consider a hypothetical example of an extremely small crystal of smectite with cumulative layer charge of –6 (equivalent to 20 smectite unit cells, each with –0.3 charge). Envision that this smectite crystal occurs in sodium-rich soil water, and that the –6 charge is balanced by six Na⁺¹ ions exchanged onto its exposed surfaces (mainly interlayers). If the composition of the soil water were to change to calcium-rich, an exchange reaction would take place, and Na⁺¹ would be replaced by Ca⁺². How many calcium ions would replace the six Na⁺¹ ions? In order to maintain charge balance, three Ca⁺² ions would exchange with six Na⁺¹ ions, a reaction that can be depicted as:

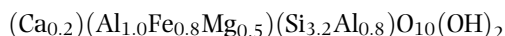


Continuing, if the soil water were to become dominated by dissolved aluminum (this occurs in soils affected by acid rain), the following exchange reaction would likely occur:



In other words, 6 mol of Na⁺ are released for each 3 mol of Ca⁺² or 2 mol of Al⁺³ adsorbed.

The role of layer charge in cation exchange can be examined for this smectite (ferruginous beidelite) from a Holocene soil in Costa Rica:



The layer charge is -0.4 and the molecular mass is 401.48 g. The CEC of this smectite produced solely by layer charge can be determined as follows:

$$\text{CEC} = |-0.4| * \text{mol}_c / 401.48 \text{ g} * 100 \text{ cmol}_c / \text{mol}_c * 1000 \text{ g/kg} = 99.6 \text{ cmol}_c / \text{kg} \rightarrow 100 \text{ cmol}_c / \text{kg}$$

Note that the result 99.6 cmol_c/kg is reported to only one significant figure, limited by imprecision in determination of the -0.4 layer charge. This value is a slight underestimate because it does not include contributions from complexation to unsatisfied bonds – the actual CEC of this mineral determined by experimental analysis is $\sim 105 \text{ cmol}_c/\text{kg}$, and this calculation illustrates the overwhelming control of layer charge on CEC in clay minerals.

One kg of smectite with $\text{CEC} = 100 \text{ cmol}_c/\text{kg}$ could adsorb 100 cmol of Na^{+1} , 50 cmol of Ca^{+2} , or 33.3 cmol of Al^{+3} . Stated in terms of mass, 1 kg of this smectite could absorb 22.98 g of Na (100 cmol $\times 0.2298 \text{ g/mol}$), 20.04 g of Ca, 8.98 g of Al, or any charge-balanced mixture of these three as long as the sum of charges amounts to 100 cmol_c/kg (e.g. 20 cmol_c each of Na and Ca and 60 cmol_c of Al).

2.6 LOW-TEMPERATURE NON-SILICATE MINERALS: CARBONATES, OXIDES AND HYDROXIDES, SULFIDES, SULFATES, SALTS

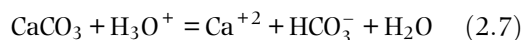
While silicates dominate most soils and sediments in terms of mass and volume, non-silicate minerals often play important roles in contaminant fate and transport, plant nutrient cycling, the carbon cycle and soil chemistry, and in many cases these

minerals are controlled by climate, and hence are important to consider with respect to analysis of paleoclimate and modern climate change.

2.6.1 Carbonates

The formative unit of carbonate minerals is the carbonate anion (CO_3^{-2}), formed by trigonal planar arrangement of one C^{+4} and three O^{-2} atoms (Fig. 2.4). The -2 charge is typically balanced by divalent cations, and by far the two most common carbonate minerals are calcite (CaCO_3) and dolomite ($\text{CaMg}[\text{CO}_3]_2$). In calcite the -2 charge on the carbonate anion is satisfied by Ca^{+2} ; in dolomite, the -4 charge of two CO_3^{-2} anions is balanced by one Ca^{+2} and one Mg^{+2} . A small sampling of the dozens of additional carbonate minerals include cerussite (PbCO_3), siderite (FeCO_3), rhodochrosite (MnCO_3), smithsonite (ZnCO_3), nahcolite (NaHCO_3), and malachite ($\text{Cu}_2\text{CO}_3[\text{OH}]_2$). There are three CaCO_3 polymorphs – calcite, aragonite and vaterite.

Carbonates are soluble in acid so tend not to occur in acidic soils like those of tropical, temperate or moist alpine regions; rather, they tend to form in semi-arid to arid soils and lakes as well as in marine environments, sedimentary and metamorphic rocks and limestone caves. The dissolution of carbonates has the potential to buffer acid deposition or acid mine drainage:



Note that the hydronium ion is consumed by calcite dissolution and the result is a circum-neutral solution with dissolved calcium buffered by bicarbonate (HCO_3^{-}).

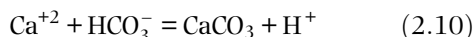
Many marine organisms scavenge Ca^{+2} and CO_3^{-2} (often in the form HCO_3^{-}) from seawater to form shells, and when these organisms die their shells fall to the seafloor to someday comprise parts of limestones. Limestones are also comprised

of carbonate mud that directly precipitates out of saturated sea water, especially in tropical and subtropical shallow seas.

The formation of carbonates effectively removes CO_2 from the atmosphere and sequesters it in sediments, soils or rocks, contributing to long-term regulation of atmospheric CO_2 . In the atmosphere, CO_2 dissolves in H_2O and falls out of the sky as carbonic acid (H_2CO_3), which then dissolves in water to produce the bicarbonate anion:



Dissolved calcium ions then combine with dissolved bicarbonate to form calcite and dissolved H^+ :



The CO_2 molecule in Equation 2.8 is transferred to carbonic acid, then to dissolved bicarbonate, and finally into a solid carbonate mineral. This reaction occurs in any solution saturated with respect to calcium carbonate, whether in a marine environment, soil or lake. Details of carbonate geochemistry are presented in Chapter 5.

Most carbonates are rhombohedral, although some, including aragonite, are in the orthorhombic crystal class. Figure 2.27 illustrates the rhombohedral crystal structure of carbonates, in this case, calcite.

Cerussite and smithsonite are ores of lead and zinc, respectively, and while effectively insoluble in neutral solutions, are soluble in weak acids. Cerussite was used as a white pigment in paint until the early 1900s. By the 1930s Pb in paint had been banned in many countries (e.g. Australia, France, Greece, Spain, Sweden, Tunisia, United Kingdom). Sale of lead in residential paint was banned in the United States in 1978. Cerussite in the form of **lead paint** is still present

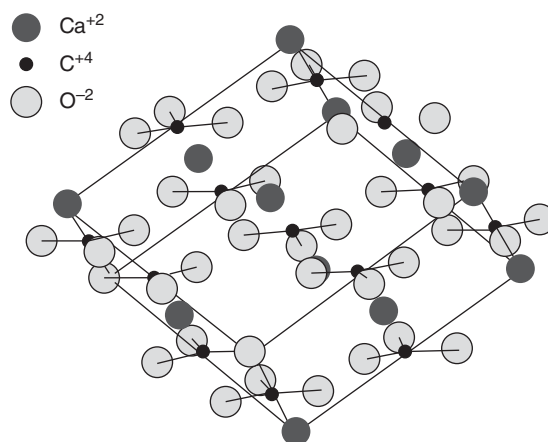


Fig. 2.27 Molecule of calcite. Note rhombohedral structure that is also expressed in macroscopic crystals. Lines between oxygen atoms are used to indicate individual CO_3^{2-} molecules within the mineral structure.

in many residential environments, where chipped and peeling paint in old buildings can cause lead poisoning (especially in children), and cerussite also occurs in lead-contaminated soils. When ingested, it dissolves in stomach acids ($\text{PbCO}_3 + \text{H}^+ \rightarrow \text{Pb}^{+2} + \text{HCO}_3^-$). From there it is incorporated into the bloodstream where it ultimately interferes with the development of the central nervous system – it is a potent neurotoxin, responsible for diminished intelligence and other neurological problems (particularly in young children), and it also impairs functioning of the kidney and blood-forming organs.

2.6.2 Oxides and hydroxides

If you have ever seen reddish or orange soils, sediments or rocks, you have seen evidence of oxides in a natural environment. These hues are produced by various configurations of iron with O and/or (OH). The most common *oxides* of iron are **hematite (Fe_2O_3)** and **magnetite**

(**Fe₃O₄**). In hematite, both Fe atoms in the unit formula are ferric (Fe⁺³); in magnetite, two Fe atoms are ferric and one is ferrous (Fe⁺²), indicating that magnetite occurs in less-oxidizing environments than hematite. Iron hydroxides include **ferrihydrite** (Fe₅HO₈•4H₂O or 5Fe₂O₃•9H₂O or Fe₂O₃•2FeOOH•2.6H₂O), **goethite** and **lepidocrocite** (both are FeOOH or Fe₂O₃•3H₂O), all of which can be considered metastable precursors to hematite. When written as Fe₂O₃•xH₂O, it becomes evident that increasing temperature and expulsion of H₂O from hydroxides will lead to the formation of hematite. Whereas hematite imparts a distinctly reddish-orange color to soils and rocks, the hydroxides impart more muted brownish yellow to brownish orange colors (Plate 1).

Many trace metals substitute for iron in oxide and hydroxide structures, so iron oxides and hydroxides exert strong controls on trace metal partitioning and mobility. Some of these trace metals also form their own oxide minerals; for example, uranium (as U⁺⁴) forms uraninite, UO₂) and copper (as Cu⁺¹) forms cuprite, Cu₂O.

In addition to iron, Al and Ti are other major elements that form oxides; for example, Al₂O₃ is the metamorphic mineral corundum and Al also forms Al(OH)₃ (gibbsite), a common mineral in well-leached soils of the tropics. There are three polymorphs of TiO₂, rutile, anatase and brookite, and FeTiO₃ is the igneous–metamorphic mineral ilmenite.

Oxides and hydroxides form in oxidized (or oxidizing) environments, which generally include well-aerated soils and sediments, including well-drained soils (as opposed to saturated wetland soils) and sandy sediments at or near the Earth's surface, or rocks with high amounts of free O₂. Oxides tend to be quite stable in these environments, but in O₂-poor reducing environments, oxides tend to dissolve, driven by reduction of iron (Fe⁺³ → Fe⁺²), a reaction often catalyzed by microbes.

2.6.3 Sulfides and sulfates

Sulfides and sulfates are sulfur-based minerals that are sensitive to redox environment and are generally highly reactive when outside of their environments of formation. Sulfides contain the *sulfide* (S⁻²) anion, a common species in chemically reduced systems such as wetlands, anoxic sediments (in deep lakes, estuaries and marine environments), and O₂-poor deeper crustal environments. Sulfates, on the other hand, contain the *sulfate* anion (SO₄⁻²) and form in oxidizing environments, especially where evaporation rates are high, as well as in certain sedimentary and low-grade metamorphic environments.

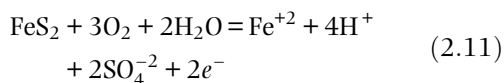
The most common sulfides are the iron sulfides like pyrite (FeS₂; Plate 2) and its polymorph marcasite, as well as pyrrhotite (~FeS) and mackinawite (~FeS), a disordered low-temperature sulfide that forms in anoxic surficial environments and imparts a black hue to chemically reduced sediments.

Given the similarity in ionic radii and charge among iron and many transition metals, the sulfide anion forms chemical bonds with a wide array of trace metal cations, resulting in:

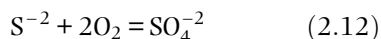
- 1 extensive substitution of metals such as As, Cu, Ni, Pb and Zn for iron in pyrite;
- 2 the occurrence of dozens of trace metal-sulfide minerals, including arsenopyrite (FeAsS), chalcopyrite (CuFeS₂), cinnabar (HgS), galena (PbS), sphalerite (ZnFeS₂) and many others.

Sulfides are common ores of heavy metals, and for millennia have been mined by civilizations ranging from Roman to Aztec (Triple Alliance) to those of the modern world. The dissolution of sulfides is a common source of trace-metal contamination and acidic drainage, whether naturally derived or exacerbated by human activity. When removed from their reducing environments of formation and exposed to oxidizing environments (most soils, river sediments, freshly

exposed rock surfaces), sulfides are unstable and oxidize rapidly. For example, when pyrite is exposed by erosion, uplift or mining to O_2 , it reacts in the following manner to release iron to solution and produce sulfuric acid:



Note that oxidation here is driven by oxidation of the S^{-2} sulfide anion to sulfate (SO_4^{-2}), a process that liberates 8 electrons from the sulfur atom, and which is worth examining individually:



The 8 electrons are consumed by the 4 oxygen atoms (or, stated differently, by the two oxygen molecules) because each O atom consumes $2e^-$ and transforms from zero-charged O_2 (i.e. O^0) to O^{-2} oxygen anions. The electron flow associated with this and other oxidation reactions is a significant source of energy to microbes, and although not depicted in this reaction, sulfide oxidation is typically linked to microbial activity, a topic that is covered in Section 2.10.

The S atom in the sulfate anion is S^{+6} . Examples of the sulfide and sulfate anions are shown in Fig. 2.28. The sulfate anion is an example of a

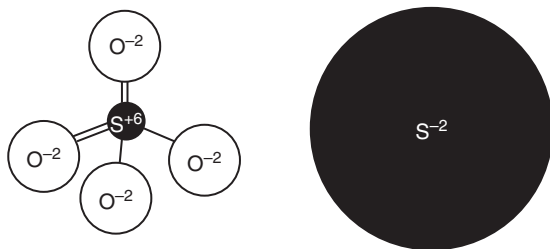


Fig. 2.28 Schematic diagram of sulfate (left) and sulfide (right) anions, with S ions drawn to scale (the radius of S^{-2} is approximately six times greater than S^{+6} in tetrahedral coordination).

polyatomic anion, because it is comprised of more than one atom.

Sulfide and sulfate minerals are structurally very different. In sulfides, the S atom occurs as the large sulfide anion (S^{-2}) with a crystal radius of 1.7 Å covalently bonded to octahedrally coordinated metal cations, whereas in sulfate minerals, sulfur is a small S^{+6} cation tetrahedrally coordinated with four O^{-2} ions. (The S^{+6} crystal radius = 0.26 Å). The SO_4^{-2} anion is octahedrally coordinated about metal cations, but the bonds are ionic, a significant difference between sulfates and sulfides (which are dominated by covalent metal-S bonds).

The metal cations in sulfates are commonly divalent (Ca^{+2} , Mg^{+2} , Fe^{+2} , Ba^{+2}) as well as monovalent (Na^{+1} , K^{+1}) and trivalent (Al^{+3} , Fe^{+3}), and some sulfates contain additional components like H_2O and OH^- . They form when solutions become saturated with respect to the metal cation and sulfate anion, environments that include arid region S-bearing soils and lakes, sulfidic mine tailings and hydrothermal systems. By far the most common sulfate is the hydrated calcium sulfate **gypsum**, $CaSO_4 \cdot 2H_2O$, a component of some arid-region soils where evaporation drives up concentrations of Ca^{+2} and SO_4^{-2} in solution to the point that gypsum precipitates directly from solution.

In acidic soils, rocks and mine tailings, where oxidation of sulfides produces dissolved SO_4^{-2} and low pH values and evaporation increases Al and Fe increase to the point of saturation, sulfates like **alunite**, **jarosite**, **melanterite** and **rozenite** and **chalcantite** precipitate out of solution.

Because the bonds between cations and sulfate anions in these minerals are ionic (and thus generally soluble in water), and also because various trace metals can substitute into sulfate mineral structures, sulfates play an important role in trace-metal transport and availability in certain environments, particularly sulfide-rich mine tailings and their surroundings. The chalcantite

depicted in **Plate 3** was photographed on a sulfide-rich rock surface in mine tailings that are rich in Cu and SO_4^{2-} (from sulfide oxidation) – chalcantite precipitates on rock surfaces on exposed tailings during periods of dry weather, but during rainy (or even very humid periods), the chalcantite dissolves, releasing copper to solution, where it eventually flows downstream and impairs aquatic habitat.

2.6.4 Halide and nitrate salts

Salts formed by ionic bonds between ions of alkali metals (e.g. Na^+ , K^+) or alkaline Earth metals (e.g. Mg^{+2} , Ca^{+2}) and the halide anion chloride (Cl^{-1}) are extremely soluble in water, so halide salts only tend to crystallize and persist in arid environments. The most common halide salt is **halite**, which forms when salty seas evaporate to produce sedimentary salt beds like basins of the Gulf of Mexico and Saudi Arabia. The salty crusts that form when playa lakes evaporate in arid regions are halite (and/or carbonates or sulfates) in a natural surficial environment (Fig. 2.29).

Nitrate salts contain a metal cation and the nitrate anion NO_3^- – they are very soluble at the Earth surface and thus only exist in very arid environments like the Atacama Desert and Death Valley. Examples include KNO_3 (known variously as potassium nitrate, niter or saltpeter) and NaNO_3 (known variously as sodium nitrate, sodium niter or chile saltpeter), both of which are used in explosives and fertilizers.

2.7 MINERAL GROWTH AND DISSOLUTION

Crystallization and dissolution of minerals occur in the presence of solutions, most commonly aqueous solutions in pores among mineral grains in soils and sediments, or water in fractures of crystalline bedrock. In some cases minerals grow

directly from aqueous solutions, a process referred to as mineral **precipitation** or **neof ormation**. In other cases, minerals form via transformation from pre-existing minerals and incorporate parts of crystallographic structures inherited from the parent mineral, a process referred to as **transformation**. Mineral dissolution can be **congruous**, the term for the complete dissolution of a mineral, or **incongruous**, a partial dissolution process where only part of a mineral goes into solution and the remainder persists in the solid state, forming a new mineral via transformation.

There are numerous external factors that influence mineral formation or dissolution, including:

1 Temperature – Minerals such as smectite, goethite and kaolinite form in low-temperature environments, i.e. soils, sediments and sedimentary rocks. Minerals like olivine, pyroxene and Ca-feldspar form at elevated temperatures associated with igneous or metamorphic environments. Some



Fig. 2.29 Salts on dessicated shoreline of Salt Creek, Death Valley, USA. The white crusty coating is halite and other soluble salts precipitated at the surface when evaporation caused solute concentrations to rise above the saturation point. Dessication polygons are ~ 0.5 m in diameter. (Courtesy of the US Geological Survey.)

minerals (e.g. quartz, hematite, calcite) can form across a wide range of temperature, from surface conditions to higher-temperature crustal regimes.

2 Redox conditions – Ferrous iron (Fe^{+2})-bearing minerals (including many silicates) typically are stable in reducing environments of wetlands or deeper crustal regimes, but are unstable when exposed to the atmosphere in soils or surficial sediments. Sulfide minerals will crystallize in reducing systems but are unstable (and thus will dissolve) in oxidizing systems; the opposite is true for iron oxides, which tend to be stable in soils but unstable in reducing environments of wetlands and O_2 -poor groundwater systems.

3 Time – Mineral growth and dissolution are always controlled to some extent by kinetic factors, some of which are inherent to specific mineral characteristics, some of which are related to the interplay of temperature, solution chemistry, redox conditions, physical conditions and biological factors.

4 Biological factors – the scale of biological control on mineral growth ranges from macroscopic to microscopic. Plant roots exude chelating compounds that extract K^+ from illite, thus enhancing its transformation to smectite, and the micro-organism *thiobacillus ferrooxidans* enhances pyrite dissolution in acidic systems by catalysing sulfide oxidation.

5 Physical factors – the porosity and permeability of sediment, soil or rock can control mineral growth by limiting (or enhancing) flow of reactants and products. For example, mineral reactivity can be retarded in clays and shales where transport of reactants to the mineral surface is limited by low permeability.

6 Solution chemistry – calcite will not form in the presence of solutions that lack dissolved calcium and carbonate (or bicarbonate), nor will calcite form in acidic systems.

Ultimately, mineral growth or dissolution is controlled by conditions at the interface of mineral surface and aqueous solution. **Crystal growth** can be viewed as a three-step process

involving: (1) movement of reactants (ions) to surface of crystal growth by advection or diffusion; (2) reaction on the mineral surface, involving, sequentially, attachment/adsorption of ions to surface, diffusion or ions to bonding sites, and formation of chemical bonds with reaction surface; and (3) desorption and diffusion or advection of reaction products away from reaction surface. Mineral dissolution similarly involves advection or diffusion of species such as H^+ or H_2O to mineral surface, attachment of H^+ or H_2O to cations (and breaking of cation – mineral chemical bond), and diffusion or advection of complexed cations away from mineral surface. Chemical reactions that foster mineral growth and dissolution clearly require movement of aqueous species (e.g. cations and anions), and the main processes governing ion transport are advection and diffusion.

Advection is the movement of matter by physical flow in response to a force. A good example of advective transport is flow of water in a river channel that transports silt particles carried in suspension – the water and silt are moving by advection, flowing down a gradient in response to gravity. Dissolved ions in the same river are also moved by advection, as are water vapor and dissolved CO_2 in moving air masses, or immiscible organic liquid contaminants (or ions or particles) in migrating groundwater.

In terms of mineral growth and dissolution, advection is important because it is a mechanism for migration of dissolved ions to and from the mineral surface, whether the aqueous solution is seeping through soil pores and carrying Ca^{+2} and HCO_3^- to a surface where calcite can precipitate, or whether it is water emanating from a groundwater spring that delivers dissolved Fe^{+2} to the land surface, where it can oxidize to form iron hydroxide.

Examples of flow rates in various geological materials are presented in Table 2.6.

Examples of crystalline rocks include basalt, granite, gneiss, marble, crystalline (metamorphosed) limestone, slate and schist. Overlap of K

Table 2.6 Hydraulic conductivity (K) of selected rocks and sediments (from Heath, 1983). Values are in units of meters per day (m/d).

Type	K
Shale	10^{-8} – 10^{-4}
Sandstone	10^{-4} – 10^0
Porous Limestone	10^{-3} – 10^0
Cavernous Limestone	10^1 – 10^4
Crystalline rocks (unfractured)*	10^{-8} – 10^{-6}
Crystalline rocks (fractured)*	10^{-3} – 10^{-1}
Clay	10^{-7} – 10^{-3}
Silt	10^{-3} – 10
Sand	1 – 10^2
Gravel	10^2 – 10^4

values of different grain sizes is the result of textural differences, mainly compaction. Well-compacted sand can be less permeable than loosely compacted silt, for example. Note that flow rates differ by many orders of magnitude as a function of grain size or rock type, and also that rates vary by orders of magnitude even within a given sediment or rock type, again accounting for textural differences.

Darcy's Law, formulated by French engineer Henri Darcy in the mid-1850s, is useful in illustrating controls on flow rate (Q) in porous media:

$$Q = K \cdot A \cdot dh/dl \quad (2.13)$$

K is hydraulic conductivity (commonly in m/s or m/day), a value inherent to the soil, sediment or rock in question (Table 2.6), A is the cross-sectional area under consideration (commonly in m^2); and dh/dl is the gradient. When combined, the terms produce units of m^3/s or m^3/day . This equation and its derivatives are commonly applied to predicting flow rates and, by extrapolation, advective transport in aquifers and other *saturated* media. It is important to note that Darcy's Law does not apply to the type of unsaturated flow that occurs in systems such as soils and sediments above the water table.

Often when analyzing advection pertaining to chemical reactions, rate of advective flow is a useful parameter. Rate of flow (V) in saturated soils and groundwaters can be determined by the following equation, where n is porosity; n = void space/total volume expressed as a decimal (after Heath, 1982).

$$V = (K/n) \cdot dh/dl \quad (2.14)$$

In saturated *sand* with $K = 50$ m/d, porosity = 0.20 (20%), and a gradient of 1.0 m per 1000 m ($dh/dl = 10^{-3}$), advective flow rate is:

$$V = (50 \text{ m} \cdot \text{d}^{-1} / 0.20) \cdot 10^{-3} = 0.25 \text{ m} \cdot \text{d}^{-1}$$

In saturated *clay* with $K = 5 \times 10^{-5}$ m/d, porosity = 0.4, and the same gradient used in the sand calculation ($dh/dl = 10^{-3}$), advective flow rate is:

$$\begin{aligned} V &= (5 \times 10^{-5} \text{ m} \cdot \text{d}^{-1} / 0.40) \cdot 10^{-3} \\ &= 1.25 \times 10^{-4} \text{ m} \cdot \text{d}^{-1} \end{aligned}$$

When ions in solution are being transported by advection and are being removed from solution by mineral growth (or adsorption), this type of transport is termed *non-conservative*. The same can be said for ions being added to solution by mineral dissolution, or degassing of CO_2 or O_2 from aqueous systems – if concentrations are altered by addition or removal of aqueous species, their behavior is considered *non-conservative*. Conversely, when ions or other species in solution are not removed or added by reactions involving minerals or other processes, the behavior of the components is termed *conservative*.

Diffusion is the migration of ions or particles from areas of high concentration to areas of low concentration. It is a spontaneous process in response to the tendency of systems to progress from ordered to more disordered, i.e. in the direction of greater entropy. Diffusion differs from advection in two main ways. The first is that movement is in response to a chemical gradient,

rather than to a physical force. The second is that it can occur in the absence of movement of the solution that contains the ion or particle undergoing diffusion. Ions can diffuse in completely stationary water bodies, or even in solids.

In some cases diffusion and advection occur simultaneously – consider advection of H_2AsO_4^- -poor groundwater through a sandy layer in contact with a clay layer containing H_2AsO_4^- -rich pore water. Because the clay is virtually impermeable, water effectively does not migrate across the sand/clay boundary, but H_2AsO_4^- ions will diffuse into the sandy layer water across a diffusion gradient. The diffusion of ions raises an important point because charge balance must be maintained, either by exchange of ions of the same charge (i.e. one cation for another cation, or one anion for another anion), or by paired migration of charge-balanced cation–anion pairs. One possible mechanism for maintaining charge balance in the sand–clay diffusion example above is that the water in the sand layer may be quite dilute compared to, e.g. Na^+ - H_2AsO_4^- -rich pore water in the clay, in which case Na^+ and H_2AsO_4^- would diffuse out of the clay into the sand, satisfying two constraints: (1) spontaneous diffusion in response to a concentration gradient, and (2) maintenance of electrical neutrality of ionic charge in solution.

If the groundwater in the sand layer is CO_3^{2-} rich compared to the clay pore water, CO_3^{2-} will likely diffuse into the clay layer and H_2AsO_4^- into the sand layer to satisfy gradients and balance charge (Fig. 2.30).

Apparently the 1850 s was a time of great thinking about flux, because in addition to the work of Henry Darcy on advection, the German physicist Adolf Fick formulated the following equation to represent the net movement of a diffusing particle proportional to the concentration gradient:

$$F = -D \cdot \frac{dc}{dx} \quad (2.15)$$

Fick's Law states that the rate of movement per unit cross-sectional area (i.e. the flux, F) toward lower an area of lower concentration is controlled by the difference in concentration ($\frac{dc}{dx}$) between two points – the greater the difference, the greater the rate of diffusion. The concentration gradient ($\frac{dc}{dx}$) is expressed in units of mol/cm^4 , which is arrived at by dividing difference in concentration (dc , which is expressed in mol/cm^3) by distance (dx , in cm). Dimensional analysis tells us that units of F are $\text{mol cm}^{-2}\text{s}^{-1}$ as follows:

$$\begin{aligned} & -D(\text{cm}^2 \cdot \text{s}^{-1}) \cdot \frac{dc}{dx}(\text{mol cm}^{-4}) \\ & = F(\text{mol cm}^{-2} \cdot \text{s}^{-1}) \end{aligned} \quad (2.16)$$

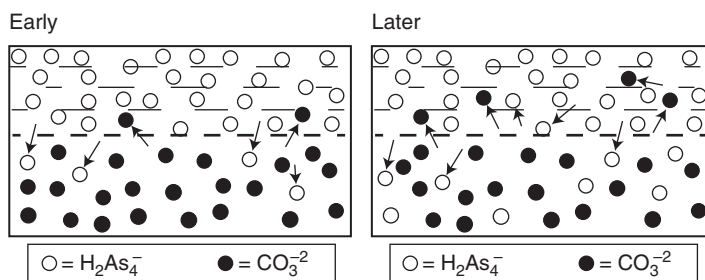


Fig. 2.30 Schematic representation of diffusion across the semi-permeable boundary between a clay-rich layer (upper) and sandy layer (lower). The pronounced difference in concentration of the two species across the boundary drives diffusion over time, from “Early” to “Later” (this system has not reached dynamic equilibrium). Loss of arsenate from the upper layer could drive further dissolution or desorption of the As-bearing minerals. According to Fick's Law, as time passes in this example, the rate of diffusion will decrease as the chemical gradient decreases. Also, note the charge balance, where two monovalent arsenate ions exchange for each divalent carbonate ion.

F is the flux of a particles or ions (molar concentration) through a cross-sectional area (expressed in cm^2) per second. The negative sign on the right side of equation 2.15 reflects the flux in the opposite direction of the concentration gradient, that is, toward areas of lower concentration.

If there is no difference in concentration, there will be no net transfer. There will still be motion, but migration of ions or molecules from point A to point B will equal migration from B to A, the classic case of a system in a state of dynamic equilibrium. The term D is the diffusion coefficient, values of which depend on the ion or molecule in question. In aqueous solutions, D is typically on the order of $10^{-5} \text{ cm}^2/\text{s}$; in solids, D is on the order of $10^{-10} \text{ cm}^2/\text{s}$. Ions with higher charge tend to diffuse more slowly than similar ions with higher charge (e.g. K^{+1} vs. Ca^{+2} ; HPO_4^{-2} vs. $\text{H}_2\text{PO}_4^{-1}$).

Ostwald ripening is a process whereby finer-grained crystals dissolve and the dissolution products re-crystallize on the surfaces of larger, compositionally similar crystals. The greater surface area : volume ratio of finer-grained crystals makes them less stable than coarser crystals. Ostwald ripening tends to occur as systems progress from relatively disordered to more ordered, a common example being the transformation over time from the disordered iron hydroxide ferrihydrite ($\sim 5\text{Fe}_2\text{O}_3 \cdot 9\text{H}_2\text{O}$) to more-ordered goethite ($\text{Fe}_2\text{O}_3 \cdot \text{H}_2\text{O} = 2\text{FeOOH}$) and eventually hematite (Fe_2O_3). The disordered, kinetically favored ferrihydrite tends to sorb trace elements (e.g. arsenate, phosphate, trace metals) to a much greater extent than goethite or hematite; thus, in this case the result of Ostwald ripening of iron oxyhydroxides from ferrihydrite $\rightarrow$ goethite $\rightarrow$ hematite tends to result in release of potential contaminants to soil water or groundwater.

2.8 BIOMINERALIZATION

Biomineralization is the term for mineral growth aided by organisms. It has been known for well

over a century that marine organisms such as bivalves, corals, gastropods, microscopic coccoliths and many other organisms remove dissolved calcite and carbonate from sea water to form shells or exoskeletons – this is the origin of much of the calcite in limestone. More recently research has focused on the role of microbes in biomineralization and this is the focus of this section.

The most-abundant biominerals are carbonates followed by phosphates, iron hydroxides and iron sulfides, sulfates (e.g. gypsum), and various other less-common groups. Amorphous silica ($\text{SiO}_2 \cdot n\text{H}_2\text{O}$) also forms in the tests of diatoms, but the amorphous structure means that it is technically not a mineral; in reality, biominerals are often very disordered and lack the crystal morphology (Fig. 2.31.a) that tends to characterize minerals formed by more “typical” inorganic processes (Fig. 2.31.b).

Biomineralization is typically divided into two main types, **biologically induced mineralization (BIM)** and **biologically controlled mineralization (BCM)**. BIM describes mineralization that is mediated by an organism but occurs outside of the organism; BCM, on the other hand, occurs when mineral growth takes place within the organism (e.g. within vesicles inside cells, as in many marine organisms that foster calcite or aragonite growth).

Given that BIM occurs outside of organisms, it is more similar to inorganic mineral growth, and the main mechanism by which organisms foster mineral growth is by locally causing supersaturation of elements; in other words, they produce elevated concentrations of mineral-forming elements in solution (e.g. immediately outside of a cell of a bacterium) to the extent that chemical bonds form and crystal lattices (or disordered solids) begin to take shape. Organisms have the potential to alter pH (e.g. by exuding H^+ or OH^-), to alter redox conditions (e.g. by exuding O_2 into pore water as a product of photosynthesis), to exude potential constituents of minerals (e.g. cations or CO_2), and to provide a surface

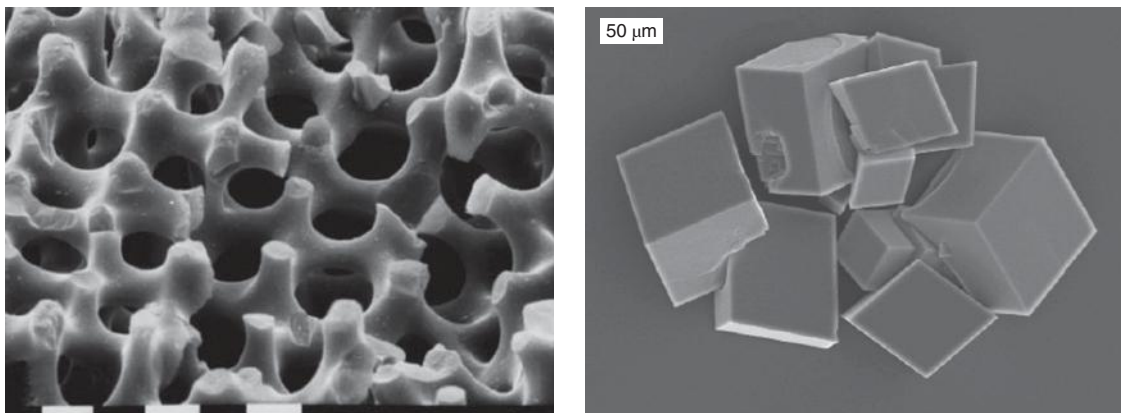
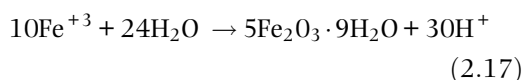


Fig. 2.31 Biogenic calcite (left, formed by an echinoderm) and inorganically precipitated calcite (right). (Weiner and Dove 2003. Reproduced with permission of the Mineralogical Society of America.)

(e.g. cell wall or biofilms exuded by organism) upon which nucleation can occur – often the lack of a condensation nucleus will limit crystal growth, so cell surfaces or exudates can provide the starting point for crystal growth.

An example of BIM concerns alteration of redox conditions by the bacteria *Galljanella* and *Leptothrix* that results in crystallization of ferrihydrite. The bacteria obtain energy by causing oxidation of iron from ferrous to ferric ($\text{Fe}^{+2} \rightarrow \text{Fe}^{+3} + e^-$). In reducing/anaerobic sediments, iron oxidation is limited by lack of available O_2 , but like all oxidation reactions this one releases an electron, an important energy source at the microbial level; thus, the promotion of iron oxidation by these bacteria provides a source of energy while simultaneously facilitating ferrihydrite formation, which can be represented as:



According to Le Chatelier's principle, increasing a reactant such as Fe^{+3} (or O_2 or H_2O) will drive the reaction to the right, so effectively what these iron-oxidizing bacteria do is serve to increase

concentrations of reactants, pushing the system (at a small scale, i.e. adjacent to the cell) to supersaturation. The ferrihydrite forms around the perimeter of the cell and over time may evolve to more-crystalline minerals such as goethite and hematite, perhaps by Ostwald ripening as described above.

Two sources of much more information and case studies of biomineralization are the *Reviews in Mineralogy and Geochemistry* issue on this topic edited by Dove et al. (2003) and the issue of *Elements* magazine devoted to this topic and edited by Lu and Dong (2012). The latter of these two focuses more explicitly on environmentally related topics.

REVIEW QUESTIONS

1. Diagram schematic structures of olivine, a trioctahedral mica (2:1 phyllosilicate) and a pyroxene. For each, indicate points or planes where water can attack the crystal and cause dissolution. Why are these points prone to attack? To help you get started, compare the bond types between Si–O and (larger cations) –O (e.g. Fe–O, Mg–O).

2. Determine the cation exchange capacity for two different soils:
 - A. Soil that contains 25% quartz, 20% $\text{Fe}(\text{OH})_3$, 50% kaolinite and 5% humus (organic material).
 - B. Soil that contains 35% quartz, 10% illite and 55% smectite.
 - C. Given that many plant nutrients exist as cations in solution, what implications does this have for the fertility of soil A vs. soil B?
3. List chemical elements that can substitute for Fe and Mg in the octahedral layer of chlorite or smectite. Constrain yourself to rows 3 and 4 of the periodic table.
4. What is the overall charge on the 2:1 layer in a smectite with a 2:1 layer composition of $\text{Al}_{1.8}\text{Mg}_{0.2}\text{Si}_{3.8}\text{Al}_{0.2}\text{O}_{10}(\text{OH})_2$? What is the overall charge on the 1:1 layer of kaolinite with a composition of $\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$? What implications does this have for the cation exchange capacity of each (≤ 3 sentences)?
5. What is the coordination number of a cation in octahedral coordination? Why is it termed octahedral coordination?
6. List 3 elements that could potentially substitute for Si in tetrahedral coordination. Justify your response using data and pertinent rules from Pauling and Goldschmidt.
7. Determine the structural (unit cell) formula for a smectite with the following composition (determined by TEM-AEM analysis, Ryan and Huertas, 2009). [For a hint, a structural formula for muscovite, e.g., is $\text{KAl}_2(\text{Si}_3\text{Al})\text{O}_{10}(\text{OH})_2$]

SiO_2	56.04
Al_2O_3	25.50
Fe_2O_3	11.15
MgO	4.48
CaO	2.80
Na_2O	<0.12
K_2O	<0.12

8. Express the composition of calcite in terms of wt% CaO and CO_2 , and also in terms of mg/kg and mol/kg of C, O and Ca.
9. Will a mineral dissolve more rapidly, or more slowly, if products of the dissolution reaction diffuse away from the mineral surface after dissolution? Provide an example of a mineral dissolution reaction as reference and use both Le Châtelier's Principle and Fick's Law to justify your response.
10. Describe the main differences between biologically induced mineralization (BIM) and biologically controlled mineralization (BCM) and explain three ways by which organisms can enhance mineral growth.

REFERENCES

- April, R.H., Hluchy, M.M., and Newton, R.M., 1986. The nature of vermiculite in Adirondack soils and tills. *Clays and Clay Minerals* 34: 549–556.
- Brindley, G.W., Maksimovic, Z. 1974. The nature and nomenclature of hydrous nickel-containing silicates. *Clay Mineralogist* 110: 271–277.
- Christidis, G.E. and Mitsis, I., 2006. A new Ni-rich stevensite from the ophiolite complex of Othrys, Central Greece. *Clays and Clay Minerals* 54, No. 6: 653–667.
- Dähn, R., Scheidegger, A.M., Manceau, A., Schlegel, M. L., Baeyens, B., Bradbury, M.H., Morales, M. 2002. Neoformation of Ni phyllosilicate upon Ni uptake on montmorillonite: A kinetics study by powder and polarized extended X-ray absorption fine structure spectroscopy. *Geochimica Cosmochimica Acta* 66: 2335–2347.
- Dove, P.M., De Yoreo, J.J. and Weiner, S. (eds.), 2003. *Biomineralization*. Reviews in Mineralogy and Geochemistry 54, Mineralogical Society of America.
- Drever, J.I., 1997. *The Geochemistry of Natural Waters: Surface and Groundwater Environments*, 3rd edn. Simon and Schuster, Upper Saddle River, New Jersey, USA.
- Gaudin, A., Petit, S., Rose, J., Martin, F., Decarreau, A., Noack, Y., and Borschneck, D. 2004. Accurate crystal chemistry of ferric smectites from the lateritic

- nickel ore of Murrin Murrin (Western Australia); I, XRD and multi-scale chemical approaches. *Clay Mineralogist Min.* 39: 301–315.
- Hattori, K., Takahashi, Y., Guillot, S., and Johanson, B., 2005. Occurrence of arsenic (V) in forearc mantle serpentinites based on X-ray absorption spectroscopy study. *Geochim. Cosmochim. Acta* 69: 5585–5596.
- Hay, R.L., Hughes, R.E., Kyser, T.K., Glass, H.D. and Liu, J., 1995. Magnesium-rich clays of the meerschau mines in the Amboseli Basin, Tanzania and Kenya. *Clays Clay Miner.* 43: 455–466.
- Heath, R.C., 1983. Basic ground-water hydrology. United States Geological Survey — Water Supply Paper 2220. Accessed on-line 5 December 2013: <http://pubs.usgs.gov/wsp/2220/report.pdf>.
- Hillier, S., 1994. Pore-lining chlorites in siliciclastic reservoir sandstones: Electron microscope, SEM, and XRD data, and implications for their origin. *Clay Minerals* 29: 665–679.
- Klein, C., Hurlbut, C.S., 1993. *Manual of Mineralogy: (after James D. Dana)*. 21st edn, Wiley, New York, 681 p.
- Langmuir, D., 1997. *Aqueous Environmental Geochemistry*, Prentice-Hall, Upper Saddle River, New Jersey, USA, 600 p.
- Lu, A., Dong, H. (eds.), 2012. *Minerals, Microbes, and Remediation*. Elements: An International Magazine of Mineralogy, Geochemistry and Petrology, vol. 8.
- Manceau, A., Marcus, M.A., Tamura, N., Proux, O., Geoffroy, N., and Lanson, B. 2004. Natural speciation of Zn at the micrometer scale in a clayey soil using X-ray fluorescence, absorption, and diffraction. *Geochimica Cosmochimica Acta* 68: 2467–2483.
- McBride, M.B., 1994. *Environmental Chemistry of Soils*. Oxford University Press, New York. 406 p.
- Moore, D. and R.C. Reynolds, Jr., 1997. *X-Ray Diffraction and the Identification and Analysis of Clay Minerals*, 2nd edn, Oxford University Press, New York.
- Newman, A.C.D., and Brown, G. 1987. The Chemical Constitution of Clays. In: *The Chemistry of Clays and Clay Minerals*. Newman ACD, ed. Mineralogical Society, London, 1987, pp. 1–128.
- Oze, C., Fendorf, S.E., Bird, D.K., and Coleman, R., 2004. Chromium geochemistry in serpentinized ultramafic rocks and serpentine soils in the Franciscan Complex of California. *American Journal of Science* 304: 67–101.
- Pauling, L., 1960. *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*. 3rd edn. Ithaca, NY: Cornell University Press.
- Petit, S. and Decarreau, A., 1990. Hydrothermal (200 °C) synthesis and crystal chemistry of iron-rich kaolinites. *Clay Minerals* 25: 181–196.
- Ryan, P.C. and Huertas, F.J., 2009. The temporal evolution of pedogenic Fe-smectite to Fe-kaolin via interstratified kaolin-smectite in a moist tropical soil chronosequence. *Geoderma* 151: 1–15.
- Ryan, P.C. and Reynolds, R.C., Jr., 1996. The origin and diagenesis of grain-coating serpentine/chlorite in Tuscaloosa Formation sandstone, U.S. Gulf Coast. *American Mineralogist* 81: 213–225.
- Ryan, P.C., Kim, J., Wall, A.J., Moen, J.C., Corenthal, L. G., Chow, D.R., Sullivan, C.M., and Bright, K.S., 2011. Ultramafic-derived arsenic in a fractured bedrock aquifer. *Appl. Geochem.* 26: 444–457.
- Shannon, R.D., 1976. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Cryst.* A32: 751–767.
- Sparks, D.L., 1995. *Environmental Soil Chemistry*, Academic Press, San Diego, USA.
- Thompson, G.R. and Hower, J., 1975. The mineralogy of glauconite. *Clays and Clay Minerals* 23: 289–300.
- Weiner, S. and Dove, P.M., 2003. An Overview of Biomineralization and the Problem of the Vital Effect. In *Biomineralization*. eds. P.M. Dove, S. Weiner and J.J. De Yoreo. Mineralogical Society of America, Washington, D.C., 54: p. 1–31.

Organic Compounds in the Environment

3.1 INTRODUCTION TO ORGANIC CHEMISTRY: CHAINS AND RINGS, SINGLE, DOUBLE, AND TRIPLE BONDS, FUNCTIONAL GROUPS, CLASSES OF ORGANIC COMPOUNDS, ORGANIC NOMENCLATURE

Organic geochemistry applies to two important issues in environmental science. The first is the fate, transport, and degradation of synthetic organic contaminants in soils, surface waters, aquifers and the atmosphere. These synthetic compounds are typically petroleum-based fuels, solvents, pesticides and pharmaceuticals. The second involves understanding the origins and behaviors in the environment of a wide array of natural organic compounds including (but not limited to) humic and fulvic substances; methane and other non-industrial hydrocarbons (from deep crustal and mantle origins to atmospheric reactions); dimethyl sulfide and other natural marine carbon–sulfur compounds; amino acids in sedimentary rocks; and charcoal produced

during forest fires. It is an interdisciplinary field including contributions from geology, soil science, biogeochemistry, environmental geochemistry, hydrogeology, limnology, oceanography, atmospheric chemistry, and others. At its most fundamental level, organic geochemistry is concerned with carbon-based compounds that are studied in organic chemistry classes, so the first part of this chapter will focus on selected concepts from the field of organic chemistry, including the structures and compositions of some compounds that are environmental contaminants.

3.1.1 Definition of organic compounds

Organic compounds are based on chains or rings of carbon atoms that are commonly bonded to hydrogen, nitrogen, oxygen, sulfur, chlorine, phosphorous, and less commonly silicon, iron, trace metals or other elements. For reasons based on historical classification and their often-inorganic origins, there are a few carbon-based compounds that are considered “inorganic

carbon”, including the carbonate minerals (e.g. calcite, dolomite), bicarbonate and carbonate anions (HCO_3^- and CO_3^{2-}), carbon dioxide, and the carbon polymorphs diamond and graphite – they are both pure carbon (allowing for trace impurities) but they possess different atomic structures. Diamonds occur in kimberlite, a rare igneous rock that has its origins in the upper mantle, so its origins are inorganic. Graphite in metamorphic rocks is produced by burial, compaction and heating of algae and other organisms that existed millions to billions of years ago. Although the origin of the carbon in graphite is often anaerobic decay of organisms (clearly organic), the mineral is considered inorganic because of its metamorphic origin.

3.1.2 Hybridization of carbon atoms in organic compounds

Carbon atoms have four valence electrons, which can be configured in three different geometric arrangements known as sp^3 , sp^2 and sp **hybridizations**. Sp^3 -hybridized carbon atoms form 4 single bonds known as σ bonds, and to minimize repulsion, the electrons spread themselves as far apart as possible, meaning they are tetrahedrally coordinated around the nucleus (Fig. 3.1). Sp^2 carbon forms a double bond comprised of σ and π bonds (using two of carbon’s four valence electrons) as well as two single bonds (each using one valence electron), and the structure of the atom is trigonal planar (triangular). Carbon atoms can also occur in sp hybridizations that form linear arrangements and feature a sigma bond on one side and a triple bond (comprised of one σ bond and two π bonds; three valence electrons) on the other. Each of these three C hybrids is shown in Fig. 3.1 using Lewis electron-dot diagrams and simplified Lewis structure diagrams featuring single lines for single bonds, double lines for double bonds and triple lines for triple bonds. Examples of simple organic compounds, each featuring

	sp^3	sp^2	sp
Lewis electron dot diagrams			
Lewis structures			
	<i>Ethane</i>	<i>Ethene</i>	<i>Ethyne</i>
Structural formulas of representative compounds			
Chemical formulas	C_2H_6	C_2H_4	C_2H_2

Fig. 3.1 Hybridization of carbon atoms, with ethane, ethene (ethylene), and ethyne (acetylene) shown as examples of compounds containing the different hybridizations.

two C hybridized atoms of either sp^3 , sp^2 or sp arrangements are also diagrammed.

Ethane is comprised of two sp^3 atoms joined by a single bond. Note in the Lewis diagrams that each of the two sp^3 atoms has three unpaired electrons that can be satisfied by bonding to the single valence electron in three H atoms, producing a chemical formula of C_2H_6 . Ethene is comprised of two sp^2 atoms joined by a double bond; the Lewis diagrams depict two unpaired electrons in each C atom that are free to bond with two H atoms to make ethene, producing the chemical formula C_2H_4 . Finally, ethyne consists of two sp hybridized atoms joined by a triple bond, and in the Lewis diagrams you will see that each C atom has only one unpaired electron available to bond to H, producing the ethyne chemical formula C_2H_2 .

Hydrocarbons like alkanes with only sp^3 orbital arrangement and single bonds are sometimes referred to as **saturated** hydrocarbons. Alkenes,

which contain at least one double bond, are considered **unsaturated**, as are alkynes. Here, the term saturation refers to the capacity for bonding to H atoms – ethane is saturated with respect to H atoms but the electrons in ethene could be rearranged to accommodate two more H atoms (in effect transforming ethene to ethane). The concept of saturation becomes important when considering fate and transport because it affects natural degradation of organic pollutants in soils, groundwater and other environmental systems. Saturated organic compounds tend to degrade more slowly, meaning that their toxic effects can persist for longer periods of time in the natural environment.

3.1.3 Alkanes

We will begin a systematic progression through some common organic compounds with the group of hydrocarbons (i.e. those compounds that contain only C and H) known as **alkanes**, a

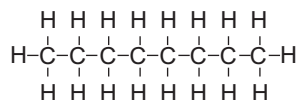


Fig. 3.2 Structural formula of octane.

relatively simple group of hydrocarbons in terms of composition and structure. Alkanes, also known as paraffins, are comprised of a single chain of sp^3 -hybridized carbon atoms bonded to each other and to H atoms by single (σ) bonds; there are no sp^2 - or sp -hybridized C atoms.

Ethane, with its short single chain of two sp^3 -hybridized C atoms bonded to six H atoms, is a good example of an alkane. Its **structural formula**, which is a graphical representation of structure and composition, is shown in Fig. 3.1, as is its **chemical formula** (or molecular formula), C_2H_6 . **Octane**, a common word in the transportation lexicon, is another good example of an alkane. Its chemical formula is C_8H_{18} and it consists of eight linearly arranged sp^3 C atoms bonded to eighteen carbon atoms. In Fig. 3.2 octane is presented in simplified Lewis structure form and does not depict the kinks and bends in the structure imparted by the tetrahedral nature of the sp^3 -hybridized C atoms.

The general chemical formula for the alkane group is $\text{C}_n\text{H}_{2n+2}$, where n represents the number of C atoms in the chemical formula. So, for example, the alkane with 15C atoms will have 32H atoms and produce the chemical formula $\text{C}_{15}\text{H}_{32}$. Table 3.1 presents some common alkanes. From this table you will discover that alkanes are primarily used as fuels. The prefixes

Table 3.1 Prefixes used in IUPAC nomenclature applied to common alkanes.

No. of C atoms	Prefix	Name	Chem. Formula	Common use
1	Meth-	Methane	CH_4	Natural gas
2	Eth-	Ethane	C_2H_6	Natural gas
3	Prop-	Propane	C_3H_8	Liquefied petroleum gas (LPG)
4	But-	Butane	C_4H_{10}	LPG - lighter fluid
5	Pent-	Pentane	C_5H_{12}	Solvent, fuel
6	Hex-	Hexane	C_6H_{14}	Solvent, gasoline
7	Hept-	Heptane	C_7H_{16}	Gasoline/petrol
8	Oct-	Octane	C_8H_{18}	Gasoline/petrol
9	Non-	Nonane	C_9H_{20}	Gasoline/petrol
10	Dec-	Decane	$\text{C}_{10}\text{H}_{22}$	Gasoline/diesel
16	Hexadeca-	Hexadecane	$\text{C}_{16}\text{H}_{34}$	Diesel, jet fuel, heating oil

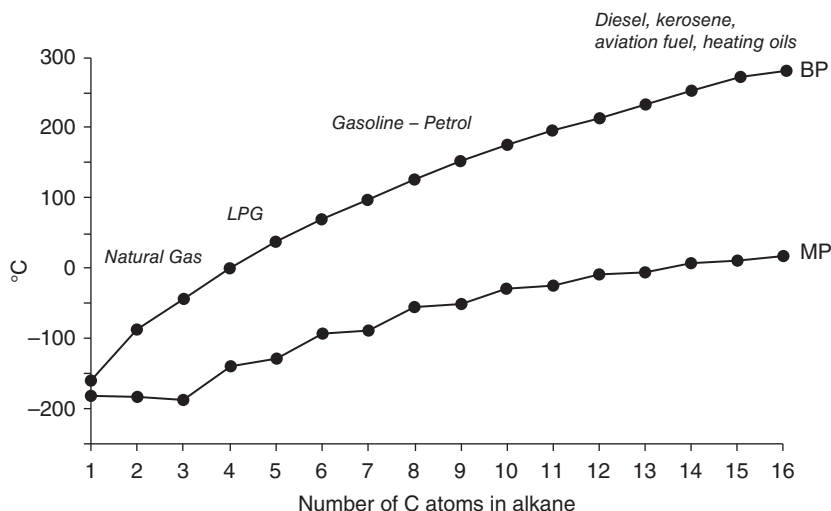


Fig. 3.3 Melting point (MP) and boiling point (BP) both increase with increasing length of alkane chain. LPG = liquid petroleum gas. Methane would plot as “1” on the x -axis, octane would plot as “8”, and so on.

presented in this table apply to all forms of hydrocarbons, particularly alkenes and cyclic hydrocarbons in addition to alkanes.

The physical properties of alkanes are closely related to the length of the hydrocarbon chain. For example, melting and boiling points both increase as a function of the number of C atoms with the exception of the slight irregularities in the melting point curve, which occur because even-numbered alkanes pack more closely in the solid phase, and thus more energy is needed to break bonds in the solid $\rightarrow$ liquid transition.

A few interesting observations from Fig. 3.3 include (1) alkanes with $C \leq 3$ occur in the gaseous state at effectively all Earth-surface conditions, (2) butane and pentane have boiling points in the range of typical air temperatures, making them transitional between gas and liquid (hence the need to compress them to produce LPG), and (3) components of diesel and jet fuel (approximately from $C = 12$ through 16) will solidify at temperatures of around 0°C . While this can create problems with these types of “heavier” fuels in cold climates, this information is useful when it comes to the refining of crude petroleum

(crude oil), which consists of a wide array of alkanes of various length as well as alkenes, aromatic hydrocarbons and other compounds. The basic principle of petroleum refining involves first heating crude oil and then cooling it; as it cools, and with the use of a complex piping system, progressively lighter compounds (smaller molecules) condense into liquid (Fig. 3.3) and are separated from the remaining crude. The first compounds to be removed are the tars and heaviest oils, progressing through lighter and lighter compounds, ultimately leaving only methane and ethane.

When methane or ethane are reacted with Cl_2 gas, the product is chlorofluorocarbons (CFCs), which are inert, stable, chemically and structurally simple, relatively non-toxic liquids used in air conditioners, refrigerators, aerosol cans and foam plastic (fast-food boxes, e.g.). Shown in Fig. 3.4 are two representative examples, CFC-12 (Freon-12) and CFC-113 (Freon-113).

The problem with CFCs is their tendency to leak into the atmosphere as gases, where their stability gives them long residence times. In the presence of ultraviolet radiation, chlorine atoms break apart from the CFC molecule, where they become

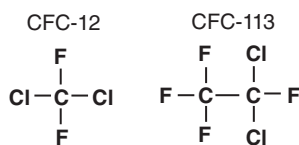


Fig. 3.4 Structural formulas for two common chlorofluorocarbons.

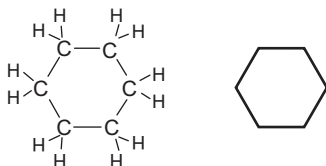


Fig. 3.5 Structural formula for cyclohexane (left) and the hexagon that is the commonly used symbol for cyclohexane.

catalysts in feedback loops that destroy stratospheric ozone (O_3) – see **Chapter 9**.

Cycloalkanes are alkanes consisting of a hydrocarbon ring rather than a single chain. Also sometimes known as naphthenes, cycloalkanes occur in crude petroleum and have the general formula C_nH_{2n} . Their nomenclature is similar to alkanes but with the prefix “cyclo-”, so the six-C cycloalkane is known as cyclohexane (C_6H_{12}).

Note that the abbreviated symbol for cyclohexane is a hexagon (Fig. 3.5) – similarly, cyclopentane would be represented by a pentagon and cyclooctane by an octagon.

3.1.4 Alkenes

A hydrocarbon chain with at least one double bond is termed an alkene, or, by the older yet still relatively common term, an olefin. You are already familiar with the common alkene known as ethene, which is commonly referred to as **ethylene** in environmental geochemistry. Similarly, propene, diagrammed (Fig. 3.6) with ethylene, is more

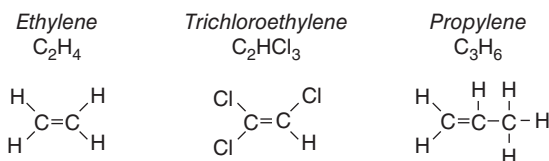


Fig. 3.6 Common ethenes.

commonly known as **propylene**. Industrial-scale reaction of ethylene with Cl_2 gas produces trichloroethylene (TCE), a solvent used to clean circuit boards and in dry cleaning, among other uses. It is a common groundwater contaminant and probable carcinogen.

Alkenes are derived directly from crude oil and by transformation from alkanes in petrochemical plants – they are used as reactants (industrial feedstock) in the production of many petroleum-based products including solvents like TCE and perchloroethylene (PCE; C_2Cl_4 , ethylene with 4 Cl replacing the 4H in Fig. 3.6) as well as polymers such as polystyrene.

3.1.5 Functional groups

Functional groups are small groups of atoms that are attached to organic molecules and impart important characteristics to the molecule to which they are attached. In particular, functional groups influence many properties of compounds, including solubility, toxicity, volatility, and reactivity. Examples of functional groups include hydroxyls, alkanes, alkenes, aromatic hydrocarbons, metals, halides, and many other components. Functional groups typically are comprised of atoms such as hydrogen, oxygen, nitrogen, chlorine and sulfur, and carbon is often a constituent as well (e.g. the two C atoms in an ethyl group).

Many functional groups, particularly those containing oxygen, nitrogen or halides (F, Cl, Br, I), can create polar molecules from compounds that are otherwise non-polar (and hence insoluble in water) – this clearly has industrial

applications (e.g. increased effectiveness of solvents), but equally as important can have environmental consequences related to contaminant fate and transport. As a rule, organic compounds are insoluble in water (picture an oil slick on the ocean), but increased solubility produced by many functional groups means that organic compounds with O, N, and Cl (certain pesticides and solvents, for example) will be more soluble in groundwater and present greater challenges for remediation. A set of selected functional groups is shown in Table 3.2.

We will return to functional groups in the following sections, but for purposes of illustration, let us consider the effect of two functional groups on an ethane skeleton (Fig. 3.6). In the first, one H atom is replaced with a hydroxyl (OH^-) functional group, producing the compound known as **ethanol** that possesses different properties from those of ethane (Fig. 3.7) – unlike ethane,

it is highly soluble in water (the OH makes ethanol a polar molecule), less volatile and, in moderate doses, a non-toxic component of beverages like wine, beer and other alcohol-bearing drinks.

If the carbon on the left side of the ethane skeleton is part of a carboxyl group (COOH) rather than a methyl group (CH_3 , as in ethane), the result is acetic acid (also known as ethanoic acid), a compound with markedly different

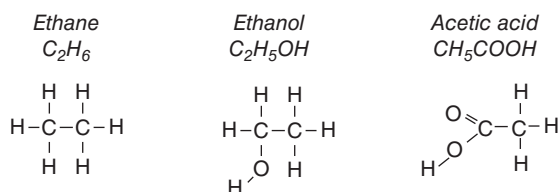


Fig. 3.7 Examples of two functional groups attached to an ethane chain. Ethanol contains a hydroxyl group (OH^-) in place of a hydrogen atom. Acetic acid contains a carboxylic acid functional group (COOH^-).

Table 3.2 Examples of some relatively common functional groups.

Group Name	Formula	Structure	Prefix	Suffix	Example
Methyl	CH_3^{-1}	$\begin{array}{c} \text{H} \\ \\ \text{R}-\text{C}-\text{H} \\ \\ \text{H} \end{array}$	Methyl-	-methane	Methyl bromide
Ethyl	$\text{C}_2\text{H}_5^{-1}$	$\begin{array}{c} \text{H} \quad \text{H} \\ \quad \\ \text{R}-\text{C}-\text{C}-\text{H} \\ \quad \\ \text{H} \quad \text{H} \end{array}$	Ethyl-	-ethane	Ethylbenzene
Hydroxyl (alcohol)	OH^{-1}	$\text{R}-\text{O}-\text{H}$	Hydroxy-	-ol	Ethanol
Carboxyl	COOH	$\begin{array}{c} \text{O} \\ \\ \text{R}-\text{C} \\ \\ \text{OH} \end{array}$	Carboxy-	-oic acid	Acetic acid
Phenyl	$\text{C}_6\text{H}_5^{-1}$		Phen- Phenyl-	-benzene	Polychlorinated biphenyls (PCBs)
Chloro	Cl^{-1} , Br^{-1} , F^{-1} , I^{-1}		Chloro- Bromo- Fluoro- Iodo-	-chlorine -bromine -fluorine -iodine	Trichloroethylene (TCE)
Amine	NH_2	$\begin{array}{c} \text{H} \\ \diagup \\ \text{N} \\ \diagdown \\ \text{H} \end{array}$	Amino-	-amine	N-nitrosamine
Sulphide	S^{-2}	$-\text{S}-$	–	-sulphide	Dimethylsulphide (DMS)

properties from those of ethane. For example, the H atom bonded to the O atom can dissociate from the rest of the compound according to this reaction, which produces the acetate anion (abbreviated OAc):



Carboxylic acids such as formic (methanoic) acid, acetic (ethanoic) acid and amino acids occur naturally (ants inject formic acid when they bite), and carboxylic groups are the source of the acidity in humic and fulvic acids in soils and surface waters. Typical pK_a values of carboxylic acids are -2 to -5 meaning that they are relatively weak acids similar in strength to carbonic acid. In natural waters where inorganic acids (e.g. nitric acid or sulfuric acid) are lacking or are low in abundance, carboxylic acids may buffer aqueous systems at pH of approximately 4 to 5.

3.1.6 Aromatic hydrocarbons and related compounds

Aromatic hydrocarbons are cyclic hydrocarbons characterized by six C atoms bonded to each other by alternating sigma and pi bonds. The alternating

bonds were once thought to be fixed as illustrated below in Fig. 3.8, but it is now known that the bonds between carbon atoms are formed by delocalized electrons that effectively produce alternation of single and double bonds. Effectively, the bonds in aromatic hydrocarbons can be viewed as a hybrid of σ and π bonds.

The symbol shown in example C of Fig. 3.8 is the currently accepted symbol for the benzene ring. If no atoms are shown attached to the ring, it is assumed that the only atoms bonded to the carbon atoms are hydrogen atoms, as is shown in example D. Examples A through D all essentially depict benzene (C₆H₆), with A, B and D each depicting the compound at an instant in time.

Many types of functional groups can attach to the benzene ring to produce a variety of compounds, some of which are covered in the following section.

Benzene rings can also polymerize to produce the family of compounds known as polycyclic aromatic hydrocarbons (Fig. 3.9), or polyaromatic hydrocarbons, or polynuclear aromatic hydrocarbons (abbreviated PAHs in any case). Naphthalene, a two-ring PAH, is a component of coal tar and occurs as a soil and aquifer contaminant near coal-tar facilities. Benzo(a)pyrene, a five-ring PAH, is a product of incomplete combustion and a

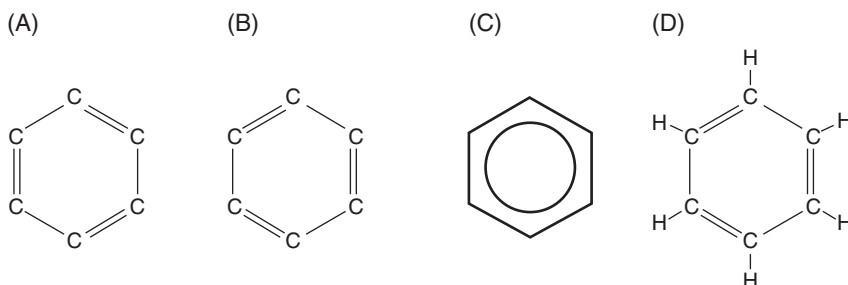


Fig. 3.8 Representations of benzene rings. A and B indicate alternation of sigma and pi bonds. At one instant in time, the double bonds are located as is shown in example A; the next instant in time they switch to the configuration in example B. This rapid alternation that characterizes the delocalized single and double bonds in aromatic hydrocarbons. C is the commonly used, pared-down structural formula for benzene, and a more-detailed structural formula is presented in D.

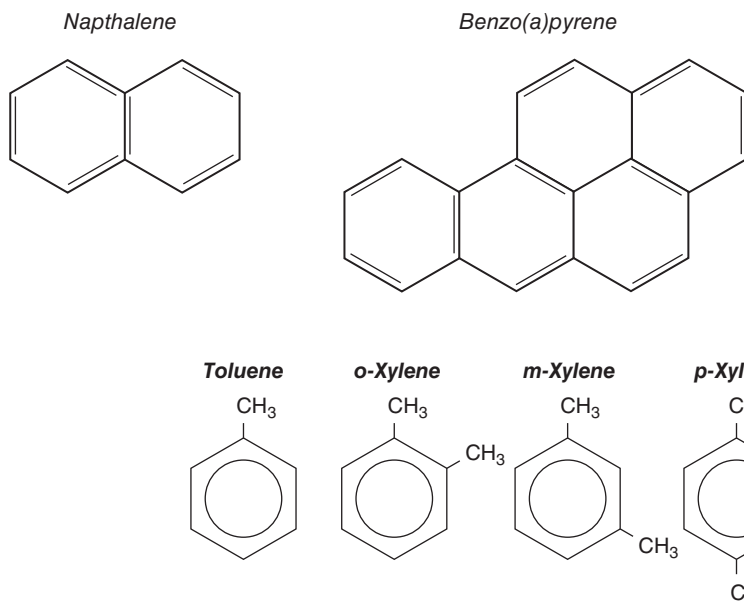


Fig. 3.9 Structural formulas of naphthalene and benzo(a)pyrene.

Fig. 3.10 Structural formulas for toluene, the three isomers of xylene, and ethylbenzene. Composition, abundance and positioning of functional groups distinguish these compounds.

known indoor air contaminant, particularly in buildings with cigarette smoke or wood-stove smoke.

Finally, as you might suspect, they are called “aromatic hydrocarbons” because of the sweet odor that they emit, one that you have probably caught a whiff of at a gasoline filling station. However, not all aromatic hydrocarbons produce those distinctive aromas – the term originated in the 19th century – and technically what makes a compound “aromatic” is the cyclic ring structure with delocalized σ and π bonds.

Addition of a methyl group to a benzene ring produces **toluene**. **Xylene** is produced if two methyl groups occur on a benzene ring – note the different geometric arrangements of xylene, where o signifies the ortho position, m signifies the meta position and p signifies the para position. **Ethylbenzene**, as the name strongly suggests, contains an ethyl group on a benzene ring.

Benzene, toluene (also known as methylbenzene), ethylbenzene and xylene (also known as

dimethylbenzene) comprise a group of aromatic compounds known in the environmental circles as **BTEX** (Fig. 3.10) – they occur naturally in crude petroleum, and because they are of similar molecular mass to octane, they behave similarly during petroleum refining and are partitioned into gasoline. They are more soluble than octane and other alkanes in gasoline and can present significant groundwater contamination problems in areas with leaking fuel-storage tanks.

Phenols (short “e”, accent on second syllable) are a group of compounds derived from benzene rings by addition of at least once hydroxyl group in place of an H atom (Fig. 3.11). If all five H atoms are replaced by Cl atoms, the result is a toxic liquid known as pentachlorophenol (and by now you should be starting to appreciate the systematic nomenclature for organic compounds). “Penta”, as it is often known, is used as a wood preservative and is the target of many groundwater remediation efforts.

Two phenyl groups, each with one chlorine atom in place of a hydrogen atom, attached to

a chlorinated ethane chain, comprise the molecule known variably as dichloro-diphenyl-trichloroethane (Fig. 3.12), **DDT**, or 4,4'-(2,2,2-trichloroethane-1,1-diyl)bis(chlorobenzene), its IUPAC name.

The Swiss chemist P.H. Muller was awarded a Nobel Prize in 1948 for his discovery of the usefulness of DDT as an insecticide, which

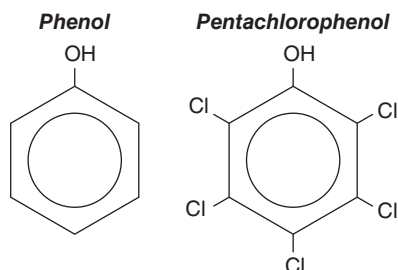


Fig. 3.11 Structural formulas of phenol (C_6H_5OH) and pentachlorophenol (C_6Cl_5OH).

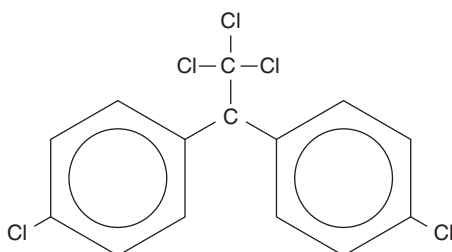


Fig. 3.12 Structural formula of the compound variably known as dichloro-diphenyl-trichloroethane (**DDT**) and 4,4'-(2,2,2-trichloroethane-1,1-diyl)bis(chlorobenzene).

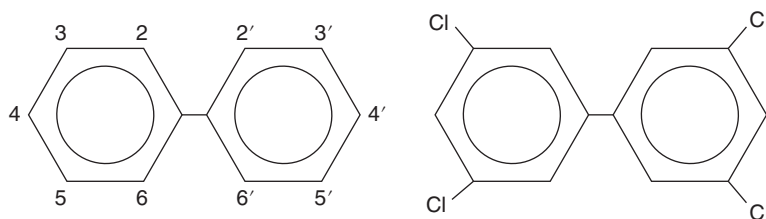


Fig. 3.13 Structural formula of the class of compounds known as Polychlorinated biphenyls (PCBs). The numbered sites on the left correspond to potential bonding sites of functional groups (in PCBs, mainly H or Cl, sometimes OH). The example on the right is 3,3',5,5'-tetrachlorobiphenyl.

resulted in great success at combating insect-borne diseases like malaria. However, DDT is a persistent organic pollutant that decomposes very slowly in the natural environment, and thus bioaccumulates and biomagnifies up the food chain. It was responsible for drastic declines in eagle populations in the 1960s and 1970s because of the way that it causes thinning of egg shells, early hatching and high mortality rates of chicks. Because of their slow decomposition rate, DDT and its metabolites dichlorodiphenyldichloroethylene (DDE) and dichlorodiphenyldichloroethane (DDD) remain pollutants in many aquatic ecosystems. In fact, DDE and DDD are more toxic and more resistant to decomposition than DDT.

Polychlorinated biphenyls (PCBs, Fig. 3.13) were used extensively in the 20th century because they are very stable liquids that are excellent electrical insulators. They are relatively simple chemically, containing two phenyl groups bonded across C atoms with variable amounts of Cl in place of H at any of the sites 2 through 6 and 2' through 6' on the benzene rings (phenyl groups).

PCBs are carcinogenic, and although no longer used, their resistance to biodegradation means they are persistent organic pollutants that remain trapped in sediments in aquatic environments and also are transported atmospherically – they have been detected in regions far from manufacture and use, including the polar regions. The Great Lakes and the Hudson River in North America are well-known problem areas, and PCBs in

fish, poultry, pork and beef have caused scares in Europe and other sites globally.

Dioxins are a family of compounds produced as unintended by-products of pesticide manufacture, but they also occur naturally in trace amounts – humans have probably been producing dioxins ever since they could control fire because combustion of wood or peat with trace amounts of natural Cl can lead to chlorinated dioxin production. For example, dioxin has been discovered in bog sediments in Canada presumably derived from combustion well before the pre-industrial era (Silk et al., 1997). A similar occurrence has been noted in England, where the pre-industrial era dioxin is a different isomer from modern industrially associated versions (Green et al., 2001). Anthropogenic sources of dioxins include pulp mills that use chlorine to bleach paper – during this process, Cl from the bleaching agents combines with organic matter from wood pulp to produce dioxins, and for this reason there is a move to limit the amount of chlorine used and to capture dioxins and related compounds produced during this process. Polychlorinated dibenzo-*p*-dioxin (PCDD) is a common by-product of pesticide production, and similar to PCBs, chlorine atoms can occur at any of the sites numbered (in the case of dioxin) 1 through 4 and 6 through 9.

In the example in Fig. 3.14, chlorine atoms occupy positions 2, 3, 7 and 8 and the compound is given the name 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), a common form of dioxin in the environment. Its presence in agent orange (see

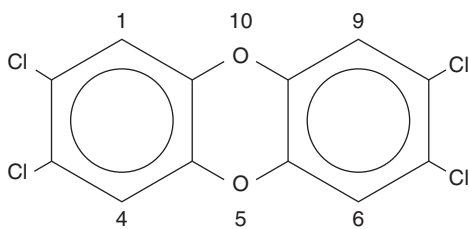


Fig. 3.14 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), a member of the dioxins. Sites labelled with 1, 4, 6 and 9 are occupied by H.

2,4-D and 2,4,5-T below) used as a defoliant during the Vietnam-US war in the 1960s–1970s has been attributed to **neurological** problems in exposed populations. Other selected incidents of human illness and ecological contamination include well-documented exposures at Times Beach, Missouri, where dioxin-contaminated oil was used to control dust, and at Love Canal, New York, where chemical waste containing dioxin was buried under soil under what was to become a subdivision. Both of these incidents occurred in the late 1970s. An accident at an agent orange manufacturing plant in Spolana, Czechoslovakia (1968) caused short-term problems (human poisoning) and long-term ecological problems from wastes washed into rivers and soils by subsequent floods. In 2004, the then-candidate for the presidency of the Ukraine, Viktor Yushchenko, was poisoned by very high doses of dioxin (allegedly added to a drink or food) that produced **chloracne**, a characteristic symptom of chlorinated hydrocarbon poisoning.

The herbicide **2,4-dichlorophenoxyacetic acid** (2,4-D) is a broad-leaf specific herbicide that usually does not kill grasses. For this reason, it is used to kill unwanted flowering and leafy plants, including noxious weeds in grasslands and other areas. It is also a useful example of a synthetic organic compound because it contains numerous functional groups, including carboxyl, phenyl and chlorine.

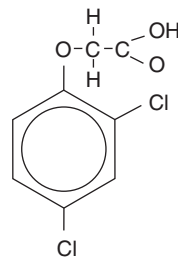


Fig. 3.15 Structural formula of 2,4-dichloroethylene. Note the carboxyl (COOH) functional group in the upper right of the diagram and two Cl functional groups.

In the absence of dioxin, 2,4-D (Fig. 3.15) has relatively low toxicity according to the USEPA, UKACP and other organizations. It is relatively soluble due to the presence of the carboxyl, chlorine and phenoxy groups and is prone to seepage to groundwater. Yet at the same time, its high solubility means that it is available to microbes and hence it decomposes relatively rapidly in well-aerated soils. Adding a third Cl atom to 2,4-D produces 2,4,5-trichloroethylene (2,4,5-T), the defoliant known as Agent Orange.

3.1.7 Nitrogen, phosphorus and sulfur in organic compounds

Dimethyl sulfide (DMS) is a natural compound produced by organic decay in marine waters that plays a role in global climate. Released to the atmosphere as DMS gas, it oxidizes to produce sulfate (SO_4^{2-}) particles that scatter sunlight, limiting the amount of solar radiation that reaches the Earth's surface.

Glyphosate (Fig. 3.16) is a commonly used **organophosphate** herbicide that, unlike many organophosphate insecticides, may not affect the human nervous system in notably deleterious ways. It is the active ingredient in commercially available weed-killers like Roundup and Rodeo, and is a useful example because of its widespread use and the fact that it contains both N and P functional groups. The N occurs in a NH ketimine group (analogous to an amine group lacking

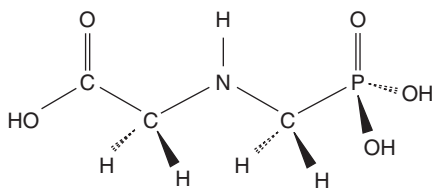


Fig. 3.16 Structural formula of glyphosate, known as N-(phosphonomethyl) glycine in IUPAC nomenclature. Note carboxyl group (far left), amine group (center), and phosphate (far right).

one H atom) while the P occurs in a phosphono group, which is basically a phosphate group with a C atom in place of an O atom (note tetrahedral configuration). Hydroxyl groups are sometimes written as HO rather than OH to illustrate bonding geometry; for example, in Fig. 3.17 the hydroxyl on the left contains an O atom bonded to C, hence the HO on the left side of the molecule.

The presence of carboxyl and hydroxyl groups enhances the solubility of glyphosate (these two functional groups can lose a proton, especially in neutral to alkaline waters), and it tends to decompose more rapidly than most herbicides. It is also desirable because it is believed to be less toxic than other herbicides (e.g. those in the organochlorine family), yet it still presents short-term respiratory risk and long-term risk to human kidneys and reproductive systems, according to the USEPA. Some genetically modified (GM) crops have been engineered to be resistant to glyphosate, so fields are sprayed with glyphosate in the hope that the glyphosate kills all plants but the GM crop.

Malathion is an organophosphate insecticide that has been used to spray wide areas for mosquito eradication, as was the case in New York City during the West Nile Virus scare of 1999–2000. It is also used as a topical remedy for killing head lice. Its relatively complex structure is

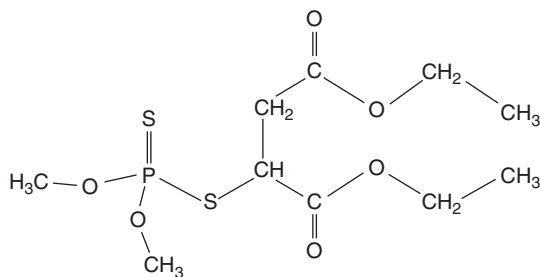


Fig. 3.17 Structural formula of malathion, which is known as O,O-dimethyl dithiophosphate or diethyl mercaptosuccinate in IUPAC nomenclature. Note the numerous functional groups, including methyl (CH_3), phosphate with sulfide in place of an O atom (thiophosphate), and esters (region with O double bonded to C–O).

illustrated below. Note that the geometry is presented in a more abbreviated form than is shown for glyphosate in Fig. 3.17.

It is appreciably less toxic than parathion, diazinon and other related organophosphates, yet still it is a potential carcinogen (the USEPA is still weighing evidence), and in the human body and in chlorinated water supplies it transforms to the **metabolite** malaoxon, which is toxic in elevated amounts.

Picloram, marketed under the trade name Tordon (IUPAC name: 4-amino-3,5,6-trichloro-2-pyridinecarboxylic acid), and **atrazine** (1-chloro-3-ethylamino-5-isopropylamino-2,4,6-triazine), are also common herbicides. Picloram is somewhat selective in that it is toxic to woody and broadleaf plants but not to most grasses, so is commonly applied to rangelands to control noxious weeds. It is relatively non-toxic according to the USEPA. Atrazine is less selective and used for control of early-emergent woody, broad-leaved and grassy plants, and it is believed to be toxic to wildlife (e.g. causing hermaphroditism in amphibians and reptiles) and potentially carcinogenic to humans. It is banned in the EU but used widely elsewhere, including the US, where it is applied to combat the spread of weeds in cornfields.

Methyl tert-butyl ether (Fig. 3.18) is a gasoline additive used as an oxygenate to improve combustion and prevent formation of volatile organic compounds produced by incomplete combustion. 1990 amendments to the US Clean Air Act required addition of oxygenates like MTBE in regions with smog problems, including large cities like LA, Houston and Denver, but also in smaller cities in the intermountain American west like Missoula, Montana

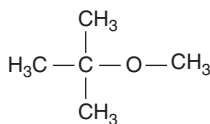


Fig. 3.18 Methyl tert-butyl ether (MTBE), or 2-Methoxy-2-methylpropane according to IUPAC nomenclature.

and Boise, Idaho. While MTBE contributed to improved air quality, its high solubility makes it a groundwater contaminant, and it imparts an unpleasant, turpentine-like taste to water at concentrations $\sim < 20$ ppb and may be a carcinogen.

3.1.8 Pharmaceutical compounds

Numerous pharmaceutical compounds including antibiotics, synthetic hormones, pain-killers and antidepressants have emerged in the past two decades as threats to drinking water and aquatic ecosystems. If they are incompletely adsorbed in the bodies of humans or livestock, these compounds may enter aquatic systems. A more obvious pathway is haphazard disposal of pharmaceuticals as waste. Many are compounds that degrade slowly as they migrate through municipal wastewater systems to sewage treatment plants, or leak out of landfills into groundwater systems, so they tend to travel through channels and by overland flow into streams, lakes, estuaries and marine waters.

Many synthetic organic compounds are endocrine disruptors that interfere with the synthesis and cycling of natural hormones in the bodies of humans and other animals. As a result, endocrine disruptors can negatively impact reproduction, development, the neurological system, metabolism and behavior. The effect of DDT on development of bird eggs is a clear result of endocrine disruption, and there are many other examples, including feminization of fish, mammals, amphibians and reptiles, decreased sperm counts in human males, cancers in wildlife and humans, and much more. The list of compounds is very extensive, and ranges from pharmaceuticals, dioxin and dioxin-like compounds, polychlorinated biphenyls, DDT and other pesticides, and plasticizers such as bisphenol A, a component of many plastics (including food and bottled-water packaging).

A classic case of environmental trade-offs is the case of the common birth control drug

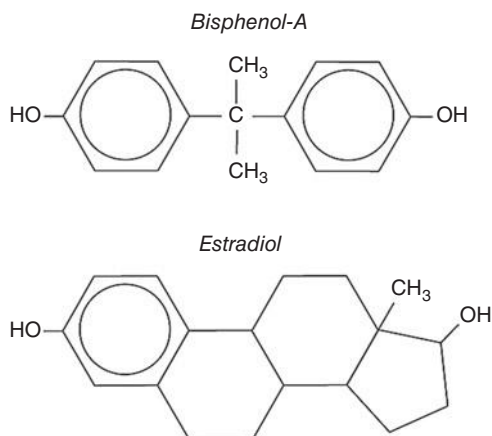


Fig. 3.19 Top = structural formulas of bisphenol-A (top), known as BPA and also as 4,4'-(propane-2,2-diyl) diphenol in IUPAC nomenclature. Bottom = structural formula of estradiol, also known in IUPAC nomenclature as (17 β)-estra-1,3,5(10)-triene-3,17-diol.

ethinylestradiol (17 α -ethinylestradiol), which is a potent synthetic estrogen-based endocrine disruptor present in sewage effluent, but also is at least partially responsible for tactics to control human population growth. As a synthetic hormone, it is not hard to envision how it affects the endocrine system – it is intended to do so. Bisphenol-A (Fig. 3.19), on the other hand, is not designed to affect the endocrine system, but by comparing the similarities in composition and structure between natural sex hormones like estradiol (Fig. 3.19) and the many synthetic organic compounds presented in the previous pages, it is not difficult to envision how the endocrine systems of humans and other animals can be affected, often during fetal development (e.g. Colborn and Thayer, 2000).

3.2 NATURAL ORGANIC COMPOUNDS AT THE EARTH SURFACE

Natural organic compounds occur in soils, surface water and groundwaters, some in dissolved forms referred to as dissolved organic

carbon (DOC) and others in solid forms, commonly termed particulate organic carbon (POC). Concentrations vary widely, from <1 mg/l (ppm) in oceans and rain water to ~100 ppm in some marsh and soil waters. The distinctive brownish tea-like color of streams in humid regions is due to DOC.

Humic and fulvic substances are among the most common forms of natural organic carbon in surface environments; combined, humic and fulvic substances are known as humus (Brady and Weil 2002). Examples include:

1 Humin, the insoluble component of humus, is derived from decomposition of the more-resistant plant-derived organic components including carbohydrates, aromatic and alkane groups, proteins, and organic acids (sugars and starches are rapidly consumed by organisms).

2 Humic acid, the soluble form of humin that occurs abundantly in soil waters.

3 Fulvic acid, the most soluble form of humus that tends to occur in surface waters along with humic acid leached from soils. Plate 4 shows tea-colored river water with dissolved organic carbon (DOC) in the form of humic and/or fulvic acids.

Humin, humic acid and fulvic acid do not have fixed chemical compositions, but rather consist of a varied array of alkane chains and rings, aromatic, carboxyl and hydroxyl groups, and some S- and N-based functional groups. Humic acid is richer in C and contains less O than fulvic acid (McBride, 1994).

Table 3.3 Compositional ranges of some natural organic compounds (by wt%).

	Humic	Humic Acids	Fulvic Acids
Carbon	>62%	50–60%	40–50%
Hydrogen	<3	3–6%	4–7%
Oxygen	<30%	30–40%	40–50%
Nitrogen	>5%	2–5%	1–4%
Sulfur	<0.5	0.1–1%	0.1–4%

Natural organic compounds play important roles in carbon cycling and climate forcing, chemical weathering, storage and release of plant nutrients, complexation of Al, Fe and trace metals (which enhances metal mobility), and contaminant fate and transport (i.e. the behavior in soils, air and water of contaminants). The flux of large quantities of iron from tropical forests to the oceans is enhanced by complexation of Fe to organic matter, and this relationship becomes important when considering that iron is a limiting nutrient in the oceans that plays an important role in phytoplankton productivity and hence fixation of CO₂ from atmosphere to ocean. Similarly, formation and degradation of POC and DOC in soils is closely linked to C cycling and climate – flux of C to and from soil organic matter can be either a source or sink of atmospheric CO₂. Fine-grained particulate organic matter in soils strongly adsorbs many non-polar organic contaminants, limiting their mobility but also their degradation rates. Humus exerts strong controls on developing and preserving soil aggregates, and thus exerts controls on water movement through soils as well as resistance to erosion.

3.2.1 Fossil fuels

While relatively unstable in oxidizing environments, organic matter tends to persist in reducing environments like swamps and sediments on the deep ocean floor. Swampy sediments and ocean floor, respectively, are where the fossil fuels **coal** and **petroleum** originate: “petro” comes from the Greek word for rock and “oleum” is derived from the latin word for oil. When buried under layers of sediment, marine organisms (mostly algae and zooplankton) transform to hydrocarbons and plant-derived terrestrial organic carbon transforms to coal – these post-burial transformations are forms of **diagenesis** (as are mineralogical transformations in sediments and sedimentary rocks such as smectite → illite).

Kinetic and thermodynamic factors dictate that deep, long-term burial will transform immature organic matter into coal and petroleum hydrocarbons. (The origin of fossil fuels is covered in greater detail in the section on the carbon cycle in Chapter 8, but the following section presents some basic concepts to facilitate subsequent discussion of fate and transport of organic compounds).

Swamp-derived humic organic matter transforms to coal through the maturation sequence shown below:

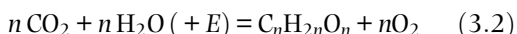
peat → lignite (low – grade coal) → bituminous → anthracite (high – grade coal), and eventually → graphite (in metamorphic environments).

Coal consists of a mix of alkane chains and rings and aromatic groups, and with increasing grade progresses from relatively low-carbon lignite (70–80 wt% C) to high-carbon anthracite (>90% C) and eventually to graphite, which is nearly purely C other than impurities such as sulfur, which comprises 1–3% of most coals and is the source of sulfuric acid in acid rain.

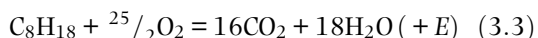
Soon after burial in anoxic deep-ocean sediments, remains of microscopic marine organisms begin to decompose, and the lack of O₂ gas prevents oxidation of organic matter to CO₂ gas. Kerogen, a waxy compound with long-chain alkanes and various other immature hydrocarbons, forms in the early stages of maturation. The immature high molecular weight alkane chains undergo catagenesis, a process whereby long alkane chains are transformed to progressively shorter chains (this is also known as cracking), transforming waxy kerogen into tar-like hydrocarbons and eventually into liquid and gaseous hydrocarbons. The end products of catagenesis are the gases methane and ethane. Liquid and gaseous hydrocarbons typically mature in organic-rich shales (known as source rocks) but typically are forced out of the high-pressure pores in deeply buried shales and into permeable sandstones

and limestones (reservoir rocks) with lower pore pressures. From there, extracting liquid hydrocarbons is almost as easy as pumping water out of a groundwater well, not accounting for complications produced by geological structures, clogging of pores by clays, destruction of drilling rigs by high pore pressures, and various other issues.

The C stored in sedimentary rocks as coal and hydrocarbons was first fixed into plants through photosynthesis (the E represents solar energy).



The oxygen depicted in the sugar $\text{C}_n\text{H}_{2n}\text{O}_n$ is likely to be consumed by mineralogical or biochemical reactions in the O_2 -poor conditions of anoxic sediments, leaving C and H to form hydrocarbons or coal. So, when we burn fossil fuels, we effectively are transferring carbon from a long-term geological repository where it is stored in reduced form to the atmosphere in the form CO_2 .



One way to think of combustion (of octane in this case) is as the reverse of photosynthesis, where the reactants (instead of products) are O_2 and reduced C, and products (instead of reactants) are CO_2 , H_2O and energy (heat and light energy).

3.3 FATE AND TRANSPORT OF ORGANIC POLLUTANTS, CONTROLS ON BIOAVAILABILITY, BEHAVIOR OF DNAPLS AND LNAPLS, BIODEGRADATION, REMEDIATION SCHEMES

Thousands of refined and synthetic organic compounds have been introduced to the environment, and in the past five decades many have emerged as some of the most problematic sources of pollution in air, soil and water. The most common

contaminants are pesticides (including chlorinated hydrocarbons and organophosphates), organic solvents and pesticides and fossil fuels and associated aromatic hydrocarbons – sources are varied and include industry, agriculture, golf courses, the military and residential lawns. While it is difficult to generalize the behavior of such a large group of compounds, organic contaminants tend to possess low solubility in water and high vapor pressures (volatility). Their fate and transport is largely controlled by the following factors (which, to some extent, are interdependent):

- 1 Partitioning of the compound into gaseous, liquid or solid phases.
- 2 The solubility of the compound, a property that is generally inversely related to adsorption.
- 3 The degree to which the organic contaminant in question will adsorb to minerals or organic matter in soils and sediments.
- 4 The density of the liquid organic compounds relative to water.
- 5 The extent to which the compound will undergo degradation (decomposition into CO_2 , O_2 and other products), a process that depends on factors including biological controls (the composition of the natural microbial community, for example), physical properties (e.g. temperature, moisture content, porosity and permeability of soil or sediment), chemical characteristics (e.g. reduction-oxidation potential, pH, salinity) and kinetic controls.

3.3.1 Solid-liquid-gas phase considerations

Most organic contaminants are liquids, often with high vapor pressures (think of the odorous vapors that escape from gasoline-petrol), so we must consider the potential of such compounds to vaporize when spilled or applied to the land surface or subsurface. Vapor pressure is defined as the pressure of the gas phase (vapor) in equilibrium with a liquid at a given temperature in a closed system (where the vapor cannot continuously evaporate out of

Table 3.4 Data on various properties of organic compounds. Sources are Drever (1997), Girard (2005) and US Centers for Disease Control (<http://www.cdc.gov/niosh/>).

Compound	Composition	Log Kow	Solubility (mg/L)*	Specific Gravity	Vapor Pressure (Pa)
Aldrin	C ₁₂ H ₈ Cl ₆	6.5	0.03	1.6	0.00008
p,p'-DDT	C ₁₄ H ₉ Cl ₅	6.2	<0.1	1.6	0.000025
2,4,2',4'-tetrachlorobiphenyl	C ₁₂ H ₆ Cl ₄	6.0	0.1	1.2	0.13
Hexachlorobenzene	C ₆ Cl ₆	5.6	0.006	1.2	0.001
Chlordane	C ₁₀ H ₆ Cl ₈	5.5	<0.1	1.6	0.0013
Heptachlor	C ₁₀ H ₅ Cl ₇	5.3	<0.1	1.7	0.04
Octane	C ₈ H ₁₈	5.2	0.7	0.7	1500
Pentachlorophenol	C ₆ Cl ₅ OH	5.0	1.0	1.3	2.7
Endrin	C ₁₂ H ₅ Cl ₆ O	4.6	0.2	1.7	0.000027
Heptane	C ₇ H ₁₆	4.5	3.0	0.70	5300
Phenanthrene	C ₁₄ H ₁₀	4.5	1.2	1.2	0.13
Dieldrin	C ₁₂ H ₈ Cl ₆ O	4.3	0.2	1.7	0.000024
1,2,4-trichlorobenzene	C ₆ H ₃ Cl ₃	4.2	35	1.5	40
1,4-dichlorobenzene	C ₆ H ₄ Cl ₂	3.6	120	1.3	160
Ethylbenzene	C ₈ H ₁₁	3.2	170	0.87	950
Xylene	C ₈ H ₁₀	3.1	180	0.88	1070
Perchloroethylene (PCE)	C ₂ Cl ₄	3.0	200	1.6	1870
Chlorobenzene	C ₆ H ₅ Cl	2.8	450	1.1	1170
Toluene	C ₇ H ₈	2.7	500	0.87	2930
Trichloroethylene (TCE)	C ₂ HCl ₃	2.4	1200	1.5	7800
Benzene	C ₆ H ₆	2.1	1800	0.88	10000
Phenol	C ₆ H ₅ OH	1.5	8400	1.05	48

* in water at 25 °C.

the system). The higher the vapor pressure of a liquid, the lower the boiling point of the liquid, and the more likely it is to vaporize. In general, organics with high vapor pressures at normal temperatures, and that are mainly insoluble in water and have boiling temperatures < 200 °C, are classified as volatile organic compounds (VOCs), a group that include the alkanes, aromatic hydrocarbons, some chlorinated hydrocarbons (e.g. TCE, PCE, pentachlorophenol) and aldehydes (e.g. formaldehyde).

3.3.2 Solubility considerations

Organic compounds are generally hydrophobic (that is, they are insoluble in water), but the presence of certain functional groups, particularly

nitrogen-based (e.g. amine), and oxygen-based (e.g. carboxylic and hydroxyl) groups, and sometimes halide groups (e.g. Cl), enhances solubility. The term commonly used to express the solubility of organic compounds is the **octanol-water partition coefficient**:

$$K_{OW} = C_{\text{octanol}}/C_{\text{water}} \quad (3.4)$$

It is a simple expression – the K_{OW} is defined as the solubility of an organic compound in *n*-octanol (C₈H₁₇OH) divided by its solubility in water. (In many cases, octanol is used as a proxy for fatty tissues in organisms.) Non-polar organic compounds tend to be soluble in other organic compounds (like octanol) and insoluble in water, so K_{OW} values for organic compounds tend to be

greater than 1, commonly by orders of magnitude, and for this reason, values of K_{OW} are often reported as $\log [K_{OW}]$. Refer to Table 3.4 for values of representative compounds. Any compound with $K_{OW} > 1$ will preferentially partition into organic compounds; this is the case for basically all potential organic pollutants, so the real difference in assessing mobility is the extent of the K_{OW} . The greater the K_{OW} , the less soluble in water and more likely to dissolve in (or adsorb onto) organic compounds, including natural organic matter in soils (see section 3.34 on adsorption). Compounds in the upper rows of Table 3.4, especially those with $\log [K_{OW}] > 4$ are very strongly sorbed onto/into particulate organic carbon (e.g. humus, decaying vegetation); those compounds with $\log [K_{OW}]$ values < 4 are generally more polar and their fates in the environmental become increasingly controlled by the higher aqueous solubility and potential to be sorbed onto surfaces of clays or other minerals.

3.3.3 Interactions of organic compounds and organisms

Considering human health and ecological impacts, the K_{OW} is generally proportional to the potential of an organic compound to **bioaccumulate (or bioconcentrate)**, provided that the compound does not rapidly degrade. (bioaccumulation is the process by which a substance is incorporated into an organism and, over time, increases in concentration because the rate of removal is lower than rate of intake; bioconcentration is largely the same process but does not include ingestion, and instead only includes absorption through skin or via respiration). Stable compounds with high K_{OW} values are soluble in fatty tissues of organisms and are only sparingly removed from organisms in urine. This point is illustrated in Fig. 3.20, which plots K_{OW} versus bioaccumulation factor (BAF), where BAF is the

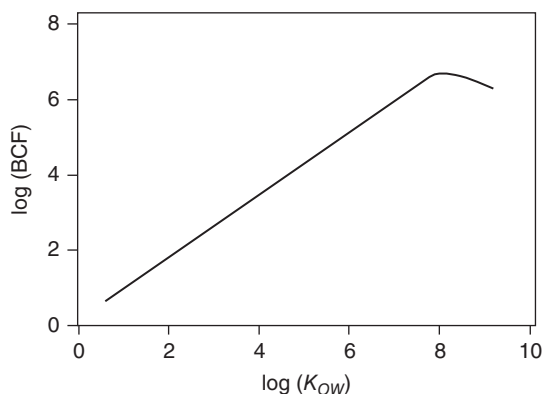
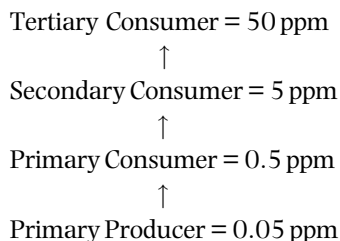


Fig. 3.20 Plot of $\log (K_{OW})$ against $\log$ (bioconcentration factor, BCF), showing trend for a range of organic compounds in middle trophic level species (BCF generally increases with trophic level). Note the linear relationship for most of the plot (decreasing solubility in water = increased BCF), and the curious dip at high values of K_{OW} . Created from data in Banerjee and Baughman (1991) and Arnot and Gobas (2006).

ratio of the concentration of the organic compound in the organism (or fatty tissue of the organism) divided by the concentration of the organic compound in the environment (e.g. water or sediments).

While bioaccumulation (or bioconcentration) pertains to increasing concentration of a substance (with time) in an organism, **biomagnification** occurs across trophic levels when secondary and tertiary consumers consume organisms lower on the food chain with relatively low concentrations of hydrophobic substances, and because the rate of ingestion is greater than the rate of excretion (resulting in bioaccumulation at the organismal level), the result to the ecosystem is biomagnification. Biomagnification tends to occur with insoluble organic compounds, and the classic example is concentration of DDT up the food chain in aquatic ecosystems, where DDT concentrations commonly increase by orders of magnitude from low trophic level primary producers (plants) to primary consumers

(herbivores) and eventually to high trophic level tertiary consumers (carnivores) according to the following scenario:



Probably the most well-known consequences of this phenomenon were the precipitous drops in populations of eagles, peregrine falcons and other predatory birds in the latter part of the 20th century due to biomagnification of synthesized organochlorine compounds such as DDT and PCBs. While predatory bird populations have begun to rebound in response to restrictions on DDT, PCBs and other such compounds, biomagnification of dimethylmercury ($[\text{CH}_3]_2\text{Hg}$) is a problem that appears to be increasing in scope and magnitude. Hg in the environment is derived from numerous sources, some like coal combustion and waste incineration that are broadcast over wide areas from multiple sources, and others, such as industrial and mining operations, that are localized **point sources** of contamination. (More information on Hg is presented in Chapter 8 in the context of air quality.)

Carbon filters are generally effective at removal of organic compounds from drinking water – the non-polar compounds are attracted to carbon, whether in designed filters or in natural organic matter in soils or sediments.

3.3.4 Adsorption of organic compounds

The physical or chemical attraction of organic pollutants to the surfaces of particles tends to limit pollutant mobility while also diminishing availability to decomposing microbes. **Adsorption** is

a surface effect characterized by attachment of compounds or ions to particle surfaces by relatively strong covalent or electrostatic forces or relatively weak van der Waals forces. Viewed in this way, adsorption can play a positive role in an environmental sense by limiting bioavailability of toxic organic compounds, but at the same time adsorption minimizes degradation rates, meaning that the contaminants will persist for longer periods of time. At this point it is worth making a couple of comments about terminology. The first is that adsorption is bonding of a compound or ion to a particle surface (refer to description in Section 2.5), whereas absorption is incorporation of a compound into another substance.

Adsorption and **absorption** are types of **sorption**. Because organic molecules often diffuse into the structure of humus, rather than merely become attached to surfaces of humus, and also because it is difficult to discern whether organic molecules undergo adsorption, absorption, or a combination of the two, the more general term sorption is often applied. A sorbed organic compound may be held by van der Waals on the surface of a particle of humus, or it may have diffused into the structure of the particle. Lastly, the term desorption describes release of a compound back into the fluid phase from a particle, i.e. it the reverse of adsorption (or absorption).

Adsorption produced by an electrostatic or covalent attraction between adsorbent and adsorbate is termed **chemisorption**. This is the process that governs ion exchange with surface of minerals or humus particles, and it is much stronger than **physisorption (physical adsorption)** that occurs when insoluble organic compounds form relatively weak van der Waals bonds with particles of generally any type. In chemisorption, the chemical properties of the solid absorbent (e.g. layer charge in clay minerals, surface charge in hydroxides, functional groups in humus) and adsorbate (is it cationic? anionic? a polar compound?) play an important role in the adsorption process and often selectively partition

certain phases out of the fluid and into a sorbed form; the proportion of a chemisorbed compound (on a particle surface) relative to other phases in the aqueous system may be significantly greater than its proportion in the fluid (often dissolved) phase. Physisorption generally does not result in any change in proportions of adsorbed species vs. those that remain in the fluid phase – it is relatively indiscriminate and independent of solution chemistry and solid phases.

Adsorption of organic compounds is controlled by the properties of soil solids (the adsorbents) and the organic molecules themselves (the adsorbates).

Adsorbates: Non-polar compounds like alkanes, DDT and PCBs tend to sorb strongly to natural organic matter and not to clays, hydroxides or other minerals (<http://www.epa.gov/nrmrl/lrpcd/wm/projects/135925.htm>).

Conversely, the relatively soluble organic compounds, particularly ionic forms of organic compounds and, to a lesser extent, polar organic compounds, are attracted to ion exchange sites in minerals. Figure 3.21 presents solubilities and K_{OW} data for selected organic compounds – recall that the higher the K_{OW} , the more likely it will be that an organic molecule will sorb to organic material like soil humus.

Adsorbents: Adsorption of organic contaminants by soil organic matter (humus) is generally greater than adsorption by inorganic clays, especially in organic-rich or silty-to-sandy soils, but in clay-rich soils with typical organic matter content (0.5 to 5%), clay minerals play an important role in organic matter adsorption, sometimes exerting greater control than humus by virtue of the abundance of clays. Clearly, minerals with high

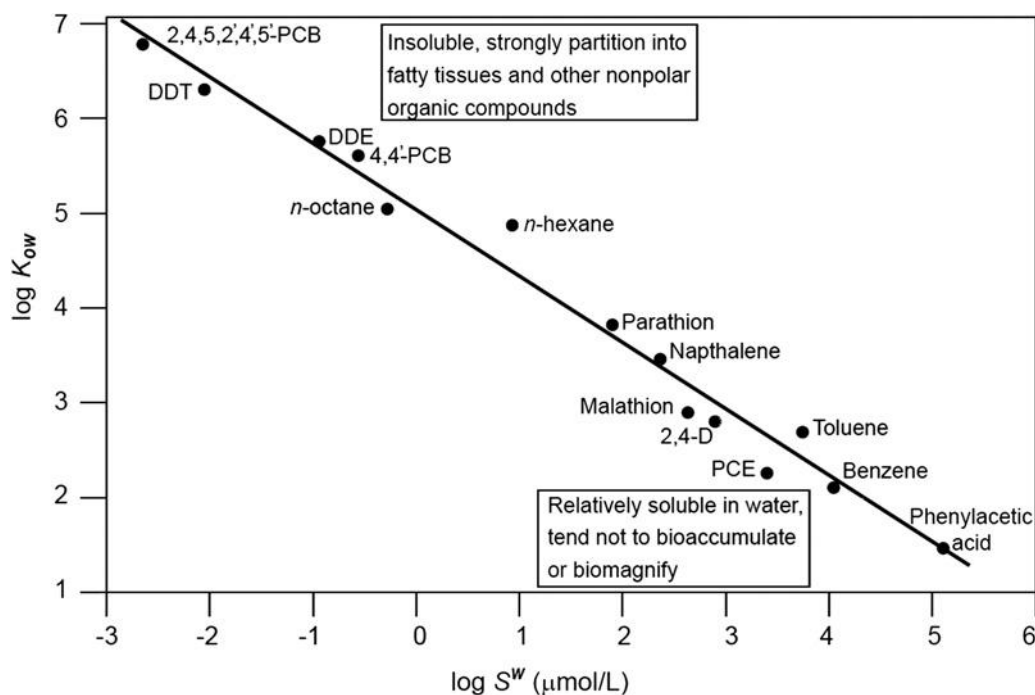


Fig. 3.21 Correlation between the solubility in water (presented as $\log S^w$, the aqueous solubility) and octanol–water coefficient (presented as $\log K_{OW}$) of organic compounds. Low-solubility compounds such as DDT and PCBs plot in the upper left of the line and relatively soluble compounds such as benzene, PCE and 2,4-D plot in the lower right of the line. (Data from Chiou et al. 1977, 2005.)

cation-exchange capacity such as smectites and vermiculites will more readily adsorb cationic (e.g. diquat, paraquat) and polar (e.g. atrazine, TCE) organic compounds than will those with low CEC (e.g. clays with no layer charge like kaolinite and coarse-grained neutral minerals such as quartz and feldspars). In acidic to neutral soils, iron hydroxides possess positive surface charge and are capable of attracting anionic compounds, notably the acidic herbicides such as 2,4-D and picloram and the insecticide pentachlorophenol, compounds that dissociate to produce H^+ and an anionic conjugate base in solution. This can be seen using 2,4-D as an example (Fig. 3.22).

In the following example (equation 3.5), R^- is any anion that might be dissolved in water, and on the reactants side, a chloride anion is sorbed to ferrihydrite; on the products side, the anion symbolically represented as R^- has undergone an exchange reaction with chloride. Examples of anions include species such as NO_3^- ; the 2,4-D anion and other organic anions; arsenite or arsenate anions; phosphate anions; and many more.



Water content is also an important consideration when predicting adsorption of organic matter. In general, the moister the soil, the lower the

extent of organic contaminant adsorption due to strong attraction of the polar water molecule to mineral surfaces (e.g. Sawhney and Singh, 1997). To extend this further, saturated groundwaters with low organic matter content tend to exhibit poor retention of organic contaminants by adsorption, resulting in high mobility (as dissolved phase, fine-grained mobile particulate or NAPL).

Adsorption isotherms are a common graphical means of representing adsorption data, containing information about the type of interaction between adsorbate and adsorbent. They are called isotherms because adsorption is temperature dependent (sorption decreases with increasing temperature), and thus isotherms present adsorption features at fixed temperature. In these plots, the x -axis is the concentration of an adsorbate in a fluid (or gas), and the y -axis is the amount absorbed to a solid, when the system is at equilibrium. Two examples are presented in Fig. 3.23, the Langmuir isotherm, which describes chemisorption, and a linear C-type (constant partitioning) isotherm that describes physisorption. Isotherms are presented here as they pertain to organic compounds and are referred to again in greater detail in Chapter 4 (Section 4.10).

Chemisorption is a monolayer process where the number of potential adsorption sites exerts a strong control on the amount of adsorbate that can bond to the particle surface, so as the equilibrium concentration (of adsorbate) increases, the amount adsorbed increases logarithmically (the rate of increase diminishes). In physisorption, there is no monolayer restriction like that associated with chemisorption, and the number of available adsorption sites on the particle surface is no longer a controlling variable. The result is a linear correlation between increasing concentration (e.g. of an organic compound in water) like that depicted in the C-type isotherm. At low concentrations, chemisorption can also produce a linear isotherm. This is because at very low

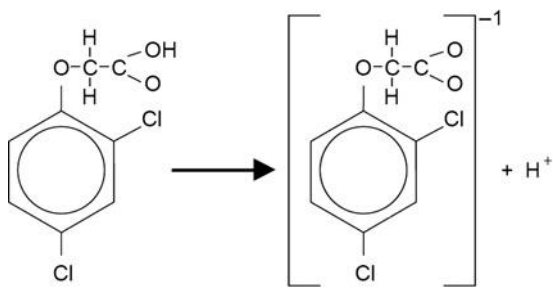


Fig. 3.22 Formation of 2,4-D anion by deprotonation. The resulting anion may be sorbed to anion exchange sites.

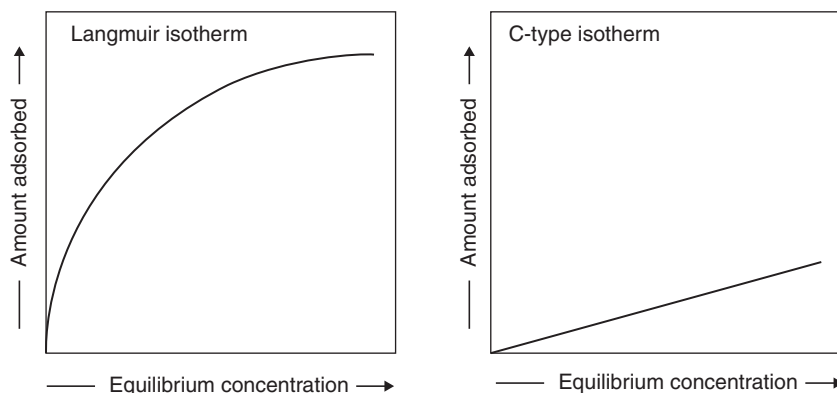


Fig. 3.23 Isotherms describing chemisorptions (left) and physisorption (right) or organic compounds.

concentrations of adsorbate, the number of available adsorption sites will not be a limiting factor on adsorption; only at higher concentrations where molecules must compete for adsorption sites will the isotherm become non-linear (logarithmic). You can conduct experiments using adsorbents and adsorbates, measuring sorption as a function of adsorbate concentration, and the resulting curve will suggest (but not prove) potential sorption mechanisms. Proof only comes with direct measurement using molecular-level methods such as infrared spectroscopy, often combined with indirect techniques such as chromatography (e.g. Johnston et al., 2002).

The Langmuir isotherm can be derived from the Law of Mass Action (Chapter 1.6), presented here as a summarized version of the derivation given by McBride (1994). If A represents an adsorbate, and S represents potential absorption sites, we can envision adsorption as follows:



where S^*A indicates sites (S) that have adsorbed the compound (adsorbate A). (To be clear, S^*A does not represent the mathematical expression S times A .) Recall that at equilibrium, adsorption – desorption is a dynamic process and this reaction

proceeds in both directions at equal rates. Thus, this equation can be expressed as an equilibrium constant:

$$K_{\text{ads}} = [S^*A]/([S] \cdot [A]) \quad (3.7)$$

As usual, brackets indicate activities (effective concentrations) of species in solution or on surfaces, where $[S]$ and $[S] \cdot [A]$ are assumed to represent mole fractions of vacant adsorption sites and occupied sites, respectively.

The proportion of sites occupied by A (θ_A) can be written as:

$$\theta_A = [S^*A]/[S^*A] + [S] \quad (3.8)$$

where $[S^*A] + [S]$ is the total number of possible sites, occupied (S^*A) and vacant (S). The proportion of vacant sites can be expressed as:

$$1 - \theta_A = [S]/[S^*A] + [S] \quad (3.8)$$

It then follows that the number of occupied sites relative to unoccupied sites can be expressed as this ratio:

$$\theta_A/(1 - \theta_A) = [S^*A]/[S] = K_{\text{ads}}(A) \quad (3.10)$$

$K_{\text{ads}}(A)$ is the equilibrium constant that governs adsorption of A. Rearranging produces the standard Langmuir equation:

$$\theta_A = K_{\text{abs}}(A)/(1 + K_{\text{abs}}(A)) \quad (3.11)$$

This equation produces the Langmuir isotherm and describes chemisorption.

For more detailed discussions of this topic, the reader is referred to thorough texts (e.g. Chiou, 2002) and chapters (e.g. McBride, 1994; Langmuir, 1997) on organic matter adsorption.

3.3.5 Non-aqueous phase liquids (NAPLs) in the environment

Non-polar organic liquids are commonly referred to as non-aqueous phase liquids (NAPLs) because they can occur in various environments as liquid phases that are not dissolved in the aqueous phase. For example, octane and water are immiscible fluids; the log K_{OW} for octane is approximately 10^5 , indicating that it is 100 000 times more soluble in octanol than water. Clearly, octane is quite insoluble in water, and with a density of 0.7 g/cm^3 , liquid octane will float on top of

water ($\rho = 1.0 \text{ g/cm}^3$), whether the water occurs in a surface body like a lake or ocean or whether it occurs in an aquifer, where octane will saturate pores atop the water table.

Light non-aqueous phase liquids (LNAPLs) are water-insoluble liquid compounds like octane and other fuels with densities less than 1.0 g/cm^3 – these generally are the alkanes (aliphatic hydrocarbons) and aromatic hydrocarbons (including BTEX and PAHs). Water-insoluble liquids with density > 1 are known as **DNAPLs**, where the D stands for dense. DNAPLs generally include chlorinated hydrocarbons like the solvent TCE, the insecticide pentachlorophenol, PCB oils and PAHs like those in creosote distilled from coal tar.

In terms of fate and transport (Fig. 3.24), NAPLs tend to:

1 Migrate downward through soil toward the water table, where their behavior is dictated by solubility and density:

A LNAPLs tend to fill pores between sediment grains and pool up atop the unconfined aquifer, and depending on their solubility, a certain amount will become dissolved in the aquifer below the water table and be transported laterally with groundwater flow. The dissolved

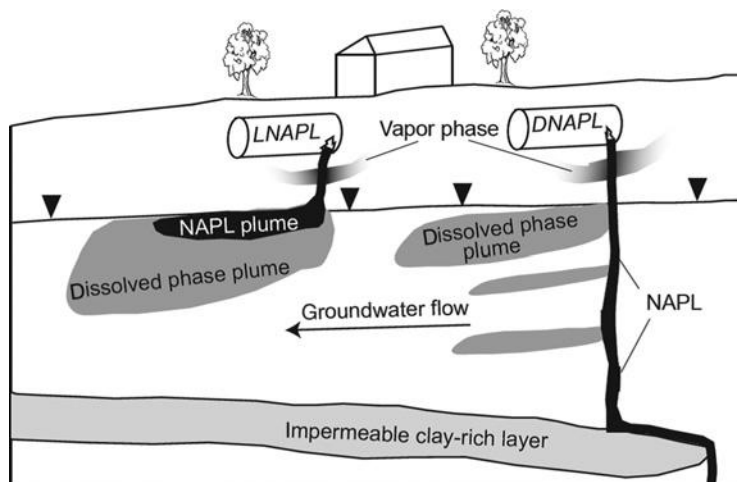


Fig. 3.24 Schematic diagram of behavior of NAPLs in the subsurface. Shown are two leaking underground storage tanks (USTs). Triangles indicate water table. The scale of the diagram does not permit depiction of sorption in the soil (refer to Chapter 9, Section 9.8).

phase plume is often rich in benzene, toluene, ethylbenzene and xylene (BTEX).

B DNAPLs sink to the bottom of the aquifer and forms pools atop clay or rock confining layers. Any fractures in the confining layer permit further seepage into the deeper confined aquifer. Of course, depending on solubility, some DNAPL will dissolve in aquifer water and migrate with groundwater flow (Fig. 3.24).

2 Partially adsorb to soil particles (particularly soil organic matter and reactive clays like smectites).

3 Partially evaporate, a process that is most important for compounds with high vapor pressures (Table 3.4).

4 Undergo degradation, to the extent permitted by natural microbial communities and physico-chemical conditions in the soil or aquifer (see the following section).

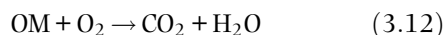
The multiple pathways of NAPLs in soils and aquifers makes remediation (and modeling of behavior) difficult. Many questions must be addressed, for example: What percent of the spilled NAPL will sorb to soil particles? (many factors must be considered, including solubility of the compound and composition of soil solids). How much can be expected to vaporize? How much will be decomposed by the natural soil microbial community before it reaches the water table? How much will dissolve in the groundwater and how much will remain as a non-aqueous liquid (i.e. how much will be either pooled atop or settled at the base of the aquifer)? Appendix I presents a case study that examines leakage of PCE into an alluvial aquifer and attempts to remediate.

3.3.6 Biodegradation

Because they are composed largely of C and H, organic compounds are potential energy sources for microbial organisms, particularly bacteria.

The C atoms in organic compounds occur in reduced form (C^{-4} to C^0), and oxidation releases electrons, the basic energy source for microbial life. In fact, the transition from reduced to oxidized carbon releases up to 8 electrons per atom (or 8 moles of electrons per mole of C), making C one of the most energy-rich sources for microbial metabolism (reduced nitrogen is also very energy-rich when oxidizing from N^{-5} to N^{+3}).

Respiration is a common example of oxidation of organic matter that releases electrons. The reactants are organic matter and O_2 ; the products are CO_2 and H_2O . The carbohydrates consumed by animals contain reduced carbon that reacts with oxygen to release energy (electron flow) to fuel life. Decomposition of organic compounds in soils and sediments is similar, but by virtue of their abundance relative to vertebrates, the microbial community is the main agent of decomposition (a cubic centimeter of humus-rich soil typically contains somewhere in the range of 100 million to 1 billion bacteria). A generic reaction for decomposition of organic matter (OM) is shown below.



Note that this reaction typically is bacterially mediated, but organic compounds can also decompose abiotically by oxidation, a process that generally occurs at rates that are orders of magnitude slower than their microbially mediated counterparts, but abiotic decomposition can be an important process in systems with relatively low bacterial populations. One such example is the atmosphere, where strong oxidizers such as the hydroxyl (OH^*) and hydroperoxy (HO_2^*) radicals are very effective at oxidizing reduced carbon in organic compounds.

Reaction 3.11 is a simplified example that likely skips many steps of intermediate decomposition products known as **metabolites**. The toxicity and the behavior in the environment of a

metabolite may differ appreciably from the original compound; for example, the metabolite may be more soluble or less soluble, or more likely (or less likely) to vaporize out of soil, than the original organic compound. The metabolite may decompose much more rapidly in the presence of soil microorganisms, or it may be more persistent than its parent compound (such as DDE and DDD compared to parent DDT).

Vinyl chlorite (IUPAC name: chloroethene) is an interesting compound related to decomposition and metabolism because it can occur as a metabolite (e.g. of TCE) and it also decomposes to metabolites. It is carcinogenic and is a very common pollutant, occurring in 35% of the US EPA's National Priorities List (NPL) of contaminated sites. It is synthesized and used to make products such as polyvinyl chloride (PVC), so it may occur in the environment as a primary pollutant or as a metabolite. It may form as a metabolite from TCE via DCE, as shown in Figure 3.25.

In shallow soil, vinyl chloride tends to vaporize due to its very high vapor pressure (340 000 Pa); if it does not vaporize, its high solubility (2700 mg/L) means that it may leach into groundwater, and if the groundwater is anaerobic (reducing), decomposition to the metabolite chloroethylene

oxide tends to occur very slowly (note that oxygen is a reactant (Fig. 3.25)).

3.3.7 Remediation

Common approaches to remediation of organic pollutants in soils, groundwater and surface water as summarized as follows:

Bioremediation typically involves augmenting natural, *in situ* microbial communities to enhance microbial decay, from adding nutrients to foster increased populations and effectiveness of indigenous microbial populations to pumping O₂ into soil or shallow groundwater to enhance aerobic activity.

Soil vapor extraction (SVE) is an *in situ* approach that takes advantage of the high vapor pressure of many organic pollutants. Vacuums are inserted into soil and the goal is to create a negative pressure gradient that causes desorption of VOCs and transport of vapors toward extraction wells, where VOCs can be extracted and treated.

Air sparging involves injection of air into soil or groundwater to cause desorption of VOCs from solids or transfer from dissolved state to a vapor phase. SVE is then used to remove VOC-containing air.

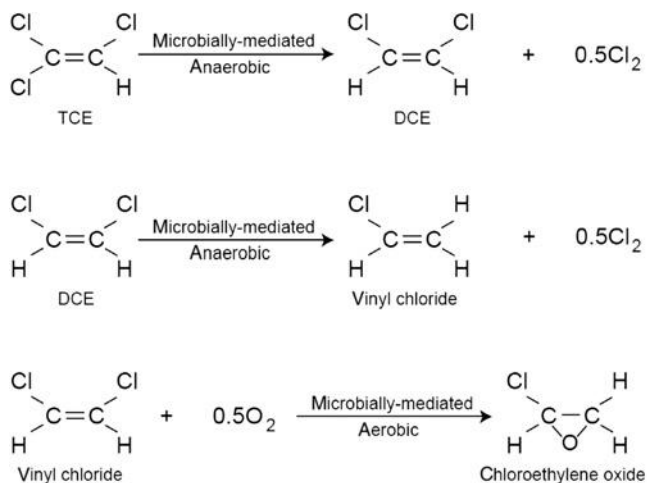


Fig. 3.25 Decomposition of trichloroethylene (TCE) to the metabolite DCE, then to the metabolite vinyl chloride, under (most likely) microbially mediated reducing conditions and. Decomposition of vinyl chloride to the metabolite chloroethylene oxide tends to occur rapidly via microbial mediation in aerobic (oxidizing) waters but very slowly in anaerobic waters, and thus may persist in groundwater for years or decades.

Chemical oxidation involves addition of strong oxidizers such as hydrogen peroxide (H_2O_2) to enhance oxidation.

Reductive dechlorination of organochlorine compounds may involve addition of a reducing agent (e.g. molasses, zero-valent iron) that aids in reduction of Cl that eventually (ideally) produces less-toxic metabolites (where $\sim \text{Cl}^+$ in organochlorine compounds is reduced to Cl^- in solution). Figure 3.25 depicts this approach and Appendix I presents a case study involving reductive dechlorination of PCE.

Pump and treat technology is often applied to groundwater where the above-listed approaches are ineffective – contaminated groundwater is pumped to the surface, treated to remove contaminants and then (often) reinjected.

Land farming involves extraction of contaminated soil and transport (by truckloads) to rural regions where soils are spread out at the surface and natural vaporization, oxidation and microbial decay remove or transform contaminants.

3.4 SUMMARY

Given the wide range in composition and structure of synthetic and natural organic compounds that occur as environmental contaminants, there is no simple rule of thumb that allows us to fully understand their behavior. They are generally insoluble and tend to sorb to organic matter, but as Table 3.4 shows, solubility varies widely. Behavior is often site specific (i.e. dependent on the percent soil organic matter, mineralogy, pH, microbial community, temperature, moisture), and studies of organic compounds in field-based and experimental research continue to expand our knowledge in this critical area of environmental science.

QUESTIONS

1. What volume of groundwater containing 2 mg/L dissolved oxygen (DO) would be

required to oxidize 23 L of octane that leaked out of a UST and into an unconfined aquifer? List 5 factors that could affect reaction rate, and create a schematic diagram to show potential pathways of octane as it leaks out of the UST. List and briefly describe three potential remediation strategies.

2. If humus is transported by erosion from a field into a lake, will this result in increased DO, or decreased DO, in the lake? Support your response with a chemical equation.
3. Which is more likely to be adsorbed by ferrihydrite in a soil at pH = 7.0, benzene or 2,4,5-T? Create a sketch to demonstrate. Also consider the potential for ferrihydrite to adsorb 2,4,5-T at pH = 5 relative to pH = 7; e.g., considering only ferrihydrite and 2,4,5-T, what might enhance 2,4,5-T adsorption at pH 5, and what might limit its adsorption at pH = 5?
4. Which is more likely to respond to air sparging as a remediation tool, octane or pentachlorophenol? Explain using pertinent data.
5. Reductive dechlorination of PCE or TCE in groundwater may result in breakdown to less-chlorinated metabolites, including vinyl chloride, which is not responsive to reductive dechlorination (Fig. 3.25). Describe potential strategies for remediation of groundwater with elevated vinyl chloride.
6. Which compound is more prone to bioaccumulation, toluene or 2,4,2',4'-tetrachlorobiphenyl? Cite pertinent data in your response.

REFERENCES

- Arnot, J.A. and Gobas, F.A.P.C., 2006. A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms. *Environmental Reviews* 14(4): 257–297.
- Banerjee, S. and Baughman, G.L. 1991. Bioconcentration factors and lipid solubility. *Environmental Science and Technology* 25: 536–539.

- Brady, N.C., Weil, R.R., 2002. *The Nature and Properties of Soils*, 14th edn. Prentice Hall. Upper Saddle River, New Jersey.
- Chiou, C.T., 2002, *Partition and Adsorption of Organic Contaminants in Environmental Systems*. Hoboken, NJ, John Wiley & Sons, 257 p.
- Chiou, C.T., Schmedding, D.W., and Manes, M., 2005, Improved prediction of octanol-water partition coefficients from liquid-solute water solubilities and molar volumes: *Environmental Science and Technology*, 39: 8840–8846.
- Chiou C.T., Freed V. H., Schmedding D. W. and Kohnert R. L. 1977, Partition coefficient and bioaccumulation of selected organic chemicals. *Environmental Science and Technology* 11: 475–478.
- Colborn, T and Thayer, K. 2000. Aquatic ecosystems: harbingers of endocrine disruption. *Ecological Applications* 10: 949–957.
- Drever, J.I., 1997. *The Geochemistry of Natural Waters: Surface and Groundwater Environments*, 3rd edn. Simon and Schuster, Upper Saddle River, New Jersey, USA.
- Girard, J.E., 2005. *Principles of Environmental Chemistry*. Jones and Bartlett Publishers, Sudbury, Massachusetts, USA.
- Green, N. J. L., Jones, J.L. and Jones, K.C., 2001. PCDD/F deposition time trend to Esthwaite Water, UK, and its relevance to sources. *Environmental Science and Technology*, 35 (14): 2882–2888.
- Johnston, C.T., Agnew, S.F., Schoonover, J.R., Kenney, J.W., Page, B., Osborn, J. and Corbin, R. 2002. Raman study of aluminum speciation in simulated alkaline nuclear waste. *Environmental Science and Technology*. 36 (11): 2451–2458.
- Langmuir, D., 1997, *Aqueous Environmental Geochemistry*, Prentice-Hall, Upper Saddle River, New Jersey, USA, 600 p.
- McBride, M.B., 1994. *Environmental Chemistry of Soils*. Oxford University Press, New York. 406 p.
- Sawhney, B. L., Singh, S.S. 1997. Sorption of atrazine by Al- and Ca-saturated smectite. *Clays and Clay Minerals* 45: 333–338.
- Silk P., Lonergan G.C., Arsenault T.L. and Boyle D.C. 1997. Evidence of natural organochlorine formation in peat bogs, *Chemosphere*, 35, no 12: 2865–2880.

Aqueous Systems – Controls on Water Chemistry

4.1 INTRODUCTION TO THE GEOCHEMISTRY OF NATURAL WATERS

The composition of natural waters is largely controlled by two main factors, changes of state and biogeochemical processes. An example of a common change of state is evaporation, which is responsible for the low amounts of dissolved substances in atmospheric water; biogeochemical factors include exchange reactions, mineral crystallization, mineral dissolution and decomposition of organic matter that occur in water in all reservoirs within the hydrologic cycle (Fig. 4.1). The following section examines changes of state, beginning with evaporation and its effect on atmospheric water and surface water, followed by the effect of precipitation and infiltrated water on chemical weathering and the composition of surface water and groundwater.

4.1.1 Geochemistry and the hydrologic cycle

Evaporation is effectively a distillation process in which dissolved ions and particles are left behind in the surface water body, and the water droplets that form in the atmosphere (and fall as precipitation) are the closest thing to pure H_2O on Earth. In order to understand this distillation process, consider that water molecules (except for those rare ones with the heavier isotopes of oxygen or hydrogen, i.e. ^{17}O , ^{18}O or ^2H) have a mass of 18 Daltons (or 18 g per mol), and typical dissolved ions such as Na^+ , Ca^{+2} , Cl^- and HCO_3^- have masses on the order of 23 to 61 Daltons. Thus, the probability of a water molecule evaporating with an attached, electrostatically bonded ion or molecule tagging along is quite low. Their masses are just too high, and the result is that these solutes are left behind in surface water. As a result,

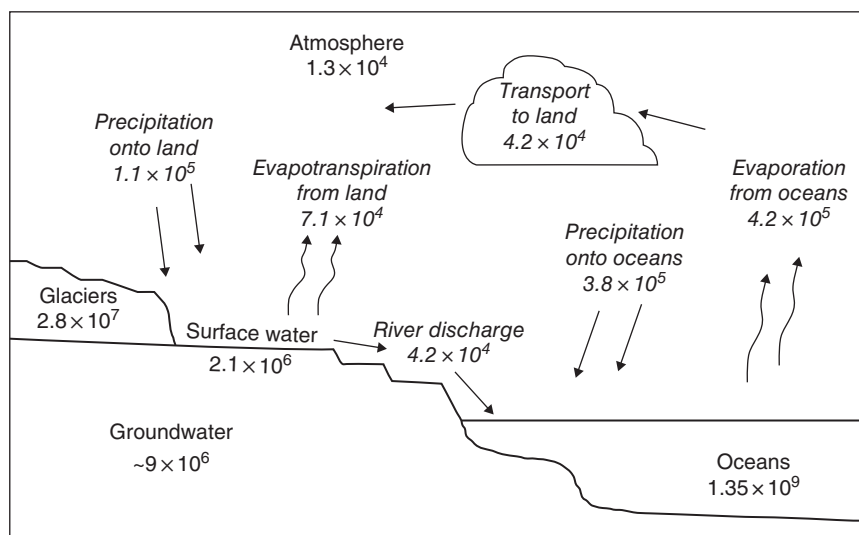
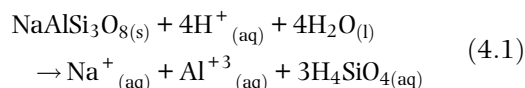


Fig. 4.1 Hydrologic cycle with volumes of **reservoirs** given in units of km³ and **fluxes** shown in units of km³/year. Surface water as shown above refers to fresh water on the continents, and the overwhelming majority of this reservoir is accounted for in a few large lakes. Also, less than one-half of groundwater is accessible for use by humans; some is too deep and salty. (Modified from Winter et al. 1998.)

atmospheric water is very dilute, typically with total dissolved solids (TDS) on the order of <10 mg/L (Table 4.1).

While evaporation tends to result in extremely dilute water vapor, it has the opposite effect on surface water bodies; in extreme scenarios (hot and arid), surface water tends to become progressively more concentrated in dissolved ions with progressive evaporation, which helps to explain why lakes or seas in isolated arid basins become so rich in salts – good examples include the Great Salt Lake in Utah (USA), the Aral Sea (or what is left of it) in central Asia, Lake Magadi (east Africa) and the shrinking Dead Sea (Middle East), water bodies with some of the highest TDS in the world. Evaporation of many such arid region lakes and seas is often so pronounced that it produces water that is oversaturated with respect to various dissolved ions, resulting in crystallization of evaporite salts that occur as crusty mineral deposits on shorelines (Fig. 4.2) or as chemical sedimentary deposits in lake beds or the sea floor.

Once atmospheric water condenses and falls to the surface as **precipitation**, **infiltration** into soils and rocks tends to foster chemical weathering of minerals (a topic that is covered in detail in Chapter 9). Precipitation is generally slightly acidic (natural pH is generally 5.0 to 5.5) and this factor, combined with low concentrations of dissolved ions typical of precipitation water (Table 4.1), facilitate the chemical decomposition of minerals. The following reaction depicts the dissolution of albite (sodium plagioclase feldspar) in the presence of water and H⁺:



In this example, chemical weathering releases dissolved Na, Al and Si into solution.

Dissolution of calcite can be depicted as:

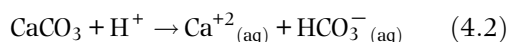


Table 4.1 Representative values for natural waters (in mg/L) from various reservoirs within the hydrologic cycle. Data are from Livingstone (1963), Holland (1978), Meybeck (1979), Mayo and Loucks (1995), Berner and Berner (1997) and other sources cited therein.

Ion	Precipitation		Surface Water			Groundwater	
	Continental	Coastal	Tropical River	Arid River	Seawater	Carbonate	Granitic
H ⁺ (as pH)	5.0–6.0*	5.0–6.0	6.5–7.0	7.0–8.5*	8.1**	7.3	7.1
Na ⁺	0.2–1.0	1.0–5.0	1.5	117	10770	6.1	0.8
Mg ⁺²	0.05–0.5	0.4–1.5	1.0	24	1290	33.5	6.0
K ⁺	0.02–0.3	0.2–0.6	0.8	6.7	399	2.5	0.7
Ca ⁺²	0.02–3.0*	0.2–1.5	5.2	109	412	70.6	25.6
Si	0.1	0.02	7.2	30	2.8	9.6	8.0
HCO ₃ ⁻	0.1	0.1	20	183	122	268	113
Cl ⁻	0.2–2.0	1.0–10	1.1	171	19500	9.0	1.7
SO ₄ ⁻²	1.0–3.0	1.0–3.0	1.7	238	2712	22.6	5.7
NO ₃ ⁻	0.4–1.3	0.1–0.5	< 1	< 1	< 1	3.3	0.8

* higher values correspond to windblown dust in arid regions. Areas with acid rain have pH < 5.

** increasing CO₂ in ocean water is causing pH to decrease.

Nitrate values are highly variable due to cycling of organic matter and climate factors

The example of tropical river water is the Amazon, the arid river is the Rio Grande, and seawater is global ocean mean.

A quick glance at Table 4.1 reveals that the most abundant species in natural waters include Na⁺, Ca⁺², Si (often as H₄SiO₄) and HCO₃⁻ (but not Al: its low solubility means that Al tends to become incorporated into minerals or disordered solids rather than remaining in solution). Mg⁺², K⁺ and SO₄⁻² are also mainly derived from weathering of bedrock, while NO₃⁻ is mostly derived from decomposition of organic matter. For this reason the composition of **surface waters** and **groundwater** will depend at least in part on the type of bedrock in the watershed; for example, note the compositional differences in groundwater in carbonate regions as compared to granitic regions (Table 4.1).

The high residence time of water in the saturated zone (Fig. 4.3) and the lack of dilution by precipitation tends to result in higher TDS for groundwater compared to surface water bodies (except where strongly influenced by evaporation). Return flow of water from the subsurface (vadose or saturated zone) into surface-water bodies (Fig. 4.3) is the main source of dissolved

species in surface water. In humid tropical or temperate regions, abundant precipitation tends to result in lakes and streams that are dilute relative to arid streams and lakes.

The saturated zone occurs below the water table and includes groundwater aquifers as well as saturated clays and bedrock units that do not produce water at sufficient rates to qualify as aquifers (4 L per minute is typically the minimum production rate required for domestic use, and much more is required for agricultural, municipal or industrial applications). The only gases that occur within pores in the saturated zone occur as small amounts of dissolved gases such as O₂, N₂, H₂S or CO₂. The saturated zone is less oxidizing than most surface water, ranging from chemically reducing to intermediate between oxidizing and reducing.

Groundwater contains relatively low dissolved O₂ concentrations except in near-surface levels or in cases where surface water rapidly infiltrates, resulting in pores that contain high amounts of dissolved O₂. In contrast to the saturated zone, the vadose zone is usually chemically oxidizing



Fig. 4.2 The white deposits surrounding the shallow pond in the foreground are evaporite salts that crystallized because evaporation of spring water raised dissolved ion concentrations to saturation. Evaporite minerals in this area include halite, gypsum, calcite and borax. Death Valley, USA. (Courtesy of the US Geological Survey.)

except in cases where abnormally high amounts of organic matter (e.g. below landfills or wetlands) consume O_2 to produce reducing conditions. Ultimately, controls on the composition of surface water and groundwater are diverse and include climate, bedrock and soil type, topography, residence time, crystallization-dissolution of minerals, redox, acid-base reactions, ion exchange, biotic factors and anthropogenic factors (e.g. land use and contamination).

Returning to evaporation, the effect of this process on surface water is often depicted by plotting the ratio of $Na/(Na + Ca)$, where concentrations are expressed in mg/L on the x -axis against the log of the concentration of total dissolved solids

(TDS, mg/L) on the y -axis (Gibbs, 1970). Halite ($NaCl$) is much more soluble than calcite ($CaCO_3$), and when the concentration of dissolved solids is driven up by evaporation, the remaining surface water eventually becomes saturated with respect to calcite, causing calcite to crystallize and taking Ca out of solution, and resulting in increased ratio of Na/Ca in solution. Consequently, streams and lakes from arid regions are characterized by high Na/Ca ratios and high TDS. So too is ocean water. These waters plot in the upper right of Fig. 4.4. One environmental consequence of the high $Na:Ca$ ratio and high TDS of surface waters (and therefore also often groundwaters) in many arid regions is sodification of agricultural soils

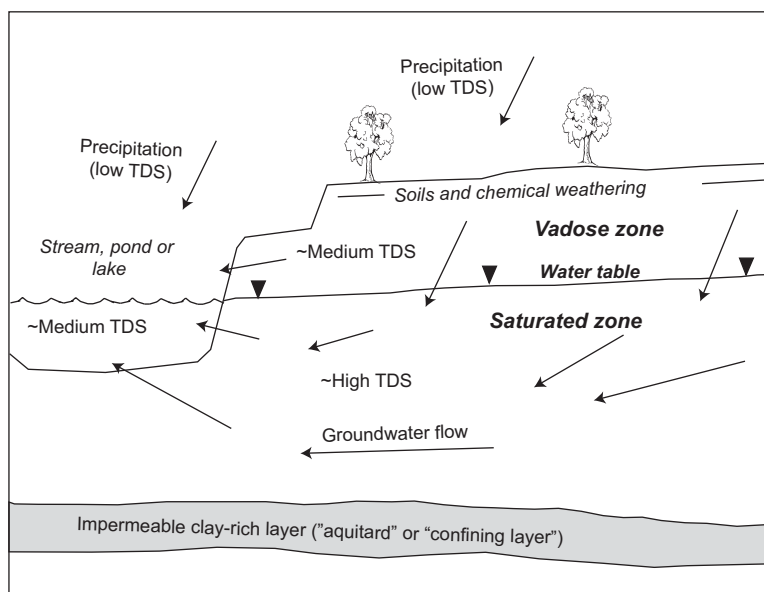


Fig. 4.3 Sketch showing, from top to bottom, the vadose zone (unsaturated zone), which includes soil and any sediment or rock above the water table) and the shallow part of an aquifer (the saturated zone). Connections between groundwater and surface water are also shown by arrows indicating, typical groundwater flow paths. Relative total dissolved solids (TDS) of precipitation, soil water, groundwater and surface water for a temperate region are shown.

when high-Na waters are used for irrigation. High Na destroys soil structure, reduces permeability and alters nutrient availability in negative ways.

The points in the lower right of Fig. 4.4 mostly represent headwater tropical streams, mountain streams and meltwater streams with low TDS caused by dilution by rainfall and snowmelt. Streams in temperate climates where water composition is dominated by weathering of bedrock plot in the center left of the diagram. This is because limestones are the most rapidly weathering rocks in most regions of the world, and their dissolution releases Ca ions into solution. The Onyx River of Antarctica is a great example of a stream whose initial (headwater) composition has very low TDS and a high Na/(Na + Ca) ratio; it originates from glacial meltwaters. As it flows downstream, it becomes progressively more influenced by chemical weathering, which increases TDS and lowers the Na/(Na + Ca) ratio. Another stream that illustrates these concepts

well is the Rio Grande of the southwestern USA, which originates in the Rocky Mountains, where chemical weathering initially dominates stream composition, and then flows into arid basins where evaporation eventually is the dominant control on water composition. Conceptually, streams will trend from lower right to center left to upper right as high mountain streams fed by snowmelt (lower right) transition into more-evolved streams with a Ca-rich bedrock weathering signature (center left) and eventually, further downstream, towards the Na-rich, high-TDS signature (upper right) produced by evaporation in arid environments.

This background on processes and reservoirs within the hydrologic cycle provides a framework for understanding some of the large-scale natural controls on water chemistry. The following section steps down to the atomic scale to explore some fundamental aspects of water molecules as they pertain to geochemistry.

4.2 THE STRUCTURE OF WATER – GEOMETRY, POLARITY AND CONSEQUENCES

The structure and geometry of the water molecule explain many of the chemical and physical aspects of water that influence its behavior in the environment. Water is a covalently bonded molecule with a single oxygen atom in a -2 valence state and two $+1$ -charged hydrogen

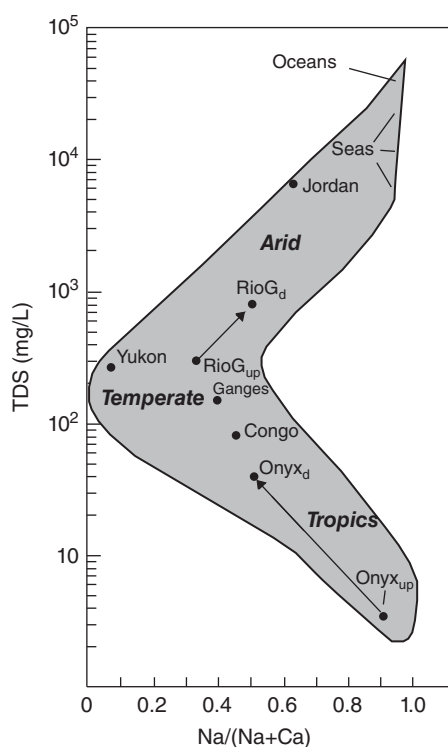


Fig. 4.4 Changes in the ratio of $\text{Na}/(\text{Na} + \text{Ca})$, based on units of mass (e.g. mg/L), plotted against total dissolved solids (TDS, in mg/L), for streams from different climate zones, and also for seas and oceans. Arrows indicate changes in stream composition that occur progressively downstream; for example, with increasing contributions from chemical weathering, tropical soils increase in TDS and the $\text{Na}/(\text{Na} + \text{Ca})$ ratio decreases. (Modified from Gibbs 1970 and Andrews et al. 2004.)

atoms; it is not a linear molecule, and this is crucial to its physical and chemical behavior. The valence electrons of the oxygen atom are arranged in a tetrahedral configuration, where two points of the tetrahedron contain paired electrons unable to form chemical bonds, and the other two points are unpaired electrons that form bonds with H atoms. The result is a dipolar molecule with a negative pole on the side with the O atom and a positive pole on the side with the 2H atoms. The angle between the H atoms is approximately 105° , and the O–H bond distance is $0.97 \text{ \AA}$ (0.097 nm) (Fig. 4.5).

The dipolar nature of H_2O makes it a very effective solvent that can form electrostatic attractions between anions the H atoms of the positive pole, and between cations and the O atom of the negative pole. Figure 4.6 depicts a schematic

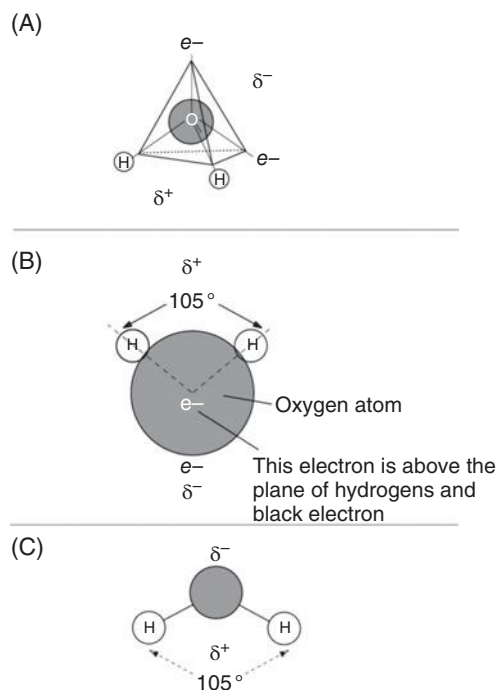


Fig. 4.5 The structure of the H_2O molecule.

(A) emphasizes the tetrahedral nature of the molecule, and (B) and (C) show the structure and dipolar nature in two dimensions.

example of H_2O dipoles oriented to facilitate solution of Na^+ and Cl^- ions – the dipolar character of water molecules, particularly their orientation with respect to dissolved ions, helps to increase the distance between cations (e.g. Na^+) and anions (e.g. Cl^-), thus decreasing the force of attraction between cation and anion, enhancing solubility and decreasing the likelihood of crystallizing a solid such as NaCl .

The geometry of the H_2O molecule, combined with hydrogen bonding between adjacent molecules, is responsible for the hexagonal open structure of H_2O_s (ice), resulting in the unusual situation that produces a solid phase ($\rho[\text{H}_2\text{O}_s] = 0.92 \text{ g/ml}$ at 1 atm) with lower density than its

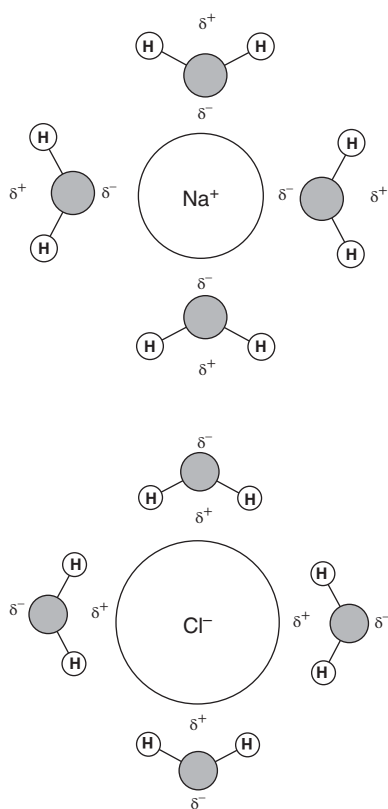


Fig. 4.6 Schematic sketch of the relationship of polar water molecules to a sodium cation (above) and chloride anion (below) in solution.

liquid phase ($\rho[\text{H}_2\text{O}_l] = 1.0 \text{ g/ml}$ at 1 atm). For this reason, ice floats on water and lakes do not freeze solid in cold winters, thus enabling complex aquatic ecosystems to exist in cold climates.

Hydrogen bonds are a type of van der Waals bond that form as the result of dipole–dipole attraction between a hydrogen atom and unshared electron pairs from anions with high electronegativity (e.g. O, N, F, and C, to a lesser extent) in polar molecules. The attraction between an O atom in one H_2O molecule and an H atom in an adjacent H_2O molecule is an example of a hydrogen bond (Fig. 4.7). Although intermolecular hydrogen bonds are only $\sim 5\%$ as strong as the intramolecular O–H covalent bonds (O–H hydrogen bond length is $1.97 \text{ \AA}$), they are responsible for many of the unique qualities of water, including its anomalously high melting and boiling points compared to other liquids with similar mass, its ability to dissolve certain compounds of N and C, and high surface tension that causes capillary rise of water above the water table (Fig. 4.3) and into the vadose zone.

4.3 DISSOLVED VERSUS PARTICULATE: EXAMPLES OF SOLUTIONS AND SUSPENSIONS

Aqueous systems are mixtures that contain a liquid phase (H_2O) with dissolved and suspended constituents. **Dissolved constituents** generally occur as ions that are electrostatically attracted to

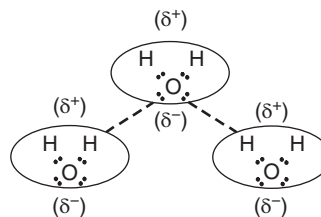
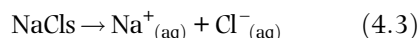


Fig. 4.7 Hydrogen bonds (dashed lines) between adjacent water molecules.

one of the positive or negative poles of water molecules, and are uniformly distributed throughout the liquid solvent (in an aqueous solution, the solvent is water). They are mobile and tend to diffuse readily and reach relatively constant concentrations within a water body, at least over reasonable spatial and temporal scales.

Dissolution of the ionic solid halite by water into Na^+ and Cl^- ions produces a simple aqueous solution, which can be represented as:



Water-soluble substances like halite are often referred to as “hydrophilic” (water-loving) – they dissolve because the attraction between the H atom (or O atom) in a water molecule and an anion (or cation) in the solid substance is greater than the attraction between the cation and anion in the solid compound. This is especially true for ionic bonds, which are notoriously unstable in the presence of water. Most salts of Cl^- , SO_4^{2-} , and NO_3^- are easily dissolved; carbonate minerals contain a mixture of bonds that range from covalent (C–O in CO_3^{2-} anion) to ionic (the bond between Ca and the CO_3^{2-} anion) and are moderately soluble over time spans of months to thousands of years or more, depending on many factors (e.g. climate, composition and acidity of aqueous solution, bacterial activity and flow rate).

Covalently bonded solids are far less soluble than ionic solids – a good example is quartz (SiO_2). The polar covalent bonds between Si and O atoms in quartz are far greater than the attraction that might be produced by, e.g., hydrogen bonds between H in H_2O and O in SiO_2 . Similarly, many organic compounds are effectively insoluble in water, including some organic solids as well as liquids such as octane (C_8H_{18}), for similar reasons. Thus, quartz and many other silicate minerals including feldspars and phyllosilicates (clays and micas) and some oxides and organic particles tend to commonly exist in aqueous systems as **particulates**, sediments or as **suspended material**,

rather than as dissolved ions. Ions can also exist as part of the particulate load when they are adsorbed onto mineral surfaces. Grains of quartz sand and phyllosilicates are common components of the suspended load in lakes, rivers and oceans, as are the ions adsorbed onto their surfaces.

Octane and other liquid hydrocarbons tend to exist as **immiscible liquids** aqueous systems, often floating atop surface water due to their lower density ($\leq 1 \text{ mg/L}$). Such insoluble compounds are commonly referred to as “hydrophobic” or in environmental science, as non-aqueous phase liquids (NAPLs), a topic covered in Chapter 3.

4.3.1 Dissolved vs. particulate vs. colloidal

Aqueous systems consist of a mixture of dissolved ions and suspended particles, and in some cases, immiscible liquids too. Understanding whether or not a certain element or compound exists in a dissolved state or as particulate matter, or whether an organic liquid has partially dissolved, is often crucial in understanding environmental availability, fate and transport. Typically, concentrations of soluble ions (e.g. Ca^{+2} or NO_3^-) or soluble liquids (e.g. benzene) in a wetland or lake will be relatively well mixed throughout the water body, whereas particulates and immiscible liquids are much more heterogeneous in their distribution. Particulates tend to get deposited out of suspension in slow-moving currents or stationary water, meaning any elements associated with those particulates would be concentrated in a certain part of the bottom sediments – this is the typical behavior of phosphorous and lead, for example. The behavior of other elements strongly depends on redox conditions, dissolved oxygen, the presence or absence of various other ions in solution, biotic constituents, the presence of sorptive solids, and more.

Uranium is an example of an element that, depending on conditions, can exist in soluble or

insoluble particulate states. If uranium in groundwater occurs in a particulate form it is less likely to be transported by slow-moving groundwater flow than would be dissolved uranium, so we would expect less uptake of $U_{\text{particulate}}$ into residential groundwater wells than U_{aq} . Furthermore, particulate uranium is much more likely to be removed by simple filters than the more highly mobile dissolved phase. Low-density immiscible liquids like octane and other fuels are largely partitioned atop the water body, whereas high-density immiscible liquids like chlorinated solvents tend to sink to the bottom of lakes, rivers and aquifers. However, understanding partitioning of elements and compounds into dissolved vs. particulate phases is often not a simple task, complicated by the need to quantify kinetic factors, aqueous system composition, mineralogical content and microbial activity, for example.

An additional factor to consider related to the behavior of dissolved and particulate matter is that some particulates are extremely fine grained, with particle diameters (according to IUPAC nomenclature) ranging from 10 Å (the thickness of a single unit cell of smectite clay) to 100 000 Å (=1 µm = mid-sized clay). These ultrafine-grained particles are known as **colloids**, and in the realm of environmental geochemistry include members of the smectite clay mineral group, hydroxides of iron and aluminum, and humic organic compounds. Particles of such fine grain size tend to remain suspended in water regardless of flow rate and hence are far more mobile than typical particulates, only settling out of solution when the colloidal particles flocculate under certain geochemical conditions (e.g. seawater can cause clays to flocculate and hence precipitate onto the seafloor). Furthermore, their surfaces are highly reactive exchange sites, so colloids are often responsible for enhanced mobility of otherwise immobile elements (e.g. As, Pb or U sorbed onto mobile, colloidal iron hydroxide in soil or sediments).

4.4 SPECIATION: SIMPLE IONS, POLYATOMIC IONS AND AQUEOUS COMPLEXES

Not all ions in solution exist as simple **monoa-tomic ions** like Ca^{+2} or Cl^- but rather occur charged molecules known as **polyatomic anions**. You have already seen some examples of polyatomic anions – representative members of this group include:

- carbonate (CO_3^{-2});
- bicarbonate (HCO_3^-);
- nitrate (NO_3^-);
- hydroxide (OH^-);
- phosphate (PO_4^{-3}), hydrogen phosphate anion (HPO_4^{-2}), dihydrogen phosphate anion (H_2PO_4^-);
- sulfate (SO_4^{-2});
- cyanide (CN^-);
- arsenate (AsO_4^{-3}), arsenite (AsO_3^{-3});
- dichromate ($\text{Cr}_2\text{O}_7^{-2}$);
- aluminum tetrahydroxide ($\text{Al}[\text{OH}]_4^-$).

Polyatomic cations are less abundant than their anionic counterparts. Examples include:

- aluminum monohydroxide ($\text{Al}[\text{OH}]^{+2}$) and aluminum dihydroxide ($\text{Al}[\text{OH}]_2^+$);
- ammonium (NH_4^+);
- uranyl (UO_2^{+2}).

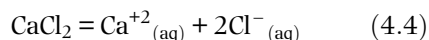
Complexes are species formed by bonding between two simpler substances that typically results in greater solubility of one of the substances; for example, relatively insoluble trace metals that normally don't occur in solution can be dissolved when they form complexes with species such as organic matter or certain anions. A good example is uranium (in the U^{+6} form), which forms soluble complexes with the hydroxyl and carbonate anions, enhancing its solubility in water: two common uranium complexes are $\text{UO}_2(\text{OH})_2$ and $\text{UO}_2(\text{CO}_3)_2^{-2}$. Organic compounds such as formic acid, humic acid and oxalate also often form complexes that enhance trace metal solubility in soils, surface water and groundwater.

It is also important to recognize that there are some dissolved species that are not ionic. For example, Si commonly exists in solution as silicic acid, H_4SiO_4 , which can also be written $\text{SiO}_2 \cdot 2\text{H}_2\text{O}$, illustrating that dissolved silica is effectively hydrated SiO_2 . In fact, the best way to envision dissolved silica is as a SiO_4^{4-} tetrahedron bonded to four H^+ , a sketch of which is presented in Fig. 9.2 in reference to products of chemical weathering of silicate minerals.

The **speciation** of a given element in an aqueous system refers to the form in which it exists – does it occur as a simple dissolved ion, or as a polyatomic ion, or does it occur as a particulate? According to IUPAC terminology, **chemical species** describe the specific form of an element defined with respect to its isotopic composition, electronic or oxidation state, and/or complex or molecular structure. Speciation of an element is the distribution of the element among dissolved and particulate phases in a system. In the study of arsenic speciation, for example, we might inquire whether As occurs in one of the soluble anionic forms (including H_2AsO_4^- , HAsO_4^{2-} , AsO_4^{3-} , H_2AsO_3^- , HAsO_3^{2-}), as a solid arsenic sulfide (e.g. As_2S_3 , AsS) or iron–arsenic sulfide (e.g. FeAsS), as a complex formed between As and colloidal iron hydroxide or As and dissolved (or particulate) organic matter, or as an ion sorbed (i.e. exchanged) onto the surface of a hydroxide or clay mineral. Considered in this way, speciation analysis involves the determination of chemical species present in solution and in the solid state, and the association of elements with those chemical species. Given the abundance of minerals and related solids in most systems, precise speciation analysis commonly requires determination of the chemical composition of minerals as well as amorphous solids and organic matter. Yet given the greater availability of species in solution, in some cases speciation analysis is more focused on dissolved constituents compared to solid forms.

4.5 CONTROLS ON THE SOLUBILITY OF INORGANIC ELEMENTS AND IONS

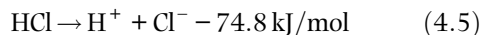
Perhaps the simplest means by which ions become dissolved in aqueous solutions is through the congruous (complete) dissolution of a mineral like the ionic salt calcium chloride, CaCl_2 , for which dissolution is represented as:



There are numerous factors that control the solubility of solids such as the mineral CaCl_2 and aqueous species like the dissolved Ca and Cl ions shown above. We will first consider a pair of important physical controls, temperature and residence time.

Perhaps the most important physical control on the solubility of a given compound or element is **water temperature**, where increasing temperature generally increases solubility of ions due to the increased kinetic energy of the water molecules. However, there are two important exceptions. One pertains to solids that release heat when they dissolve. In such exothermic dissolution reactions, heat is a product of dissolution, so if we consider dissolution in terms of Le Châtelier's principle, adding heat to the solution will drive the reaction in the direction of the reactants, thus inhibiting dissolution.

A good example of an **exothermic dissolution reaction is HCl in water – the ΔH** (change in enthalpy) for $\text{HCl} \rightarrow \text{H}^+ + \text{Cl}^-$ is -74.8 kJ/mol . What this means is that 74.8 kJ of heat are released by each mole of hydrochloric acid that dissolves in water. The negative sign for ΔH signifies that heat is released to the surroundings (in this case, the aqueous solution):



The enthalpy change associated with dissolution of NaCl is $+3.9 \text{ kJ/mol}$, i.e. 3.9 kJ of heat are

consumed per mol of NaCl that dissolves; the reaction is endothermic, and predictably, its solubility increases with increasing temperature. The dissolution of KNO_3 in water absorbs 35 kJ of heat per mole ($\Delta H = +35 \text{ kJ/mol}$), so (1) KNO_3 solubility increases with increasing temperature, and (B) the temperature dependence for KNO_3 dissolution is greater than it is for NaCl.

Like solids for which dissolution is exothermic, the solubility of gases is also inversely proportional to temperature, a phenomenon described by the temperature dependence of the constant in Henry's Law, and also evident through the lens of Le Châtelier's principle, where heat produced by gas dissolution means that **gas solubility decreases as water becomes warmer**. Examples of this phenomenon at work in Earth surface systems include (1) thermal pollution of surface waters that causes decreases in dissolved oxygen that can impair aquatic ecosystems, and (2) warming oceans that will have diminished capacity to retain dissolved CO_2 . Examples of ΔH values for selected gases include $\text{CO}_2 = -19 \text{ kJ/mol}$, $\text{O}_2 = -12 \text{ kJ/mol}$ and $\text{SO}_2 = -39 \text{ kJ/mol}$.

While it does not affect the inherent solubility of a given element or compound, **residence time** is an important control on the extent of dissolution because dissolution progresses parabolically (Fig. 4.8), where longer residence (exposure) times produce progressively more dissolution, although the rate decreases with time as the system approaches saturation (Fig. 4.8).

Residence time describes the amount of time a representative molecule of water will spend in a reservoir, where reservoirs can range from very small (e.g. pore spaces in soils or sediments) to large (e.g. lakes, the atmosphere or the ocean). At steady state, i.e. where rate of inflow = rate of outflow, residence time is calculated as follows:

$$t_{\text{res}} = \frac{\text{volume (or mass) of reservoir (in L)}}{\text{flux into (or out of) reservoir (in L/min)}} \quad (4.6)$$

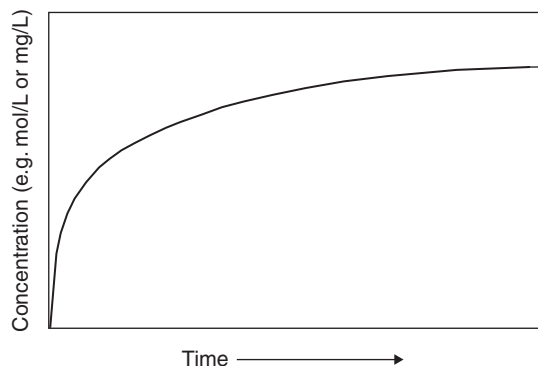


Fig. 4.8 Decreasing rate of dissolution of a mineral (e.g. sodium feldspar) in a closed system as measured by the concentration of a dissolved constituent (e.g. Na^+) released during dissolution. Note that concentration in solution increases until the solution reaches saturation, at which time the system reaches dynamic equilibrium. In an open system, leaching of Na^+ would likely prevent attainment of dynamic equilibrium. In terms of residence time, the concentrations of dissolved constituents will increase with time in this type of example.

A pair of examples serves to illustrate the effect of residence time. The first is the global ocean, with an approximate mass of $1.37 \times 10^{21} \text{ kg}$ of water and a flux of $3.9 \times 10^{17} \text{ kg}$ of water per year. The flux assumes a steady state where the cumulative discharge of all of the world's rivers into the ocean plus precipitation into the ocean is $3.6 \times 10^4 \text{ km}^3/\text{yr}$, a value that is equaled by water lost from the ocean by evaporation and seepage into subduction zones, sediments and other sinks. This yields a mean residence time of water molecules in the oceans of:

$$\begin{aligned} t_{\text{res(ocean)}} &= 1.35 \times 10^{21} \text{ kg} / 3.9 \times 10^{17} \text{ kg/yr} \\ &= 3.5 \times 10^3 \text{ years} \end{aligned}$$

In other words, water molecules spend an average of 3500 years in the ocean before being cycled out (e.g. by evaporation, burial in sediments or

injection into subduction zones). Conversely, a small lake with a volume of 9.3 km^3 and a flux of $1.2 \text{ km}^3/\text{yr}$ would have a residence time of 7.8 years and, given less time for mixing, will likely be characterized by greater heterogeneity than the well-mixed global ocean. The high residence time of the ocean means that it is much more likely to reach (or approach) equilibrium with respect to dissolved constituents, but it will also take much longer for natural processes to flush contaminants out of the ocean than the rapidly overturning lake.

4.5.1 The ratio of ionic charge: ionic radius and its effect on solubility

The ratio of ionic charge : ionic radius ($|z|/r$), or ionic potential, is useful in predicting solubility of ions because in general, the greater this ratio, the lower the solubility. So, we would predict that

high-charge, low-radius ions like Al^{+3} , C^{+4} , Ti^{+4} , Si^{+4} , Zr^{+4} , N^{+5} , P^{+5} and S^{+6} would be relatively insoluble and that relatively low-charge, high-radius ions like Na^+ , K^+ , Ca^{+2} , Pb^{+2} , and Br^- and Cl^- would be soluble, and we would be mostly correct. However, in some cases, the formation of polyatomic anions can greatly enhance the solubility of low-radius, high-charge ions like C^{+4} (HCO_3^-), N^{+5} (NO_3^-), S^{+6} (SO_4^{2-}), and although not an anion, the formation of the polyatomic aqueous species H_4SiO_4 greatly enhances the solubility of Si^{+4} . Figure 4.9 depicts numerous elements plotted as a function of $|z|/r$ vs. the proportion present as particulate matter compared to the proportion present in solution (the higher this value, the less soluble). Notable features include the high solubility of low-charge, large-radius ions like Cl^- , Na^+ , Ca^{+2} and other halogens, alkalis and alkaline-earth metals; the intermediate solubilities of divalent transition metals (Zn^{+2} , Co^{+2} , etc.); low solubilities of

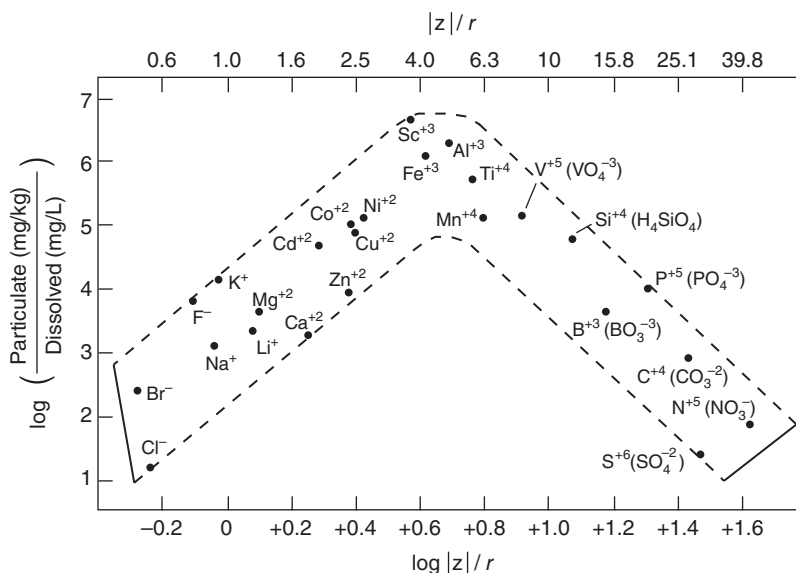


Fig. 4.9 Prediction of solubility of aqueous ions based on the ratio of ionic charge (z) to ionic radius (r , in Å), a ratio known as ionic potential. In the case of polyatomic molecules (e.g. H_4SiO_4), radius (r) is for the elemental ion shown (e.g. Si^{+4}), not for the entire molecule. The z values on both upper and lower x -axes are absolute values, and the y -axis data are based on concentrations in river water determined by Martin and Whitfield (1983); the plot shown is based on a graph in Andrews et al. (2004).

high-charge, intermediate-size cations such as Al^{+3} , Fe^{+3} and Ti^{+4} ; and relatively high solubility of high-charge, small-radius cations like C^{+4} , N^{+5} and S^{+6} which form the large radius polyatomic anions carbonate, nitrate and sulfate.

4.5.2 Reduction–oxidation reactions

Chromium serves as a good example of an element whose oxidation state exerts a strong control on its behavior in solution. As would be predicted based on z/r rationale, the trivalent ion (Cr^{+3}) is relatively insoluble (it is also non-toxic), yet when **oxidized** to the hexavalent state (Cr^{+6}), chromium is soluble as an oxyanion (soluble CrO_4^{-2} or $\text{Cr}_2\text{O}_7^{-2}$) and it is also toxic. Predictably, if hexavalent chromium is reduced to Cr^{+3} , it becomes less soluble and less toxic. Thus, **reduction** and **oxidation** reactions, often referred to informally as **redox** reactions, are additional important controls on the solubility and bioavailability of many elements, particularly those that can exist in more than one oxidation state.

Iron is very sensitive to reduction and oxidation – ferrous iron (Fe^{+2}) is relatively soluble in reducing (O_2 -poor) aqueous systems, while ferric iron (Fe^{+3}) is largely insoluble in oxidizing (O_2 -rich) systems. Numerous elements have two or more common oxidation states, and changes in oxidation state produced by transfer of valence electrons is one of the most important concepts to consider when assessing elemental behavior in aqueous systems. In addition to Cr and Fe mentioned above, numerous other elements exist in various oxidation states in nature, including the following:

- manganese (Mn^{+2} , Mn^{+4} , Mn^{+7});
- copper (Cu^0 , Cu^{+1} , Cu^{+2});
- uranium (U^{+4} , U^{+6});
- arsenic (As^{+3} , As^{+5});
- carbon (C^{-4} , C^0 , C^{+2} , C^{+4});
- nitrogen (N^{-3} , N^0 , N^{+3} , N^{+5});
- sulfur (S^{-2} , S^0 , S^{+4} , S^{+6}).

There are also many elements in nature that have only one oxidation state, including hydrogen (H^{+1}), the alkali (Na^{+1} , K^{+1}) and alkaline-earth (Mg^{+2} , Ca^{+2}) metals. When the highly electronegative elements oxygen and fluorine form chemical bonds in compounds, they have predictable charges – oxygen is always O^{-2} in compounds (except in rare cases like peroxides, where it is partially reduced to O^- or a mixture of O^0 and O^{-2}) and fluorine (F^{-1} in compounds). Redox reactions can also affect the elements that exist with only one valence state in nature. Examples include Ni^{+2} adsorbed to iron hydroxide that gets released to solution when the iron hydroxide dissolves when iron is reduced ($\text{Fe}^{+3} \rightarrow \text{Fe}^{+2}$), or Pb^{+2} that is released to solution when galena (PbS) dissolves via sulfide oxidation. Elements that occur in their native state, including metals like gold and silver (Au , Ag), the carbon minerals diamond and graphite (both are pure C), and the diatomic gases O_2 , N_2 , S_2 , have oxidation states of zero.

Environments with abundant O_2 , like the atmosphere, bubbly streams, aerated soils and shallow groundwater are typical oxidizing environments, whereas O_2 -poor environments like anoxic wetlands, deep groundwater and water-logged soils are typical reducing environments.

4.5.3 Half-cell reactions

Redox reactions are often considered in terms of half-cell reactions for individual elements involved in the reaction. The electron transfer associated with redox reactions can be measured (reduction–oxidation involves transfer of electrons and thus produces a current), and by convention, values are reported relative to the value for the oxidation of H_2 gas:

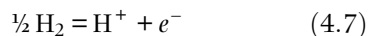


Table 4.2 contains examples of half-cell reactions, where the values assigned to each half-cell reaction

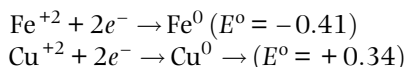
Table 4.2 Half-cell reactions and standard electrode potentials (25 °C, 1 atm) presented from strongest oxidizing agents (top) to weakest oxidizing agents (strongest reducers, bottom).

Half-reaction				Standard Electrode Potential, E° (V)
Oxidizing Agent		Reducing Agent		
F_2	+ $2e^-$	$\rightarrow$	$2F^-$	+2.88
Au^+	+ e^-	$\rightarrow$	Au^0	+2.64
Cl_2	+ $2e^-$	$\rightarrow$	$2Cl^-$	+1.36
$O_2 + 2e^-$	+ $4H^+$	$\rightarrow$	$2H_2O$	+1.23
Br_2	+ $2e^-$	$\rightarrow$	$2Br^-$	+1.08
Hg_2^{+2}	+ $2e^-$	$\rightarrow$	$2Hg^0$	+0.85
I_2	+ $2e^-$	$\rightarrow$	$2I^-$	+0.53
Cu^+	+ e^-	$\rightarrow$	Cu^0	+0.52
Cu^{+2}	+ $2e^-$	$\rightarrow$	Cu^0	+0.34
$2H^+$	+ $2e^-$	$\rightarrow$	H_2	0.00
Pb^{+2}	+ $2e^-$	$\rightarrow$	Pb^0	-0.13
Ni^{+2}	+ $2e^-$	$\rightarrow$	Ni^0	-0.24
Cd^{+2}	+ $2e^-$	$\rightarrow$	Cd^0	-0.40
Fe^{+2}	+ $2e^-$	$\rightarrow$	Fe^0	-0.41
S^0	+ $2e^-$	$\rightarrow$	S^{-2}	-0.44
Cr^{+3}	+ $3e^-$	$\rightarrow$	Cr^0	-0.74
Zn^{+2}	+ $2e^-$	$\rightarrow$	Zn^0	-0.76
U^{+4}	+ $4e^-$	$\rightarrow$	U^0	-1.38
Al^{+3}	+ $3e^-$	$\rightarrow$	Al^0	-1.70
La^{+3}	+ $3e^-$	$\rightarrow$	La^0	-2.36
Mg^{+2}	+ $2e^-$	$\rightarrow$	Mg^0	-2.36
Rb^{+1}	+ e^-	$\rightarrow$	Rb^0	-2.60
Na^{+1}	+ e^-	$\rightarrow$	Na^0	-2.71
Ca^{+2}	+ $2e^-$	$\rightarrow$	Ca^0	-2.87
Sr^{+2}	+ $2e^-$	$\rightarrow$	Sr^0	-2.90
Ba^{+2}	+ $2e^-$	$\rightarrow$	Ba^0	-2.91
K^+	+ e^-	$\rightarrow$	K^0	-2.94
Li^+	+ e^-	$\rightarrow$	Li^0	-3.04

are defined relative to the standard hydrogen electrode (SHE), which is given the arbitrary standard electrode potential (E°) value of zero. Note that E° values in Table 4.2 are presented in order from strong oxidizing agents (e.g. F_2) to weak oxidizing agents (strong reducers, e.g. Li^+).

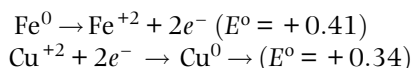
Elements with positive E° values are electron recipients (oxidizing agents) likely to cause oxidation of less-electronegative elements (those lower in the table) – these elements (or ions) with positive E° are the oxidizing agents that include the gases oxygen, fluorine and chlorine, which respectively undergo reduction to O^{-2} , F^{-1} and Cl^{-1} . Note that elements with negative E° values are mostly alkali, alkaline-earth and trace metals, elements that we know are likely to undergo oxidation to forms ions such as Na^{+1} , Fe^{+2} , Ni^{+2} and Al^{+3} . Positive E° values indicate that the reaction will occur spontaneously as written in Table 4.2, and the strongest oxidizing agents (those most likely to undergo reduction) have the most positive E° values. By this convention, all reactions are written as reduction reactions, so keep in mind that half-cell reactions with negative E° are not spontaneous as written and are likely to proceed in the reverse direction (the reaction of Na^+ to Na^0 will not occur in natural waters). You will also observe that some metals are intermingled with the strong oxidizers O_2 , F_2 and Cl_2 – these are metals that can occur in nature in their native states (e.g. Au^0 , Cu^0 , Hg^0 , Pt^0). ***Note: in some tables you will find that the signs are opposite of those presented here, in which case negative values are associated with oxidizers like O_2 and F_2 and positive values with elements likely to be oxidized (e.g. alkalis, alkaline earths, trace metals).*

Standard electrode potential values are useful because they allow us to compare the tendency of one element to undergo oxidation or reduction in the presence of another element. Consider the following two half-cell reactions:

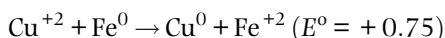


Both half-cell reactions show reduction of divalent cations to native elements. For Fe, the negative E° of this reaction indicates that it is not spontaneous as written, but rather is spontaneous in the opposite direction. The negative E° of the

Cu reduction reaction indicates that it is spontaneous in the direction written, so we can re-write the equations in the directions of spontaneity:



Note that it is necessary to change the sign of E° for the Fe half-cell reaction when the direction is reversed. The sum of the two reactions produces the following reduction–oxidation reaction (and E° for the reaction):



What this means is that in an aqueous system with copper and iron, copper will tend to oxidize iron to produce divalent iron ions and native copper metal, helping to explain why in some ore bodies, Cu exists as pure native copper (Cu^0) while iron exists as Fe^{+2} in pyrite (FeS_2).

4.5.4 Redox reactions in the environment

In addition to the inorganic half-cell reactions shown above, organisms and organic matter often strongly influence Eh; for example, decomposition of organic matter typically consumes O_2 and results in reducing conditions. One classic example is eutrophication, where degradation of algae in lakes, ponds and shallow marine waters reduces dissolved O_2 in the water. The microbial organisms responsible for degradation (consumption) of algae are mainly aerobes that consume O_2 , and the decrease in O_2 ultimately alters community structure for the worse and in some cases increases mobility of redox-sensitive metals such as Cr, Fe and Mn.

Another example of the effect of organic matter on Eh and trace-element mobility is petroleum leaks into groundwater, where microbial degradation of petroleum compounds consumes O_2

and reduces Eh. This can result in reduction of ferric to ferrous iron and dissolution of iron hydroxides, which in turn releases arsenic into solution – aquifers where dissolved arsenic was once well below drinking-water standards (e.g. 10 ppb in the USA) can experience substantial increases of dissolved arsenic that exceed standards following petroleum leaks.

A few additional examples of how redox conditions control fate and transport of aqueous species include:

- Nitrogen is less mobile in soils when it exists as N^{-3} in the reduced polyatomic ammonium cation (NH_4^+) as compared to when it exists as N^{+5} in the highly mobile, oxidized nitrate anion (NO_3^-) – this is an important consideration because migration of N compounds into surface water can lead to eutrophication (toxic algal blooms and hypoxia), a threat to many aquatic ecosystems.
- Metal-bearing sulfide minerals such as pyrite (FeS_2), arsenopyrite (FeAsS) and galena (PbS) are stable in reducing environments but decompose in oxidizing environments driven by oxidation of the sulfide anion (S^{-2}) to S^{+6} (which occurs in the sulfate anion, SO_4^{2-}) – in this case, oxidation can result in release of trace metals and arsenic to soils and aquatic systems.
- Organic compounds, including both natural (e.g. humus in soil) and synthetic/ refined (e.g. fuels, solvents, pesticides and pharmaceuticals), are relatively stable in reducing environments but tend to decompose to products like CO_2 and H_2O in oxidizing environments.

Redox conditions in most waters are controlled by the quantity of **dissolved oxygen (DO)** (as O_2 gas dissolved in water). Highly oxygenated surface waters in cool streams and lakes, e.g., contain approximately 7 to 9 mg/L of dissolved oxygen, and comprise the most oxidizing of aqueous systems at and near the Earth surface. In such high-DO environments, redox-dependent elements occur in oxidized states (e.g. Fe^{+3} , N^{+5} in NO_3^- , S^{+6} in SO_4^{2-} , Cr^{+6} in $\text{Cr}_2\text{O}_7^{2-}$). Wetlands and groundwaters contain naturally lower amounts

of DO than do surface waters; for example, groundwaters generally contain DO from 1–6 mg/L, with concentrations decreasing with depth, and wetlands contain similarly low values (wetland sediments are often anoxic, with 0 mg/L DO). Accordingly, oxygen-poor environments of deep groundwater and stagnant wetlands (especially wetland sediments) contain reduced species such as S^{2-} (which occurs in sulfide minerals like pyrite or as H_2S , hydrogen sulfide gas), Fe^{+2} (in sulfide and silicate minerals) and N^{-3} (in NH_4^+ , ammonium, or NH_3 , ammonia gas). It is worth noting that values of DO below 5 to 6 mg/L generally do not foster healthy aquatic ecosystems (with the exception of natural swamps or deep marine realms that are adapted to these conditions).

Analysis of reduction–oxidation on water chemistry will continue again in Section 4.9 in terms of Eh–pH (pe–pH) diagrams, an approach that considers the paired role of redox and pH on the mobility of elements and compounds.

4.5.5 pH and acid–base consideration

Solution pH exerts a strong control on the composition of aqueous systems, influencing the speciation and solubility of many elements and compounds. For example, metals are generally more soluble in acidic water than in neutral or basic water, and highly basic waters can enhance the solubility of compounds of elements such as aluminum, arsenic, and selenium.

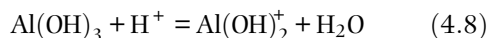
The main control on the pH of natural waters is the carbonate system, a topic that is covered in detail later in Chapter 5 (Section 5.1). Important points to consider here are (1) that the presence of dissolved CO_2 (as H_2CO_3 , carbonic acid) in atmospheric water tends to result in pH values ~ 5.0 – 5.5 , and (2) in soils, chemical weathering reactions tend to consume H^+ from carbonic H_2CO_3 (delivered in precipitation) and produce surface water and groundwater with pH between 6.5–8.5. In some cases stronger acids may control

pH, one example being the role of sulfuric acid (H_2SO_4) in creating acid mine drainage surface waters (often with pH = 2–3) and acid precipitation with pH ~ 3.0 – 4.5 . Evaporation in arid environments may produce briny sea or lake waters with pH = 9–12.

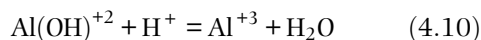
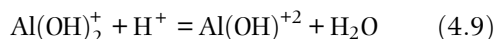
Aluminum is an excellent example of an element whose solubility in solution is strongly influenced by pH. At neutral pH, aluminum is nearly insoluble and exists as the Al^{+3} ion in silicates such as kaolinite ($Al_2Si_2O_5[OH]_4$) and hydroxides such as gibbsite ($Al[OH]_3$). However, with increasing acidity, the affinity of hydrogen ions (H^+) for the hydroxyl anion (OH^-) in the mineral is greater than the bond between the OH^- and the Al^{+3} cation in the mineral lattice. This process, known as **hydrolysis**, can be defined as the dissolution of minerals and other solids caused by bonding of H^+ to an anion in the solid.

Two important consequences of hydrolysis are:
1 consumption of H^+ ions that lowers solution acidity (i.e. buffers the system by consuming H^+ , thus increasing pH);
2 dissolution of solids that releases ions into solution.

Getting back to aluminum, and using the mineral gibbsite as our example, we can examine how variations in pH of aqueous solutions control solution composition. In an acidic solution, hydrolysis of gibbsite can be depicted as follows:



The products of this reaction are dissolved aluminum as a polyatomic aluminum hydroxide cation, $Al(OH)_2^+$, and water. Further increases in acidity can result in the following sequence of reactions:

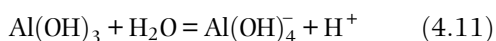


where dissolved aluminum hydroxide anions are progressively stripped of OH^- by hydrolysis,

ultimately resulting in dissolved aluminum as the simple Al^{+3} cation in highly acidic waters (Fig. 4.10).

Figure 4.10 shows dissolved aluminum species as a function of pH, ranging from particulate $\text{Al}(\text{OH})_3$ at neutral pH to $\text{Al}(\text{OH})_2^+$, $\text{Al}(\text{OH})^{+2}$ and Al^{+3} with progressively lower pH.

It is also clear in Fig. 4.10 that, in addition to increasing solubility at low pH, Al solubility also increases with increasing pH, an effect caused by reaction of $\text{Al}(\text{OH})_3$ with H_2O in basic solutions.



In basic solutions where $[\text{H}^+]$ is orders of magnitude less than $[\text{OH}^-]$, aluminum hydroxide (as gibbsite or as the $\text{Al}(\text{OH})_4^-$ anion) reacts with water to produce hydrogen ions in order to bring the system closer to equilibrium, and the resulting $\text{Al}(\text{OH})_4^-$ anions are soluble in alkaline water.

An alternative way to visualize $\text{Al}(\text{OH})_3$ solubility in basic solutions is as follows:

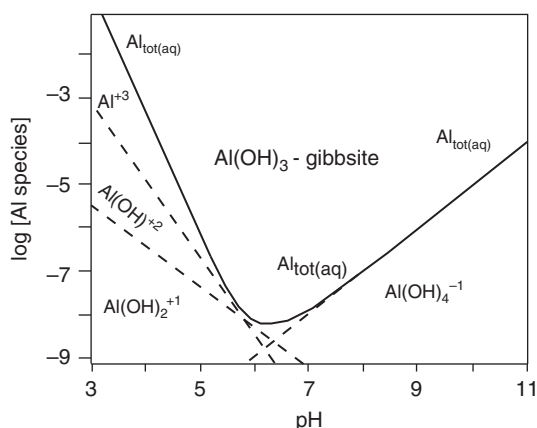
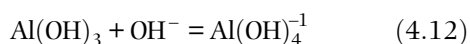


Fig. 4.10 Solid curve is sum of all dissolved Al species and dashed lines correspond to Al species indicated. Note low solubility of Al between pH = 6 to 7 and exponential increase in solubility with decreasing or increasing pH.

Many other metals in addition to Al exhibit **amphoteric** behavior in solution, including arsenic (technically a metalloid rather than a metal), cadmium, iron, lead and zinc – such amphoteric metals are able to react as either an acid or base, meaning that they can have higher solubilities at acidic and basic pH values as compared to neutral pH.

4.5.6 Ligands and elemental mobility

Metal mobility is often enhanced by the presence of **ligands**, which are defined as neutral molecules or anions (often organic) that can combine with a metal cation to form a **complex** that is usually more soluble than the un-complexed metal cation. Inorganic anion ligands include many of the examples presented in previous sections, including hydroxyl (OH^-), bicarbonate (HCO_3^-), sulfate (SO_4^{2-}) and phosphate (PO_4^{3-}). Common organic ligands include acetate (CH_3COO^-), oxalate ($\text{C}_2\text{O}_4^{2-}$), and the high molecular weight anion ethylenediaminetetraacetate (EDTA) (Fig. 4.11).

EDTA is special type of ligand known as a **chelate**, a molecule that is able to form multiple bonds around a metal cation in a cage-like structure. Note the coordination of the metal cation (M^+) to multiple bonding sites within the chelate.

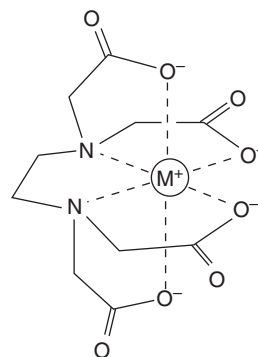


Fig. 4.11 Structural formula of ethylenediaminetetraacetate (EDTA) molecule, a good example of a chelating compound.

In soils, the large molecular weight humic and fulvic compounds, products of decomposition of organic (~plant) matter, are important chelating agents that enhance migration of Al and Fe.

4.6 ION ACTIVITIES, IONIC STRENGTH, TDS

In concentrated solutions (e.g. seawater, deep groundwater brines, acid mine drainage), the spacing between dissolved ions is low enough so that nearby cations and anions are electrostatically attracted to each other. These attractions diminish the ability of the ions to participate in chemical reactions (e.g. mineral growth, ion exchange), and because their effective concentration is lower than their actual concentration in solution, the concentration (of species i) is multiplied by an activity coefficient (γ) to produce a value known as the ion activity:

$$a_i = [i] * \gamma \quad (4.13)$$

where a_i is activity of ion i , $[i]$ = concentration of ion i , and γ_i is the activity coefficient of ion i .

In dilute solutions, γ is approximately 1 and thus activity is nearly equal to concentration. In more-concentrated solutions such as saline waters, many groundwaters or acid mine waters, $\gamma < 1$ and the value of γ decreases with increasing concentration and temperature, reflecting the fact that higher temperatures and ion concentrations cause divergence of concentration and activity. In most fresh surface waters and groundwaters, γ values are ~0.8–1.0.

Ion activity coefficients can be calculated from this Debye–Huckel equation as follows:

$$\log \gamma_i = -A * z_i^2 * \sqrt{I} \quad (4.14)$$

A is a constant that accounts for pressure and temperature. At 1 bar of pressure (acceptable

for surface water and most groundwater), values of A range from 0.4913 at 0 °C to 0.5042 at 20 °C and 0.5221 at 40 °C. The term z_i is the ionic charge and I is ionic strength (covered below). This form of the Debye–Huckel equation is valid for ionic strengths (I) $< 10^{-2}$, and thus applies to most fresh waters, but not brines or sea water. For reference, arid streams with high dissolved solids are in the $I = 10^{-2}$ range and most groundwater is similar or more dilute; temperate and tropical surface waters tend to have ionic strengths that are lower (more dilute) than 10^{-2} . For briny solutions a more-detailed version of the Debye–Huckel equation that takes into account the finite size of ions is needed, but for the types of water present in most streams, lakes and aquifers, the form presented here is sufficient.

A pair of calculations help to illustrate the importance of considering ion activity. In both calculations temperature = 20 °C. The variables are the cations (Na^+ and Ca^{+2}) and the ionic strength (10^{-2} and 10^{-4}). In order to solve for γ_i the equation can be rearranged as:

$$\gamma_i = 10^{\wedge}(-A * z_i^2 * \sqrt{I}) \quad (4.15)$$

For Na^+ at $I = 10^{-2}$, $\gamma_i = 0.89$.

For Na^+ at $I = 10^{-4}$, $\gamma_i = 0.99$.

For Ca^{+2} at $I = 10^{-2}$, $\gamma_i = 0.62$.

For Ca^{+2} at $I = 10^{-4}$, $\gamma_i = 0.95$.

These examples help to illustrate that, in dilute solutions (e.g. $I = 10^{-4}$), activity coefficients are nearly 1.0 and the activity of a given ion is almost equal to its concentration. However, as ionic strength increases, ion activity decreases, and this effect is more pronounced for higher-charge ions. In effect, for $I = 10^{-2}$ at 20 °C, only ~62% of the calcium in solution is actually available to react. The remainder is likely bound in aqueous complexes such as CaCO_3^0 or CaSO_4^0 .

4.6.1 Ion activity product

Now, consider a general reaction at equilibrium where:



A and B, and C and D are reactants and products, respectively, and the lowercase a , b , c and d represent molar fractions of reactant or product. Recall from chapter one the equation for the equilibrium constant:

$$K_{eq} = [C]^c * [D]^d / [A]^a * [B]^b \quad (4.17)$$

where $[C]^c$ represents the concentration of reactant C raised to the c power, and so on. The term on the right side of the equal sign is the equilibrium product. We can also write this equation using ion activities rather than concentrations.

$$K_{eq} = a_C^c * a_D^d / a_A^a * a_B^b \quad (4.18)$$

If the system is not at equilibrium, i.e. if the expression on the right side is greater than or less than the K_{eq} , then the expression on the right side is an **ion activity product (IAP)**, where:

$$IAP = a_C^c * a_D^d / a_A^a * a_B^b \quad (4.19)$$

If the IAP is less than the K_{eq} , then the reaction will progress in the direction of products (i.e. to the right). Conversely, if the $IAP > K_{eq}$, the reaction will proceed to the left (forming A and B from C and D). The saturation state of the solution is represented by the symbol Ω (omega), where

$$\Omega = IAP / K_{sp} \quad (4.20)$$

The following equations describe how the IAP can be used to examine saturation state:

- If $\Omega > 1$, the solution is oversaturated (or supersaturated);
- If $\Omega = 1$, the solution is at saturation;
- If $\Omega < 1$, the solution is undersaturated.

This should make sense intuitively, for if a system contains high concentrations (or activities) of reactants and low concentrations (or activities) of products, the reaction will proceed in the direction of products.

4.6.2 Ionic strength

The **ionic strength** of a solution is determined by:

- 1 – concentrations of ions in solution;
- 2 – ionic charges of the ions in solution.

The equation for ionic strength is defined as follows:

$$I = 0.5 \sum [i] * z_i^2 \quad (4.21)$$

The units of ion concentration $[i]$ are mol/L and z_i is the charge on ion i . The ionic strength of a solution is thus the sum of the products of the concentrations of each ion in solution times their ionic charge, for all ions in solution. Typically, I for fresh-water is on the order of 10^{-3} to 10^{-4} mol/L, whereas seawater has a typical value of 0.7 mol/L. If we consider the dissolved major ions in a sample of groundwater from a bedrock aquifer in Vermont with the following composition, we can compute ionic strength (Table 4.3).

Here, the ionic strength of the sample is 6.32×10^{-3} , typical of groundwater from deep aquifers of relatively high ionic strength. (note that values are often presented unitless although technically they could be expressed as mol charge²/L).

4.6.3 Total dissolved solids

The sum of the concentrations of all dissolved ions in solution is termed the **total dissolved solids**

Table 4.3 Computation of ionic strength from dataset with concentrations in mg/L.

	mg/L	mmol/L	z^2	$[i] * z_i$
pH = 7.5	—	—	—	—
Na ⁺	4.79	0.21	1	0.21
Mg ⁺²	21.8	0.90	4	3.59
K ⁺	0.61	0.02	1	0.02
Ca ⁺²	46.8	1.15	4	4.59
HCO ₃ ⁻	205.0	3.36	1	3.36
Cl ⁻	0.90	0.03	1	0.03
SO ₄ ⁻²	20.3	0.21	4	0.85
SUM	300.2	5.87		12.63
I =	6.32 × 10 ⁻³			

Table 4.4 Ranges of total dissolved solids for different classifications of waters (after Gorrell 1953).

Type	TDS (mg/L)
Fresh	<10 ³
Brackish	10 ³ –10 ⁴
Saline	10 ⁴ –10 ⁵
Brine	>10 ⁵

(TDS), and values are typically presented in terms of mg/L or mmol/L. For the groundwater example above, TDS is the sum of the concentrations of all cations (Na⁺, Mg⁺², K⁺, Ca⁺²) and anions (HCO₃⁻, Cl⁻, SO₄⁻²) in solution, which gives a value of 300.2 mg/L or 5.87 mmol/L. You will note that, unlike ionic strength, TDS does not account for ionic charge. TDS can be used to classify waters as shown in Table 4.4.

Brackish waters occur in mixing zones (e.g. estuaries) between fresh water and marine water, while brines occur where hydrothermal processes or evaporation produce waters with extremely high TDS (> 35 g/L).

The interactions among solutes (mainly dissolved ions) and solids (mainly minerals) exert a strong control on water composition. Ions can be removed from solution by sorption to the surfaces of solids such as clay minerals, hydroxides and organic matter, or by

mineral crystallization, where mineral growth proceeds by precipitation of ions from solution.

4.7 SOLUBILITY PRODUCTS, SATURATION

The solubility of a compound is the amount of the compound that dissolves to form a saturated solution. Solubility and saturation are two important concepts in the realm of aqueous geochemistry – highly soluble minerals will result in high concentrations of dissolved ions in solution, so clearly, considerations of solubility and saturation have applications to questions of contaminant behavior. Consider the case of three possible solid forms of lead, each presented in Table 4.5 with their respective K_{sp} values.

It is clear from these data that lead chloride and lead sulfate are many orders of magnitude more soluble than lead carbonate, emphasizing the importance of mineralogical analysis in studies of aqueous systems. A soil containing 100 ppm Pb in PbCO₃ may yield less dissolved Pb in solution water than would a soil with only 10 ppm Pb contained within PbCl₂. Recognition of the speciation of trace metals in minerals or other solids can clearly exert a strong control over their mobility, and given enough data can allow us to quantitatively predict equilibrium concentrations.

The solubility of a mineral can be determined by examining the dissolution reaction at equilibrium, the point where the amount of reactant (the mineral) and products (dissolved ions) are constant. How could we do this? One way would be to immerse a crystal of a mineral such as

Table 4.5 Examples of solubility products for three lead (Pb) minerals (25 °C, 1 atm).

Formula	Mineral name	K_{sp}
PbCO ₃	Cerussite	7.4×10^{-14}
PbCl ₂	Cotunnite	1.7×10^{-5}
PbSO ₄	Anglesite	1.6×10^{-8}

angelsite into deionized water and measure the concentrations of Pb^{+2} and SO_4^{-2} until they no longer increase. At this point, a dynamic equilibrium will have been established where any new dissolution of the angelsite is matched by an equal amount of precipitation of PbSO_4 . The solution is saturated with respect to lead and sulfate ions, and no net dissolution or precipitation occurs. The reaction that describes this process is:



The reaction can be quantified by the equilibrium constant

$$K_{\text{eq}} = a_{\text{Pb}^{+2}} * a_{\text{SO}_4^{-2}} / a_{\text{PbSO}_4} \quad (4.23)$$

The solid crystal of angelsite (or what remains of it) cannot be represented in terms of activity, so by convention its activity is given the value of 1, and the equation now appears as

$$K_{\text{eq}} = a_{\text{Pb}^{+2}} * a_{\text{SO}_4^{-2}} \quad (4.24)$$

In reality, as long as there is solid angelsite, the amount of it will not affect the reaction between Pb^{+2} and SO_4^{-2} ions and the solid angelsite – it is the concentrations of the ions that are most important, at equilibrium.

The K_{eq} for a solid and its associated dissolved ions is known as the **solubility product** (K_{sp}), which has the following formula and K_{sp} value for angelsite at 1 atm and 25 °C (Table 4.5):

$$K_{\text{sp}} = a_{\text{Pb}^{+2}} * a_{\text{SO}_4^{-2}} = 1.6 \times 10^{-8} \text{mol}^2/\text{L}^2 \quad (4.25)$$

The activity of Pb^{+2} is equal to the activity of SO_4^{-2} , and the value of the activity of each is

$$\sqrt{1.6 \times 10^{-8} \text{mol}^2/\text{L}^2} = 1.3 \times 10^{-4} \text{mol/L}$$

Thus, in this equilibrium situation, the concentrations of Pb^{+2} and SO_4^{-2} are each $1.3 \times 10^{-4} \text{mol/L}$, and these are the predicted concentrations at

equilibrium that would be produced by dissolution of angelsite.

Recall that if the ion activity product is greater than the K_{sp} ($\Omega > 1$), the solution is oversaturated and angelsite will precipitate from solution. If the $\Omega < 1$ (the $\text{IAP} < K_{\text{sp}}$), the solution is undersaturated and angelsite will dissolve. If $\Omega = 1$ ($\text{IAP} = K_{\text{sp}}$), the system is saturated and at equilibrium, and rates of mineral dissolution and mineral crystallization will be equal. So a system with $[\text{Pb}^{+2}] = 0.7 \times 10^{-4} \text{mol/L}$ and $[\text{SO}_4^{-2}] = 1.7 \times 10^{-4} \text{mol/L}$ will be undersaturated with respect to angelsite because the IAP will be $< K_{\text{sp}}$.

However, a system with $[\text{Pb}^{+2}] = 0.9 \times 10^{-4} \text{mol/L}$ and $[\text{SO}_4^{-2}] = 6 \times 10^{-4} \text{mol/L}$ will be oversaturated with respect to angelsite, because the $\text{IAP} = 5.4 \times 10^{-8} \text{mol/L}$, which is greater than the K_{sp} . Lead and sulfate ions will precipitate out of solution to form angelsite.

Given this approach, it is possible to predict whether a solution is saturated, oversaturated, or undersaturated with respect to angelsite or any other mineral. A detailed example pertaining to Ca^{+2} , HCO_3^- and the mineral calcite is presented later in Chapter 5.

And finally, for comparison, if dissolution of cerussite into ions of Pb^{+2} and CO_3^{-2} was the process controlling the amount of dissolved lead, predicted equilibrium concentration would be far lower than what was determined above with respect to angelsite: obtaining the K_{eq} for cerussite from Table 4.5, we determine that $\sqrt{7.4 \times 10^{-14} \text{mol}^2/\text{L}^2} = 2.7 \times 10^{-7} \text{mol/L}$ of Pb^{+2} . In other words, if dissolution (or precipitation) of cerussite is controlling Pb^{+2} in solution, concentrations will be approximately 1000 times lower than if the more-soluble angelsite is the corresponding mineral.

4.8 CO-PRECIPITATION

Co-precipitation describes the process in which a mineral (calcite, e.g.) crystallizes from solution and incorporates ions into its structure in addition

to the dominant ions (e.g. calcium and carbonate). One of the most important examples of co-precipitation is the uptake of trace metals and metalloids such as arsenic into the structure of iron oxy-hydroxides that range from generally well-crystalline oxides hematite (Fe_2O_3) and magnetite (Fe_3O_4) to less-crystalline hydroxides such as ferrihydrite (written as $5\text{Fe}_2\text{O}_3 \cdot 9\text{H}_2\text{O}$ or $\text{Fe}_5\text{HO}_8 \cdot 4\text{H}_2\text{O}$) and the polymorphs goethite ($\alpha\text{-FeO}[\text{OH}]$) and lepidocrocite ($\gamma\text{-FeO}[\text{OH}]$).

Incorporation of trace metals into Fe-hydroxides is often responsible for minimizing trace metals in solution or in exchange sites, making them less environmentally available. Once co-precipitated, their fate will be dictated by the fate of the solid oxy-hydroxide – if it moves as a particulate, so will co-precipitated constituents, and if it dissolves, it will release trace metals back into solution, at which point they may be adsorbed onto exchange sites, or remain in solution, or become incorporated into a new solid. For example, consider an iron hydroxide that incorporated arsenic into its structure when it crystallized in an oxidizing soil, but then was eroded into a stagnant wetland. The reducing conditions of the wetland sediment likely will result in dissolution of the hydroxide and release of iron and arsenic into solution. However, if sufficient sulfur is available in the wetland sediment it will probably exist as the sulfide anion (S^{2-}) and if it reacts with the liberated iron, could result in crystallization of pyrite (FeS_2) or a related sulfide (potentially with co-precipitated arsenic).

Another example of co-precipitation is the uptake of radium into relatively insoluble barium sulfates (which can range from low-order hydrated forms to the mineral barite, BaSO_4). While Ra^{+2} is a relatively soluble cation, it fits quite well into the BaSO_4 structure (substituted for Ba^{+2}), and thus can be removed from solution by co-precipitation when BaSO_4 crystallizes. Many trioctahedral clay minerals contain metals such as Cu, Ni and Zn incorporated into octahedral sheets and there is emerging evidence that

As(III) and As(V) can substitute for Si^{+4} in the tetrahedral sheets of clay minerals. There is no question that in order to fully understand controls on aqueous systems, geochemists must identify the composition, abundance, mineralogical form, potential mobility (as particulates) and solubility of solids in the aqueous system.

4.9 BEHAVIOR OF SELECTED ELEMENTS IN AQUEOUS SYSTEMS

Alkali and alkaline-earth elements, including those commonly referred to as the **base cations** (Na, K, Mg, Ca), are among the most soluble metals in aqueous systems. Their high solubility is enhanced by low ionic charge (e.g. Na^{+1} , Ca^{+2}), relatively large ionic radii and the fact that, often, the minerals in which these elements reside are soluble (e.g. halite, NaCl ; sylvite, KCl) or moderately soluble (e.g. calcite, CaCO_3 , or other carbonates). However, in some cases, the solubility of the base cations will be limited by their sequestration in minerals of relatively low solubility, including albite ($\text{NaAlSi}_3\text{O}_8$), microcline (KAlSi_3O_8) and other feldspars; micas such as muscovite ($\text{KAl}_3\text{Si}_3\text{O}_{10}(\text{OH})_2$) and biotite ($(\text{KAlMg}_2\text{Si}_3\text{O}_{10}(\text{OH})_2$); and other silicate minerals such as the amphibole hornblende ($\text{Ca}_2(\text{Mg,Fe,Al})_5(\text{Si,Al})_8\text{O}_{22}(\text{OH})_2$). Again, it is clear that in order to fully understand the solubility of base cations, or any element, it is important to understand the types of minerals present in the system. More detailed information on mineral solubility and dissolution rates is presented in Chapter 9 on soils and chemical weathering.

4.9.1 Examples of heavy metals and metalloids

Heavy metals general refer to metallic elements with atomic number greater than 20, including the transition-group metals (comprising the

central rows of the periodic table), the lanthanides (La to Yb) and actinides (elements heavier than Fr). These elements are far less abundant than the alkali and alkaline-earth metals, and hence are often referred to as **trace metals**. They occur as trace constituents of various rock types, and thus can occur naturally in soils and water, but often they are present at higher concentrations due to human activities like mining, metal refining, industrial activities and waste disposal.

The classic situation where heavy metals are most mobile is acid mine drainage settings (**Plate 12**) (or their natural acid drainage analogs; **Plate 5**), where (1) O_2 -rich streams produce oxidizing conditions, and (2) the chemical weathering of sulfides produces solutions of low pH. At higher pH, heavy metal solubilities are limited by coprecipitation with solid phases such as oxides, hydroxides and carbonates, and under reducing conditions they tend to bond with the sulfide anion (S^{2-}) to form sulfides that are insoluble in O_2 -poor waters (e.g. PbS) (*note: It is difficult to use the term “mineral” to describe the solid phases that may form in these situations because some of them are amorphous, and therefore they lack the internal atomic structure required to be minerals*). Heavy metals that exhibit this typical “soluble in acidic, oxidizing water” behavior include cadmium, cobalt, copper, lead, nickel and zinc. These metals commonly occur together in base-metal sulfide ore deposits in minerals such as covellite (CuS), chalcopyrite (CuFeS₂), galena (PbS) and sphalerite (ZnS) along with pyrite and other sulfides.

Given their high solubilities in acidic, oxidizing conditions, many heavy-metal remediation schemes rely on constructing alkaline wetlands to remove metals from rivers and other aquatic systems – such wetlands are characterized by high pH and low Eh (low dissolved oxygen) environments that facilitate precipitation of metals as carbonates, sulfides or other solid phases. The low kinetic energy of the wetland is beneficial because sediments settle out of suspension, so the wetland acts as both a chemical filter and

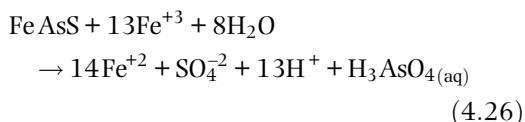
physical filter, causing chemical precipitation of metal sulfides and physical settling of most particles that might remain in suspension.

Not all heavy metals behave according to the pattern described above. As mentioned above with reference to redox reactions, chromium is an example of a trace metal that can exist either as metal cation (Cr^{+3}) or, in the Cr^{+6} state, as chromate (CrO_4^{2-}) or dichromate ($Cr_2O_7^{2-}$) oxyanions. As you might expect from the high cation charge, hexavalent chromium tends to occur in highly oxidized environments. It is also the main form of Cr in waters with pH > 7, meaning that unlike the base metals described above, Cr will be very soluble in basic/alkaline waters.

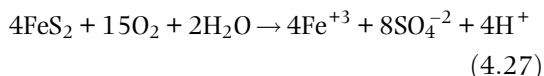
Uranium also tends to occur as a hexavalent cation (U^{+6}), but unlike Cr, U forms polyatomic anions commonly containing the carbonate and hydroxide anions in neutral to alkaline waters. Examples include $UO_2(CO_3)_2^{2-}$ and $UO_2CO_3(OH)_3^{-1}$, and these U-carbonate-hydroxide complex anions greatly enhance U mobility in aqueous systems. At low pH (< 6), U primarily occurs as UO_2^{+2} , a relatively insoluble cation (commonly termed the uranyl ion) that is likely to be sorbed by minerals or organic matter. It is the absence of available carbonate and hydroxide anions at pH < 6 prohibits that formation of soluble $U-CO_3-OH$ anions and renders U relatively insoluble in slightly acidic to highly acidic waters.

Metalloids are elements that possess metallic and non-metallic properties, and for this reason one of their most well-known uses is as semiconductors. They are located on the right side of the periodic table between the metals and gases, and include boron, silicon, germanium, arsenic, selenium and antimony. The metalloid that generates the greatest amount of concern and attention in environmental geochemistry is **arsenic**. It often occurs in base metal deposits or gold deposits, typically as arsenic sulfides such as realgar (AsS) or orpiment (As₂S₃), as the iron–arsenic sulfide arsenopyrite (FeAsS), or as arsenic coprecipitated in other sulfides (e.g. in pyrite).

Arsenic is often released to surface waters or soils by oxidation of arsenic-bearing sulfides. The example below is for arsenopyrite:



Note that in this reaction, the oxidation of sulfur is driven by the reduction of ferric iron (Fe^{+3}) to ferrous iron (Fe^{+2}). This is a typical process in iron-rich sulfide deposits, where the oxidation of pyrite (here shown as a single reaction that summarizes two or more steps) produces high quantities of available ferric iron (note: this equation describes formation of sulfuric acid associated with acid mine drainage; Chapter 9):



In natural systems arsenic occurs in three oxidation states that span a wide redox range, from -3 to $+3$ and $+5$; common forms of As in rocks and surficial environments are $+3$ and $+5$, often written As(III) and As(V). Arsenic in the trivalent state can substitute for ferric iron given their identical charge and similarity in radii in octahedral coordination ($\text{As}^{+3} = 0.58 \text{ \AA}$; $\text{Fe}^{+3} = 0.65 \text{ \AA}$) (Table 2.1); As(V) is also sorbed into Fe-hydroxides, so As mobility tends to be lower in oxidizing waters where Fe-hydroxides are stable.

In solution, arsenic exists in two oxidation states, As^{+3} and As^{+5} , both of which form oxyanions, arsenite in the case of As^{+3} and arsenate in the case of As^{+5} . The general form of arsenite is $\text{H}_n\text{AsO}_3^{3-n}$, and arsenate is $\text{H}_n\text{AsO}_4^{3-n}$. Over pH ranges in most natural settings, arsenite occurs as arsenious acid (H_3AsO_3) and the arsenite anion $\text{H}_2\text{AsO}_3^{-1}$, whereas arsenate occurs as two arsenate anions, $\text{H}_2\text{AsO}_4^{-1}$ and HAsO_4^{-2} . Arsenic acid (H_3AsO_4) occurs in oxidizing conditions at a pH < 2 and thus might be found in very acidic mine drainage situations. Lower pH means greater activity of H^+ , so acidic

waters tend to contain the more H-rich species H_3AsO_3 and $\text{H}_2\text{AsO}_4^{-1}$. In fact, high concentrations of H^+ in solution drive arsenite (shown below) and arsenate anions towards more protonated states:



In high-pH solutions, arsenic species tend to lose H^+ to solution, as is illustrated for arsenate anions:



Arsenite predominates in reducing solutions, whereas arsenate is more common in oxidizing solutions, a predictable relationship when considering that arsenite contains the more reduced As(III) and that arsenate contains the most oxidized form of arsenic, As(V). In this sense, arsenic would be termed a **redox-sensitive element** that commonly occurs as oxyanions.

Arsenic is generally soluble under all conditions except for acidic, reducing environments that contain available sulfide anions which foster formation of arsenic sulfides, or coprecipitation of As into iron sulfide. This can be problematic when trying to remediate sites that are contaminated with heavy metals and arsenic. Employing the strategy of creating alkaline, reducing conditions will immobilize base metals such as Cu, Ni, Pb and Zn, but will likely create high dissolved arsenic concentrations. Also, because of its anionic character in solution, it is generally not attracted to cation-exchange sites on surfaces of clay minerals or organic matter. At low pH, when the surface of iron oxides and hydroxides are positively charged (by H^+), arsenate is sorbed, forming relatively strong inner-sphere complexes. However, at high pH (> 10), sorption is inhibited by the strong negative charges on the surfaces of iron oxyhydroxides (caused by deprotonation) that results in repulsion of arsenate and arsenite anions.

The metalloid selenium takes on two oxidation states, $+4$ and $+6$, and like arsenic, also occurs as

oxyanions. The selenite anion, SeO_3^{2-} , is the dominant species in mildly reducing waters, whereas the selenate anion, SeO_4^{2-} , is the dominant form under oxidizing conditions. In acidic conditions, the selenite anion reacts with protons to form HSeO_3^- . While selenium in general is soluble, the more-reduced form, selenite, is strongly sorbed to iron oxyhydroxides at circum-neutral pH, whereas the more oxidized selenate is only very weakly sorbed. In this way its behavior is the opposite of arsenic, where the more oxidized arsenate anion is strongly sorbed but arsenite is not.

4.9.2 Eh–pH diagrams

Speciation of trace metals is commonly illustrated using **Eh–pH diagrams** (also known as Pourbaix diagrams after pioneering work in the 1940s by Belgian chemist M. Pourbaix), which are sometimes plotted as **pe–pH diagrams**. These diagrams were developed to depict the influence of redox conditions and pH on speciation. The flux of electrons associated with redox reactions can be measured, and while the tendency to gain (or lose) electrons does not correspond to a concentration of electrons in the same way that acidity corresponds to a concentration of H^+ ions, redox potential nonetheless can be envisioned as an activity of electrons in solution (a_{e^-}), where a_{e^-} really represents the tendency of a system to provide (or consume) electrons. From this reasoning, the unit pe can be defined in a manner analogous to pH (recall that $\text{pH} = -\log_{10} a_{\text{H}^+}$):

$$\text{pe} = -\log_{10} a_{e^-} \quad (4.30)$$

The pe scale in natural Earth systems ranges from -15 to $+20$, where the lower limit corresponds to reduction of H_2O to H_2 gas and O^{2-} , and where the upper limit corresponds to oxidation of $2 \text{H}_2\text{O}$ molecules to 4H^+ ions and O_2 gas.

The activity of electrons in solution is often expressed relative to the standard hydrogen

electrode by the unit Eh, where the conversion from pe to Eh is:

$$\text{Eh} = \text{pe} * 2.303RT/F \quad (4.31)$$

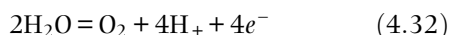
Thus, $\text{Eh} = \text{pe} / 16.9$. In the above equation, the gas constant $R = 8.314 \times 10^{-3} \text{ kJ/deg}$, $T = 298.15 \text{ K}$ (i.e. 25°C), $F = 96.5 \text{ kJ/volt-gram equivalent}$ and 2.303 is the conversion factor for $\log_e$ to $\log_{10}$.

Many field probes measure redox conditions using oxidation–reduction potential (ORP) meters that report results in units of volts (V) or millivolts (mV). These are the same units of measure as Eh, but the difference between ORP and Eh is that ORP measured in the field is often determined relative to a Ag–AgCl reference electrode, whereas Eh is defined with respect to the standard hydrogen electrode (SHE). However, the SHE is not convenient for field use, so field probes use other standard references. Accordingly, ORP values obtained in the field must be converted to Eh, which can be done by correcting (adding) the ORP readings obtained with a practical reference electrode for field work (e.g. Ag–AgCl) to the SHE-normalized value to account for the difference between the reference electrode and the SHE. For example, a measurement of the redox potential of the ZoBell solution used to calibrate the ORP meter using an Ag/AgCl reference electrode in 4 M KCl is approximately $+230 \text{ mV}$, whereas the same ZoBell solution measured against the SHE will read approximately $+430 \text{ mV}$. So, to convert ORP readings to Eh, 200 mV must be added to the ORP voltage, and a field ORP measurement of -53 mV (using Ag–AgCl as reference) would convert to an Eh value of $+147 \text{ mV}$.

Typically, Eh–pH diagrams depict systems at 25°C and 1 atm with components that commonly include H_2O , CO_2 , O , S and the metal or metalloid in question. They also include information on concentrations or activities of the various components in the system. Alternatively, a user may specify these variables if using software such

as Geochemist's Workbench or similar software (e.g. PHREEQE) to produce an Eh–pH diagram. Examples of where selected natural environments plot in Eh (pe) – pH space are shown in the diagram in Fig. 4.13.

The upper limit on the stability of water can be expressed by this equation:



Eh can be calculated from the Nernst equation expressed as follows:

$$\text{Eh} = E^0 + (RT/nF) * \ln(K_e) \quad (4.33)$$

where E^0 is the standard potential for the half-cell reaction, R is the gas constant, T is temperature (Kelvin), n is the number of moles of electrons transferred and K_e is the equilibrium constant for the reaction.

Factoring in constants R , T (298.15 K, i.e. 25 °C), F and 2.303 (to convert $\ln$ to $\log_{10}$), for a generic reaction as shown in equations 4.16 and 4.17, the equation becomes:

$$\text{Eh} = E^0 + 0.0592/n * \log\left([C]^c * [D]^d / [A]^a * [B]^b\right) \quad (4.34)$$

So for equation 4.30,

$$\text{Eh} = E^0 + 0.0592/4 * \log\left(\text{PO}_2 * [\text{H}^+]^4 / [\text{H}_2\text{O}]^2\right) \quad (4.35)$$

However, given that PO_2 and $[\text{H}_2\text{O}]$ are both 1, the equation can be simplified to:

$$\text{Eh} = E^0 + 0.0592/4 * \log\left([\text{H}^+]^4\right) \quad (4.36)$$

And considering that $\text{pH} = -\log[\text{H}^+]$ and that $\log([\text{H}^+]^4) = 4\log[\text{H}^+]$,

$$\text{Eh} = E^0 - 0.0592 * \text{pH} \quad (4.37)$$

What this indicates is that the upper limit of the stability of water is a straight line with a slope of -0.0592 . The y -intercept of E^0 , which applies to equation 4.30, is $+1.23 \text{ V}$, the same as the value listed in Table 4.2 for this reaction relative to the standard hydrogen electrode.

A similar calculation for the lower limit on the stability of water gives us the same equation as 4.32:

$$\text{Eh} = E^0 - 0.0592 * \text{pH}$$

In other words, the lower limit of the stability of water is also a straight line with a slope of -0.0592 . The difference is in the value of E^0 , which for the standard hydrogen electrode (equation) is 0.0 V – this is by definition, given that the standard hydrogen electrode is the reference for all other redox reactions and is assigned the value of zero. Therefore, the y -intercept for the lower

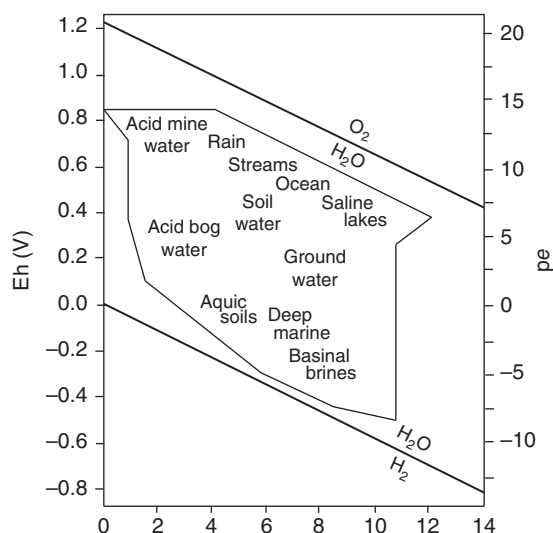


Fig. 4.12 Approximate locations of selected natural systems as a function of reduction–oxidation potential and pH. Note that redox units are given in terms of Eh (left) and pe. The thin line bounding the natural environments indicates the limits of nearly all natural waters (after Bass Becking et al. 1960).

stability limit of water is zero, as it appears on the Eh *y*-axis in typical Eh–pH diagrams (Fig. 4.12).

While they are very useful in depicting potential speciation, the disadvantage of Eh–pH diagrams is that they do not illustrate the tendency of certain assemblages to exist in various states of disequilibrium, nor do they reflect the fact that kinetics (or reaction rates) often dictate mineral assemblages and dissolved species in low-temperature geological environments such as soils, surface waters and groundwaters. We might not consider the potential for As uptake by Fe oxyhydroxides in an anoxic wetland based on equilibrium considerations alone. $\text{Fe}(\text{OH})_3$ and FeOOH are not stable in anoxic settings; however, it might take years or decades or longer for Fe oxyhydroxides to dissolve in this type of setting, meaning that they could exert a control over arsenic mobility in spite of the fact that they are out of equilibrium with their environment.

Eh–pH diagrams also do not provide information on the potential of a given ion to be sorbed to mineral surfaces, nor do they consider coprecipitation reactions that might remove ions from solution, a good example being the tendency

of Ba to remove Ra from solution upon crystallization of BaSO_4 . It is rare that Eh–pH diagrams include Si, and even when they do, they do not portray incorporation of trace metals into silicate clays or other minerals as trace or minor components. Nonetheless, they are very useful at providing simplified graphic representations of elemental speciation, and they are widely used in environmental geochemistry.

Eh–pH diagrams of copper and lead are presented in Fig. 4.13. The diagrams for these two metals are shown here because they depict the main patterns and types of species that occur for many other heavy metals, including Cd, Co, Ni and Zn. In both diagrams, the system includes the heavy metal (Cu or Pb) plus oxygen, sulfur, water and carbon dioxide (meaning that we can assess the role of carbonates). It is important to be aware of the system composition presented in such diagrams because it will affect speciation.

Eh–pH diagrams for arsenic and uranium are presented in Fig. 4.14 to illustrate patterns for two of the elements that tend to form oxyanions, and also to provide data for elements that are inherently toxic. Compared to copper, lead and

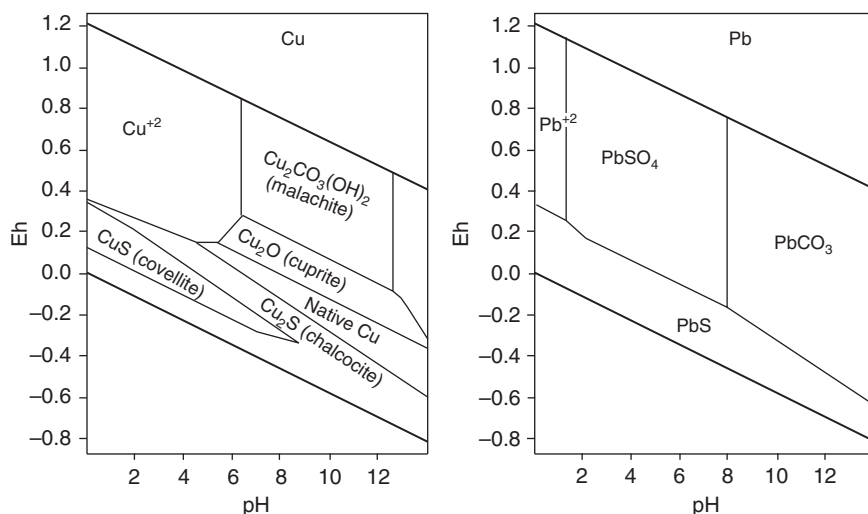


Fig. 4.13 Copper and lead Eh–pH diagrams at 25 °C and 1 atm pressure. Activities of Cu and Pb are 10^{-6} mol/L, $p\text{CO}_2 = 400$ ppmv ($10^{-3.4}$ atm) and total S = 10^{-2} mol/L.

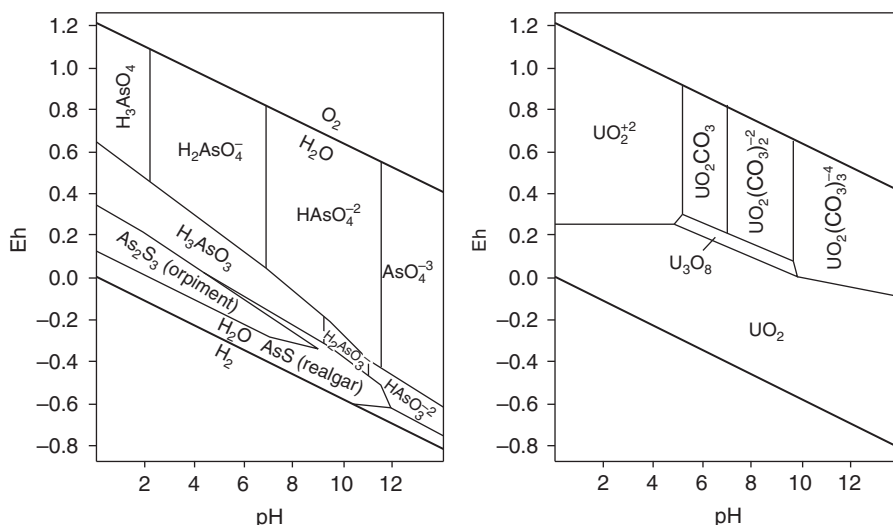


Fig. 4.14 Arsenic and uranium Eh–pH diagrams for systems at 25 °C and 1 atm pressure. For the arsenic diagram, activity of $\text{As}_{\text{total}} = 10^{-6}$ and S activity = 10^{-2} mol/L. For the uranium diagram, activity of $\text{U} = 10^{-8}$ mol/L and C (as CO_2) is $10^{-3.4}$ atm.

other trace metals that are effectively insoluble at neutral or higher pH, arsenic and uranium possess broad aqueous fields at neutral and higher pH. In fact, they are soluble over most Eh and pH ranges, merely changing forms and oxidation states with changing redox and pH values.

Aluminum and iron are useful to consider in conjunction because they behave similarly under many natural geochemical conditions. Both are largely insoluble in neutral waters and become progressively (and exponentially) more soluble with increasing acidity as well as increasing alkalinity. They are both abundant in silicate minerals, and they commonly precipitate out of soil solutions, surface waters and groundwaters as hydroxides, e.g. $\text{Fe}(\text{OH})_3$ and $\text{Al}(\text{OH})_3$. However, there are important differences.

- First, aluminum exists only in the Al^{+3} state (in compounds and as a simple ion in solution) and is not redox sensitive, whereas iron (as Fe^{+2} and Fe^{+3}) is very redox sensitive. This difference is apparent in the Eh–pH diagrams, where species boundaries for aluminum are vertical, meaning that changing Eh conditions do not result in a

change in Al speciation. In a system controlled by Eh and pH, only changing pH will affect Al speciation. Conversely, changing Eh can result in a change in iron speciation, reflecting the fact that iron is sensitive to redox conditions (although in reality, if Al is co-precipitated in a redox-sensitive mineral like $\text{Fe}(\text{OH})_3$ or Fe_2O_3 , its behavior will be in part controlled by redox reactions affecting the iron hydroxide or oxide). For example, decreasing Eh (reflecting transition to more-reducing conditions) at constant pH of 6 from Eh = 0.6 V to 0.2 V will cause reduction of Fe^{+3} to Fe^{+2} , resulting in dissolution of $\text{Fe}(\text{OH})_3$ and leaving Fe^{+2} in solution. Note that at low to intermediate Eh, iron exists as the Fe^{+2} cation in solution or as Fe^{+2} in minerals such as pyrite and siderite. At high Eh values, iron occurs either as the Fe^{+3} simple cation in solution (in highly acidic, oxidizing water) or as Fe^{+3} in ferrihydrite and other iron oxides and hydroxides.

- Secondly, iron forms compounds with sulfur and carbonate, whereas Al does not (note presence of pyrite and siderite in Eh–pH diagram, Fig. 4.15).

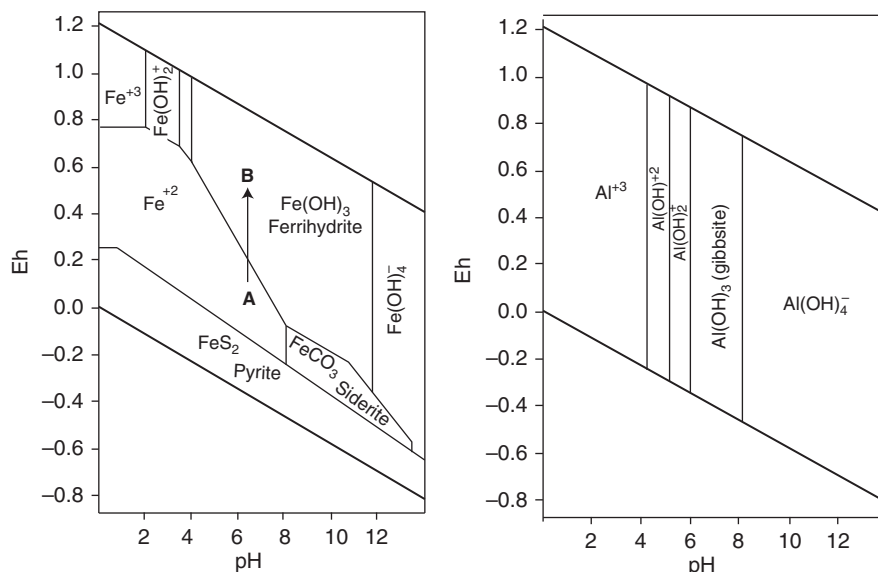


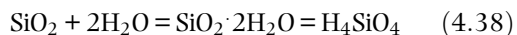
Fig. 4.15 Eh–pH diagram for iron and aluminum. For the iron diagram, activity of $\text{Fe} = 10^{-6}$ mol/L, $S = 10^{-2}$ mol/L and $\text{CO}_2 = 10^{-3.4}$ atm. For aluminum diagram, $\text{Al} = 10^{-9}$ mol/L. Note that iron speciation is influenced by both Eh and pH; conversely, aluminum speciation is controlled by pH but not Eh. The arrow from A to B represents oxidation as shown in Plate 5.

- Thirdly, based on the sizes of the stability fields at high pH, it appears that Al is more soluble in high pH solutions than in iron.

Recall that Fig. 4.10 presents a diagram of aluminum speciation as a function of pH, which facilitates a more quantitative analysis than is possible with Eh–pH diagrams.

4.9.3 Silicon in solutions

The behavior of silicon in solution is nearly the complete opposite of most trace metals because silicon reaches its greatest solubility under high pH. At pH values < 9.8 , the dominant silicon species is H_4SiO_4 , alternately referred to as hydrated silica or silicic acid. This species can be produced in a simple sense by dissolution of quartz by hydration:



If dissolved silica concentrations (measured with respect to SiO_2) below pH = 9.8 are 10^{-4} M, the solution is in equilibrium with quartz. If dissolved silica concentrations are $> 10^{-4}$ M (6 mg/L) at pH < 9.8 , then the solution is oversaturated with respect to quartz and we would expect quartz to precipitate out of solution. Alternatively, if the solution is in equilibrium with respect to amorphous silica, which is far more soluble than quartz, the equilibrium concentration of H_4SiO_4 (as SiO_2) would be approximately 2×10^{-3} M (120 mg/L). These values compare well with concentrations of dissolved silica in ocean water, streams, groundwaters and soil waters, which generally range from 10 to 100 mg/L. Values < 1 mg/L can be produced in lakes and marine environments where diatoms (siliceous phytoplankton) sequester Si from solution to form cell walls (tests), and values > 100 mg/L occur in hydrothermal systems where high temperatures enhance silica solubility.

Above a pH of 9.8, high concentrations of OH^- cause deprotonation of H_4SiO_4 as follows:



This reaction accompanies exponential increases in silica solubility at $\text{pH} > 9$ (Fig. 4.16), and with further increases in pH, deprotonation produces $\text{H}_2\text{SiO}_4^{2-}$ to the extent that at the extremely high pH of 13.1, the concentration of $\text{H}_2\text{SiO}_4^{2-}$ in solution is greater than H_3SiO_4^- :

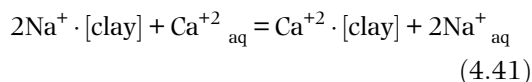


The total amount of dissolved silica will be the sum of the three hydrated silica species. At $\text{pH} < 9$, nearly all dissolved silica occurs as H_4SiO_4 . Between $\text{pH} = 9.8$ and 13.1, the main dissolved species is H_3SiO_4^- , with lesser amounts of H_4SiO_4 and $\text{H}_2\text{SiO}_4^{2-}$.

4.10 EFFECT OF ADSORPTION AND ION EXCHANGE ON WATER CHEMISTRY

This topic was first introduced in Chapter 2 (Section 2.5) where the focus was the properties of minerals as they influence ion exchange, and also of the role of pH in ion exchange (e.g. PZC, Table 2.4); here, the focus is equilibrium exchange and its effects on the composition of aqueous solutions.

The exchange reaction of Ca^{+2} from aqueous solutions into an exchange site (on a mineral or organic matter, e.g.) occupied by Na^+ can be represented as:



Why might such a reaction occur, i.e. why might Ca^{+2} replace Na^+ on a mineral surface exchange site? One possible reason is that the attraction of

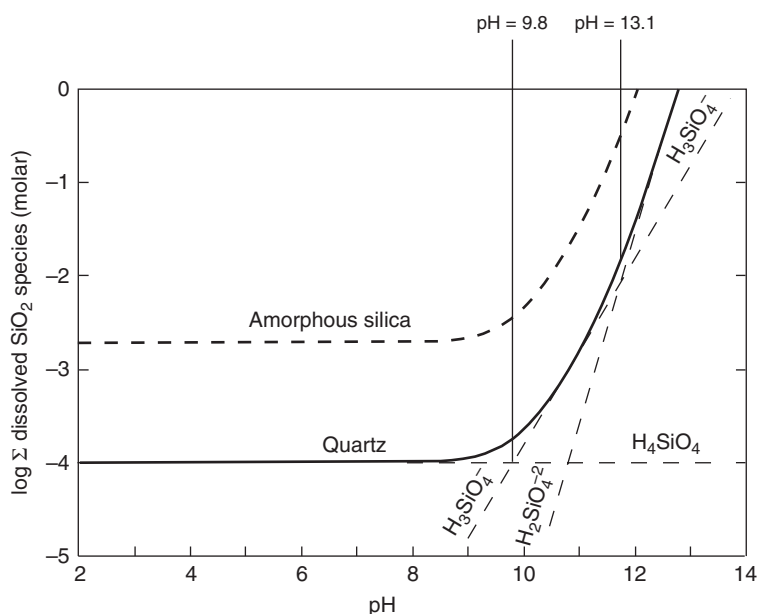


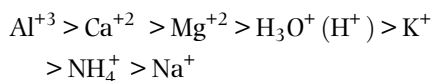
Fig. 4.16 Solubility of silica as a function of pH in equilibrium with amorphous silica and with quartz. Note that if amorphous silica is the solid phase in equilibrium with silica, dissolved silica concentrations are approximately an order of magnitude greater than when quartz controls silica solubility.

divalent Ca^{+2} to the mineral surface is greater than the attraction of Na^+ to these sites, which would result in release of Na^+ to solution and adsorption of Ca^{+2} .

4.10.1 Ionic potential, hydration radius and adsorption

The ratio of ionic charge:ionic radius (**ionic potential**) for an element helps to determine how strongly it will be attracted to exchange sites – high ratios indicate strong attraction. Although ionic radius is often used, the *hydration* radius of a cation is a better predictor of its attraction to exchange sites – this is because, in solution, ions exist within a shell of water molecules that are closely bonded to the ion. The hydration shell causes the charge on the ion to be more diffuse (in an extreme case, the hydration shell of Al^{+3} is almost ten times greater than its ionic radius). Table 4.6 presents values of ionic radii and well as hydration radii for a group of cations.

Other factors, including the composition of the aqueous solution, are also important, but in a general sense, order of decreasing affinity for cation exchange sites (in ~ dilute fresh water systems) among common cations, determined by ratio of hydration radius to ionic charge, is:



The high charge and low radius of Al^{+3} , whether in ionic form or hydrated (Table 4.6), means that it will out-compete Na^+ and Ca^{+2} for exchange sites in acidic soils (e.g. at $\text{pH} < 5$). This is a problem in acid rain-impacted soils prone to loss of calcium, a crucial limiting nutrient in many forest ecosystems. The strong attraction of certain smectite interlayers for K^+ make it more strongly sorbed than would be predicted based on ionic

Table 4.6 Hydration radii relative to ionic radii and charges of selected cations (Volkov et al., 1997). R_{ionic} is ionic radius and R_{hydr} is hydration radius. Lower ratios of $R_{\text{hydr}} : \text{chg}$ (chg = ionic charge) correspond to stronger adsorption to cation exchange sites. H_3O^+ is the hydronium ion form of H^+ .

Cation	R_{ionic} (Å)	R_{ionic} : chg	R_{hydr} (Å)	R_{hydr} : chg
H_3O^+	1.15	1.15	2.80	2.80
NH_4^+	1.48	1.48	3.31	3.31
Li^+	0.94	0.94	3.82	3.82
Be^{+2}	0.31	0.16	4.59	2.30
Na^+	1.17	1.17	3.58	3.58
Mg^{+2}	0.72	0.36	4.28	2.14
Al^{+3}	0.53	0.18	4.80	1.60
K^+	1.49	1.49	3.31	3.31
Ca^{+2}	1.00	1.00	4.12	2.06
Rb^+	1.63	1.63	3.29	3.29
Cs^+	1.86	1.86	3.29	3.29
Zn^{+2}	0.74	0.37	4.30	2.15
Cd^{+2}	0.97	0.49	4.26	2.13
Pb^{+2}	1.32	0.66	4.01	2.01

potential alone. In acidic soils, H^+ (as H_3O^+) is strongly attracted to mineral surfaces, behavior that is strongly controlled by the abundance of H^+ in acidic solutions. In alkaline soils, H^+ is likely to be desorbed to neutralize pH. Also, as is discussed in Chapter 9 (section 9.6), Na^+ is more strongly adsorbed than Ca^{+2} in water with high ionic strength, so exchange sites on particles in desert soils and marine sediments will likely be dominated by Na^+ . In temperate soils and fresh waters, Ca^{+2} is likely to be the dominant adsorbed cation.

4.10.2 Law of mass action and adsorption

Another important control on exchange is the chemistry of the aqueous solution, both in terms of concentrations of potential adsorbates (in this case cations like Na^+ and Ca^{+2}), as well as overall

ionic strength of the solution. In a general sense, the quantity of Ca^{+2} removed from solution will occur in approximate proportion to its abundance in solution accounting for competition with other ions (e.g. Mg^{+2}) and the composition of ions in exchange sites (e.g. Na^{+}).

Exchange of Ca^{+2} for Na^{+} as shown in equation 4.41 can be described by this formulation of the law of mass action:

$$K = a_{\text{Na}^{+}}^2 * a_{\text{Ca}^{+2} \cdot [\text{Clay}]} / a_{\text{Ca}^{+2}} * a_{\text{Na}^{+} \cdot [\text{Clay}]}^2 \quad (4.42)$$

K in this equation is not a constant, but rather is a coefficient (often referred to as the “mass action coefficient” or “selectivity coefficient”) that varies, especially as a function of changing ionic strength. The higher the value of K , the more likely that Ca^{+2} will be more-strongly adsorbed to the particle (clay in this case) than Na^{+} .

Given that exchange often involves ions of different charge (e.g. Na^{+} vs. Ca^{+2}), a term known as the equivalent fraction (X_M^{+}) is introduced, where M^{+} is a generic cation. X_M^{+} is the fraction of the cation exchange capacity occupied by cation M^{+} , where a divalent cation will occupy twice the capacity per mole as a monovalent cation, and $\Sigma X_M^{+} = 1.0$. Equation 4.42 can now be expressed as:

$$K = a_{\text{Na}^{+}}^2 * X_{\text{Ca}^{+2} \cdot [\text{Clay}]} / a_{\text{Ca}^{+2}} * X_{\text{Na}^{+} \cdot [\text{Clay}]}^2 \quad (4.43)$$

This is the Kerr approach to expressing the selectivity coefficient. There are a few other approaches, including the Gaines–Thomas, Gapon and Vanse-low, but the relatively subtle differences among them do not merit detailed analysis given the scope of this book.

The relative abundances of Na and Ca in exchange sites can be determined by assuming that $K = 1$ (i.e. that there is no strong preference by the clay for either Na^{+} or Ca^{+2} , which is clearly a simplification) and by realizing that, in a simplified system of only two cations (Na^{+} and Ca^{+2}),

$X_{\text{Na}[\text{Clay}]}^{+} + X_{\text{Ca}[\text{Clay}]}^{+2} = 1.0$ (i.e. that Na^{+} and Ca^{+2} occupy 100% of exchange sites). This being the case, we can examine two scenarios, one in high ionic strength water and one in low ionic strength water.

If we set activities of Na^{+} and Ca^{+2} in solution to a very high end-member value of 1.0 M each, the equation becomes:

$$K = 1.0^2 * X_{\text{Ca}^{+2} \cdot [\text{Clay}]} / 1.0 * X_{\text{Na}^{+} \cdot [\text{Clay}]}^2 = 1$$

Rearranging gives:

$$X_{\text{Na}^{+} \cdot [\text{Clay}]}^2 - X_{\text{Ca}^{+2} \cdot [\text{Clay}]} = 0$$

Given that $X_{\text{Na}^{+} \cdot [\text{Clay}]} + X_{\text{Ca}^{+2} \cdot [\text{Clay}]} = 1$, we can set:

$$X_{\text{Na}^{+} \cdot [\text{Clay}]} = x$$

$$X_{\text{Ca}^{+2} \cdot [\text{Clay}]} = x - 1$$

This results in an equation $x^2 + x - 1 = 0$, and using the quadratic formula, we find that $x = 0.618$, so of the total exchange sites, 61.8% are occupied by Na^{+} ($X_{\text{Na}^{+} \cdot [\text{Clay}]} = 0.618$), and the remaining 38.2% of exchange sites are occupied by Ca.

If we set activities of Na^{+} and Ca^{+2} in solution to a low end-member value of 0.001 M each, the equation becomes:

$$K = 0.001^2 * X_{\text{Ca}^{+2} \cdot [\text{Clay}]} / 0.001 * X_{\text{Na}^{+} \cdot [\text{Clay}]}^2 = 1$$

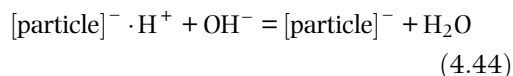
Solving as above, ($X_{\text{Na}^{+} \cdot [\text{Clay}]} = 0.031$), indicating that only 3.1% of exchange sites are occupied by Na^{+} , whereas 96.9% are occupied by Ca^{+2} . This emphasizes the fact that, at low ionic strength (e.g. temperate river, lake and soil water or groundwater), particles have a strong preference to adsorb Ca^{+2} over Na^{+} ; conversely, particles in sea water, salty lake water and certain hydro-thermal waters of high ionic strength will have a

preference to incorporate Na^+ rather than Ca^{+2} into exchange sites. Performing these calculations for other monovalent–divalent pairs will produce similar results. For example, in fresh waters, toxic metals such as Pb^{+2} and Cd^{+2} will sorb more strongly to clays than will Na^+ , but if eroded and transported into the marine environment we would predict exchange reactions involving adsorption of Na^+ and desorption of Pb^{+2} or Cd^{+2} .

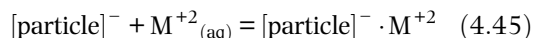
4.10.3 Adsorption edges

One of the main controls on adsorption of ionic species to surfaces of minerals and organic matter is the charge on the particle surface, and solution pH exerts an important influence on surface charge. At low pH, the high concentrations of H^+ in solution mean that in acidic solutions, particle surfaces will be dominated by H^+ , producing a positively charged particle surfaces. The result is that (1) metal cations will not be attracted to these surfaces, and (2) anions will be attracted to the positive charge imparted by the H^+ -rich surface. At high pH, desorption of H^+ leaves

negatively charged particle surfaces that can be illustrated by a simple equation.



The negatively charged, deprotonated surface favors adsorption of cations:



In this case, a divalent metal cation in solution is adsorbed to the surface of a negatively charged particle above the adsorption edge for that particle and that metal cation.

Accordingly, adsorption of heavy metal cations (Cu^{+2} , Pb^{+2} , Zn^{+2} , etc.) or base cations (Na^+ , K^+ , Ca^{+2} , Mg^{+2}) decreases with decreasing pH, and for most metal cations adsorption will shift from nearly complete to nearly zero adsorption over a narrow pH range, typically 1–2 units of pH (Fig. 4.17). This threshold is known as the **adsorption edge**, or in more general terms, the sorption edge. The opposite is true for adsorption of anions – adsorption will decrease with

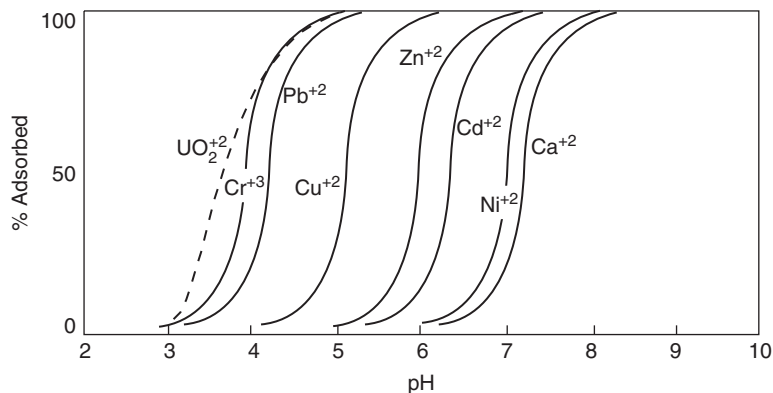


Fig. 4.17 Adsorption of metal cations onto ferrihydrite as a function of pH. Data are from Dzombak and Morel (1990) except for uranyl, which is from Hsi and Langmuir (1985). The curvature in the adsorption edges indicate that, at low pH, adsorption is relatively weak; with increasing pH, cations are more-strongly adsorbed. For example, at pH = 4, Cu^{+2} is weakly adsorbed (i.e. only a small proportion of Cu^{+2} ions in solution will be adsorbed at pH = 4), but at pH = 5.5, nearly all Cu^{+2} in solution will be adsorbed onto ferrihydrite. The boundary between pH 4 and 5.5 is the adsorption edge for Cu^{+2} .

increasing pH as H^+ is released from surfaces to produce negatively charged surfaces that repel anions. The adsorption edge is controlled by pH, characteristics of the metal undergoing adsorption (or desorption), and the structure and charge distribution of the adsorbing particulate. The majority of adsorption edge data have been determined for iron hydroxides given their near ubiquity in surficial systems.

Absorption of Cr^{+3} and Pb^{+2} occurs at the lowest pH of the cations indicated in Fig. 4.17, reflecting their low solubilities in solution and corresponding stronger likelihood of adsorption. The more-soluble metal cations (Ni^{+2} , Zn^{+2}) exhibit adsorption edges at higher pH values, with the highly soluble Ca^{+2} shown for reference (Ca^{+2} is not adsorbed to ferrihydrite at $pH < 6$). So, for example, at a pH of 5, we would predict that all Cr^{+3} and Pb^{+2} ions in an aqueous system would be adsorbed to surfaces of iron hydroxides (assuming sufficient $Fe(OH)_3$ is present in the system), and that Ni^{+2} and Zn^{+2} ions would be present in solution. If we assume that the $Fe(OH)_3$ occurs as fine-grained colloidal particles suspended in the water, filtration of this water using a standard

0.45- μm filter (Drever, 1973) before chemical analysis should reveal the presence of Ni and Zn in solution, but not Cr and Pb, which should be retained with the particles on the filter.

Adsorption of representative anions by $Fe(OH)_3$ in Fig. 4.18 illustrates decreasing adsorption with increasing pH.

Arsenate, AsO_4^{-3} , exhibits an adsorption edge at a pH of approximately 11, meaning that $> 50\%$ of arsenate in solution will be adsorbed at $pH < 11$, and nearly all arsenate will be adsorbed at $pH < 9$ (provided sufficient amounts of iron hydroxide occur in the soil or aquifer). It forms strong inner-sphere complexes with iron hydroxide over a broad range of pH values, from slightly basic ($pH = 9$) through neutral and very acidic. Arsenite (AsO_3^{-3}) on the other hand is never completely adsorbed. Predictably, it is adsorbed in progressively lower amounts at $pH > 8$ due to increasing negative charge of the $Fe(OH)_3$ surface, yet it also exhibits progressively lower adsorption at pH values < 8 . The differing behaviors of arsenate versus arsenite are related to speciation, specifically the extent of protonation. Recall that arsenate mainly occurs as two anionic species, $H_2AsO_4^{-1}$ and $HAsO_4^{-2}$ —these

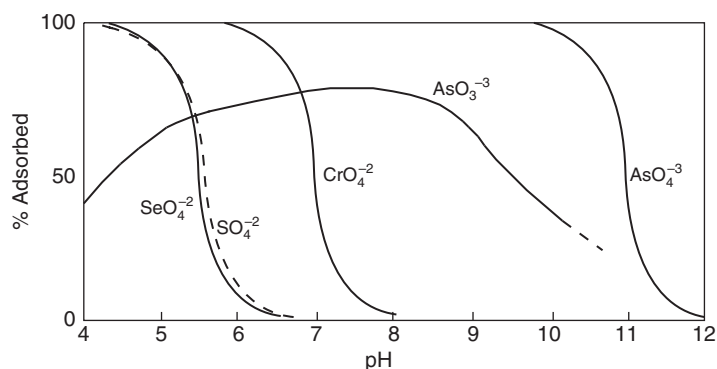


Fig. 4.18 Adsorption of the polyatomic anions selenate, sulfate, chromate, arsenite (AsO_3^{-3}) and arsenate (AsO_4^{-3}) onto ferrihydrite as a function of pH. Data are from Dzombak and Morel (1990). Note that for all anions shown but arsenite, the shape of the curve is the inverse of the curves for metal cations shown in Fig. 4.19, reflecting the fact that, as pH decreases, surfaces of ferrihydrite become progressively more occupied by H^+ , creating positive surface charge needed for adsorption of anions. Thus, adsorption increases to the left of this diagram. Although not shown, decreasing pH causes protonation of the anions (e.g. to CrO_4^{-2} to $HCrO_4^{-}$, AsO_4^{-3} to $HAsO_4^{-2}$ to $H_2AsO_4^{-}$, etc.).

will both be attracted to positively charged particle surfaces at $\text{pH} < 9$. However, as pH decreases towards more acidic values, anionic arsenite (H_2AsO_3^-) reacts with H^+ to form the uncharged H_3AsO_3 species (arsenious acid), which is not electrostatically attracted to particle surfaces and thus largely remains in solution.

One final issue to address regarding adsorption edges is the relationship between the adsorbing particle (i.e. the sorbent) and the potentially adsorbed ion (i.e. the sorbate). From Chapter 2 (Table 2.4), you will recall that different minerals possess different points of zero charge (PZCs) and isoelectric points (IEPs), both of which pertain to pH thresholds at which the charge on a mineral surface (i.e. the sorbent) changes. Below this value, anions are likely to be adsorbed, and above this value, cations are likely to be adsorbed. However, properties of individual ions in combination with the PZC or IEP of the sorbent ultimately determine the pH of the adsorption edge. While PZCs and IEPs of iron oxides and hydroxides generally occur in the pH range of 6 to 8, adsorption edges of cations to the iron hydroxide ferrihydrite occur over a pH range from 3 to 8 (Fig. 4.17). Particles with higher PZCs or IEPs (e.g. iron and aluminum hydroxides) are more suited to anion adsorption, especially at neutral to acidic pH , whereas particles with low PZCs or IEP (e.g. smectites and other silicate clays) are more suited to cation adsorption (except at extremely low pH). The factors that control adsorption are multiple and include mineral content (Is ferrihydrite present? Or is smectite the main sorbent in the system?), pH and aqueous composition as well as kinetics, residence time, temperature and competition among ions.

4.10.4 Adsorption isotherms

Adsorption isotherms describe the extent of adsorption of a given substance as a function of its abundance in solution, at a constant temperature (explaining the term “isotherm”),

accounting for variability in adsorption capacity of different solids (e.g. smectite vs. quartz). All adsorption isotherms show increasing adsorption with increasing concentration (or activity) or a substance in solution – the more of a given substance (e.g. simple or complex ion, organic liquid, gas) there is in solution, the more of it will adsorb to the surface of a solid, based on basic equilibrium considerations.

The simplest isotherm is a **linear** adsorption isotherm (or linear distribution coefficient), which is described by this equation:

$$C_{i(\text{ads})} = K_d * a_{i(\text{aq})} \quad (4.46)$$

$a_{i(\text{aq})}$ is the activity of a given species (i) in solution, K_d is an empirically derived equilibrium constant that quantifies the tendency of species i to adsorb onto a given solid (relative to the amount that remains in solution), and $C_{i(\text{ads})}$ is the concentration of species i adsorbed to the solid (typically in moles of i per kg of solid). The linear relationship shown in Fig. 4.19A reflects a couple of simplistic and unrealistic assumptions, that (1) the solid has an unlimited number of adsorption sites, and (2) that the rate of adsorption with increasing concentration of species i in solution does not decrease as adsorption sites become occupied by i . Two other models that deal with these overly simplistic assumptions are the Langmuir and Freundlich models for adsorption.

The **Langmuir isotherm** (Fig. 4.19B) accounts for the fact that there is a finite number of adsorption sites and as extent of adsorption increases and adsorption sites become saturated, the curve flattens and takes on a logarithmic form. It can be expressed as:

$$C_{i(\text{ads})} = C_{i(\text{ads,max})} * K_L * a_{i(\text{aq})} / (1 + K_L * a_{i(\text{aq})}) \quad (4.47)$$

In this case, $C_{i(\text{ads,max})}$ is the maximum quantity of species i that can be adsorbed to the solid surface and K_L is the Langmuir equilibrium

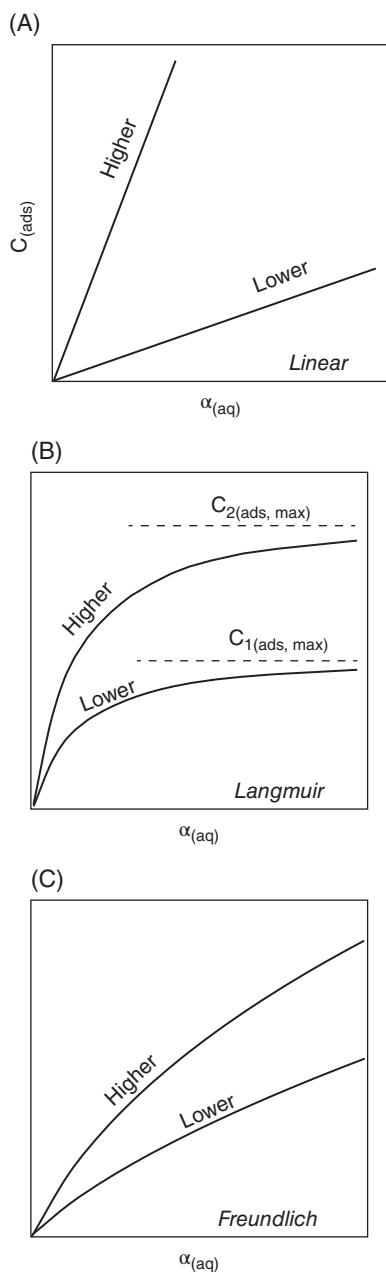


Fig. 4.19 Adsorption isotherms. A = linear isotherm, B = Langmuir isotherm and C = Freundlich isotherm. The x-axis is activity of sorbate (e.g. ion) and y-axis is extent of adsorption. Scenarios for lower and higher sorption potential are shown in each case, where controls are either characteristics of adsorbent (i.e. solid, e.g. CEC) or adsorbate (e.g. ionic charge), or both.

adsorption coefficient. No matter how high the activity of i becomes in solution, the mass of i adsorbed ($C_{i(ads)}$) cannot exceed available adsorption sites ($C_{i(ads, max)}$).

The **Freundlich isotherm** (Fig. 4.19C) also accounts for the fact that adsorption is not linear and thus also reflects the effect of decreasing adsorption rate with increasing activity of i in solution; however, the Freundlich isotherm does not account for a finite number of adsorption sites, and instead reflects the potential that species can become progressively more-weakly adsorbed in outer-sphere sites. It can be expressed as:

$$C_{i(ads)} = K_L * C_{i(ads)}^n \quad (4.48)$$

When the term n is < 1 , the curve becomes logarithmic.

4.11 OTHER GRAPHICAL REPRESENTATIONS OF AQUEOUS SYSTEMS: PIPER AND STIFF DIAGRAMS

Various graphical representations of aqueous geochemistry have already been presented, including Eh–pH diagrams, adsorption edge diagrams, the Bjerrum plot of H_2CO_3 – HCO_3^- – CO_3^{2-} , and pH vs. activity. Two additional diagrams that are commonly used to present aqueous chemistry data are Piper plots and Stiff diagrams.

Piper diagrams plot the major aqueous cations Ca, Mg, and (Na + K) on one ternary diagram (Fig. 4.20, lower left), and the main aqueous anions chloride (Cl^-), sulfate (SO_4^{2-}), and alkalinity (mainly HCO_3^- ; refer to carbonate geochemistry in Chapter 5 for definition of alkalinity) on a second ternary diagram (Fig. 4.20, lower right). Hydrochemical facies as defined by Back (1966) are presented in the upper right and are used to delineate water compositions into different classifications.

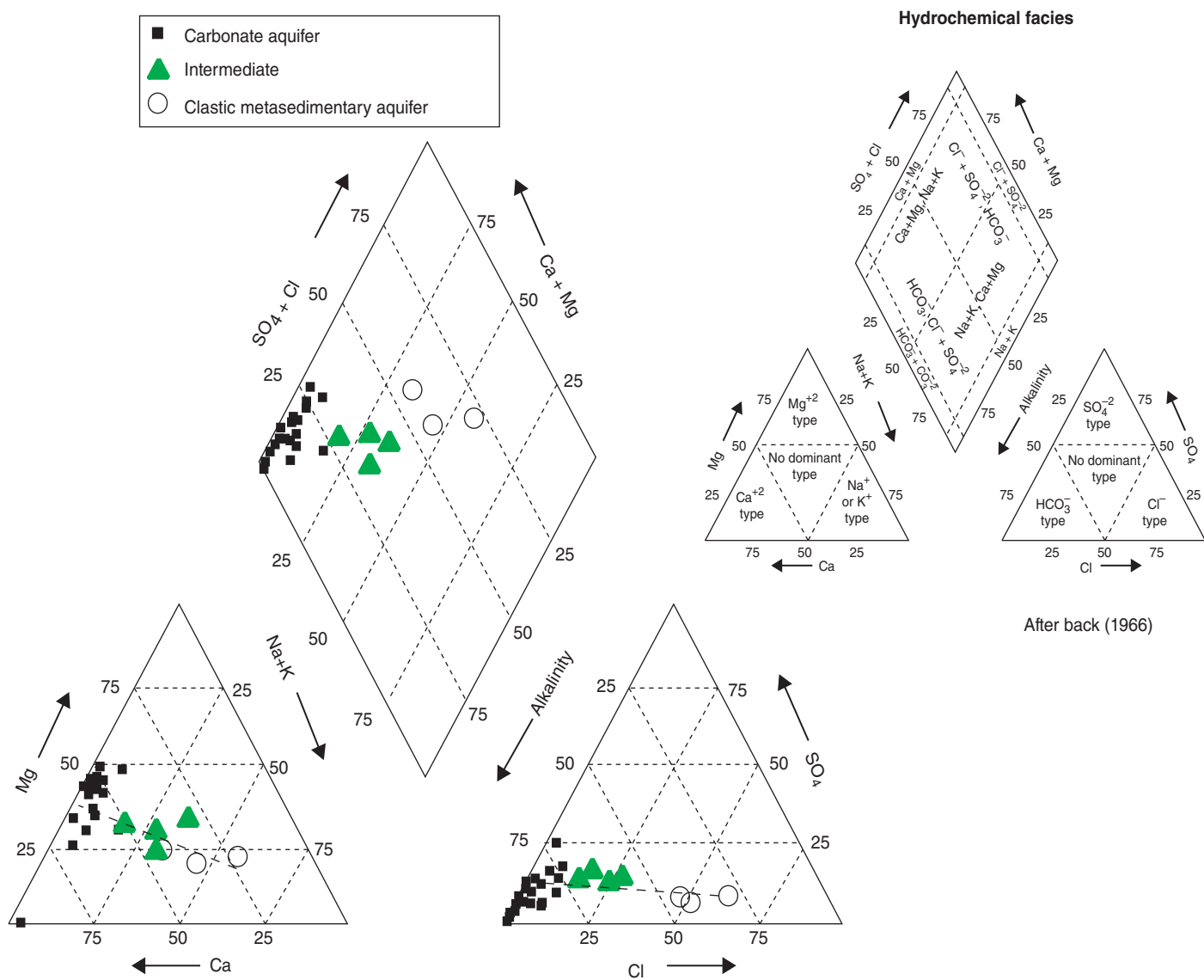
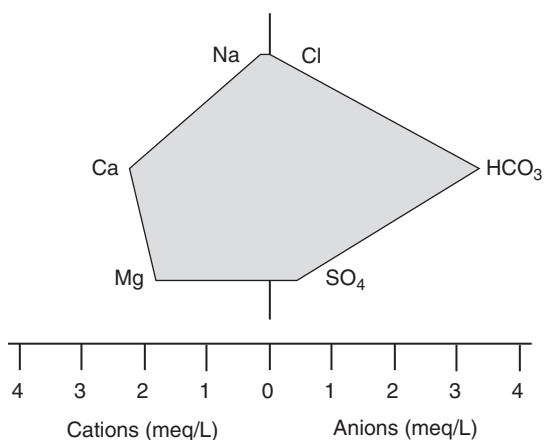


Fig. 4.20 Piper plots of groundwater from aquifers of different bedrock type. Note the Na + K–Cl composition of the clastic metasedimentary aquifer (metamorphosed shales and sandstones) compared to the Ca– HCO_3 composition of the carbonate aquifer (mainly limestone and dolostone). In most groundwater systems, alkalinity is dominated by the bicarbonate anion (HCO_3^-).

Table 4.7 Determination of values plotted on a Stiff diagram (Fig. 4.22).

Element	mg/L	Conversion	meq/L
Mg ⁺²	21.8	* 2 meq/mmol ÷ 24.31 mg/mmol	= 1.79
Ca ⁺²	46.8	* 2 meq/mmol ÷ 40.08 mg/mmol	= 2.33
Na ⁺	4.79	* 1 meq/mmol ÷ 22.98 mg/mmol	= 0.21
SO ₄ ⁻²	20.3	* 2 meq/mmol ÷ 96.06 mg/mmol	= 0.42
HCO ₃ ⁻	205.0	* 1 meq/mmol ÷ 61.02 mg/mmol	= 3.35
Cl ⁻	0.90	* 1 meq/mmol ÷ 35.45 mg/mmol	= 0.025

**Fig. 4.21** Stiff diagram of water dominated by dissolved Ca, Mg and bicarbonate.

Data are plotted in units of milliequivalents per liter (meq/L), which can be obtained from the more common mg/L unit in this way:

$$[C] \text{ (meq/L)} = [C] \text{ (mg/L)} * \text{mmol/mg} * \text{ionic charge on C (meq/mmol)} \quad (4.49)$$

where [C] is concentration of ion C. Note that if data are already present in units of mmol/L, the conversion to meq/L is much simpler:

$$[C] \text{ (meq/L)} = [C] \text{ (mmol/L)} * \text{ionic charge on C (meq/mmol)} \quad (4.50)$$

Data must then be normalized so that the sum of the 3 components, e.g. Ca, Mg and (Na + K), equals 100%. The data presented in Fig. 4.20 are from an aquifer system that consists of carbonate and clastic metasedimentary bedrock units.

Note that waters from the carbonate aquifer predominantly plot as calcium bicarbonate facies, whereas the clastic meta-sedimentary aquifer waters plot within or near to the field for (Na + K) chloride facies. Furthermore, mixing between these two end-members is indicated by the points (representing wells producing from both rock types) that fall along the mixing line (dashed line). As a reminder, the abundance of bicarbonate (HCO₃⁻) versus carbonate (CO₃⁻²), the two dominant species that comprise alkalinity, is controlled by pH. In waters of pH < 8, HCO₃⁻ is the dominant species, whereas in waters of pH > 10, the carbonate anion (CO₃⁻²) is the dominant carbonate species (Drever, 1997).

Stiff diagrams are polygons that plot concentrations (in meq/L) of the cations Na + K, Ca and Mg to the left of a center line, and the anions Cl⁻, HCO₃⁻ + CO₃⁻² and SO₄⁻² to the right of the center line. The farther from the center line, the greater the concentration of the anion or cation (Fig. 4.21). Table 4.7 presents calculations for a Stiff plot using values presented earlier in Table 4.3.

Stiff polygons are often used to show how water chemistry varies spatially by plotting them on a map or schematically, as is shown in Fig. 4.22. In this way, variability in the composition of

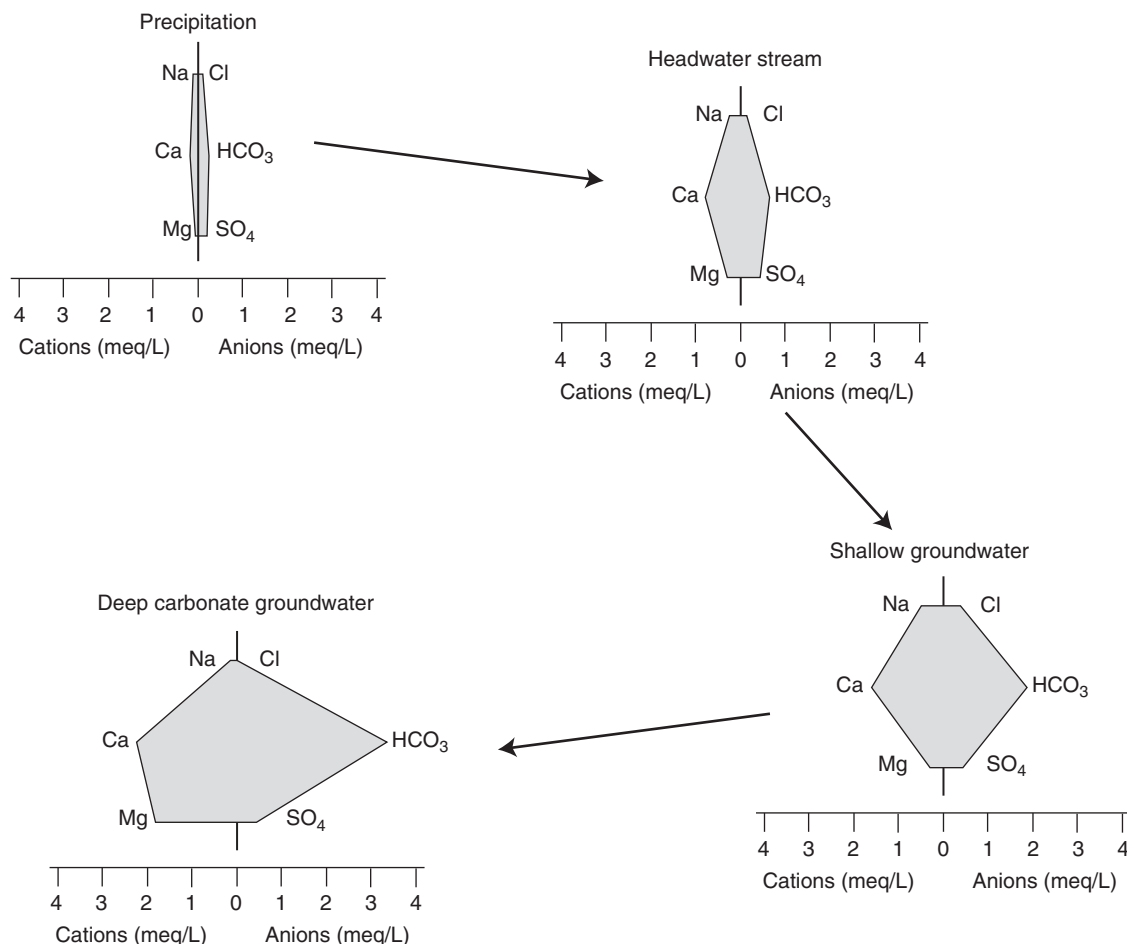


Fig. 4.22 Stiff plots showing water composition at different stages in the hydrologic cycle, emphasizing increasing TDS and alkalinity from atmosphere to surface water and eventually to deep (e.g. 150 m) carbonate aquifer groundwater.

streams, aquifers and other water bodies can be summarized visually.

4.12 SUMMARY

In summary, many factors control the chemistry of natural waters and contamination-affected waters. Evaporation effectively produces distilled water with very low TDS and solution of CO_2 into atmospheric water results in naturally acidic precipitation. These factors help to produce water that has the capacity to

dissolve minerals and organic matter at or below the surface. The composition of bedrock and natural communities combined with climate strongly influence the natural composition of surface water and groundwater. In addition to pH (and the important role of carbonate chemistry), reduction–oxidation reactions impart a strong control on mobility and stability of inorganic ions, organic compounds and minerals; for example, oxidizing waters foster breakdown of sulfide minerals and organic compounds, whereas reducing

environments foster breakdown of iron hydroxides. Sorption is strongly influenced by the mineralogy and organic matter content of the aqueous environment, so understanding the interrelationship of these factors allows us to predict and understand the fate and transport of natural compounds and contaminants in the environment.

Appendix I presents a case study that provides the opportunity to examine integrated concepts across aqueous geochemistry. Also, Appendix II presents information on instrumental analysis in environmental geochemistry, including approaches used to analyze the composition of water.

QUESTIONS

1. Convert values listed for the arid river in Table 4.1 to mmol/L and cmol_c/L (mmol of charge/L). Is this water body electrically neutral (i.e. do + and – charges balance?). Should it be electrically neutral? Determine the ionic strength of this water body and also determine TDS (in mg/L), and classify the water based on values in Table 4.4. Plot this water body on the graph in Fig. 4.4 – does it fall into the classification range of arid surface water bodies?
2. Predict the relative solubilities of Na⁺, K⁺, Mg⁺², Ca⁺², Co⁺², Fe⁺², Fe⁺³ and Al⁺³ considering ionic charge and ionic radius, then compare your findings to the graph in Fig. 4.9.
3. Provide a chemical reaction that shows the change in speciation of Al from pH = 4 to pH = 7.
4. Use the Debye–Huckel equation (4.10 or 4.11) to determine the ion activity coefficients (20 °C) for K⁺ and Al⁺³ at ionic strengths of 10^{–2} and 10^{–4}. Why is this type of calculation important? List two examples of non-ionic aqueous complexes that might form involving K⁺ and Al⁺³ (two each). Are these types of complexes more common or less common when ion activity coefficients are low?
5. Predict concentrations of Pb⁺² and CO₃^{–2} in a solution that would be produced by dissolution of cerussite.
6. Is uranium more soluble in oxidizing or reducing water? What are examples of species of U in solution? How does pH and presence of carbonate species affect U solubility in water?
7. Use the Fe and Al Eh–pH diagrams to explain controls on solubility and speciation of those two elements in aqueous systems.
8. Use the Law of Mass Action to predict relative abundances of Ca⁺² and Al⁺³ in exchange sites for two waters, one of ~ high ionic strength (10^{–1} M) and one more dilute (10^{–3} M). List pertinent assumptions and simplifications.
9. Compare linear vs. Langmuir vs. Freundlich adsorption models, including descriptions of their assumptions, graphical examples of adsorption isotherms, applicable equations and general applicability, i.e. which do you think best models adsorption in aqueous systems?

REFERENCES

- Andrews, J.E., Brimblecombe, P., Jickells, T.D., Liss, P. S., and Reid, B.J., 2004. *An Introduction to Environmental Chemistry*, 2nd edn, Blackwell Publishing, Oxford.
- Back W., 1966. Hydrochemical facies and ground-water flow patterns in northern part of Atlantic Coastal Plain. US Geol. Surv. Professional Paper 498-A. US Government Printing Office, Washington, DC.
- Baas Becking, L.G.M., Kaplan, I.R., and Moore, D., 1960. Limits of the natural environment in terms of pH and oxidation-reduction potentials. *Journal of Geology* 68: 243–284.
- Berner, E.K. and Berner, R.A., 1996. *Global Environment: Water, Air and Geochemical Cycles*. Prentice Hall, Upper Saddle River, New Jersey, USA.
- Drever, J.I. 1973. The preparation of oriented clay mineral specimens for X-ray diffraction analysis by a filter membrane peel technique. *American Mineralogist* 58: 553–554.

- Drever, J.I., 1997. *The Geochemistry of Natural Waters: Surface and Groundwater Environments*, 3rd edn. Simon and Schuster, Upper Saddle River, New Jersey, USA.
- Dzombak, D.A. and Morel, F.M.M., 1990. *Surface Complexation Modeling: Hydrated Ferric Oxide*, Wiley-Interscience, New York, 393 pp.
- Gibbs, R. J. 1970. Mechanisms controlling world water chemistry. *Science* 170: 1088–1090.
- Gorrell, H.A., 1953. Classification of formation waters based on sodium chloride content. *American Association of Petroleum Geologists Bulletin* 42: 2513.
- Hsi, C.D., Langmuir, D., 1985. Adsorption of uranyl onto ferric oxyhydroxides; application of the surface complexation site-binding model, *Geochimica Cosmochimica Acta*, 49(9): 1931–1941.
- Holland, H.D., 1978. *The Chemistry of the Atmosphere and Oceans*. Wiley, New York.
- Livingstone, D.A., 1963. Chemical composition of rivers and lakes, in *Data of Geochemistry*, United States Geological Survey Professional Paper: 440-G, U.S. Government Printing Office, Washington, D. C. online: <http://pubs.er.usgs.gov/publication/pp440G>.
- Martin, J. M. and Whitfield, M., 1983. The significance of the river input of chemical elements to the ocean. p. 265–296. In *Trace Metals in Seawater*, ed. by C. S. Wong, E. Boyle, K. W. Bruland, J. D. Burton and E. D. Goldberg, Plenum, New York.
- Mayo, A.L. and Loucks, D.L., 1995. Solute and isotopic geochemistry and ground water flow in the central Wasatch Range, Utah. *Journal of Hydrology* 172: 31–59.
- Meybeck, M., 1979. Concentration des eaux fluviales en éléments majeurs et apports en solution aux océans. *Review de Géologie Dynamique et de Géographie Physique* 21: 215–246.
- Volkov, A.G., Paula, S., and Deamer, D.W., 1997. Two mechanisms of permeation of small neutral molecules and hydrated ions across phospholipid bilayers. *Bioelectrochemistry and Bioenergetics* 42: 153–160.
- Winter, T.C., Harvey, J.W., Franke, O.L., and Alley, W. M., 1998. *Ground Water and Surface Water, a Single Resource: the Hydrologic Cycle and Interactions of Ground Water and Surface Water*. USGS Circular 1139, 79 p.

Carbonate Geochemistry and the Carbon Cycle

The high abundance, solubility and reactivity of the carbonate species H_2CO_3 – HCO_3^- – CO_3^{2-} in air, surface water, soils and groundwater means that this system exerts a very strong control on aqueous chemistry. Solution of CO_2 in H_2O creates H_2CO_3 and reactions among H_2CO_3 , HCO_3^- and CO_3^{2-} that absorb or release H^+ are important in controlling pH; furthermore, the capacity of carbonate species to form soluble complexes enhances the solubility of certain elements (e.g. U, as covered in Chapter 4). Cycling of carbon among the terrestrial biosphere, the atmosphere, the oceans (as biotic C and inorganic C) and the crust (carbonate rocks, petroleum and coal) has profound impacts on climate, marine resources and energy sources. Accordingly, Chapter 5 focuses on essential topics in carbonate geochemistry and the carbon cycle, first focusing of carbonate geochemistry and its impact on aqueous chemistry, then in the second half focusing on the cycling of carbon among reservoirs in the atmosphere, hydrosphere and lithosphere.

5.1 CARBONATE GEOCHEMISTRY: INORGANIC CARBON IN THE ATMOSPHERE AND HYDROSPHERE

In the atmosphere, CO_2 dissolves in water to produce carbonic acid, a relatively weak acid that is responsible for many important processes at/near the Earth surface. It is the most abundant acid in natural waters, and it:

1 controls the pH of natural rain (at pH = 5.6 in equilibrium with CO_2) as well as many other natural waters via the action of carbonic acid plus the bicarbonate and carbonate anions; and

2 is the driving force behind the chemical weathering reactions that (A) release dissolved ions into streams and soils and (B) also produce the minerals that form the solid component of soils.

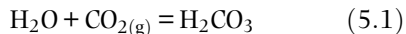
Here is an example of the crucial role that carbonate can exert on water chemistry – the chemistry of aquifers or soils with only 1% calcite will be dominated by a “carbonate signature” due to the high rate of calcite dissolution and the

production of dissolved bicarbonate (Langmuir, 1997). These waters will be close to neutral pH and have high buffering capacity. Because the bicarbonate anion (HCO_3^-) can lose H^+ to buffer basic solutions and absorb H^+ to buffer acidic solutions, it is responsible for buffering most natural aqueous systems, and it is the main contributor to alkalinity in aqueous systems.

5.1.1 Atmospheric CO_2 , carbonate species and the pH of rain

CO_2 gas will dissolve into water (e.g. droplets in air, surface water or groundwater) until equilibrium is reached. Most of the $\text{CO}_{2(\text{aq})}$ that diffuses into water remains as discrete CO_2 molecules dissolved in water, but by convention, all dissolved CO_2 in water, whether it is discrete $\text{CO}_{2(\text{aq})}$ or whether it is truly carbonic acid (H_2CO_3), is typically referred to as H_2CO_3 .

In the absence of nitrogen oxides (NO_x) and sulfur oxides (whether anthropogenic or locally derived from volcanoes and other isolated natural sources), the pH of rainfall is predominantly controlled by the reaction of water with carbon dioxide. This is because when molecules of CO_2 interact with water (e.g. droplets in the atmosphere), the CO_2 will dissolve into the water to a small extent, and the extent to which CO_2 will combine with water to form carbonic acid is controlled by the equilibrium established between the two components H_2O and CO_2 .



If what we really seek to know is the pH of natural rain controlled by equilibrium with CO_2 , we must know the following parameters:

1 The amount of CO_2 in the atmosphere – this is because the greater the amount of CO_2 in air, the greater the extent of contact between carbon dioxide and water to produce carbonic acid

(to take an extreme example, if these molecules never interact, carbonic acid will not form).

2 The equilibrium constant for the reaction of CO_2 and H_2O to H_2CO_3 – the mere presence of CO_2 and H_2O does not guarantee that they will react to form H_2CO_3 , so we need a term that quantifies the extent to which CO_2 will dissolve in H_2O (i.e. the probability that these two components will react with each other to form H_2CO_3 given a certain concentration of CO_2 and assuming abundant H_2O).

3 The extent to which carbonic acid will dissociate into H^+ and HCO_3^- (bicarbonate) once CO_2 has dissolved in H_2O . This term will be the K_a for carbonic acid.

For the calculations to follow we will assume a CO_2 concentration of 400 ppm, the approximate atmospheric concentration in the years 2013 and 2014. 400 ppm is equal to 400×10^{-6} atm.

First, the equilibrium constant for the reaction of CO_2 and H_2O to H_2CO_3 is:

$$K_{\text{CO}_2} = [\text{H}_2\text{CO}_3]/[\text{H}_2\text{O}] * p_{\text{CO}_2} \quad (5.2)$$

Given that the concentration (or activity) of water is assumed to be 1, the equation simplifies to:

$$K_{\text{CO}_2} = [\text{H}_2\text{CO}_3]/p_{\text{CO}_2} \quad (5.3)$$

Rearranging to solve for the concentration of carbonic acid in water droplets in the atmosphere:

$$[\text{H}_2\text{CO}_3] = p_{\text{CO}_2} * K_{\text{CO}_2} \quad (5.4)$$

This should intuitively make sense because what this equation is saying is that the amount of carbonic acid in atmospheric water is controlled by the amount of CO_2 in air (p_{CO_2}) and by the probability that CO_2 and H_2O will react to form H_2CO_3 (K_{CO_2}).

The partial pressure of CO_2 at 400 ppm is 400×10^{-6} atm (or $10^{-3.4}$) and the K_{CO_2} at 20°C is $10^{-1.41}$ (or 0.0389). The K_{CO_2} term is the Henry's gas constant, and the value of this

constant is temperature dependent. At 25 °C, the value is $10^{-1.47}$, whereas at 0 °C the K_{CO_2} is $10^{-1.11}$ and at 60 °C it is $10^{-1.78}$ (Plummer and Busenberg, 1982). The variability in the K_{CO_2} reflects the fact that the solubility of CO_2 in a liquid increases with decreasing temperature, a phenomenon common to all gases.

Thus, $[\text{H}_2\text{CO}_3] = 400 \times 10^{-6} \times 3.89 \times 10^{-2} = 1.56 \times 10^{-5}$ (or $10^{-4.8}$).

Carbonic acid is a weak acid that only dissociates to a small extent, and its dissociation:



is represented by this equilibrium constant:

$$K_1 = [\text{H}^+] * [\text{HCO}_3^-] / [\text{H}_2\text{CO}_3] \quad (5.6)$$

and rearranging gives us this equation:

$$[\text{H}^+] * [\text{HCO}_3^-] = K_1 * [\text{H}_2\text{CO}_3] \quad (5.7)$$

Given that dissociation of H_2CO_3 produces equal amounts of H^+ and HCO_3^- , the left side of the equation above is mathematically identical to $[\text{H}^+]^2$.

$$[\text{H}^+] * [\text{HCO}_3^-] = [\text{H}^+]^2 \quad (5.8)$$

Now, in order to determine the pH of atmospheric water we first need the concentration of hydrogen ions $[\text{H}^+]$, so we can write the equation containing the K_1 term as follows:

$$[\text{H}^+]^2 = K_1 * [\text{H}_2\text{CO}_3] \quad (5.9)$$

The K_1 is a constant value of $10^{-6.38}$ (or 4.17×10^{-7}) at 20 °C – for comparison, the value at 10 °C is $10^{-6.46}$ and at 30 °C it is $10^{-6.33}$. We determined above that $[\text{H}_2\text{CO}_3]$ in equilibrium with 400 ppm $\text{CO}_2 = 1.56 \times 10^{-5}$, so we can now determine the concentration of H^+ squared as follows:

$$\begin{aligned} [\text{H}^+]^2 &= (4.17 \times 10^{-7}) \times (1.56 \times 10^{-5}) \\ &= 6.49 \times 10^{-12} \end{aligned}$$

Taking the square root of both sides we find that

$$[\text{H}^+] = 2.55 \times 10^{-6} = 10^{-5.59}$$

Also, $\text{pH} = -\log [\text{H}^+] = -\log [10^{-5.59}] = 5.59$, which is the predicted pH of rain water in equilibrium with CO_2 at 20 °C at 400 ppm CO_2 . By this point you should be able to repeat this calculation under different scenarios, including varied temperature and p_{CO_2} . If, for example, you are curious how projected CO_2 concentrations of 500, 750 or 1000 ppm might affect the pH of rain, you now have the tools to perform such calculations.

5.1.2 Speciation in the carbonate system as a function of pH

In order to assess carbonate speciation as a function of pH in a more quantitative manner (i.e. which of the following species is the dominant dissolved form of carbonate at a given pH: H_2CO_3 , HCO_3^- or CO_3^{2-} ?), we can re-arrange the equilibrium constants as follows. First, we will work with K_1 , and assuming an aqueous solution that is not exchanging CO_2 with the atmosphere, we can determine the pH at which H_2CO_3 and HCO_3^- will exist in equal concentrations. At 25 °C,

$$K_1 = [\text{H}^+] * [\text{HCO}_3^-] / [\text{H}_2\text{CO}_3] = 10^{-6.35} \quad (5.10)$$

(at 20 °C, $K_1 = 10^{-6.38}$). This low K value is typical of weak acids that only partially dissociate in solution, especially when compared to the K values for the strong acids HCl and H_2SO_4 , each of which are approximately 10^3 , indicating nearly complete dissociation. In the case of strong acids like these, the product of the concentrations of H^+ and the conjugate base is thousands times greater than the concentration of undissociated acid (i.e. HCl , H_2SO_4).

The above equation can be re-arranged as follows:

$$[\text{H}^+]/10^{-6.35} = [\text{H}_2\text{CO}_3]/[\text{HCO}_3^-] \quad (5.11)$$

and when the concentrations of H_2CO_3 and HCO_3^- are equal, the right side of the equation will equal 1:

$$[\text{H}^+]/10^{-6.35} = 1$$

(this is true when $[\text{H}_2\text{CO}_3] = [\text{HCO}_3^-]$)

Thus, in the case where $[\text{H}_2\text{CO}_3] = [\text{HCO}_3^-]$, the concentration of hydrogen ions $[\text{H}^+]$ is $10^{-6.35}$, and $\text{pH} = 6.35$. When $\text{pH} < 6.35$, H_2CO_3 is the dominant species of carbonate in solution, and at $\text{pH} > 6.35$, HCO_3^- is the dominant dissolved species.

However, as pH rises above 6.35 towards more basic conditions, HCO_3^- is likely to lose a hydrogen ion to form CO_3^{2-} , and we can also determine the pH at which $[\text{HCO}_3^-] = [\text{CO}_3^{2-}]$:

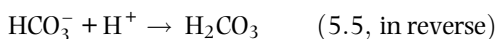
$$K_2 = [\text{H}^+] * [\text{CO}_3^{2-}]/[\text{HCO}_3^-] = 10^{-10.33} \quad (5.12)$$

Also, when $[\text{CO}_3^{2-}] = [\text{HCO}_3^-]$:

$$[\text{CO}_3^{2-}]/[\text{HCO}_3^-] = 1 = 10^{-10.33}/[\text{H}^+] \quad (5.13)$$

In order to satisfy this equation, $[\text{H}^+] = 10^{-10.33}$, and the pH at which aqueous bicarbonate is equal to aqueous carbonate ion is 10.33 (at 25 °C).

It is important to emphasize that, in the K_1 and K_2 equations, $[\text{H}^+]$ is a variable, and this means that the ratios of $\text{H}_2\text{CO}_3/\text{HCO}_3^-$ and $\text{HCO}_3^-/\text{CO}_3^{2-}$ will be controlled by pH . So in a qualitative sense, H_2CO_3 will dominate under acidic conditions because the high concentrations of H^+ will drive the system to produce carbonic acid from bicarbonate.



In basic solutions, bicarbonate will release H^+ to neutralize OH^- , in the process producing the carbonate anion and water.

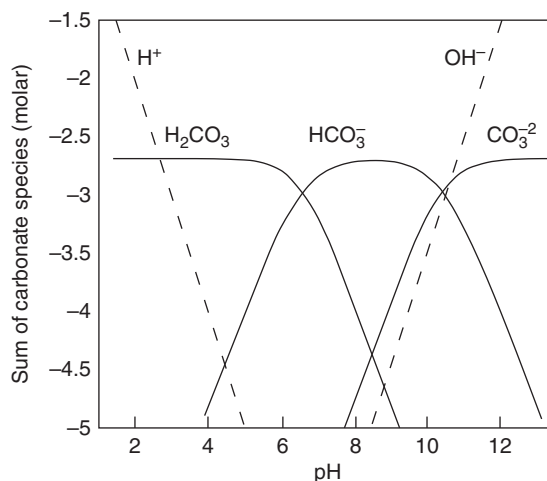
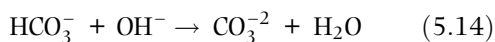


Fig. 5.1 Carbonate species as a function of pH (at 25 °C). Concentrations of H^+ and OH^- are also shown. This type of figure is known as a Bjerrum plot.



You can begin to see the manner by which bicarbonate tends to buffer solutions.

To summarize, for the system $\text{H}_2\text{CO}_3\text{--HCO}_3^-\text{--CO}_3^{2-}$, bicarbonate is the dominant species in solution between $\text{pH} = 6.4$ and $\text{pH} = 10.3$. Below a pH of approximately 6.4, carbonic acid is the dominant carbonate species in solution, and the carbonate anion dominates above $\text{pH} = 10.3$. When different values of $[\text{H}^+]$ are substituted into the above equations, the ratios of $[\text{H}_2\text{CO}_3] : [\text{HCO}_3^-]$ and $[\text{HCO}_3^-] : [\text{CO}_3^{2-}]$ can be determined over a range of pH values. These relationships are displayed graphically in Fig. 5.1.

5.1.3 Alkalinity

Alkalinity is the measure of the ability of a water body to neutralize acids, and in most waters, alkalinity is determined by the amount of dissolved bicarbonate and carbonate anion in solution as depicted below (where concentrations are represented as molar values):

$$\text{Carbonate alkalinity} = [\text{HCO}_3^-] + 2 * [\text{CO}_3^{2-}] \quad (5.15)$$

The molar (or millimolar) concentration of the carbonate anion is multiplied by 2, which means that alkalinity is presented as equivalents per liter (typically meq/L or $\mu\text{eq/L}$). The reason for the doubling is that the carbonate anion can neutralize twice as much H^+ as can the bicarbonate anion. As suggested above, in the pH range of 6.4 to 10.3, alkalinity will be dominated by bicarbonate, and above pH = 10.3, alkalinity will be dominated by the carbonate anion.

While technically the above expression is the *carbonate alkalinity*, only in rare cases do ions other than HCO_3^- or CO_3^{2-} contribute appreciably to alkalinity. Examples of other dissolved species that can contribute to alkalinity include dissolved silica (as H_3SiO_4^-), dissolved boron (as $\text{B}(\text{OH})_4^-$), the hydroxyl anion (OH^-) and various organic

anions, but typically the concentrations of HCO_3^- or CO_3^{2-} are far higher than these other species, especially in fresh waters.

Because alkalinity reflects the ability of a water body to buffer acids, it plays a crucial role in determining how aquatic ecosystems will respond to acidification from acid rain, acid mine drainage and other acidic wastes. Regions that lack carbonate rocks or carbonate minerals in soils and are exposed to acid rain are far more prone to soil and lake acidification than are regions containing carbonates (Fig. 5.2).

Given that alkalinity is effectively a measure of the buffering capacity of an aqueous solution, alkalinity is typically measured by titration using an acid like HCl, where HCl is added until the pH drops to 4 – greater amounts of bicarbonate and carbonate anions will require greater amounts of acid to lower the solution to pH = 4. At a pH of 4, effectively all HCO_3^- and CO_3^{2-} will have been converted to H_2CO_3 (or $\text{CO}_{2(\text{aq})}$) (Fig. 5.3).



Fig. 5.2 Areas shaded in light and dark gray indicate ecosystems in the continental United States that are prone to soil and lake acidification. The sensitive areas are underlain by granites or carbonate-poor rocks and soils with low buffering capacity (in response to acid deposition). Coastal waters and estuaries at risk from N and S emissions associated with acid deposition are also shown. (US EPA 2005.)

Alkalinity is commonly determined by titration, starting with a filtered water sample of a precisely measured water volume (e.g. 100 mL) and an acid of known concentration (as a *titrant*). Common titrants are H_2SO_4 and HCl at strengths of 0.02 to 1 N; for reference, 0.1600 N H_2SO_4 is a titrant commonly used by the US Geological Survey to determine alkalinity in natural waters with normal alkalinity (e.g. 50 to 200 mg CaCO_3/L).

1 Determine initial pH of the water sample and record the value.

2 Slowly add small amount of titrant (e.g. 0.5 mL if using 100 mL of water) and allow to equilibrate by gentle, continuous stirring (a magnetic stirrer is commonly used). It is better to add titrant too slowly or in aliquots that are too small rather than too rapid or too voluminous. Record the pH and volume of titrant added frequently during this process e.g. every 0.5 mL of titrant.

3 Continue to slowly add titrant and allow the sample to equilibrate at each step until a pH of 4.0 is reached. Be sure to continually record data. One way to do this is to set up a spreadsheet with volume of titrant added in one column and pH in another column.

4 As the solution approaches pH = 4.0, be sure to add progressively smaller volumes of titrant and verify that pH reaches a steady, non-fluctuating value (i.e. equilibrium) before adding the next aliquot of titrant. Avoid overshooting this pH range – pH will rapidly drop between pH = 5 and 3 as all remaining HCO_3^- is converted to H_2CO_3 .

5 Prepare a titration curve that plots volume of titrant added (in mL) on the *x*-axis and pH on the *y*-axis, or vice versa. For neutral to slightly alkaline waters the curve will typically have the form shown in Fig. 5.3.

6 Determine the values of pH that correspond to the inflection points. The first inflection should occur at a pH of 8.3 which, for most terrestrial waters (streams, lakes, groundwater), corresponds to conversion of CO_3^{2-} to HCO_3^- . This

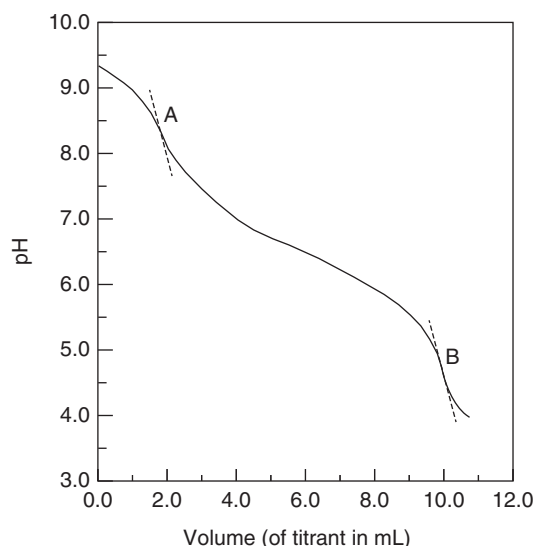


Fig. 5.3 Representative example of a titration curve. The steep slope at point A corresponds to conversion of CO_3^{2-} to HCO_3^- , and the steep slope at point B corresponds to conversion of HCO_3^- to H_2CO_3 . At point B, all carbonate alkalinity has been converted to H_2CO_3 .

inflection point is sometimes referred to as “P alkalinity”. The second point, which will occur around a pH of 4.3, corresponds to the conversion of HCO_3^- to H_2CO_3 . This inflection point is sometimes referred to as “T alkalinity”.

7 Determine the volume of titrant (in mL) consumed in order to reach the P and T alkalinities, respectively.

Alkalinity of samples with an alkalinity > 20 mg/L (i.e. most natural waters) is calculated using the following formula, where alkalinity is expressed in terms of meq/L:

$$\text{Alkalinity} = (A_{\text{mL}} * A_{\text{N}} * 1000 \text{ mL/L}) / \text{H}_2\text{O}_{\text{mL}}$$

- A_{mL} = volume of acid added
- A_{N} = normality of acid used as titrant
- $\text{H}_2\text{O}_{\text{mL}}$ = initial volume of water to be titrated.
- 1000 ml = conversion from mL to L

To report results in units of mg CaCO_3/L , convert the meq/L result as follows:

$$\frac{x \text{ meq/L} * 1 \text{ mmol/2 meq} * 100.1 \text{ mg/mmol}}{(\text{CaCO}_3)} \quad (5.16)$$

The rationale for this conversion is that Ca^{+2} contains 2 meq of charge per millimole and the mass of 1 mmol of CaCO_3 is 100.1 mg.

A concept somewhat related to carbonate alkalinity is dissolved inorganic carbon (DIC), which is the sum of all dissolved carbonate species:

$$\text{DIC} = [\text{H}_2\text{CO}_3] + [\text{HCO}_3^-] + [\text{CO}_3^{2-}] \quad (5.17)$$

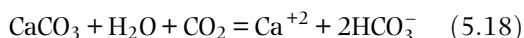
Given that CO_2 in solution is effectively synonymous with H_2CO_3 , the above equation is sometimes written with $\text{CO}_{2(\text{aq})}$ instead of H_2CO_3 . The designation DIC differentiates inorganic carbon from dissolved organic carbon (DOC) that is typically comprised of soluble fractions of humic and fulvic acids and potentially also soluble hydrocarbons such as benzene and toluene.

5.1.4 Carbonate solubility and saturation

At this point we will apply our understanding of carbonate species and calcite dissolution to examine the degree of saturation of calcium and bicarbonate in a carbonate aquifer with respect to calcite as the mineral controlling equilibrium concentrations of Ca^{+2} and HCO_3^- .

For carbonate groundwater with $\text{pH} = 7.5$ (a typical value for a carbonate aquifer), we first recognize that the concentration of bicarbonate is orders of magnitude greater than carbonic acid and the carbonate anion (Fig. 5.1); > 99% of carbonate in this system will be HCO_3^- . At this pH, we can ignore the presence of H_2CO_3 and CO_3^{2-} and this simplifies the calculation.

The dissolution of calcite can thus be represented as:



This reaction can be quantified by the following equilibrium constant (note that the concentration of bicarbonate is squared to reflect its molar abundance):

$$K_{\text{eq}} = [\text{Ca}^{+2}] * [\text{HCO}_3^-]^2 / [\text{CaCO}_3] * [\text{H}_2\text{O}] * [\text{CO}_2] \quad (5.19)$$

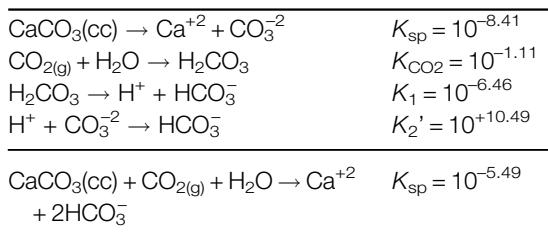
The K_{eq} for this reaction is not typically listed in tables of thermodynamic data, so we need to combine K_{eq} values of reactions listed in Table 5.1. The solid crystal of calcite in equilibrium with dissolved species cannot be represented in terms of concentration, so by convention its concentration is given the value of 1. This is because, in reality, as long as there is solid calcite, the amount of it will not affect the equilibrium between Ca^{+2} and HCO_3^- ions and the reactants carbonic acid (or $\text{H}_2\text{O} + \text{CO}_2$) and solid calcite – it is the constituents dissolved in water that are most important, at equilibrium. In dilute solutions (e.g. fresh water, groundwater), $[\text{H}_2\text{O}]$ also is assigned a value of 1, and the concentration of CO_2 is represented by the partial pressure of CO_2 (P_{CO_2}), so the equation for the equilibrium constant is now:

$$K_{\text{eq}} = [\text{Ca}^{+2}] * [\text{HCO}_3^-]^2 / P_{\text{CO}_2} \quad (5.20)$$

Table 5.1 Equilibrium reactions and constants for the carbonate system in fresh water at 10 °C (K_{10}) and 25 °C (K_{25}). Note that (cc) refers to calcite (rather than CaCO_3 polymorphs aragonite or vaterite).

Reaction	K_{10}	K_{25}
$\text{CO}_{2(\text{g})} + \text{H}_2\text{O} \rightarrow \text{H}_2\text{CO}_3$	$K_{\text{CO}_2} = 10^{-1.11}$	$K_{\text{CO}_2} = 10^{-1.47}$
$\text{H}_2\text{CO}_3 \rightarrow \text{H}^+ + \text{HCO}_3^-$	$K_1 = 10^{-6.46}$	$K_1 = 10^{-6.35}$
$\text{HCO}_3^- \rightarrow \text{H}^+ + \text{CO}_3^{2-}$	$K_2 = 10^{-10.49}$	$K_2 = 10^{-10.33}$
$\text{CaCO}_3 (\text{cc}) \rightarrow \text{Ca}^{+2} + \text{CO}_3^{2-}$	$K_{\text{sp}} = 10^{-8.41}$	$K_{\text{sp}} = 10^{-8.48}$
$\text{H}_2\text{O} \rightarrow \text{OH}^- + \text{H}^+$	$K_{\text{w}} = 10^{-14.53}$	$K_{\text{w}} = 10^{-14.00}$

The reactions in Table 5.1 for a system at 10 °C (K_{10}) are suitable for cool temperature climates where groundwater is typically in the range of 9–12 °C. The net sum of these four reactions produces the equation for calcite dissolution in the presence of carbonic acid.



Combining these equations occurs as follows

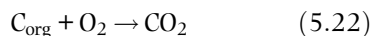
$$K_{\text{eq}} = K_{\text{CO}_2} * K_1 * K_2' * K_{\text{sp}} = 10^{-1.41} * 10^{-6.46} * 10^{+10.49} * 10^{-8.41} = 10^{-5.49}$$

This results in equation 5.21:

$$10^{-5.49} = K_{\text{eq}} = [\text{Ca}^{+2}] * [\text{HCO}_3^-]^2 / P_{\text{CO}_2} \quad (5.21)$$

Note that because the reaction for the dissociation of bicarbonate needs to be expressed in the reverse direction of what is typically presented in tables of thermodynamic data (e.g. Table 5.1), its value in the calculation is the inverse of what is presented in Table 5.1.

Groundwater typically contains dissolved CO_2 at levels greater than what would be predicted by equilibrium with the atmosphere. This is because, as surface water percolates downward toward the saturated zone, aerobic decomposition of organic matter produces dissolved CO_2 that drives partial pressures of CO_2 in groundwater up into the range of $\sim 10^{-2}$.



Note that in the absence of CO_2 produced by decomposition of soil organic matter, natural

waters in equilibrium with the atmosphere correspond to a $P_{\text{CO}_2} = 10^{-3.4}$ (using an atmospheric CO_2 concentration of 400 ppm).

In a groundwater setting where organic matter decay (equation 5.22) has produced $P_{\text{CO}_2} = 10^{-2}$, we can express the equilibrium constant as:

$$K_{\text{cc}} = [\text{Ca}^{+2}] * [\text{HCO}_3^-]^2 = 10^{-7.49} \quad (5.23)$$

Assuming that activities (or concentrations) of calcium and bicarbonate are controlled by calcite dissolution, i.e. $2 * [\text{Ca}^{+2}] = [\text{HCO}_3^-]$, and also assuming that in these relatively dilute fresh waters the activities are $\sim$ equal to concentrations: (see companion website)

$$[\text{Ca}^{+2}] = (10^{-7.49} / 4)^{0.33} = 2.13 * 10^{-3} \text{ mol/L}$$

Then, given that calcite decomposition in the presence of carbonic acid yields two moles of bicarbonate per mole of calcium,

$$2 * [\text{Ca}^{+2}] = [\text{HCO}_3^-] = 4.26 * 10^{-3} \text{ mol/L}.$$

Converting molar values to mg/L, $[\text{Ca}^{+2}] = 2.13 * 10^{-3} \text{ mol/L} * 40.08 \text{ g/mol} * 1000 \text{ mg/g} = 85.2 \text{ mg/L}$. By the same approach, $[\text{HCO}_3^-] = 260 \text{ mg/L}$. In other words, at 10 °C with $P_{\text{CO}_2} = 10^{-2} \text{ atm}$, groundwater with $[\text{Ca}^{+2}] = 85.2 \text{ ppm}$ and $[\text{HCO}_3^-] = 260 \text{ ppm}$ is saturated.

Recall that if the ion activity product is greater than the K_{sp} ($\Omega > 1$), the solution is oversaturated and calcite will precipitate from solution. If the $\Omega < 1$ (the $\text{IAP} < K_{\text{sp}}$), the solution is undersaturated and calcite will dissolve. If $\Omega = 1$ ($\text{IAP} = K_{\text{sp}}$), the system is saturated and at equilibrium, and rates of mineral dissolution and mineral crystallization will be equal.

So for a groundwater system at 10 °C that contains 46.8 mg/L Ca^{+2} ($1.17 * 10^{-3} \text{ mol/L}$), 205.0 mg/L HCO_3^- ($3.36 * 10^{-3} \text{ mol/L}$) and $P_{\text{CO}_2} = 10^{-2} \text{ atm}$, the $\text{IAP} = 1.17 * 10^{-3} \text{ mol/L} * (1.68 * 10^{-3} \text{ mol/L})^2 = 3.30 * 10^{-9}$ (or $10^{-8.48}$). The value of Ω is:

$$\Omega = 10^{-8.48} / 10^{-7.49} = 0.10.$$

The fact that $\Omega < 1$ indicates that these groundwaters are undersaturated with respect to calcite. Under these conditions, calcite will dissolve, and if, as is the case in some limestone aquifers, the calcite contains uranium, we would predict that uranium will also be released to solution.

However, a system at 10 °C and pH = 7.2 containing $[\text{Ca}^{+2}] = 123 \text{ mg/L} = 3.07 \times 10^{-3} \text{ mol/L}$ and $[\text{HCO}_3^-] = 0.5 \times 432 \text{ mg/L} = 3.54 \times 10^{-3} \text{ mol/L}$ will be oversaturated with respect to calcite because the $\text{IAP} = 3.85 \times 10^{-8} \text{ mol}^2/\text{L}^2 = 10^{-7.41}$; in this case,

$$\Omega = 10^{-7.41} / 10^{-7.49} = 1.20$$

Given that the IAP is greater than the equilibrium constant, this solution is slightly oversaturated and we predict that calcium and bicarbonate ions will precipitate out of solution to form calcite.

5.1.5 The effect of CO₂ partial pressure on stability of carbonate minerals

From the K_{CC} reaction derived above (and shown again below), it is clear that the amount of carbon dioxide dissolved in water will exert some control over the solubility (or stability) of calcite (and other carbonate minerals too):

$$K_{\text{CC}} = [\text{Ca}^{+2}] * [\text{HCO}_3^-]^2 / P_{\text{CO}_2} = 10^{-5.49} \quad (5.24)$$

Qualitatively, it is evident that increasing P_{CO_2} will cause an increase in the concentrations of dissolved calcium and bicarbonate, meaning that increasing CO₂ causes decreased stability of calcite.

As we saw above, at $P_{\text{CO}_2} = 10^{-2}$, the value of K_{CC} (at 10 °C) = $10^{-7.49}$. If the CO₂ content of surface water or groundwater is solely controlled by equilibrium with the atmosphere, $P_{\text{CO}_2} = 10^{-3.4}$ and the value of K_{CC} (at 10 °C) = $10^{-8.89}$.

When $P_{\text{CO}_2} = 10^{-2}$ atm (and the K_{CC} is $10^{-7.49}$), equilibrium concentrations of calcium and bicarbonate in equilibrium with calcite at 10 °C are:

$$[\text{Ca}^{+2}] = 3.19 \times 10^{-3} \text{ mol/L or } 128 \text{ mg/L}$$

$$[\text{HCO}_3^-] = 6.38 \times 10^{-3} \text{ mol/L or } 389 \text{ mg/L}$$

By comparison, when $P_{\text{CO}_2} = 10^{-3.4}$ atm (and the K_{CC} is $10^{-8.89}$), equilibrium concentrations of calcium and bicarbonate in equilibrium with calcite at 10 °C are much lower:

$$[\text{Ca}^{+2}] = 1.09 \times 10^{-3} \text{ mol/L or } 43.6 \text{ mg/L}$$

$$[\text{HCO}_3^-] = 2.18 \times 10^{-3} \text{ mol/L or } 133 \text{ mg/L}$$

These relatively low concentrations reflect greater stability of calcite at low dissolved CO₂ (e.g. $P_{\text{CO}_2} = 10^{-3.4}$ atm) compared to the lower stability of calcite at higher P_{CO_2} (e.g. 10^{-2} atm). Fig. 5.4 shows this graphically.

These types of calculations can help to understand and predict the effect of dissolved CO₂ on the stability of carbonate minerals in soils, surface waters and groundwater, and they also can help to provide insight into the effect that increasing atmospheric CO₂ will have on the stability of marine CaCO₃ (including skeletons and shells of organisms, polyp secretions in coral reefs, and other parts of the marine ecosystem). Based on these principles and the plot in Fig. 5.4, it is clear that increasing atmospheric CO₂ will diminish the stability of CaCO₃ in the marine environment, and this in turn is bound to have negative consequences for marine ecosystems.

5.1.6 The effect of mineral composition on stability of carbonate minerals

While it is clear that the often-interrelated variables pH, dissolved CO₂ and temperature are important controls on the stability of calcite (and therefore also on water chemistry), if we examine other carbonate minerals we see that

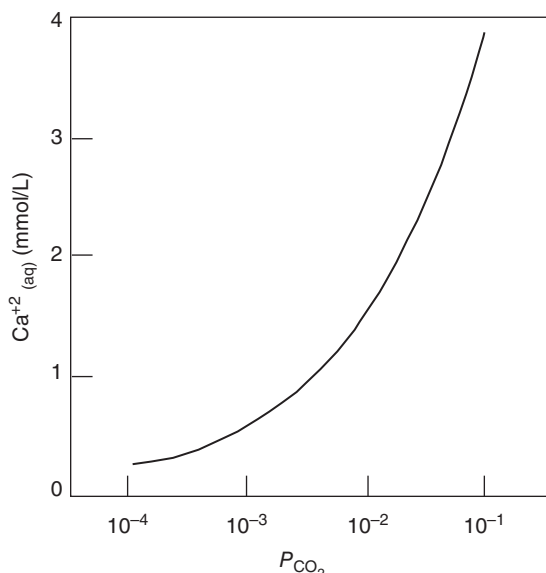


Fig. 5.4 Relationship between CO_2 dissolved in water and the stability of calcite (as indicated by the concentration of $\text{Ca}^{+2}_{(\text{aq})}$ in equilibrium with calcite). Increasing Ca^{+2} indicates decreasing calcite stability, so effectively what this graph illustrates is that increasing CO_2 enhances the solubility of calcite.

Table 5.2 Equilibrium constants of common carbonate minerals at 20 °C and 1 atm. (Morse et al. 2007 and sources cited therein).

Mineral	Formula	K_{sp}
Aragonite	CaCO_3	$10^{-8.30}$
Calcite	CaCO_3	$10^{-8.48}$
Dolomite	$\text{CaMg}(\text{CO}_3)_2$	$10^{-18.16}$
Magnesite	MgCO_3	$10^{-8.05}$
Rhodochrosite	MnCO_3	$10^{-10.09}$
Siderite	FeCO_3	$10^{-10.53}$

carbonate solubility is also controlled by mineral composition and, in the case of the CaCO_3 polymorphs aragonite and calcite, by mineral structure. For example, aragonite is less stable than calcite at standard temperature and pressure (Table 5.2); the least stable of the carbonate minerals presented in Table 5.2 is magnesite, and the most stable by far is dolomite. This

difference is very evident in weathering patterns of rocks that contain layers of calcite and dolomite, where rapidly weathering calcite is recessed and layers of dolomite form positive relief reflecting dolomite's low chemical weathering rate relative to calcite (Plate 7).

5.2 THE CARBON CYCLE

It is easy to argue that the carbon cycle is more important to understand than any other elemental cycle, at least in terms of environmental analysis. Carbon forms the basis of life, plays a major role in Earth's climate, drives chemical weathering and soil formation (as carbonic acid), and controls buffering of aquatic ecosystems (as the bicarbonate anion). Carbon occurs in organic and inorganic forms and is an important constituent in all reservoirs, including the lithosphere (as carbonate minerals and organic matter), atmosphere (mainly as CO_2 and to a lesser extent as CH_4 and CO), oceans (as dissolved CO_2 , bicarbonate, CaCO_3 in shells, and as organic carbon in plankton and other organisms), and terrestrial ecosystems (e.g. in living organisms, sediments and soils). For these reasons and more, the carbon cycle is relatively complex, yet understanding the carbon cycle is crucial to the practice of environmental geochemistry, climate analysis, ecosystem analysis and much more in the 21st century.

Basic principles related to biogeochemical cycling of elements are presented at the beginning of Chapter 6 (Section 6.1), so depending on your background it may be worthwhile reading that section before this section.

5.2.1 Oxidation states of carbon

Carbon is a redox-sensitive element, with oxidation states that range from C^{-4} to C^{+4} . Arranged in order from most-reduced to most-oxidized,

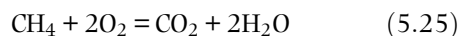
carbon occurs in the following oxidation states and compounds:

- The most-reduced form of carbon, C^{-4} , is the oxidation state of carbon in methane, CH_4 (recall that hydrogen is H^+ except in certain cases when it forms metal hydrides).
- Carbon atoms in octane (C_8H_{18}) are slightly less reduced, with an average oxidation state of $C^{-2.22}$ – this non-integer oxidation state reflects covalent bonds where carbon atoms can be thought of as a mixture of C^0 and C^{-4} , with C^{-4} slightly dominant over C^0 .
- An approximate molecular formula for humus is $C_{28}H_{26}O_{19}$ (Manahan 2000) and in this compound the average carbon atom is $C^{-0.4}$ (again, reflecting a mix of C^0 and C^{-4} , in this case dominated by C^0).
- In some types of organic matter, the average oxidation state of carbon atoms is zero. Consider glucose, $C_6H_{12}O_6$, where the charges on H^{+1} and O^{-2} balance each other out, leaving carbon atoms as C^0 . The carbon atoms in graphite and diamond, both forms of pure elemental carbon, also contain C^0 and the atoms are held together by covalent bonds (as well as van der Waals bonds in the case of graphite sheets);
- *The scenarios presented above all represent compounds with chemically reduced carbon.*
- C^{+2} occurs in carbon monoxide (CO), a compound that represents a more-oxidized form of carbon than those above, yet less-oxidized than the compounds presented below. In CO, the C^{+2} atom balances the O^{-2} atom.
- C^{+4} , the most-oxidized state of carbon, occurs in CO_2 as well as in carbonate minerals ($CaCO_3$, $MgCO_3$, etc.), the bicarbonate anion (HCO_3^-) and the carbonate anion (CO_3^{2-}).

Thus, reduced carbon including C^{-4} and C^0 and mixtures of these oxidation states occurs in organic matter and anaerobically decomposed organic matter, and oxidized carbon occurs in CO_2 and carbonates. Not surprisingly, organic matter and other forms of reduced carbon (inorganic CH_4 , for example) are stable in the anoxic to sub-oxic

environments of the Earth's crust, mantle, wetlands, marine sediments and other low- O_2 environments; conversely, carbon dioxide and carbonates are stable in O_2 -rich environments of oxidized seawater, the atmosphere, and oxidized soils and freshwater.

Given the variation in oxidation state of carbon atoms, much of the cycling of carbon is controlled by redox reactions, one of which can be summarized by the oxidation of methane (which in the atmosphere is more complicated, but this topic is covered in Chapters 7 and 8):



When viewed in reverse, this reaction would represent the reduction of carbon from C^{+4} to C^{-4} associated with formation of CH_4 . In addition to inorganic processes like oxidation (or combustion) of CH_4 , biological processes also play a significant role in carbon cycling, from photosynthesis and respiration to microbially mediated oxidation of soil organic matter to CO_2 .

The global carbon cycle (Fig. 5.5) includes all plants, animal and microorganisms, all decomposing plant remains, water bodies large (oceans) and small (lakes, ponds), soils (including various forms of organic and inorganic carbon), the crust and mantle of the Earth, products of volcanic eruptions, and more. In this section, carbon cycling is mainly presented through fluxes and reservoirs that exist on a global scale.

5.2.2 Global-scale reservoirs and fluxes of carbon

The unit of measure used to describe the mass of reservoirs at global scales is the petagram (10^{15} g). For those familiar with the unit gigaton (Gt), 1 Gt = 1 Pg. Some publications use units of Gt, others use Pg. For the purpose of this discussion we will refer to these massive quantities (10^{15} g) in terms of Pg. Also, unless otherwise indicated, values will

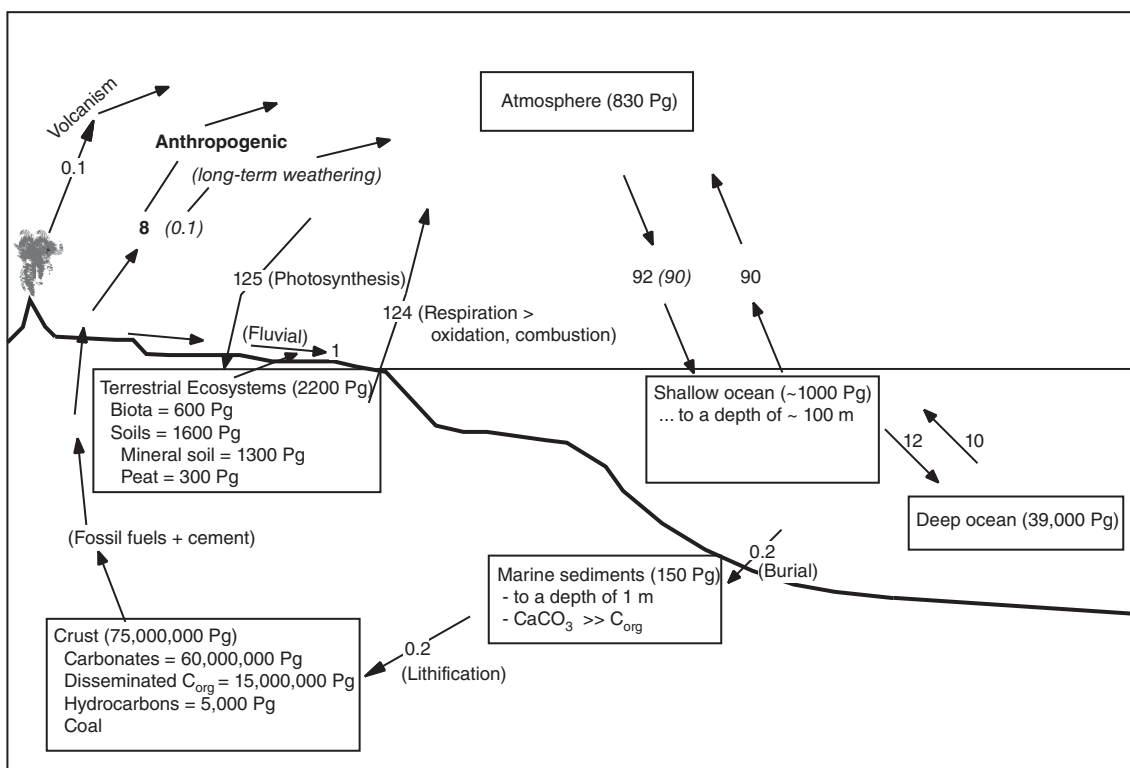


Fig. 5.5 Schematic representation of the carbon cycle. Note that anthropogenic C emissions are given as 8 Pg/yr, which is the rounded-off average for 2000–2010 (for reference, 2011 anthropogenic emissions were ~10 Pg). Sources are Sundquist and Visser (2004), Ruddiman (2007), IPCC (2007) [Solomon et al. (2007)], Freidlingstein et al. (2010), and other sources cited therein. Values in parentheses are pre-industrial estimates.

represent the mass of carbon rather than the mass of the molecule. So for a value of 1.5 Pg of carbon per year (1.5 Pg C/yr), the equivalent amount of carbon dioxide would be 5.5 Pg CO_2 /yr (where the conversion is 44 g of CO_2 per 12 g of C). Commonly, masses (or moles) of carbon are given rather than masses of molecules (e.g. CO_2 , CH_4 , C_6H_6 , etc.) to facilitate tracking of carbon.

Of the globally important C reservoirs (Table 5.3), the one with the greatest amount of carbon by far is the **mantle** of the Earth, with **340 000 000 Pg** of carbon (most of which likely occurs as graphite, diamonds, carbonate and possibly methane). Yet in spite of the enormous size of the mantle carbon reservoir, the flux between

the mantle and other reservoirs (crust, atmosphere, terrestrial ecosystems and oceans) is very slow – ~0.02 Pg/yr (Cartigny et al., 2008), a quantity that is effectively negligible for the conversation that follows – and so will not be considered in detail.

After the mantle, the next largest reservoir of carbon on Earth is the **crust**, with **75 000 000 Pg**. Of this amount, a relatively small quantity (5000 Pg) occurs as hydrocarbons, and the remainder, the vast majority, occurs either in carbonate minerals in rocks (e.g. limestone, dolostone) or as dispersed organic carbon (as kerogen, humin and other forms). The long-term flux of carbon to and from the crust at steady state is ~0.2 Pg/yr, and while this too is a low rate of exchange, it (A) is

Table 5.3 The abundance of carbon in major reservoirs globally.

Reservoir	Amount ($\times 10^{15}$ g C)
Crust	75 000 000
Carbonate rocks	60 000 000
Dispersed organic C	15 000 000
Hydrocarbons, coal	5000
Oceans	40 000
Soils	1600
Atmosphere	830
Land Plants	600

10 times the mantle flux, and (B) will serve as an important reference point in the discussion of carbon cycling, fossil fuels and climate.

The **oceans** contain **~40 000 Pg C**, the vast majority of which exists as dissolved bicarbonate in the deep ocean (**39 000 Pg C**). Lesser amounts are found as dissolved CO_2 , dissolved organic carbon and carbon in living organisms, mainly in the shallow ocean. The shallow ocean (depths < 100 m) is characterized by photosynthesis and greater mixing with the atmosphere; the deep ocean is more isolated from the atmosphere and at these depths carbon primarily occurs in inorganic forms (e.g. dissolved CO_2 , HCO_3^-). While the long-term steady-state flux between atmosphere and oceans is ~ 90 Pg/yr, this balance has shifted to a net flux into the oceans since the onset of the industrial revolution (Fig. 5.5).

Terrestrial ecosystems contain **2200 Pg C** in sub-reservoirs of biota (600 Pg C) and soils (1600 Pg C). The steady-state flux between terrestrial ecosystems and atmosphere is ~ 125 Pg/yr, but this value is strongly dependent on factors such as land use, soil temperature and other parameters that are very susceptible to short-term change.

The smallest of the global reservoirs is the **atmosphere** with **~830 Pg C**; however, the current flux of carbon from the crust of the Earth to the atmosphere (~ 8 Pg/yr) is larger than the flux of C from the atmosphere back into the crust

(~ 0.2 Pg/yr) or into marine and terrestrial reservoirs (combined = ~ 3 Pg/yr), an imbalance that is the main cause of increasing atmospheric CO_2 .

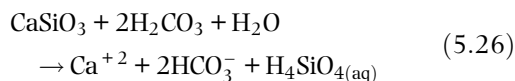
The following sections explore in detail the carbon cycle from the perspective of each of the main global-scale reservoirs.

5.2.3 Fixation of carbon into the crust

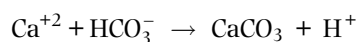
The main processes that fix carbon into the crustal reservoir occur over millions-of-years time spans and include the following two main pathways: (A) chemical weathering that facilitates transport of inorganic carbon to the ocean, where it can become buried and compacted in sediments (and form limestone, for example); and (B) photosynthesis followed by burial and compaction of organic carbon in marine sediments.

Carbonate rocks

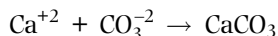
Chemical weathering of silicate rocks, using the weathering of wollastonite (CaSiO_3) as a simplified example, can be represented as follows:



Recall that the origin of carbonic acid (H_2CO_3) in the atmosphere is $\text{CO}_2 + \text{H}_2\text{O}$, so chemical weathering is one of the steps involved in the transfer of carbon from the atmospheric reservoir to the crustal reservoir. Another important step is the fluvial (i.e. by rivers) transport of dissolved Ca^{+2} and HCO_3^- to the marine environment, where dissolved calcium and bicarbonate react to form calcium carbonate (often in the form of aragonite, a polymorph of calcite), a process that is often biotically mediated when marine organisms form their shells:

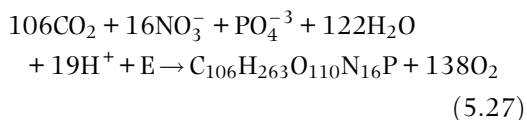


or, if the bicarbonate has been transformed to carbonate anions in the alkaline marine environment:



Burial of CaCO_3 in marine sediments and eventual compaction and cementation into carbonate-bearing rocks (e.g. as CaCO_3) provides long-term storage of carbon in the crust of the Earth. By convention, C in carbonate minerals (including calcite in fossilized shells of marine organisms) is considered **inorganic carbon**. For this reason, when calcite, dolomite or other carbonate minerals undergo dissolution, the resulting HCO_3^- (aq) or CO_3^{2-} (aq) anions are considered **dissolved inorganic carbon (DIC)**. So too is CO_2 dissolved in water, whether it is represented as CO_2 (aq) or H_2CO_3 . Conversely, dissolution of organic matter (e.g. humic acid) is characterized as **dissolved organic carbon (DOC)**. In most cases, surface water, soil water and groundwater typically contain both DIC and DOC.

Fixation of **organic carbon** into Earth's crust begins with photosynthesis, much of which takes place in the photic zone of the oceans. Alfred Redfield determined in the 1930s that the average ratio (or stoichiometry) of C : N : P in marine plankton is 106 : 16 : 1, and the photosynthetic reaction can be represented as:



At this point atmospheric CO_2 has been photochemically transformed and fixed into the biological part of the oceanic reservoir. The chemical formula $\text{C}_{106}\text{H}_{263}\text{O}_{110}\text{N}_{16}\text{P}$ represents a variety of proteins, carbohydrates, lipids, lignins and other components of biotic matter. In this compound, the oxidation states are H^{+1} , O^{-2} , N^{-3} and P^{+5} ; thus, carbon is C^0 .

If the plankton die and undergo decay in oxic (or, synonymously, *aerobic* or *oxidizing*) environments like those typified in shallow, well-circulated, oxygenated marine water or sediments, the reaction shown above will proceed in reverse and carbon will shortly be returned to the atmosphere as CO_2 – you probably will recognize this oxidation reaction as respiration ($\text{CH}_2\text{O} + \text{O}_2 \rightarrow \text{CO}_2 + \text{H}_2\text{O} + \text{E}$). One way this happens is when zooplankton consume phytoplankton; another is by aerobic decay mediated by microbes. Note also that aerobic decay also returns the vital nutrients N and P to sea water or sediments (i.e. the nutrients are “mineralized”) in the form of nitrate and phosphate.

Hydrocarbons

If the dead plankton settle through the water column to depths below the photic zone, they are more likely to end up deposited as particulate organic matter (POM) on the seafloor, where decay is likely to occur via *anaerobic* pathways in the suboxic or anoxic (chemically reducing) environment common to many marine and swamp sediments. The absence or limited amount of free O_2 in these environments limits conversion of organic C (C^0) to CO_2 (C^{+4}); instead, the residues of POM in anaerobic/reducing/anoxic environments are progressively broken down by anaerobic bacteria into progressively more hydrocarbon-like compounds.

Within the first few hundred meters of depth below the seafloor, anaerobic decomposition causes the O/C, H/C and N/C ratios of the remaining organic matter to decrease as they first result in the formation of insoluble, viscous **humins** (~ similar to humic matter in soils), and later, waxy **kerogen** (pronounced kair-o-jen). Where do the lost oxygen, hydrogen and nitrogen go? Much of these constituents are converted to H_2O , CH_4 , CO_2 and NH_3 . The CH_4 that forms in these early stages ($< 50^\circ\text{C}$) of organic maturation is known as biogenic methane. It is termed

Kerogen is a waxy mixture of organic compounds with an approximate stoichiometry of $C_{200}H_{300}O_{11}N_5S$. Type I kerogen is mainly derived from algae and they commonly mature into tar-like hydrocarbons found in oil shales. Type II kerogen is the main source of liquid hydrocarbons (their source is marine plankton), while type III kerogen is derived from higher-order plants and are the common source of coal. In addition to their position in the transformation of marine organic carbon (plankton) to hydrocarbons, kerogen is the main form of organic matter in the crustal/geological reservoir, as they comprise ~90% of organic matter in sedimentary rocks. Most organic matter is very dispersed in the crust, and only rarely does organic matter occur in the concentrations, stoichiometries and temperature range required to be considered a hydrocarbon resource.

Bitumen (pronounced “bit-you-men”) occurs

in conjunction with kerogen and also are products of the maturation of organic matter in sediments. Whereas kerogens are waxy and insoluble in organic compounds, bitumens are black and tarry and soluble in organic compounds. Both store carbon for millions of years in sediments and sedimentary rocks and given the right circumstances can transform into hydrocarbon fuel resources. At burial depths of ~1 km (~60 °C), long-chain alkanes and related compounds derived from kerogen and bitumen begin to undergo catagenesis (also termed “cracking”) that results in the formation of shorter-chain hydrocarbons. Once sediments have been buried to depths of ~2–3 km (~75–125 °C), organic matter maturation reaches the point where alkanes with ~4 to 20 carbons occur – this is the crude oil that humans pump from the ground as a fuel source.

biogenic methane to distinguish it from the inorganic methane that forms at temperatures of 150 to 200 °C at deeper crustal levels approaching the upper end of petroleum maturation. Figure 5.6 is a van Krevelen diagram showing (A) changing ratios of O/C and H/C with progressive organic maturation, and (B) differences in the types of organic matter from which hydrocarbons can form. Note the low O/C ratio of lipids compared to other components of living organic matter – it is this attribute that makes lipids most suitable for conversion to kerogen and ultimately to hydrocarbons.

Coal

Coal forms when terrestrial plants decompose in the anaerobic/anoxic environments of wetlands. The oldest coals are Devonian (~400 Ma) in age because land plants, the primary source of carbon

for coal deposits, had not sufficiently evolved and colonized the continents prior to this time. The majority of economic coal deposits are from the following time periods:

- Carboniferous and Permian (~355 to 250 Ma);
- Jurassic and Cretaceous (~205 to 65 Ma);
- Early Tertiary (Paleogene) (~65 to 24 Ma).

High global sea level during these periods led to the development of broad, gently sloping coastlines with vast forested wetlands. Modern day examples of similar swampy environments include the Pantanal in Brazil and the Okefenokee Swamp in the southeastern USA. Decomposition of fallen trees and other organic matter in anaerobic bottom waters and sediments limits loss of carbon to the atmosphere and instead fosters a decay sequence somewhat similar to hydrocarbon maturation. This sequence occurs with burial, compaction and increasing temperature and follows the maturation trend:

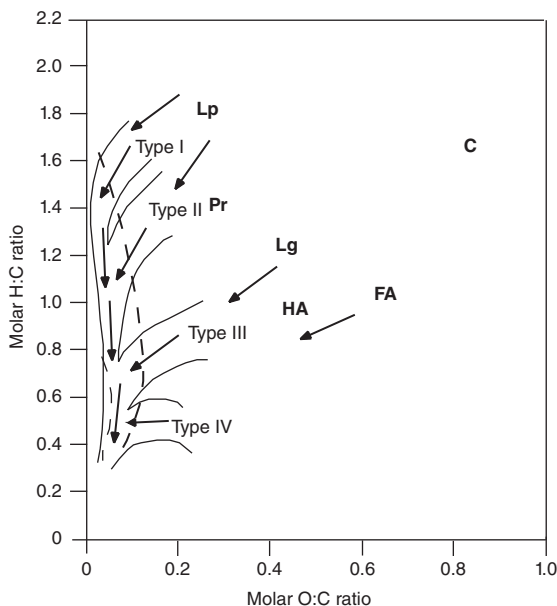


Fig. 5.6 van Krevelen diagram showing approximate composition of components of organic matter and trajectory of organic matter maturation from immature organic matter through the petroleum maturation window. Lp = lipids, Pr = proteins, Lg = lignite, C = carbohydrate, HA = humic acid, FA = fulvic acid. Types I–IV are described in the text.

peat $\rightarrow$ ($\sim 30\text{--}40^\circ\text{C}$) $\rightarrow$ lignite $\rightarrow$ ($\sim 40\text{--}60^\circ\text{C}$)
 $\rightarrow$ bituminous coal $\rightarrow$ ($\sim 200^\circ\text{C}$)
 $\rightarrow$ anthracite

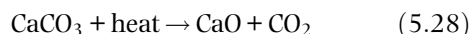
As with hydrocarbons, the carbon content of coal increases with increasing grade, from $\sim 50\%$ by mass in lignite to 90 or 95% in anthracite.

5.2.4 Rates of flux to and from the crust

Transformation of decaying organic matter to hydrocarbons or coal occurs slowly and requires tens of millions of years. The flux into the crust is slow (0.2 Pg C/yr) and most petroleum is produced from rocks whose ages range from

Pennsylvanian (~ 300 million years old [Ma]) to Miocene (~ 15 Ma); most coals are similarly old. The molecules comprising gasoline, natural gas, diesel and coal that we are burning today began their journey into sedimentary rocks as plankton or other plant remains long, long ago.

In addition to combustion of fossil fuels, **cement production** also transfers carbon from Earth's crust to the atmosphere. Cement is derived from limestone in a high-temperature process ($\sim 900^\circ\text{C}$) that results in conversion of inorganic carbon stored in the crust (as calcite) to CO_2 :



CaO is “lime”, a component of cement; carbon dioxide is released to the atmosphere as a byproduct. Approximately 10% of anthropogenic CO_2 emissions are derived from cement production.

Like the C in fossil fuels (hydrocarbons and coal), the inorganic carbon (e.g. CaCO_3) currently in carbonate rocks was fixed into the crust tens of millions to billions of years ago. The pure limestone used to make cement in the St. Lawrence River valley of southern Quebec is Ordovician (~ 460 Ma) in age; the carbon stored in this ancient calcite is being removed from the crustal reservoir at rates many orders of magnitude more rapidly than it is removed via the long-term rates of natural uplift and chemical weathering.

As a whole, the flux of carbon with respect to the crust is as follows. The transfer of carbon from atmosphere into sediments (mainly marine sediments), where they may ultimately be fixed back into the crustal reservoir, is 0.2 Pg/yr , a value that represents the combined fixation of organic and inorganic carbon. Humans are removing carbon from oil fields, coal mines and cement quarries and converting it to CO_2 at a rate of 7.8 Pg/yr (Table 5.4), the average value for 2000 to 2010 (Friedlingstein et al., 2010). This value is quantifiable because we pretty well know the amount of fossil fuels burned and cement

manufactured each year. Thus, the simple mass balance with respect to the crust is:

$$\text{Flux}_{\text{in}} = 0.2 \text{ Pg/yr}$$

$$\text{Flux}_{\text{out}} = 7.8 \text{ Pg/yr}$$

This results in an average (2000–2010) net flux of C from crust to atmosphere of 7.6 Pg/yr, ignoring uptake into other reservoirs (Flux_{out} is increasing; in 2011 the value was ~ 10 Pg/yr).

Another factor to be addressed quantitatively is the long-term flux of C from the crust (or lithosphere) to the atmosphere. Uplift and exposure of carbonates and other C-bearing rocks to the atmosphere results in oxidation of organic C and

chemical weathering of minerals, as illustrated with the weathering of wollastonite (equation 5.26) – the long-term rate prior to major human alteration was ~ 0.2 Pg/yr, representing a pre-industrial steady state where Flux_{out} and Flux_{in} were both 0.2 Pg C/yr. The current non-steady-state situation is resulting in transfer of C from crust to atmosphere at a rate ~ 40 times greater (based on ~ 8 Pg/yr $\div$ 0.2 Pg/yr) than the long-term geological (natural) flux; the current rate is also ~ 40 times greater than the flux of C to the crust. We are depleting carbon from the crust and adding it to the atmosphere at a rate 40 times greater than it is being returned to the crust.

What is the fate of the excess carbon transferred from the geological to the atmospheric reservoir?

Table 5.4 Changing fluxes of carbon among the main reservoirs for 3 time periods. In Table 5.4, negative values indicate that combustion of fossil fuel, production of cement and land-use practices (e.g. deforestation) are causing loss of C from those three reservoirs; reforestation and other changes to land are also a sink for C, and the net change in C in the terrestrial reservoir (mainly biota and soil) is the “terrestrial mass balance” – from 1980–2010, the terrestrial reservoir has experienced a net gain in C. So too has the ocean reservoir and the atmosphere. Table 5.4b depicts gross transfer of C from the lithosphere (FF/cem = fossil fuels and cement) and land-use change (LUC) to the atmosphere. Total C removed from the atmosphere is “Net sinks”. Also shown are percentages of fossil fuel and cement-derived C that remains in the atmosphere (increasing from 1960–2010), the amount of C sourced from fossils fuels and cement sequestered in the ocean (decreasing from 1960–2010), and total fossils fuels and cement sequestered (oceans + land; also decreasing with time).

Table 5.4a Fluxes of carbon (Pg/yr). Negative values indicate reservoirs that are losing C.

Time Period	Fossil fuel & cement	Land-use source	Land sink	Terrestrial mass balance	Ocean sink	Atmospheric increase
1960–1980	–3.8	–1.4	1.4	0.0	1.6	2.3
1980–2000	–5.9	–1.5	2.0	0.6	2.1	3.3
2000–2010	–7.8	–1.0	2.4	1.4	2.3	4.2

Table 5.4b Sources and sinks of carbon relative to the atmosphere (Pg/yr unless indicated).

Time Period	Gross flux in (FF/cem, LUC)	Net sinks (Ocean, Land)	Net atm increase	FF/cem C % remain in atm	FF/cem C % into ocean	% FF/cem C sequestered
1960–1980	5.3	3.0	2.3	41.9	44.1	58.1
1980–2000	7.4	4.1	3.3	44.1	35.9	55.9
2000–2010	8.9	4.7	4.2	46.8	29.9	53.2

The answer depends in large part on reservoirs (e.g. oceans, forests) that are able to fix carbon over much shorter time spans compared to the slow fixation rate into the crust. Table 5.4 delineates the main fluxes of carbon among reservoirs. In addition to the average 2000–2010 flux of 7.8 Pg/yr of carbon to the atmosphere from the crustal/lithospheric reservoir, average 2000 to 2010 uptake of carbon by the oceans was ~ 2.3 Pg C/yr and *net* transfer from the atmosphere to terrestrial reservoirs (via photosynthetic uptake into plants plus fixation into soils) was ~ 1.4 Pg/yr of carbon. This complex system is difficult to precisely quantify, but the mass balance of carbon fluxes, and measurements of atmospheric carbon concentrations, indicates that from 2000 to 2010, the atmosphere gained an average net amount of ~ 4.2 Pg of carbon per year. What this means is that $\sim 47\%$ (from $3.7 \div 7.8$ Pg C/yr) of total fossil fuel and cement-derived carbon emissions annually between 2000 and 2010 has been fixed into the ocean and terrestrial reservoirs. This is slightly higher than ocean-terrestrial uptake rate from 1960 through 1980, where $\sim 42\%$ of annual carbon emissions derived from the crust were fixed into oceanic and terrestrial reservoirs.

5.2.5 The ocean reservoir

The global ocean contains 40 000 Pg C, making it the largest of the Earth surface reservoirs. Most carbon in the oceans is derived from, and cycled back to, the atmosphere due to the vast permeable boundary between these two reservoirs. CO_2 from the atmosphere is fixed into the oceanic reservoir in two main ways. The first is via photosynthesis by phytoplankton as described above; the second is solution of CO_2 into seawater. Carbon also enters the oceans via fluvial transport, mainly as dissolved bicarbonate with lesser amounts of organic carbon, but this flux (1.2 Pg C/yr) is small compared to the amount fixed into the oceans from the atmosphere (90 Pg C/yr).

5.2.6 Fixation of C into oceans

Incorporation of carbon into the oceans by photosynthesis is generally a very short-term fixation (in other words, this carbon generally has a low residence time), especially if the phytoplankton are consumed by higher-trophic level organisms and their carbon is returned to the atmosphere as CO_2 via respiration. However, if after the plankton die they sink to great depths, the carbon will be removed from the shallow ocean, facilitating fixation of organic carbon into marine sediments. Provided that the C stored in these sediments is not oxidized to CO_2 , this process can contribute to long-term C storage.

Solution of CO_2 into sea water is the dominant means by which carbon is transferred from atmosphere to oceans, and it exists as dissolved CO_2 gas as well as carbonic acid (H_2CO_3) and the bicarbonate (HCO_3^-) and carbonate (CO_3^{2-}) anions. The difference between these two constituents, CO_2 and H_2CO_3 , is difficult to determine analytically so they are often grouped together as CO_2 (aq). As we can see from Fig. 5.5, the oceans are currently fixing 92 Pg/yr of carbon while they are emitting 90 Pg/yr, indicating that the oceans are experiencing a net increase of ~ 2 Pg/yr of carbon (the average for 2000 to 2010 is 2.3 Pg C/yr; Friedlingstein et al., 2010).

To date, approximately one-third of anthropogenic CO_2 added to the atmosphere has been absorbed by the oceans. This is a predictable consequence of the increasing CO_2 in air considered in terms of **Henry's Law**, which states that, at constant temperature, the concentration of a gas dissolved in a liquid is directly proportional to the concentration (or partial pressure) of the gas in air in contact with, and in equilibrium with, the liquid.

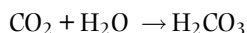
$$k_{\text{H}} = c_{\text{aq}}/c_{\text{atm}} \quad (5.29)$$

In this equation, k_{H} is the Henry's Law constant for CO_2 , c_{aq} is the concentration of CO_2 in the aqueous

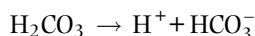
solution (e.g. sea water), and c_{atm} is the concentration of CO_2 in the atmosphere. The eloquently simple mathematics here indicates that when c_{atm} increases, and in order for k_{H} to remain constant (which it must for a given temperature), c_{aq} must also increase. So this is another aspect of the complex climate system that is not so complex. Simple first principles of chemistry dictate that increasing CO_2 in the atmosphere will increase dissolved CO_2 in sea water (note: this ignores for now the decreasing solubility of CO_2 with increasing temperature of sea water).

Ocean acidification

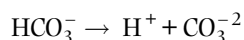
A consequence of incorporation of CO_2 into sea water is the formation of carbonic acid, as we have seen previously:



From this reaction it is evident that increased solution of CO_2 in sea water will drive the system towards increased H_2CO_3 , dissociation of which will yield H^+ and bicarbonate ions in solution.:



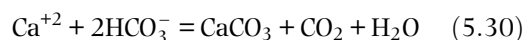
Further dissociation yields another H^+ and a carbonate anion:



Though carbonic acid is not a strong acid, its dissociation in the marine environment drives the carbonate system towards lower pH, and ocean acidification caused by increased solution of CO_2 in sea water will occur (Raven et al., 2005). Decreasing pH of sea water raises concerns about marine ecosystems, particularly pertaining to organisms that manufacture shells and other body parts from CaCO_3 . Calcium carbonate solubility is highly sensitive to pH, and ocean acidification is

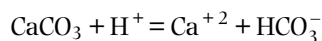
likely to diminish the stability of calcium carbonate-based organisms and ecosystems; calcareous plankton, an important component of the marine food chain, and coral reefs are expected to be particularly sensitive to lowered pH.

In sea water saturated with respect to calcium carbonate (this is the situation now and the one that has prevailed during most of Phanerozoic time, i.e. for the past 540 Ma), the following reaction occurs (equation 5.18 in reverse):

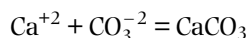


and the CO_2 produced by the reaction is in equilibrium with atmospheric CO_2 . Viewed in this way, it is possible to predict the effect of increased $\text{CO}_{2(\text{atm})}$ on the stability of CaCO_3 in sea water. One way to consider this question is in terms of Le Châtelier's principle, which indicates that increasing the concentration of any of the products will shift the equilibrium in the direction of reactants. Given that Henry's Law dictates that increased atmospheric CO_2 will result in increased CO_2 in seawater, this change in equilibrium position (an increase in one of the products of the above reaction) will drive the system towards reactants, in this case dissolved calcium and bicarbonate, and this will occur at the expense of calcium carbonate.

Another way to visualize the impact of increasing CO_2 on marine CaCO_3 -based organisms is to remember that some component of dissolved CO_2 in sea water combines with H_2O to form carbonic acid. Subsequent dissociation of carbonic acid yields H^+ and the fate of calcareous organisms in more-acidic sea water can be clearly seen as pH dependent:



We can also examine the relationship between shell formation from dissolved carbonate anions and the role of increased CO_2 in sea water:



However, excess CO_2 fosters this reaction:



thus decreasing carbonate anion concentration and inhibiting production of CaCO_3 by organisms (Orr et al., 2005).

A 2004 report by the Royal Society of London (Raven et al., 2005) indicates that global sea water pH has decreased from 8.15 to 8.05 over the past 200 years due to anthropogenic CO_2 emissions (which, due to the log scale by which acidity is measured, corresponds to a ~25% increase in H^+ in sea water). The consensus of climate scientists and oceanographers summarized in the 2007 IPCC report (Solomon et al., 2007) projects significant decreases in ocean pH with progressive increases in CO_2 . Data from the Pacific Ocean north of Hawaii (Station Aloha) depict some predictable features, including increasing inorganic carbon (i.e. dissolved CO_2) and decreasing pH, trends that are both characterized by much scatter. It also must be noted that the observed changes thus far to the pH of sea water are ≤ 0.1 pH units, a difference that is often within the margin of error of pH analysis. At the same time, as noted above, a decrease of from pH = 8.15 to pH = 8.05 represents a 25% increase in $[\text{H}^+]$.

As is the case with many aspects of the climate system, the geological record also serves as a useful tool in assessing the response of the oceans and the marine ecosystem to rapid increases in CO_2 . Volcanically derived CO_2 increased rapidly in the atmosphere during the Aptian Age of the Cretaceous Period 120 million years ago – the volcanic activity was located at and resulted in production of the Ontong-Java plateau, and this event is particularly useful to compare to the modern situation because, like now, the rate of CO_2 flux to the atmosphere was rapid, not gradual. Erba et al. (2010) examined the effect of this event on marine calcifiers, in this case their

response to increasing solution of CO_2 into sea water. In summary, their results indicate that calcareous nanoplankton responded in predictable ways, exhibiting malformations, dwarfism and decreases in abundance associated with sea water acidification. Overall flux of CaCO_3 to marine sediments also decreased, a change that Erba et al. (2010) attribute to lowered saturation state of CaCO_3 in sea water.

5.2.7 Long-term viability of oceans as C sink

An additional concern related to dissolved CO_2 in sea water is the effect of increasing ocean temperature on the solubility of CO_2 gas. If we are concerned about the effect of too much CO_2 on seawater chemistry, we also need to be concerned about the effect of rising ocean temperature on CO_2 solubility. For a given partial pressure of CO_2 in air, the amount of CO_2 in solution in a liquid (e.g. sea water) in contact with the air decreases as temperature increases (Fig. 5.7). This is a concern because the ocean has been a major sink for anthropogenic CO_2 over the past 150 years. From 1960 to 1980, 42% of fossil fuel plus cement emissions were fixed into the oceans as dissolved CO_2 ; the same value for the 2000 to 2010 time period is 30%. (it is important to note that while the proportion of CO_2 emissions sequestered into the oceans have decreased, the actual amount of CO_2 sequestered has increased, from 1.6 Pg C/yr in the 1960–1980 period to 2.3 Pg C/yr in the 2000–2010 period.) The inverse correlation between the solubility of CO_2 in seawater and temperature (Fig. 5.7) indicates that the capacity of the oceans to fix CO_2 from air will diminish as the oceans warm; however, the myriad controls on ocean chemistry and dissolved $[\text{CO}_2]$, from ocean circulation patterns to biological activity, serve to complicate precise prediction of CO_2 fixation in the oceans.

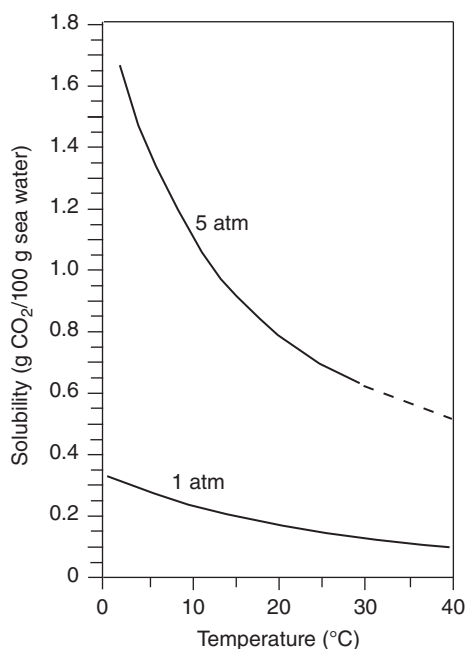


Fig. 5.7 Temperature dependence of CO_2 solubility in sea water at two different pressures. At a given pressure, increasing temperature decreases the solubility of CO_2 in sea water. Data are from Stewart and Munjal (1970) and sources cited therein.

CO_2 concentrations in sea water range from 4.4×10^{-3} g CO_2 /100 g (of sea water) up to 3.5×10^{-2} g/100 g depending on temperature, where higher values correspond to cooler, deeper waters (Berner and Berner 1996). Figure 5.7 emphasizes that (1) CO_2 is most soluble in cool sea water at great depths, (2) CO_2 is least soluble in warm surface water, and (3) at atmospheric pressure, the solubility of CO_2 in sea water decreases with increasing temperature.

Methane hydrates

An additional important and interesting reservoir of carbon that occurs primarily in marine sediments, but also in arctic soils and sediments, are the slushy-to-icy compounds known as methane hydrates (synonymous terms include gas hydrates and methane clathrates). These

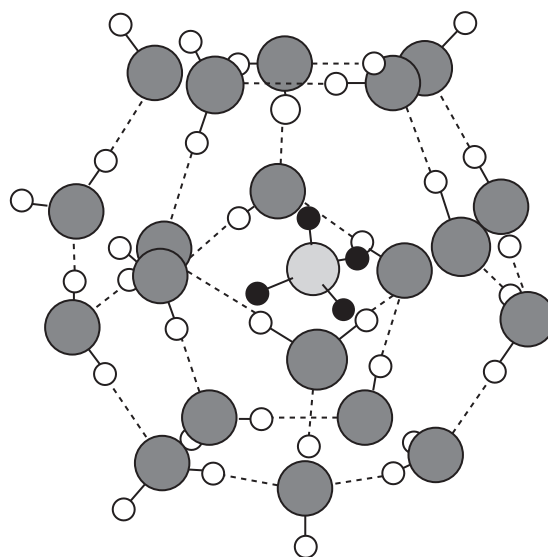


Fig. 5.8 Structure of methane clathrate (methane hydrate). Note the dodecahedral (12-sided) crystal comprised of H_2O molecules encasing a methane molecule. Oxygen atoms in water are dark gray with black outlines and the hydrogen atoms in water are white. The carbon atom in methane is light gray and the four hydrogen atoms in the methane molecule are black. Covalent bonds within the water and methane molecules are indicated by black lines, while the hydrogen bonds between adjacent water molecules are indicated by dashed lines.

compounds consist of methane encased in a crystalline dodecahedral structure formed by H_2O molecules coordinated around methane molecules (Fig. 5.8); the stoichiometry is $(\text{CH}_4)_8 \cdot (\text{H}_2\text{O})_{46}$ (Spiro and Stigliani, 2006).

The term for any such crystal where an H_2O cage-like structure traps a gas molecule is clathrate, and when methane is the gas in the structure, they are called methane hydrates or methane clathrates. The global reservoir of methane hydrates is approximately 10 000 Pg C, which is twice global hydrocarbon reserves. For this reason, energy agencies and companies are interested in them as a potential fuel source (even

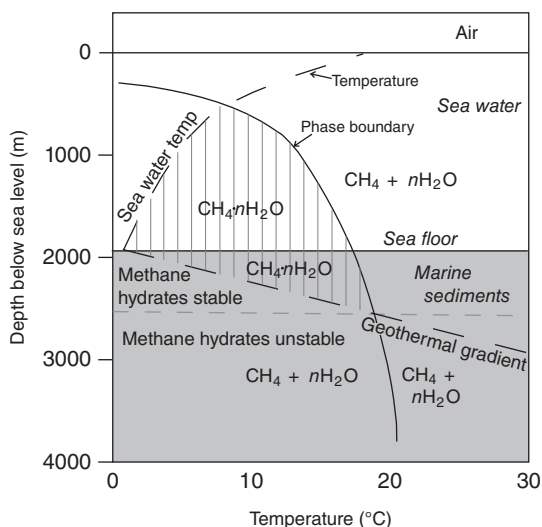


Fig. 5.9 Stability diagram for methane hydrates in marine environments. The high pressure of deep sea water is not directly shown here, but rather is indirectly indicated by water depth. The combination of high pressure and low temperature in the deep sea are the main forces that lead to formation and preservation of clathrates in marine sediments. In this example, methane hydrates are stable below 500 m in sea water and down to a depth of 500 m in marine sediments (or 2500 m below the sea floor) – the stability field is indicated by vertical lines. Warming of deep sea water would shrink the size of the stability field. (Modified from data in Caldeira et al. 2005.)

with modern technology extraction is difficult). The ability of methane to act as a strong greenhouse gas in the atmosphere raises concerns about release of methane from clathrates to air.

In marine sediments, methane hydrates typically occur within the upper 500 m of sediments (i.e. within 500 m of the sea floor); at greater depths, heat emanating from within the Earth melts the clathrate and releases methane and water. The temperature of deep sea water near the sea floor is $\sim 2^\circ\text{C}$, and the combination of low temperature and high pressure of the deep seas stabilizes these structures as lenses or pods within sediments (Fig. 5.9). In some areas,

methane hydrates even outcrop at the sea floor. Most occur on the continental shelf, and the origin of the carbon in the methane may be terrestrial organic matter transported to oceans via streams, or marine organic matter that settles through the sea-water column towards the sea floor. In either case, microbial degradation in anaerobic deep sea water facilitates formation of methane which then becomes incorporated into clathrates in cold, high-pressure deep sea water and sediment. Methane in clathrates may also be derived via outgassing from the lithosphere (e.g. gas escaping from hydrocarbons), where it can migrate upward through faults or fractures towards the sea floor to the shallow depths within marine sediments where it becomes stable in a clathrate structure.

C in cold region soils

Soils and sediments of the vast Arctic region (e.g. Alaska, Siberia, Scandinavia, northern Canada) also contain methane hydrates, approximately 1500–2000 Pg C in Arctic clathrates, a quantity that is roughly double the current amount of C currently in the atmosphere (Heimann, 2010). In the Arctic, low mean annual air temperature and the pressure of overlying sediments stabilize clathrates to depths of ~ 500 m. Methane hydrates often occur in permafrost soils and sediments and the origin is similar to the marine environment – anaerobic microbes facilitate decomposition of organic matter into methane, and in the arctic environment, anaerobic conditions are fostered by vast wetlands.

Concern with methane hydrates in terms of air quality and climate arises from the potential that global warming will decrease their stability, i.e. cause them to melt, a prospect that could release very large amounts of methane to the atmosphere. Marine sediments contain a far greater content of methane in clathrates than do Arctic soils and sediments ($\sim 10\,000$ Pg C vs. ~ 1500 Pg C); however, the great pressure of sea water

and the isolation of deep sea water from the atmosphere suggests that deep marine methane hydrates may be relatively stable. Terrestrial hydrates, on the other hand, are likely much less stable. Air temperature is very closely linked to the stability (and depth below surface) of permafrost and associated clathrates, so many climate scientists are concerned about the potential for release of methane from clathrates buried at shallow depths in Arctic soils and sediments.

5.2.8 The atmospheric reservoir

Previous sections have covered various relationships between carbon in the atmosphere and the other carbon reservoirs, so this section will focus on aspects of the C cycle related to the atmosphere not yet covered. It is important to emphasize that the concentration of CO₂ (plus CH₄ and other less-abundant forms of C) in the atmosphere is controlled by fluxes that range from very rapid – e.g. removal of CO₂ from air via photosynthesis, or via solution of CO₂ into sea water – to very slow, such as the transfer of CO₂ to the crust as described previously. At the same time, humans have greatly altered the natural cycling of C by accelerating the rate at which carbon is transferred from the crust to the atmosphere. Flux out of the atmosphere has not kept pace with anthropogenic increases, so the CO₂ concentration in the atmosphere is not in a steady state.

Changes to atmospheric C over geological time

Atmospheric CO₂ has varied by orders of magnitude over Earth's history, due in large part to evolution of life and burgeoning of photosynthesis that provided a new mechanism to fix carbon from air onto marine and eventually terrestrial reservoirs. In particular, the appearance and increase in abundance of cyanobacteria in the Early Proterozoic Era heralded the onset of

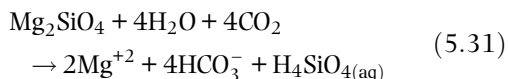
oxygenic¹ photosynthesis, facilitating the fixation of CO₂ from atmosphere to biota and the production of O₂; increasing O₂ transformed the atmosphere and shallow marine realms from reducing into oxidizing environments, a shift that occurred ~2.2 to 2.1 billion years ago and is recorded in the appearance of widespread ironstones and redbeds (Fe₂O₃-bearing sandstones) and disappearance of pyrite as a type of clast deposited in stream sediments (a good summary of early Earth C cycling is presented by Sundquist and Visser, 2003).

Atmospheric carbon likely was equal to or greater than current concentrations during most of the Phanerozoic (540 Ma to now), when values often exceeded 1000 ppm (Royer et al., 2004). What kept the Earth from being extremely warm earlier in its history is that the sun was less luminous than it is now. The early Earth had a stronger greenhouse effect but lower insolation than the modern world, so in spite of much higher CO₂ concentrations, temperatures were within ~8 °C of modern temperatures (Veizer et al., 1999), and were lower during periods of glaciations (notably parts of the late Proterozoic, Permian and, of course, Pleistocene). During most of the Pleistocene and Holocene (i.e. the past 2 Ma), atmospheric CO₂ has ranged from ~180 to 280 ppmv (as indicated by air bubbles trapped in glacial ice), and only in the past 150 years has it exceeded 300 ppmv.

Over geological time, the main sinks for atmospheric carbon have been:

- 1 Chemical weathering of silicate minerals, particularly Ca- and Mg-silicates. The reaction of wollastonite was presented in equation 5.26, and the weathering of forsterite (the magnesian variety of olivine) is shown below:

¹ Note – prior to the appearance of cyanobacteria, more-primitive types of bacteria carried out anoxygenic photosynthesis that, as the name would suggest, did not produce O₂. C fixation may have been facilitated by reactions involving H, S or some other element(s). Some of these bacteria still exist today (e.g. green bacteria).



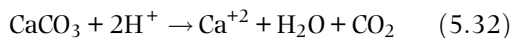
This process fosters incorporation of CO_2 into soils, streams or ocean water as the bicarbonate anion. The carbon may then be stored for long periods of time (e.g. time scales of 10^6 years and longer) if the carbonate reacts with Ca to form aragonite or calcite in marine sediments and sedimentary rocks.

2 Chemical weathering of carbonates, as shown in equation 5.18, which can result in transfer of C to oceans via fluvial transport; then, in the marine environment, if the bicarbonate anion reacts with Ca to form carbonate minerals it may become buried in sedimentary basins, where the carbon can be sequestered for millions of years.

3 Photosynthesis that facilitates transfer of carbon to organic matter followed by burial, as described above for hydrocarbons, coal and methane hydrates. In each of these three scenarios, carbon can be stored for millions to billions of years in the crust of the Earth if buried in sedimentary rock sequences.

Carbon is released from long-term geological reservoirs to the atmosphere by the following processes:

1 Liberation of CO_2 via chemical weathering of carbonates (the equation shown below is a different way to think of this reaction than what is shown in equation 5.18):



In this example, we can assume that (1) tectonic forces have caused uplift and exhumation of once deeply buried sedimentary rocks, and (2) the source of the acid driving the weathering reaction is HNO_3 or H_2SO_4 derived from volcanism, natural combustion or other sources. If the H^+ source is carbonic acid, the result may be no net addition of carbon to the atmosphere (see equation 5.18).

2 Oxidation of organic matter in the crust (e.g. hydrocarbons) and marine sediments (e.g. methane

hydrates). This process too would be facilitated by uplift and exposure of sedimentary rocks. Whether or not uplift leads to a net release of carbon (via these two examples), or whether it leads to a net removal of carbon from the atmosphere, depends on the amount of silicate weathering as compared to oxidation of organic matter and loss of carbon from carbonate rocks (Ruddiman, 2007).

3 Volcanic emissions of carbon dioxide.

4 Metamorphism of carbonate rocks at temperatures high enough to liberate CO_2 (equation 5.28). You will note that this natural process is analogous to manufacture of cement as shown earlier in this chapter. CO_2 can escape from metamorphic rocks via fractures and other pathways.

5 Metamorphism of organic matter, e.g. in meta-sedimentary rocks, a process where reduced carbon is oxidized to CO_2 and allowed to escape via fractures and other pathways.

Feedback loops

Note that the main long-term controls on atmospheric CO_2 involve transfer of carbon between the atmosphere and the main lithospheric reservoirs, namely, carbonates and organic matter in the crust. The rates of these processes are relatively slow compared to modern anthropogenic fluxes (Fig. 5.5; Table 5.4). The Phanerozoic record of atmospheric CO_2 concentration reflects the influence of feedbacks (or feedback loops), where **positive feedbacks** amplify processes (these processes or triggers are sometimes termed climate forcings) already in progress, whereas **negative feedbacks** counteract processes already under way.

An example of a geochemically related positive feedback loop is as follows:

1 increased insolation (due to sunspot activity or orbital variations) causes warming of the atmosphere;

2 the warming atmosphere heats the shallow ocean;

3 increased ocean temperature decreases the solubility of CO_2 in sea water;

4 loss of CO_2 from sea water to the atmosphere causes the atmosphere to undergo further warming;

5 the further increase in air temperature causes additional losses of CO_2 from sea water; and so on. In a positive feedback loop, the initial forcing (in this case warming of air due to increased solar energy reaching the Earth) sets a process in motion that can lead to greatly amplified results.

An example of a negative feedback in the carbon-climate system is chemical weathering. Consider again an initial forcing (e.g. increased insolation) that causes the temperature of the atmosphere to rise, liberating CO_2 from sea water and terrestrial sources. However, given that CO_2 is a reactant in chemical weathering reactions, increasing atmospheric CO_2 will drive weathering reactions to the right, transferring carbon from CO_2 to HCO_3^- , where it can be fixed into marine carbonates. This process decreases the concentration of CO_2 in the atmosphere, or at least it counters the increase caused by the initial trigger, and in doing so counteracts the initial temperature increase.

The pronounced global cooling during late Cenozoic time that led to the Pleistocene glacial epoch has been attributed to chemical weathering associated with uplift in the Himalaya, Andes and other mountain ranges that lowered atmospheric CO_2 concentrations (refer to sources in Ruddiman, 2007) – this period of elevated tectonic activity, much of which was related to contraction and uplift from western Europe to east Asia, raised mountains to high elevations and supported a greater-than-normal abundance of alpine glaciers. The mechanical action of alpine glaciers produces fine-grained glacial flour with high surface area:volume ratio, thus increasing the rate of chemical weathering. Increased rates of chemical weathering result in decreased amounts of CO_2 in the atmosphere; in turn, lower CO_2 in the atmosphere leads to global cooling which at the time may have fostered an even greater abundance and extent of alpine glaciers, which in turn would have ground up even more

rock, further enhancing chemical weathering, decreasing CO_2 in air to a greater extent, and setting in motion a positive feedback loop whose result was two million years of extensive alpine and continental glaciation.

It is the back-and-forth between the many mechanisms controlling fluxes between reservoirs, amplified at times by feedback loops in the climate system, that regulate carbon content in the atmosphere, crust, and terrestrial and marine reservoirs (the reader is referred to Ruddiman, 2007 for a more-extensive account of climate feedbacks, paleoclimate and the climate system in general). A geological time scale perspective on CO_2 is presented in Fig. 7.2, showing changes in atmospheric CO_2 over the past 550 Ma (i.e. Phanerozoic time).

Anthropogenic C and the atmosphere

Moving to more recent times, the mass of carbon in the atmosphere has been increasing progressively since the late 18th century as carbon from the crust (hydrocarbons, coal and calcite) and, to a lesser extent, carbon from terrestrial ecosystems, is converted to CO_2 . Understanding the behavior of the modern atmosphere considered as a reservoir within the carbon cycle requires the following assumptions and information:

- First, in terms of understanding the 21st century carbon cycle, transfer of carbon from the atmosphere to the crust (or mantle) by natural processes is very small compared to anthropogenic fluxes. Fixation of C into the crust occurs at rates ~ 40 times slower than the human-enhanced flux of carbon from crust to atmosphere (Fig. 5.6; Table 5.4). The only way that fixation of C into the lithosphere will become significant is if humans undertake large-scale carbon sequestration via mineral carbonation or injection of CO_2 into deep reservoirs.
- Secondly, the oceans are currently absorbing a net quantity of ~ 2 Pg C/yr (and rising) from the atmosphere, a process driven by the change in equilibrium between atmosphere and ocean.

Increasing partial pressure of CO_2 in air will force greater amounts into sea water.

- Thirdly, given that CO_2 is a reactant in photosynthesis, the equilibrium of this reaction should be shifted to the products side – Le Châtelier's Principle indicates that photosynthesis should occur at a greater rate; however, this interpretation ignores complications such as potential absence of limiting nutrients, soil erosion, and desertification, all of which would limit photosynthesis (and hence C uptake into the terrestrial reservoir) on a global scale.

- Fourthly, atmospheric CO_2 and global mean temperature have been closely linked throughout Earth history, particularly the Pleistocene (the past two million years) and Holocene (the past ~10 000 years). Regardless of whether the initial trigger of warmings in the past has been temperature increase (e.g. due to increased insolation) or addition of CO_2 to the atmosphere (e.g. via uplift and oxidation of organic matter in sedimentary rocks), temperature and CO_2 have the potential to act in concert via various feedback loops to produce rapid warming. Geologic and ice-core evidence tells us that when the Earth's atmosphere warms, it does so very rapidly, and the negative feedback loops that eventually kick in to counter warming and instigate cooling operate at much slower rates. The issues are covered in more detail in Chapter 7. Exchange between atmosphere and the oceans, and between the atmosphere and terrestrial reservoirs (soils, forests), occurs on relatively short time scales, typically days to months to years to hundreds or, in the case of deep circulating marine water, thousands of years. C that humans are extracting from the crustal reservoir will either become partitioned into terrestrial or marine reservoirs, or it will build up in the atmosphere. All of these are happening right now, and the residence time in these reservoirs is short compared to the crustal reservoir. For example, CO_2 in air is converted to organic carbon by photosynthesis over the life span of a plant, and the cycle is completed when the plant decays in the

presence of water, oxygen and microorganisms and the carbon is returned to the atmosphere as CO_2 . This occurs in marine and terrestrial ecosystems. Similarly, solution of CO_2 gas from atmosphere into ocean water or groundwater can result in storage of C as HCO_3^- for tens to thousands of years, depending on how deeply the water circulates into the marine environment or aquifer.

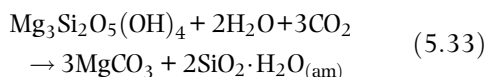
The tenuous status of carbon in terrestrial and marine reservoirs is a concern. What happens if the oceans reach saturation and do not continue to absorb roughly one-half of anthropogenic carbon emissions, as has been the case over the past 50 years? Similarly, if organic carbon and methane in warming permafrost soils are oxidized to CO_2 in air, how will society respond? Cutting emissions is clearly one strategy (although a difficult one to achieve); intentional sequestration of carbon into stable reservoirs is another strategy currently being explored.

5.2.9 Sequestration

As the concentration of atmospheric CO_2 steadily rises into the 21st century, humans have begun to explore ways to intentionally remove CO_2 from air through a group of processes that fall under the umbrella term sequestration. Many of the approaches are biological and involve fixation via photosynthesis of carbon into ecosystems, particularly marine reservoirs, e.g. via iron fertilization of sea water to accelerate marine microbial fixation, or terrestrial reservoirs, including reforestation (to store C in biomass) and improved soil management to fix and store photosynthetically fixed carbon. However, some sequestration approaches are geochemical in nature; the two we will examine in this section are mineral carbonation and injection of CO_2 into permeable rock deep in sedimentary basins.

Mineral carbonation involves a chemical reaction between a magnesium silicate mineral and

CO₂ (in the form of carbonic acid). Earlier in this section we examined weathering of forsterite (olivine) as an example of how silicate weathering can result in fixation and storage of C in carbonate rocks; shown below is the reaction of serpentine with carbonic acid (note: equation 5.33 expresses a similar process to equation 5.31):



The magnesite (MgCO₃) produced in this reaction can be sequestered as white MgCO₃ powder. There are many abandoned mines and quarries with tailings piles rich in serpentine-group minerals, and if this reaction could be carried out on a large scale, we might have a legitimate means of removing carbon from the atmosphere via a process that mimics natural chemical weathering. Some hydrocarbon reservoir sandstones contain authigenic Mg-rich berthierine and chlorite minerals that also may be suitable for mineral carbonation. The US Department of Energy has funded studies on mineral carbonation, and while in principal this approach has potential, the main barrier is that the reaction occurs very slowly at ambient temperatures and requires heat to carry out at rates that would be effective. For this reason, the process is not cost-effective with current technology and economics.

Another suggested approach to carbon sequestration with connections to geochemistry is injection of CO₂ into deep porous aquifers with impermeable caps, the classic example being sandstones beneath shale confining layers (or aquitards). In some cases the host reservoirs are sandstones from which petroleum has been produced – in fact, deep injection of CO₂ is already taking place in hydrocarbon reservoirs of the North Sea. This scenario differs from mineral carbonation in that the mechanism for retaining the CO₂ in the reservoir is via solution of CO₂ into deep basin groundwater. One of the potential

problems with this approach is that CO₂ produces carbonic acid, which has the potential to dissolve calcite in the reservoir rock.

Though beyond the scope of this text, it is also worth drawing attention to the fact that there are many other C sequestration schemes, ranging from intentionally fixing C into soil by conservation farming practices to fertilizing ocean water with Fe, where the rationale is that Fe is a limiting nutrient in the oceans and increased Fe will result in increased primary productivity, fixing C into phytoplankton.

The topic of carbon cycling is highlighted again in Chapter 7, where the focus is the atmosphere, first examining the geological topic of atmospheric change over the past ~3 billion years, and then focusing on C as a global atmospheric pollutant.

QUESTIONS

1. Calculate the pH of rain if CO₂ concentration in the atmosphere were to reach 1000 ppm.
2. Predict the concentration of Ca⁺² in solution (e.g. soil water) in equilibrium with calcite at 25 °C.
3. What is the oxidation state of C in CO₂? In CH₄? What does this difference imply about the stability of CH₄ in contact with the atmosphere?
4. Is the arid river water in Table 4.1 saturated with respect to Ca and HCO₃ (in equilibrium with calcite), or is it below saturation?
5. Determine the mean residence time of a carbon atom in the atmosphere.
6. Compare the rate of burial of C in marine sediments to the pre-industrial rate of exhumation of C (e.g. via weathering and volcanic emissions). How do these rates compare to each other, and to the modern rate of anthropogenic C emissions?
7. Describe the two main means by which C is transferred from atmosphere to ocean (i.e.

organic vs. inorganic). Provide details, and indicate which tends to result in long-term storage. Describe two long-term environmental concerns related to transfer of C from rock to ocean, using chemical equations and graphs to support your answer.

8. Warming air and seawater are likely to destabilize C stored in methane hydrates – which is likely more vulnerable, CH₄ in permafrost soils or CH₄ in deep marine environments, and why?
9. Provide an example of a mineral carbonation chemical reaction, showing how C might be fixed from the atmosphere into long-term mineral storage. What is the main factor limiting practical application of mineral carbonation as a strategy for managing carbon?

REFERENCES

- Berner, E.K. and Berner, R.A. 1996. *Global Environment: Water, Air and Geochemical Cycles*. Prentice-Hall, Upper Saddle River, New Jersey, USA.
- Caldeira, K., Akai, M., Brewer, P., et al. (2005) Ocean storage. In: *IPCC Special Report on Carbon Dioxide Capture and Storage* (eds. Metz, B., Davidson, O., de Coninck, H., Loos, M., and Meyer, L.), pp. 277–317. Cambridge University Press, Cambridge.
- Cartigny, P., Pineau, F., Aubaud, C., and Javoy, M., 2008. Towards a consistent mantle carbon flux estimate: Insights from volatile systematics (H₂O/Ce, δ D, CO₂/Nb) in the North Atlantic mantle (14 °N and 34 °N). *Earth and Planetary Science Letters* 265: 672–685.
- Erba, E., Bottini, C., Weissert, H.J. and Keller, C.E., 2010. Calcareous Nannoplankton Response to Surface-Water Acidification Around Oceanic Anoxic Event 1a. *Science* 329: 428–432.
- Friedlingstein, P., Houghton, R.A., Marland, G., Hackler, J., Boden, T.A., Conway, T.J., Canadell, J.G., Raupach, M.R., Ciais, P., and Le Quéré, C., 2010. Update on CO₂ emissions. *Nature Geoscience* 3: 811–812.
- Heimann, M., 2010. How Stable Is the Methane Cycle? *Science* 327 no. 5970: 1211–1212.
- Langmuir, D., 1997. *Aqueous Environmental Geochemistry*, Prentice-Hall, Upper Saddle River, New Jersey, USA.
- Manahan, S.E., 2000. *Environmental Chemistry* (7th edn). CRC Press, Boca Raton, Florida.
- Morse, J.W., Arvidson, R.S., and Lüttge, A., 2007. Calcium carbonate formation and dissolution. *Chemical Reviews* 107: 342–381.
- Orr, J. C., V. J. Fabry, O. Aumont, L., et al. 2005. Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature* 437: 681–686.
- Plummer, L.N. and Busenberg, E., 1982, The solubilities of calcite, aragonite and vaterite in CO₂-H₂O solutions between 0 and 90 °C, and an evaluation of the aqueous model for the system CaCO₃-CO₂-H₂O. *Geochimica et Cosmochimica Acta* 46: 1011–1040.
- Raven, J., K. Caldeira, H. Elderfield, O. Hoegh-Guldberg, P.S. Liss, U. Riebesell, J. Shepard, C. Turley, and A.J. Watson. 2005. Ocean acidification due to increasing atmospheric carbon dioxide. Policy Document. The Royal Society, London, 60 pp.
- Royer, D.L., Berner, R.A., Montanez, I.P., Tabor, N.J., and Beerling, D.J., 2004. CO₂ as a primary driver of Phanerozoic climate change. *GSA Today* 14, 4–10.
- Ruddiman, W.F., 2007. *Earth's Climate: Past and Future* (2nd edn). W.H. Freeman and Co., New York.
- Solomon, S., Qin, D., and Manning, M. (eds.), 2007. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*.
- Spiro, T.G. and Stigliani, W.M., 2006, *Chemistry of the Environment*, 2nd edn Prentice Hall, Upper Saddle River, NJ, USA, 489 p.
- Stewart, P.B. and Munjal, P., 1970. Solubility of carbon dioxide in pure water, synthetic sea water, and synthetic sea water concentrates at –50 to 250 °C and 10 to 45atm pressure. *Journal of Chemical and Engineering Data*, 15 (1): p 67–71.
- Sundquist, E.T. and K. Visser, 2004, The geological history of the carbon cycle, in Schlesinger, W.H., ed., *Biogeochemistry, Treatise on Geochemistry*, vol. 8, Elsevier, Amsterdam, The Netherlands, Ch. 9, p. 425–472.
- Veizer, J. et al., 1999, Sr⁸⁷/Sr⁸⁶, C¹³, and O¹⁸ evolution of Phanerozoic seawater. *Chemical Geology* 161: 59–88.

Biogeochemical Cycles – N, P, S

Chemical reactions cause elements to be transferred among various compounds and reservoirs that vary in form and scale. Some elements have very strong affinities for certain reservoirs; for example, silicon is strongly partitioned into minerals in the crust and in sediments at the surface of the Earth, whereas only small amounts of Si, mainly as wind-blown mineral particles, are found in the atmosphere. Conversely, nitrogen is strongly partitioned into the atmosphere and concentrations of nitrogen in the crust are very low. In some cases, elemental cycling occurs primarily by inorganic processes (e.g. mineral reactions deep within sedimentary basins or hydrothermal systems), but often elemental cycling is influenced or strongly controlled by biological processes. Prominent examples of biogeochemical processes are the cycling of carbon, where photosynthesis fixes carbon into living organisms, and the cycling of nitrogen, where soil microbial activity can help to fix nitrogen from the atmosphere (as N_2) into soils (e.g. as NH_4^+ and NO_3^-).

Systems and elemental cycles

Important terms related to elemental cycling include **reservoirs**, **fluxes**, **residence times** and **steady state**. Geochemists consider the Earth a **closed system** with respect to elemental cycling – effectively all elements are recycled within reservoirs from the atmosphere to down into the solid Earth. Clearly this ignores addition of matter by extraterrestrial impacts (proportionally very small inputs) and negligible losses from the top of the atmosphere to space, but the consideration of Earth as a closed system is the common approach because these minor fluxes are negligible at the scale of the whole Earth.

Reservoirs range in scale depending on the scope of the problem being studied. In general, any reservoir will contain a quantifiable amount of material with definable borders defined by distinct chemical, physical, or biological characteristics. Essentially, scientists who seek to quantify the amount of material in various parts of the lithosphere–hydrosphere–biosphere–atmosphere

system, and who seek to examine the movement of material among these systems, need to define various aspects of the Earth system as reservoirs that gain and lose material. Some reservoirs are global in scale, e.g. the atmosphere, oceans, soils and biota. Studies of carbon cycling often examine reservoirs at this scale (e.g. “what is the flux of carbon to the atmosphere from the crustal reservoir?”). However, if the scope of a study is cycling of phosphorus in a small lake, pertinent reservoirs might include the water column (which could contain dissolved P and P fixed in organic matter) and lake sediments (which would likely include P in minerals and P in organic matter). Scale can go all the way down to reservoirs on the molecular scale, e.g., “is the arsenic in this soil partitioned into mineral exchange sites, the aqueous solution, or organic matter?”. Each of these comprises a micro-scale reservoir in that soil system.

Unlike the closed system approach used to characterize the Earth as a whole, reservoirs are **open systems** characterized by exchange of material. The rate of exchange of material between reservoirs is termed the **flux**. Units are typically mass/time, e.g. kg/yr, or molar values such as mol/day. In some cases fluxes are directly measured (e.g. transport of particulate iron to the oceans in rivers), whereas in other cases fluxes are inferred based on measurable changes in reservoirs. Controls on flux include:

- geochemical processes such as the oxidation of CO to CO₂ in the atmosphere (which will at least partially control CO and CO₂ abundance), dissolution of minerals during chemical weathering and sorption of elements from solution into (or onto) minerals;
- biochemical and biotic processes such as photosynthesis, respiration, nutrient uptake by plants and microbially mediated decomposition of dead biota;
- physical processes such as evaporation and condensation, ocean and air currents, and river flow.

The average amount of time an element resides in a reservoir is the **residence time**, a parameter that is controlled by the volume (or mass) of the reservoir and the flux into (and out of) the reservoir (equation 4.6). Clearly, the larger the reservoir and the slower the flux, the longer the residence time; for example, residence time of an average water molecule in the ocean is 3500 years whereas in lakes or rivers, residence time may be more likely on the order of years (small lake) to days or weeks (river).

Steady state

One important assumption associated with residence time calculations is the existence of **steady state**, a situation in which the flux of a given component into the reservoir is equal to the flux out. Thus, at steady state, the concentration of a given element or component *C* will not change with time, a statement that can be represented as:

$$dC/dt = 0 \quad (6.1)$$

where *t* is time. In reality, steady state usually will exist over fixed amounts of time, but over long time periods disturbances typically disrupt steady-state situations. One such example would be the long-term concentration of CO₂ in the atmosphere – over geological time, when measured over time scales of hundreds to thousands of years, CO₂ may appear to be (or have been) at steady state; however, changes in Earth’s orbit (Milankovitch cycles), volcanism, mountain uplift and other factors (notably anthropogenic emissions) will shift the balance of atmospheric CO₂ in one direction or the other. Certainly over times scales on the order of >10⁴ years, CO₂ varies notably and would not be characterized as steady state, yet during the ~10 000 years of the Holocene, atmospheric CO₂ remained at 270 ± 10 ppm, representing a relatively prolonged steady state. Similar analyses of steady

state can be applied to sea water composition, terrestrial weathering rates, global sea level, the pH of a stream, and much more.

Steady state is not equivalent to chemical equilibrium, which is defined more strictly and requires that processes be reversible and that the forward rate is equal to the reverse rate (see Chapter 1). To probe this in more depth, an illustrative example of steady state is the concentration of calcium in sea water. Over time spans of at least thousands of years, the flux of Ca into sea water, mainly via river transport of weathered bedrock constituents, is equal to the rate at which it is removed from sea water, mainly by precipitation of CaCO_3 and incorporation into marine sediments. However, this cannot be considered as an example of *chemical equilibrium* because the reactions, particularly the way by which Ca is added

to the oceans (i.e. river flow), are not reversible, so the oceans are considered to exist “at steady state with respect to calcium concentration.”

Geochemical cycles are often represented by box models, where boxes represent reservoirs and arrows between boxes represent fluxes or, in some cases, processes that are responsible for transfer between reservoirs. We can use simple box models to illustrate the difference between steady state and chemical equilibrium (Fig. 6.1). More-complicated box plots are used to illustrate reservoirs, fluxes and processes in elemental cycling.

If we liken the steady-state scenario depicted in Fig. 6.1 to the concentration of calcium in sea water, *Flux AB* primarily represents river transport of Ca into sea water and *Flux BC* represents chemical precipitation of CaCO_3 out of sea water.

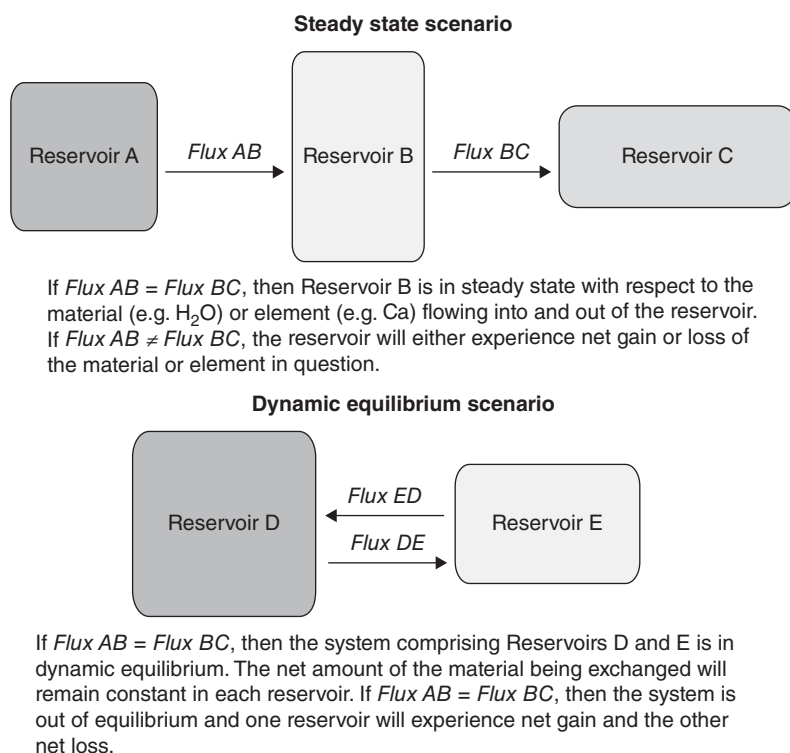


Fig. 6.1 Box model representations of the differences between steady state and dynamic equilibrium.

In this example, Reservoir A would be the terrestrial Ca reservoir, Reservoir B would be the sea water Ca reservoir, and Reservoir C would be the ocean sediment Ca reservoir. The concentration of Ca in sea water will be at a steady state if $\text{Flux AB} = \text{Flux BC}$.

In reality, steady state usually will exist over fixed amounts of time, but over long time periods disturbances, whether natural (e.g. a landslide that changes the volume of a lake) or anthropogenic (e.g. increasing atmospheric CO_2) typically disrupt steady-state situations.

The following sections first present elemental cycling as controlled by natural processes and then consider human alterations to the cycle. Evidence from the geological record provides a reference for pre-industrial reservoir sizes, fluxes and processes, and this background can be compared to measurements of modern-day fluxes to understand anthropogenic impacts. The main geological records used in quantifying the cycling of N, P and S are marine and lacustrine sediments and ice cores. Other records used include corals, paleosols, deep sea water, fossilized organisms, sedimentary rocks (especially marine carbonates and shales) and speleothems (cave deposits).

The main reservoirs to be considered for each of the cycles will be the lithosphere (mainly the crust but the mantle also plays a role in cycling of some elements), terrestrial ecosystems (mainly soils and land plants), the atmosphere and the oceans (which consist of inorganic and organic forms of various elements in the water column and in ocean-bottom sediments). In some cases we will also examine smaller reservoirs such as lakes (especially in the case of nitrogen and phosphorus cycling), rivers, soils and wetlands. For the most part terrestrial water bodies are insignificant reservoirs in terms of global-scale cycles, but can be important in terms of flux (e.g. fluvial transport of elements from terrestrial to marine reservoirs) and ecological ramifications (e.g. nitrogen and phosphorus in fresh water ecosystems).

6.1 THE NITROGEN CYCLE

Humans have profoundly altered the nitrogen cycle, perhaps more so than the cycle of any other element. By the late 1900s we had reached the point where humans had become responsible for a greater amount of reactive nitrogen in forests, lakes, farms and streams than can be produced by natural processes (Vitousek et al., 1997) and the rate of anthropogenic nitrogen fixation continues to increase (Schlesinger, 2009).

The real crux of the problem is that humans are capable of transforming unreactive N_2 (in air) to reactive species such as ammonium, nitrate and NO_x , which then are free to infiltrate terrestrial and aquatic ecosystems. Very few organisms on Earth are capable of breaking the $\text{N}::\text{N}$ bond in N_2 gas (this bond is so strong it renders N_2 nearly inert). Those that do possess this ability include certain species of prokaryotes (bacteria and archaea) to be presented below, yet in many natural systems, nitrogen is the element that limits growth due to the unavailability of reactive nitrogen. What these prokaryotes are able to do is convert stable atmospheric N_2 , a form that is unavailable as a nutrient to nearly all living organisms, to molecules of reactive nitrogen (Nr) that can be used by plants as nutrients.

Given that nitrogen is a limiting nutrient in most agricultural soils, it is added as fertilizer to increase crop yields; reactive nitrogen is also enriched in animal waste, so it also can be introduced to surface water and shallow groundwater from sewage treatment plants, septic systems and animal-feed operations. Furthermore, combustion of fossil fuels and biomass releases reactive nitrogen oxide (NO_x) species into the atmosphere that ultimately result in increased nitrogen in soils and surface waters. Consequences of this shift in nitrogen cycling include:

1 eutrophication of lacustrine and nearshore marine ecosystems and associated loss of plant and animal diversity;

2 acid rain (the nitric acid variety) and related acidification of soils and surface waters;

3 acid-rain-related losses of essential soil nutrients, particularly calcium and potassium that are particularly important in forested ecosystems;

4 increased concentrations of the potent greenhouse gas nitrous oxide (N_2O) in the atmosphere; and

5 increased concentrations of nitrogen oxides that contribute to the formation of photochemical smog.

The following sections present some of the important concepts related to nitrogen species, nitrogen cycling and the environmental impacts of human-caused alteration of the nitrogen cycle.

6.1.1 Nitrogen valence, nitrogen species

Nitrogen occurs in many oxidation states, ranging from the most-reduced form, N^{-3} , to the most-oxidized form, N^{+5} . A neutral nitrogen atom contains seven electrons, five of which are valence electrons configured as $2s^2 2p^3$, meaning that a nitrogen atom can achieve the octet rule by (1) gaining three electrons to produce N^{-3} ; (2) losing five electrons to become N^{+5} , and (3) sharing three pairs of electrons with another nitrogen atom to form N^0 in N_2 . As is shown in Table 6.1, nitrogen can occur in all oxidation states between N^{-3} and N^{+5} , although N^{-2} and N^{-1} are rare or non-existent in nature.

Reduced nitrogen (N^{-3}) occurs in organic matter where it is typically chemically bonded to carbon, and in species that are stable in reducing environments, including the ammonium ion (NH_4^+) and ammonia gas (NH_3). Wetland sediments, crustal environments with low dissolved oxygen, deep marine waters and decaying organic matter all contain ammonium and have the potential to produce ammonia gas. One example of the relationship between ammonium and ammonia occurs in alkaline soils, where NH_4^+ can lose a proton to solution, resulting in NH_3 .

Table 6.1 Some common nitrogen species and the oxidation state of nitrogen in the compounds.

Species	Name	Oxidation State
NH_3	Ammonia	N^{-3}
NH_4^+	Ammonium ion	N^{-3}
$(\text{NH}_2)_2\text{CO}$	Urea	N^{-3}
N_2H_4	Hydrazine ^a	N^{-2}
NH_2OH	Hydroxylamine ^a	N^{-1}
N_2	Molecular nitrogen (N_2 gas)	N^0
N_2O	Nitrous oxide	N^{+1}
NO	Nitric oxide	N^{+2}
HNO_2	Nitrous acid	N^{+3}
NO_2^-	Nitrite anion	N^{+3}
NO_2	Nitrogen dioxide	N^{+4}
HNO_3	Nitric acid	N^{+5}
NO_3^-	Nitrate ion	N^{+5}

^a These compounds are synthetic.

Similarly, when urea is applied as a nitrogen fertilizer, it too is prone to ammonia loss in alkaline, waterlogged soils.

N_2 gas, the most common form of nitrogen in the Earth surface/shallow crustal environment, is the least reactive of the nitrogen species, and it is this attribute that tends to limit the amount of reactive/available nitrogen in terrestrial, aquatic and marine ecosystems. Only in limited environments can N_2 be transformed by natural processes to reactive nitrogen species – these include soils and waters that contain certain bacteria and related microbes (see below), and in situations where lightning or combustion produce NO_x .

Oxidized species of nitrogen that contain N^{+1} , N^{+2} or N^{+3} tend to form in oxidized aqueous environments (e.g. streams, shallow groundwater, aerated soils, oxygenated lake or sea water, etc.), often via oxidation of organic matter or ammonium, or via microbially mediated fixation from N_2 . Of the species presented in Table 6.1, some of the more common ones in the environment include:

1 Nitrous oxide (N_2O), a greenhouse molecule ~300 times more potent than CO_2 ; one important origin of this trace atmospheric gas is denitrification

of fertilizers, notably in tropical and subtropical soils; it also forms via non-anthropogenically enhanced denitrification and nitrification (see below) in soils and marine waters.

2 Nitric oxide (or nitrogen monoxide; NO) and nitrogen dioxide (NO₂), molecules which often form during combustion and are often grouped together as NO_x;

3 Nitric acid (HNO₃), which forms in the atmosphere via the oxidation of NO and NO₂. It is also a product of volcanic eruptions.

4 Nitrate (NO₃⁻), a highly soluble ion that is a product of natural oxidation of organic matter, yet also one whose concentration at the Earth surface is greatly enhanced by anthropogenic activities, especially addition of fertilizers and combustion, where nitrate can form via oxidation of NO_x by-products (more-detailed information on nitrogen gases as atmospheric pollutants is presented in Chapter 8). Nitrate is responsible for eutrophication in many aquatic environments and it is also regulated in public water supplies because it is linked to *methemoglobinemia*, a condition linked to “blue baby syndrome”. Anaerobic bacteria in the stomach reduces nitrate to nitrite, which in turn can oxidize ferrous iron (Fe⁺²) in hemoglobin to ferric iron (Fe⁺³), preventing transport of O₂ in blood. This condition is most pronounced in children younger than six months of age.

As is the case with carbon, nitrogen concentrations are often reported based on the mass of N in a molecule, rather than the mass of the entire molecule. A good example is nitrate, which is commonly reported as NO₃-N; if a water sample contains 6.6 mg/L of dissolved nitrate, it will contain 1.5 mg/L NO₃-N. The conversion factor is based on the fact that there are 14.01 g of nitrogen in one mol of nitrate, which has a mass of 62.01 g. Therefore, to convert the mass of NO₃⁻ to NO₃-N, multiply by 0.226 [14.01 g/mol (N) ÷ 62.01 g/mol(NO₃⁻)]. The USEPA maximum contaminant level for nitrate in public water supplies is 10 mg/L NO₃-N, which corresponds to 44 mg/L NO₃⁻.

6.1.2 Processes operating within the nitrogen cycle

Given the wide range in oxidation states, a very important control on nitrogen cycling is reduction–oxidation reactions, many of which are mitigated by microorganisms listed below in this section. Figure 6.2 presents a schematic sketch of transformations that occur in the nitrogen cycle with an emphasis on soil reactions.

A good starting point for examining biogeochemical transformations in the N cycle is atmospheric N₂, the largest reservoir of nitrogen on Earth. This form of nitrogen is largely unavailable to living organisms; for example, it is not available for uptake by photosynthetic plants, and when humans and other animals inhale air containing N₂, the gas is eventually exhaled in the same, unaltered form. For this reason, animals must obtain nitrogen from proteins and other food sources. Fortunately, nitrogen-fixing microorganisms such

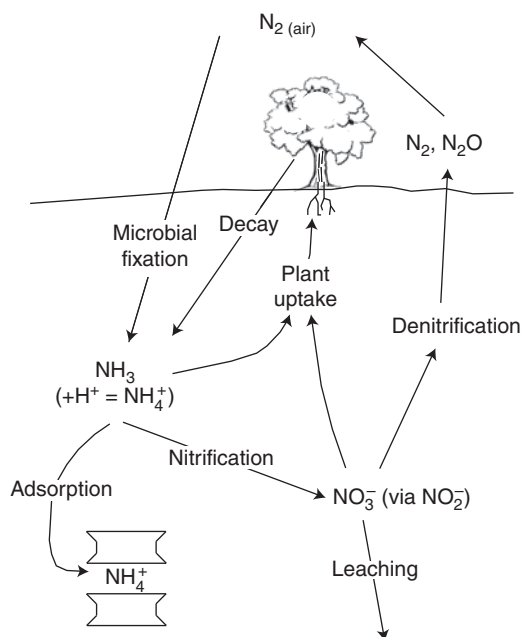
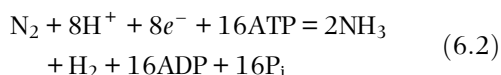


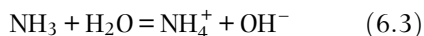
Fig. 6.2 Biogeochemical transformations in the N cycle.

as *Rhizobium*, immobile bacteria which live in root nodules of legumes (e.g. alfalfa, clover, lentils, peas, beans), and *Azotobacter*, a motile genus of soil bacteria, are capable of reducing N^0 (in N_2) to N^{-3} . Often the bacteria form a mutualistic relationship with a plant; by transforming N_2 to NH_3 or NH_4^+ , *Rhizobium* bacteria provide an available form of nitrogen as a nutrient to the plant, and in return, the bacteria obtain carbohydrates from the plant.

The term for the transformation from atmospheric N_2 to plant-available ammonium is **fixation** and it can be summarized in this reaction:

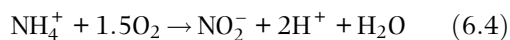


where ATP and ADP are adenosine triphosphate and adenosine diphosphate and P_i is inorganic phosphorous. Ammonia commonly reacts with water to form the ammonium ion:



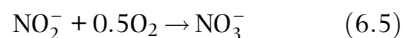
Ammonium is a plant-available form of nitrogen and it may be incorporated into plant roots in this state. If the ammonium cation persists in soil (the likely scenario in waterlogged neutral to acidic soils), it may become adsorbed (e.g. to clays or humus) or it may be transported in solution.

In aerated soils NH_4^+ may undergo oxidation to **nitrite** via a process mediated by bacteria (commonly those of the *Nitrosomonas* species), as follows:



The change in Gibbs free energy for this reaction is -66 kJ/mol ($\Delta G_R^\circ = -66 \text{ kcal/mol}$), so it is spontaneous; however, *Nitrosomonas* bacteria act as a catalyst, increasing reaction rate. What do the bacteria get from this reaction? Electrons (electron transfer is often the motivation for bacterial mediation of chemical reactions). The oxidation of one mole of ammonium to nitrite

releases 6 moles of electrons ($N^{-3} \rightarrow N^{+3} + 6e^-$), and this helps fuel bacterial metabolism. However, nearly all nitrite is further oxidized to **nitrate** in the presence of *Nitrobacter* bacteria:



The oxidation of nitrogen in this case is of a smaller magnitude than for NH_4^+ to NO_2^- – only 2 moles of electrons are released per mole of nitrite oxidized to nitrate ($N^{+3} \rightarrow N^{+5} + 2e^-$). Nitrate is typically ten times more abundant than nitrite in aerated soils and aquatic systems, and the term for the sequential transformation of ammonium to nitrite to nitrate is **nitrification**.

In chemically reducing environments such as waterlogged soils and wetland sediments, nitrates are reduced to N_2 or, less commonly, N_2O , gases that seep back into the atmosphere, effectively completing the nitrogen cycle. This process of transforming mineralized or fixed nitrogen to N_2 or N_2O gas is known as **denitrification**. As with most transformations in the nitrogen cycle, this process is facilitated by bacteria (e.g. *Pseudomonas*) that, in this case, use the nitrate as an electron acceptor during respiration. This is necessary because O_2 gas, the main electron acceptor in oxidizing environments, is largely unavailable in reducing/anaerobic/anoxic environments.

Regardless of whether a plant incorporates nitrogen as NH_4^+ , NO_3^- , or some other form, the main oxidation state of nitrogen in organic matter is N^{-3} . Thus, if nitrogen is incorporated into a plant as nitrate, the plant will reduce the nitrogen from N^{+5} to N^{-3} so that it can become a constituent of organic matter.

Once nitrogen has been incorporated into organic matter, it can either be consumed as N^{-3} (e.g. in proteins) and metabolized by animals (a process often termed **assimilation**), or it can undergo a series of transformations to other oxidation states, all of which are higher (less negative) than -3 . Thus, any reaction that

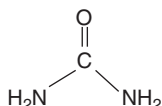
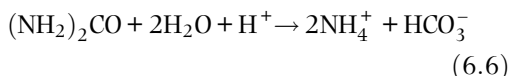


Fig. 6.3 Structural formula of urea, $(\text{NH}_2)_2\text{CO}$.

changes the oxidation state of nitrogen in organic matter will be an oxidation reaction.

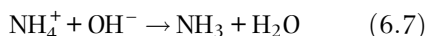
The following example summarizes the fate of organic matter undergoing decomposition in soils using the simple compound urea (Fig. 6.3) to illustrate nitrogen in organic matter. The degradation of urea and eventual oxidation of ammonium released by urea breakdown helps to illustrate the types of reactions associated with release of nitrogen from organic matter in soils and waters.

The first step, which occurs in the presence of *urease* bacteria, is hydrolysis of urea to yield two ammonium atoms plus bicarbonate. Note that in this first step, the oxidation state of nitrogen is unchanged (it is N^{-3} in both urea and ammonium).



The general term for any reaction, e.g. **decay** of plant matter, that releases available nitrogen (or any nutrient) to soil is **mineralization**; this specific case is known as **ammonification**.

At this point, the ammonium ion could lose a proton to form ammonia, which could then be volatilized into the atmosphere. Given the abundance of OH^- in alkaline solutions, this process tends to occur in alkaline soils (reaction 6.3 in reverse):



Combustion is another common process responsible for the transformation of nitrogen from reduced species to oxidized species. Reduced nitrogen in organic matter is oxidized to species such as NO and NO_2 during combustion, but this is only a

small part of nitrogen emitted by combustion (e.g. of biomass, hydrocarbons or coal). The majority of N released to air comes from N_2 in the atmosphere that combines with atmospheric O_2 in the elevated temperatures of the internal combustion engine or forest fire to form nitrogen oxides. Two examples of nitrogen – oxygen reactions that occur during combustion are:



In reaction 6.8, nitrogen is oxidized from N^0 to N^{+2} ; in reaction 6.9, nitrogen is oxidized from N^0 to N^{+4} (Table 6.1). The species produced in these reactions are generically referred to as NO_x . High temperatures associated with lightning in the atmosphere also foster formation of NO_x .

In solution, nitrogen species are typically quantified and classified as either **dissolved inorganic nitrogen (DIN)** or as **dissolved organic nitrogen (DON)**. DIN typically takes the form of low molecular weight compounds such as ammonium (NH_4^+), nitrite (NO_2^-) and nitrate (NO_3^-) whereas in DON, nitrogen is bonded to carbon in more-complex, high molecular weight organic molecules (e.g. humic acid). DIN is typically more-available for uptake by plants in terrestrial and aquatic ecosystems (and farms, lawns and golf courses), and for this reason chemical fertilizers typically contain ammonium or nitrate, or both (ammonium nitrate is a common fertilizer). Organic fertilizers (e.g. compost, manure) commonly contain more DON than DIN, meaning that the nitrogen is released slowly over time.

6.1.3 Global scale reservoirs and fluxes of nitrogen

The unit of measure used to describe the abundance of nitrogen in Earth surface reservoirs is the teragram, or Tg (10^{12} g). As was the case for carbon (Chapter 5), values represent the

mass of nitrogen rather than the mass of the molecule. Fluxes are given in teragrams of N per year (TgN/yr).

Of the globally important N reservoirs at or near the Earth's surface, the one with the overwhelming majority of nitrogen is the **atmosphere** (Fig. 6.4), which contains 4×10^9 TgN (3×10^{20} moles of N), nearly all of it as N_2 gas. Approximately 10^3 Tg of atmospheric N ($< 0.01\%$ of total N in air) exists as the more-reactive species NO, NO_2 , N_2O and NH_3 . The residence time of N_2 in the atmosphere can be computed from reservoir size (4×10^9 Tg) and flux of N_2 to terrestrial and marine ecosystems

via fixation ($\sim 3 \times 10^2$ Tg/yr), yielding a T_{res} value of $\sim 10^7$ years. Residence times of NO_x and NH_3 are far shorter and average from days to weeks; residence time of N_2O in the atmosphere is ~ 100 years.

Natural processes capable of transferring nitrogen out of the atmosphere to other reservoirs include biological fixation (mediated by bacteria and archaea) and lightning, where temperatures are sufficiently high to facilitate breaking of $N::N$ and $O::O$ bonds and fostering formation of NO_x species that can dissolve in rain and be delivered to terrestrial or marine environments. Combustion (e.g. forest fires) is also a natural means by which

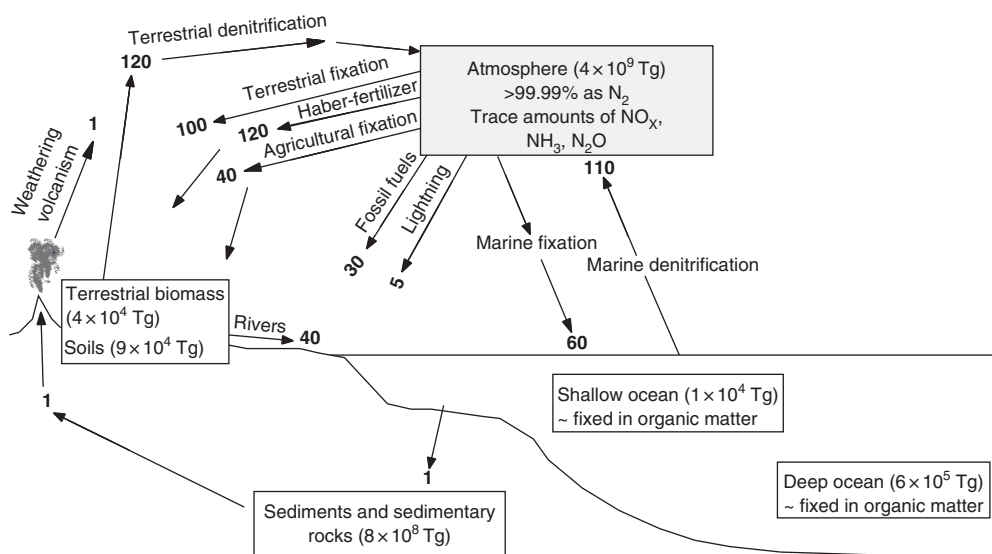


Fig. 6.4 Schematic sketch of the nitrogen cycle showing main reservoirs (in boxes) and fluxes (arrows). Numbers in bold type indicate flux between reservoirs in TgN/year, and the mass of nitrogen in each reservoir is given in Tg. “Terrestrial Fixation” represents fixation of N to natural ecosystems by *Rhizobium* and other species. “Agricultural Fixation” applies to N fixed into crops by *Rhizobium* and other species (in mutualisms with legumes). “Haber-fertilizer” represents transformation of N_2 to NH_3 via the Haber–Bosch process – this flux is often effectively synonymous with industrial N fixation. “Marine fixation” represents nitrogen incorporated into oceans by biological fixation and rainfall (e.g. dissolved NH_4^+ , NO_x); nitrogen is also introduced to marine environments from fossil-fuel combustion-derived and lightning-derived NO_x and fluvial transport of terrestrially derived N. Data are from Galloway et al. (2004) and Schlesinger (2009) and sources cited therein, and are to be considered as estimates given large uncertainties, especially in fluxes (notably marine denitrification). The net flux in the modern N cycle clearly reflects the enrichment of terrestrial ecosystems in N via transformation of N_2 to reactive nitrogen species (from Haber–Bosch process/fertilizer manufacture and formation of NO_x from combustion of fossil fuels).

N_2 can be transformed to reactive species and removed from the atmosphere. The net flux of nitrogen out of the atmosphere and into terrestrial and marine environments is causing major impacts to ecosystems, but the mass of N_2 in the atmosphere is so great that this transfer is effectively undetectable in the concentration of nitrogen in air.

After the atmosphere, the next largest reservoir of nitrogen in the Earth surface and near-surface system encompasses the sedimentary rocks of the shallow crust, with $\sim 8 \times 10^8$ Tg. The main forms of nitrogen in this reservoir are ammonium cations (e.g. NH_4^+ can substitute for K^+ in mica interlayers) and organic matter; coal, for example, typically contains ~ 0.1 to 1% N. The crustal reservoir of N is largely disconnected from the more-surficial N reservoirs (atmosphere, soils, terrestrial and marine ecosystems), but combustion of fossil fuels does release reactive nitrogen to the atmosphere and, ultimately, ecosystems.

The **oceans** contain $\sim 6 \times 10^5$ TgN, the vast majority of which exists as inorganic nitrogen and as nitrogen fixed into organic matter in the deep ocean. Lesser amounts are fixed into organic matter in the shallow ocean. Fixation of nitrogen into marine environments occurs by bacterially mediated reduction of N_2 to NH_3 , by atmospheric deposition of NH_4^+ and NO_x , and by fluvial inputs of reactive nitrogen (especially NO_3^-) from terrestrial environments. The oceans currently receive 45 to 50 TgN per year in the form of NO_3^- , NO_x and related reactive N species produced by fossil fuel combustion, fluvial transport of fertilizer run-off, and other anthropogenic sources (Duce et al., 2008).

Terrestrial ecosystems (including soils) contain $\sim 1.3 \times 10^5$ TgN in sub-reservoirs of terrestrial biota (4×10^4 TgN) and soils (9×10^4 TgN). The steady-state flux between terrestrial ecosystems and atmosphere is ~ 100 TgN/yr, but addition of fertilizers, NO_x from fossil fuel combustion and other forms of Nr have increased the flux from terrestrial systems to the atmosphere. Leaching of nitrate and other Nr species

from soils to rivers and ultimately lakes and shallow marine environments contributes to eutrophication, a topic covered in the following section.

Although not shown as a reservoir given the ephemeral nature of fluvial systems, **rivers** are an important component of the nitrogen cycle. Estimates of pre-industrial N flux via rivers from terrestrial to marine ecosystems indicate a value of ~ 20 TgN/yr; current riverine N fluxes are approximately 60 TgN/yr. Some of the reactive N transported by rivers is denitrified in the nearshore marine environments, which returns N_2 . If fertilizer use continues to increase, fluvial transport of nitrogen to lakes and shallow marine ecosystems are also likely to increase (Schlesinger, 2009).

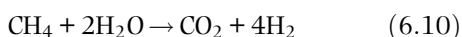
6.1.4 Human perturbation of the nitrogen cycle and resulting environmental impacts

The root of environmental problems pertaining to alteration of the nitrogen cycle all stem from the same source – transformation of molecular nitrogen (N_2) to reactive nitrogen. Three main mechanisms by which N_2 is transformed to Nr include:

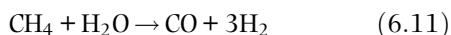
- 1** manufacture of nitrogen fertilizers by the Haber–Bosch process and then application to fields);
- 2** fixation of reactive nitrogen into agricultural soils by cultivation of legumes;
- 3** formation of nitrogen oxides as by-products of fossil-fuel combustion.

Prior to the 20th century, the main sources of nitrogen-bearing fertilizers were organic (as opposed to inorganic), including animal manure, bat and seabird guano and composted organic matter. In fact, meters-thick deposits of N-rich guano were exported globally (especially from Peru) in the 19th century to satisfy fertilizer and explosive needs. By 1909, Fritz Haber had discovered a process for obtaining ammonia from air

(N₂) and H₂. Soon after, the German chemical company BASF purchased the rights to the process and by 1913 Carl Bosch was able to increase Haber's approach from benchtop scale to industrial scale. Initially, the ammonia developed by the Haber–Bosch method was used to manufacture fertilizer, but soon after was used to make explosives for World War II. The origin of the hydrogen used in the Haber–Bosch process is methane, which is purified to remove impurities (e.g. H₂S, SO₂) and reacted as follows:



or



The hydrogen is then reacted with nitrogen gas under high temperature and pressure (~400 °C and ~200 bar) to form ammonia:



At this point, the ammonia can be converted to ammonium (by reaction with H⁺) or nitrate (by oxidation) to form N-rich fertilizers such as NH₄NO₃, KNO₃ and urea.

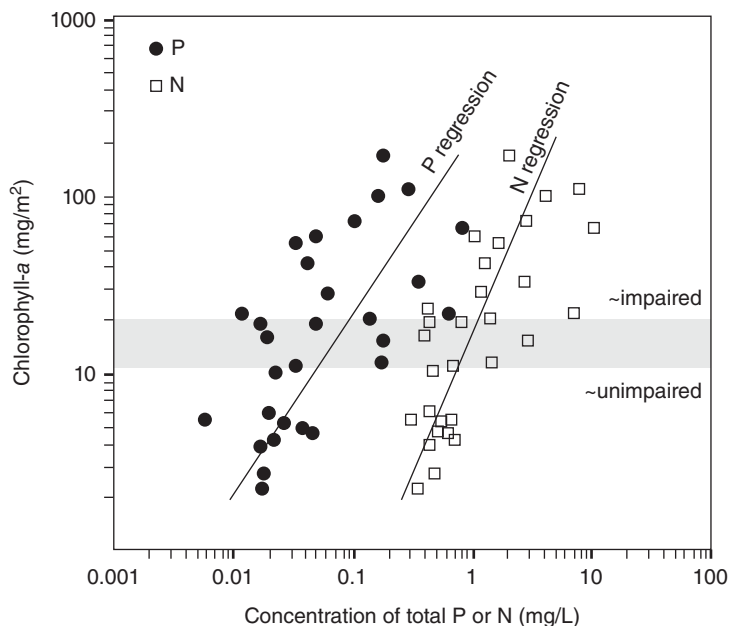
The two main limiting nutrients in soils and aquatic ecosystems are nitrogen and phosphorus – nitrogen due to the fact that the most abundant form occurs as relatively unreactive N₂ gas, and phosphorus due to its general insolubility in soils and waters – so as a result, farm fields, residential lawns, golf courses, gardens, etc. are commonly fertilized with N and P to increase productivity. Since the mid-1900s, the most common origin of the nitrogen used in fertilizers is fixation from N₂ by the Haber–Bosch process, and the N molecules added to fields are either the highly mobile nitrate anion, or are molecules such as ammonium and urea, species which are readily transformed to nitrate in the oxidizing conditions of most soils.

Due to its high solubility and low potential to be adsorbed to soil particles, nitrate is easily leached from soils and transported into surface water bodies and shallow groundwater. (It also tends to leach out of sewage treatment plants, feedlots and septic systems.) Given that low nitrogen concentrations in surface waters limits algal growth, anthropogenic activities that increase concentrations of nitrate and other Nr species in surface waters tend to trigger increased algal growth. In many lakes and shallow marine waters, increasing Nr is directly proportional to algal growth (Fig. 6.5) the result of which is a sequence of changes to the ecosystem termed eutrophication (from Greek: “eu” = good, enriched or sufficient, “trophic” = nutrient related).

When the reactive nitrogen is anthropogenically derived, the term **cultural eutrophication** is applied to distinguish this phenomenon from natural eutrophication that occurs in some wetlands, shallow lakes and coastal marine ecosystems, where deep, cold, oligotrophic (“oligo” = low) lakes can eventually evolve into eutrophic lakes or swamps as they fill in with sediment and organic matter over time.

Eutrophic lakes, bays, shallow marine waters (especially near river deltas), and rivers (especially slow-moving ones) are characterized by abundant growth of blue-green algae known as cyanobacteria that often form extensive mats on the water surface (Plate 8). The algae blooms themselves can be toxic due to the presence of various cyanotoxins including microcystins, anatoxin-a and saxitoxins (Graham et al., 2010), but in an ecological sense, the greatest damage results when the algae, which have short life-spans, die and decay in the water. Decomposition of the algae is carried out by aerobic bacteria that consume oxygen, and the greater the abundance of algae undergoing decay, the greater the amount of O₂ consumed from the water column by the aerobic decomposers. This cyclic process depletes the water of available oxygen and causes the death of other organisms, particularly fish and

Fig. 6.5 Correlation between phosphorus and chlorophyll-*a*, and between nitrogen and chlorophyll-*a*, in streams of the northeastern United States. The shaded area represents the transition zone between unimpaired streams and impaired streams in the northeastern US. Note that the axes are logarithmic and that the x-axis is in units of milligrams per liter (ppm). (Data from Risken et al. 2003.)



amphibians. In the oceans, these affected areas are known as **marine dead zones**.

Notable water bodies affected by cultural eutrophication are the delta of the Mississippi River in the Gulf of Mexico, Chesapeake Bay and San Francisco Bay in the USA, Lake Erie in the USA, shallow near-shore waters of the Baltic Sea, the Black Sea (from nutrients transported by the Danube River), parts of the South China Sea, and far too many other water bodies to list here. It is also important to remember that phosphorus is also a key cause of eutrophication, and the sources and behavior of P will be covered below in the section on the phosphorus cycle.

An example of the relationship between N fertilizer application and response was illustrated in a study carried out in Wisconsin from the 1960s through the 1990s (Portage County Groundwater Conditions, 2000), where application of fertilizer was compared to stream and groundwater nitrate concentrations. The results of this study are summarized in Fig. 6.6, which shows two graphs, one of nitrogen fertilizer sales in Wisconsin, and the other of nitrate

concentrations in baseflow of the Little Plover River. The measurements of NO_3^- in the Little Plover River were made in January when the majority of stream flow is derived from soil and groundwater influx to the stream; thus, measuring the stream during baseflow conditions provides a proxy for soil and shallow groundwater nitrate flux. The increase in nitrate shown in Fig. 6.6 is typical for streams from across the globe, and is illustrative of the changes in surface water chemistry that lead to eutrophication. It also reflects the fact that N fertilizers are being applied at rates far greater than the potential for N uptake by plants, leaving excess nitrate available for leaching into groundwater and for runoff into streams, lakes, estuaries and other surface water bodies.

Changes in global nitrogen cycling are presented in Fig. 6.7; these graphs illustrate the following:

- 1** Notable increases in nitrous oxide in the atmosphere (this record is derived from the composition of air bubbles trapped in glacial ice).
- 2** Increases in nitrogen oxides (NO_x) produced by fossil fuel combustion throughout the 20th

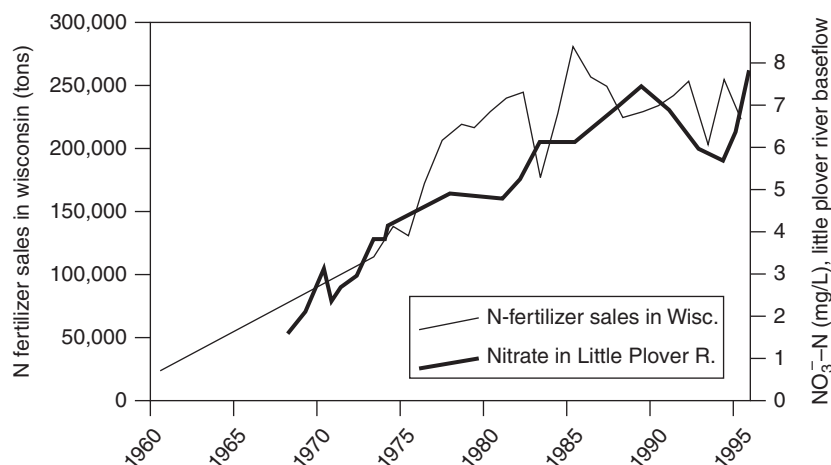


Fig. 6.6 Relationship between nitrogen fertilizer use and dissolved nitrate in the Little Plover River of central Wisconsin. Based on data from Portage County Groundwater Conditions (2000). Note that nitrate data are presented as $\text{NO}_3^- \text{--N}$, meaning that only the mass of nitrogen in nitrate is included in the calculation. Sales of nitrogen fertilizers in Wisconsin parallel national and global trends.

century, as well as a recent tapering of the curve due to emissions controls in automobiles, coal plants and other sources.

3 Increased agricultural additions of reactive N to the landscape, both from increased manure production from farm animals as well as the industrial transformation of N_2 into reactive nitrogen in fertilizer via the Haber–Bosch process. It is important to note that the curves for agricultural N are similar to population growth curves, and that additions of nitrogen to crops have far exceeded crop uptake of nitrogen since the early 20th century. The excess nitrogen (i.e. that not fixed into crops) is largely available for leaching into groundwater and surface water, where it can cause human health problems (i.e. methemoglobinemia) and ecological problems (i.e. eutrophication).

Nitrous oxide emissions are mainly derived from agricultural soils where N-bearing fertilizers are converted to N_2O – in 2008 in the United States, 68% of N_2O emissions were related to “agricultural soil management” (US EPA; <http://www.epa.gov/nitrousoxide/sources.html>), so it is not surprising that the N_2O curve and the

agricultural N production curves are similar. It must also be pointed out that N_2O forms naturally via marine nitrification and is produced in sewage treatment plant lagoons and natural wetlands. N_2O is not easily decomposed in soils or in the troposphere, and most decomposition takes place in the stratosphere; as a result, atmospheric concentrations of N_2O have increased ~20% since the mid-1800s (Fig. 6.7). In the stratosphere, N_2O is decomposed by reaction with oxygen radicals that acts to lower concentrations of beneficial stratospheric ozone.

The NO_x emissions represented in Fig. 6.7 are ultimately added to the terrestrial landscape and to nearshore marine ecosystems when the NO_x is oxidized and dissolved in rain, and then delivered to the Earth’s surface (typically) as nitric acid. The high K_a of nitric acid (2.4×10^1) means that it nearly completely dissociates, resulting in delivery of dissolved H^+ and nitrate to soils and aquatic ecosystems. Details of N_2O and NO_x reactions in the atmosphere are presented in Chapter 8, and issues associated with acid rain are presented in Chapters 8 and 9.

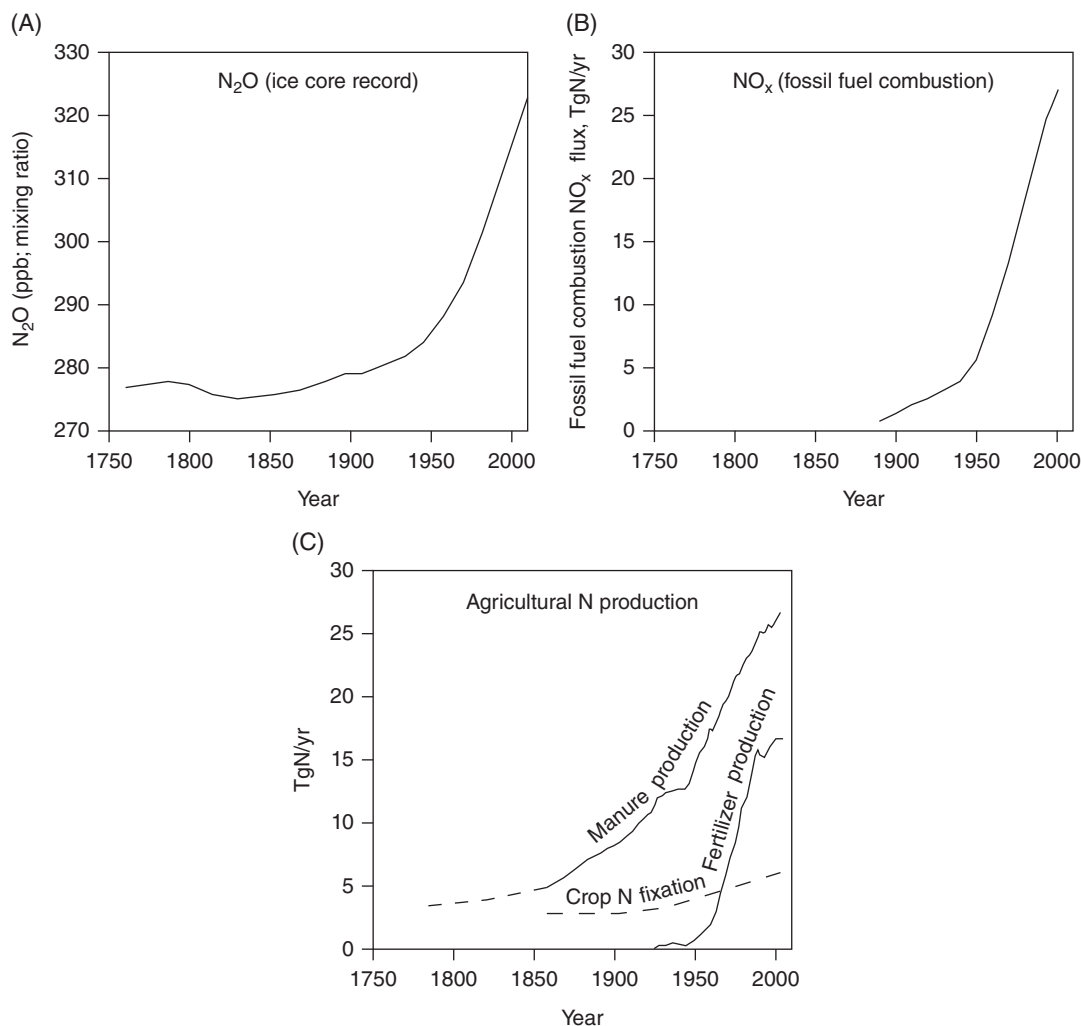


Fig. 6.7 Global trends in sources and cycling of reactive nitrogen. (A) increasing atmospheric nitrous oxide recorded in ice cores; (B) increasing emissions of nitrogen oxides from fossil fuel emissions ($\sim 2/3$ of total NO_x emissions); (C) increases in production of fertilizers derived from Haber–Bosch process and manure, as well as uptake of N by crops. Note that (A) represents atmospheric *concentration* of N_2O while (B) and (C) present *fluxes*. (MacFarling Meure et al. 2006 and Holland et al. 2005.)

6.2 THE PHOSPHORUS CYCLE

The phosphorus cycle is much simpler than the carbon and nitrogen cycles for two reasons. First, in natural environments at or near the Earth's surface, phosphorus has no significant gas phases—it occurs in solid inorganic minerals and organic solids as well

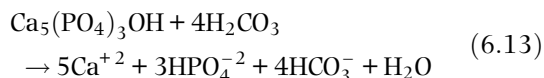
as sorbed and dissolved forms (it can also occur in a molten state in magma or lava). Second, phosphorus has only one valence state in nature, P^{+5} , so it is not prone to the wide array of reduction–oxidation reactions that occur in the C and N (and S) cycles. Like carbon and nitrogen, phosphorus is critical to living organisms; it is a crucial component of DNA

and RNA and essential in the transmission of chemical energy via adenosine triphosphate and diphosphate molecules (ATP and ADP).

The main primary source of P is apatite, $\text{Ca}_5(\text{PO}_4)_3(\text{OH}, \text{Cl}, \text{F})$, a calcium phosphate mineral that occurs in nearly all rock and sediment types at concentrations typically in the range of 0.1 to 1%. The OH^- , Cl^- and F^- anions substitute for each other and the F-rich end-member, fluorapatite, is the hardest (most resistant to chemical breakdown) of the apatite varieties. Given that tooth enamel is essentially comprised of apatite, you can surmise why fluoride is added to toothpaste. A rare but very P-rich rock type known as phosphorite, a sedimentary rock of deep marine origin, often contains 20–40% P_2O_5 (~10–20% apatite) and is mined as a source of P for fertilizers. Other phosphate minerals are far more rare than apatite and include turquoise, $\text{CuAl}_6(\text{PO}_4)_4(\text{OH})_8 \cdot 5\text{H}_2\text{O}$; the uranium-bearing mineral autunite, $\text{Ca}(\text{UO}_2)_2(\text{PO}_4)_2 \cdot 10\text{H}_2\text{O}$; and monazite (Ce, La, Nd, Th, Y) PO_4 , a mineral useful to geochronologists for dating of igneous rocks.

6.2.1 P cycling in soils

A good starting point for the P cycle is chemical weathering of hydroxy-apatite in soils.



The phosphate species produced in this reaction depend on pH (Fig. 6.8). In typical pH ranges for soils, the dominant forms are H_2PO_4^- in the pH range of 2 to 7, and HPO_4^{2-} in the pH range of 7 to 12. In very acidic situations (e.g. acid mine drainage), H_3PO_4 may occur, and under extremely alkaline conditions, the PO_4^{3-} ion is predicted to dominate.

While P is nearly ubiquitous in rocks and soils (it typically comprises 200–500 mg/kg in

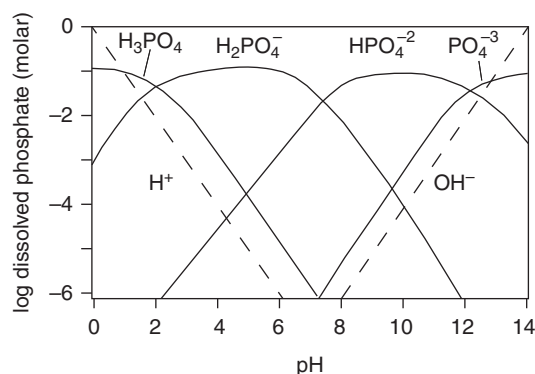


Fig. 6.8 Variation in speciation of P as a function of pH.

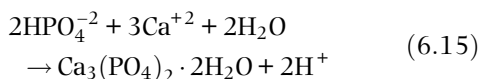
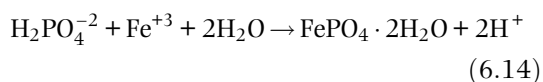
rock or soil), two factors combine to make it a limiting nutrient in soils and aquatic ecosystems: **1** Phosphorus is insoluble or only sparingly soluble under most conditions, so in soils the phosphate anion (PO_4^{3-}) tends to form poorly crystalline hydrated solids by bonding with calcium, iron or aluminum to produce species with compositions that are approximately $\text{Ca}_3(\text{PO}_4)_2 \cdot 2\text{H}_2\text{O}$, $\text{FePO}_4 \cdot 2\text{H}_2\text{O}$ and $\text{AlPO}_4 \cdot 2\text{H}_2\text{O}$. Ca-phosphate tends to occur in alkaline soils whereas Fe- and Al-phosphate are more likely to form in acidic soils. Phosphate solubility is also limited by its incorporation into hydroxides of Al and Fe (Fig. 6.9).

2 Apatite is a relatively stable mineral that is resistant to erosion, meaning that P is only slowly released to the soil solution during chemical weathering. The solubility product of apatite is $\sim 10^{-58}$, a value that would produce an equilibrium concentration of dissolved P of only $\sim 10^{-8}$ molar (Schlesinger, 1997).

Thus, the consequence of these two factors is that P is dominantly bound in the solid phase, including well-crystalline apatite and poorly crystalline soil hydroxides and phosphates. A general rule of thumb regarding speciation of P is that in soils, terrestrial surface waters and groundwater, the ratio of particulate P to dissolved P is at $\sim 10:1$ (Howarth et al., 1995), with much of the dissolved P occurring in organic acids (e.g. fulvic

acid), where the organic acid functions as a chelate. This ratio can vary, however, depending on source (e.g. P tends to be more easily dissolved from organic matter), pH (P is *least insoluble* between pH of ~ 6 and 8) and location. Open ocean water, where nearly all P is dissolved, is a major exception to this rule of thumb (recall that Fe tends to be the limiting nutrient in open marine waters).

Incorporation of P into soil solids is illustrated in the following two reactions that depict formation of pedogenic (soil-formed) iron and calcium phosphates following weathering of apatite or release of P from organic matter or other solids:



Given that P is typically adsorbed onto surfaces of iron hydroxides or incorporated into hydroxides as inner-sphere complexes, the relationship is better demonstrated schematically than with a chemical reaction (Fig. 6.9):

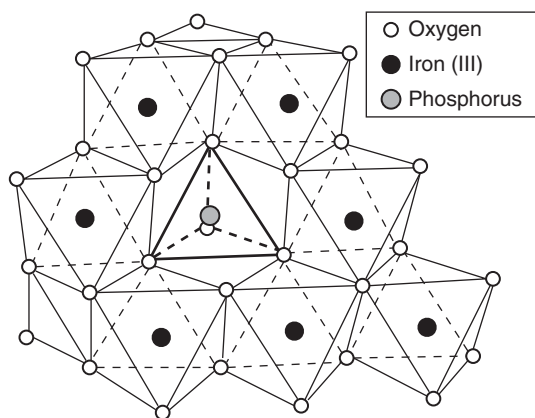


Fig. 6.9 Schematic sketch of showing tetrahedral PO_4^{-3} molecule incorporated into an iron hydroxide structure. This sketch is based on a ferrihydrite structure determined by Michel et al. (2007).

Another important point regarding P cycling between solid phases and dissolved forms is that while P itself does not undergo reduction or oxidation, $\text{Fe}(\text{OH})_3$ and $\text{FePO}_4 \cdot 2\text{H}_2\text{O}$ are prone to dissolution in reducing sediments and waters that foster reduction of Fe(III) to Fe(II). Given that the majority of phosphorus transport to lakes and estuaries occurs as P sorbed to iron hydroxides (Fig. 6.9), the behavior of P in these water bodies will largely be controlled by biogeochemical conditions that affect the stability of iron hydroxides. This was first recognized in the 1930s and 1940s (e.g. Einsele, 1936; Mortimer, 1941) and has become the paradigm for understanding P mobility in aquatic systems.

If the reduction–oxidation conditions at the sediment/water boundary are oxidizing, the P is likely to remain sorbed to the iron hydroxides; however, if the conditions are reducing (anoxic), which is commonly the case in sediments at the bottoms of lakes and estuaries, iron will be reduced, the iron hydroxide will dissolve and P will be released to solution and become available for biological uptake. $\text{Fe}(\text{OH})_3$ that was stable in the oxygenated soil and fluvial environment will be prone to $\text{Fe}(\text{III}) \rightarrow \text{Fe}(\text{II})$ reduction and release of P to solution in anoxic sediments. This mechanism is at least partly responsible for availability of P in lacustrine and estuarine sediments.

The P cycle in soils and terrestrial ecosystems is summarized in Fig. 6.10. The largest reservoir in the near-surface and Earth surface system is sedimentary rocks and sediments, but this P is largely unavailable for biological use. Most of the P available for plant uptake is contained in the organic P reservoir, i.e. it is contained within decaying organic matter, humus, fulvic acids, microbial organisms and related organic reservoirs. Any P not fixed immediately by plants is likely to be incorporated into inorganic minerals or non-crystalline solid phases such as Fe-, Al-, or Ca-phosphates and iron hydroxides; however, at the same time, many plants exude organic

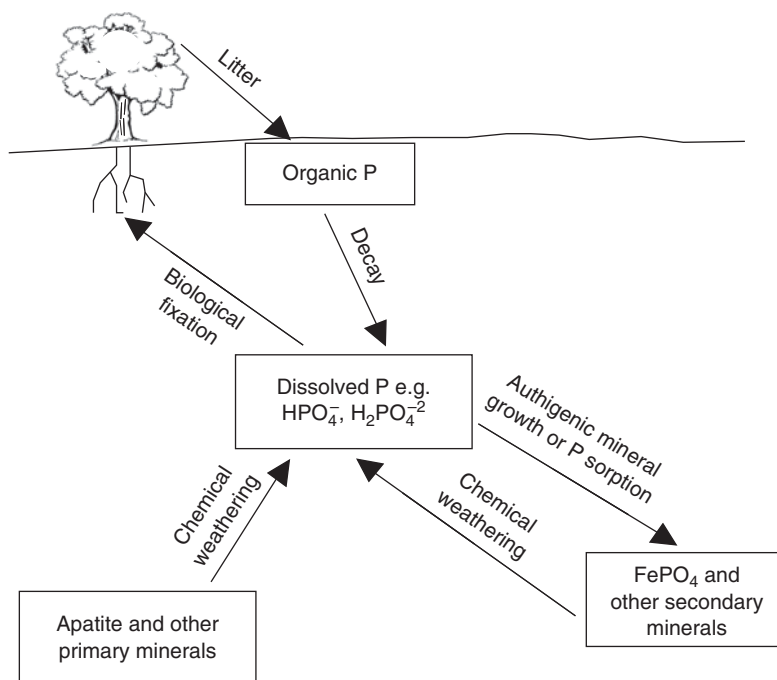


Fig. 6.10 Schematic sketch of the phosphorus cycle in soils. Due to the insolubility of phosphorus, most P cycling in terrestrial ecosystems takes place between organic P and biota.

acids, chelating compounds and enzymes that are capable of extracting P from inorganic and organic solid phases, so there certainly is some flux between secondary minerals and biota.

6.2.2 The global phosphorus cycle

While the largest reservoir of P at and near the Earth's surface is sedimentary rock and sediments, the main means by which P is transported at the Earth's surface is by stream flow, and given that P is largely insoluble, the majority of P is transported as particulate material. Approximately 22 TgP/yr are transported by rivers to lakes and to the global ocean, and ~90% is carried as adsorbed P or as P minerals (e.g. phosphates) – this component is mainly deposited in deltas and other types of marine sediments near continents. To reiterate a point made above, the

main particulates responsible for P transport in the solid state are iron hydroxides and particulate organic matter (Ruttenberg, 2007). The dissolved P carried by rivers is ultimately transported by currents into the open ocean, where it can be taken up as a nutrient or can circulate at the surface or to deeper levels.

Physical and chemical weathering of rock and sediment plus decomposition and erosion of organic matter yields ~20 TgP/yr to biota and soils, and this flux is matched by erosion of soil P and biologically bound P into streams and, ultimately, the ocean. The cycle is closed by burial of marine P into marine sediments and the eventual transformation of these sediments in rock. In general, the flux of P between the main reservoirs (sediment/rock → soils/biota → streams → ocean → sediment/rock) is ~20 TgP/yr, although as is indicated on Fig. 6.11, the riverine flux is about 10% greater, reflecting increased erosion due to

anthropogenic activities, including road building, construction, logging, agriculture and application of mined phosphorus as fertilizer. The increased fluvially transported P is responsible for much eutrophication.

In the open ocean, phosphorus is dominated by dissolved species – this is because the particulate P carried by rivers is mostly deposited in river deltas and on the continental shelf (Paytan and McLaughlin, 2007). In the shallow oceans, ~1–5% of total P is contained within biota, but the overwhelming majority of shallow marine P is dissolved (Ruttenberg, 2007). The flux of P between dissolved and biotic P reservoirs in the shallow oceans is $\sim 9 \times 10^2$ Tg/yr – this flux is not shown on Fig. 6.11. Upwelling of deep marine dissolved P (~ 60 Tg/yr) is locally responsible for productive marine ecosystems (e.g. the northwest coast of South America), and this flux is ~ in steady state with downwelling of dissolved P and biological P from the shallow ocean.

Two final points worth emphasizing are (1) unlike carbon and nitrogen, the atmospheric reservoir is very small and the only significant way that P is transported in air is as dust, and

(2) the reservoir of phosphate rock which can be mined for P comprises $\sim 1.5 \times 10^4$ Tg; of this, ~ 12 TgP are mined each year, resulting in transfer of P from the sedimentary rock reservoir to soils, biota, river flux and surface water bodies.

6.2.3 Phosphorus and eutrophication

Numerous studies throughout the 20th and 21st centuries have documented terrestrial and marine ecosystems where P is the main limiting nutrient to growth, at trophic levels ranging from phytoplankton to crops to trees. P deficiencies can be problematic in agricultural soils, and thus P is one of the target elements applied in fertilizer, whether it is chemical fertilizer derived from phosphate mining, or P contained in manure or other organic forms. However, erosion of P from stream banks, farm fields, construction sites and numerous other sources has increased input of P to aquatic systems ranging from rivers and lakes to estuaries and the oceans. The role of P as a limiting nutrient with profound effects on lake ecology is illustrated in Fig. 6.12, a graph that plots

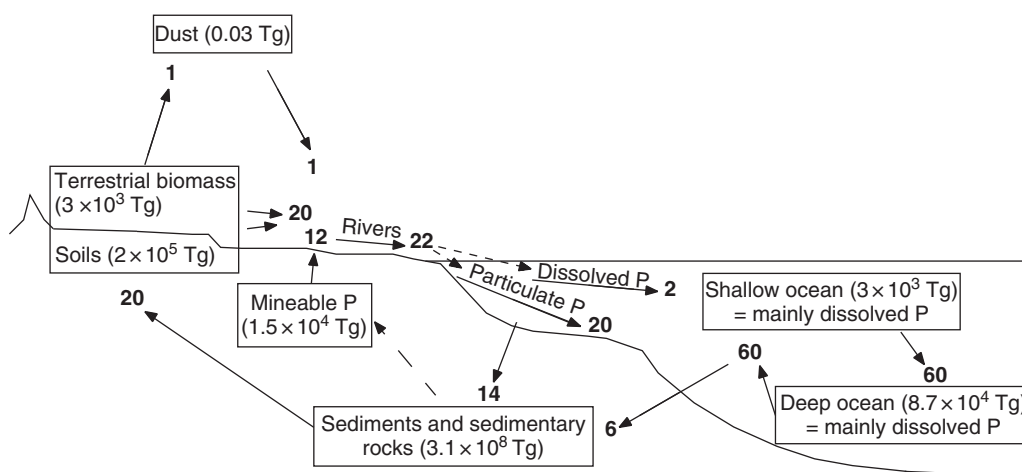


Fig. 6.11 The global P cycle. The unit for reservoir sizes is TgP and fluxes are in TgP/yr. Data sources are from Schlesinger (1997) and Ruttenberg (2007).

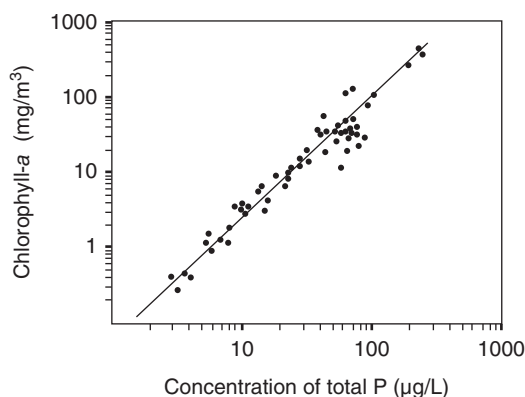


Fig. 6.12 Plot of total P vs. chlorophyll-*a* for dozens of lakes from different ecosystems globally. In each case, total P is the amount of P present in lake water at the time of spring overturn; chlorophyll-*a* concentrations reflect peak summer abundances. Data are from Sakamoto (1966), Dillon and Rigler (1974) and sources cited therein. Note that the *x*-axis in this diagram is in units of micrograms per liter (ppb).

total P against chlorophyll-*a* for dozens of lakes studied globally. What this graph shows is that, regardless of climate, geology, biota and any other factors, P concentration controls primary productivity in aquatic ecosystems.

6.3 COMPARISON OF N AND P

Data from numerous streams in New England (USA) help to illustrate the effect of P on aquatic ecosystems, and also accentuate some of the main differences between phosphorus and nitrogen in aquatic systems. Figure 6.5 plots the concentrations of total P and total N separately against the abundance of chlorophyll-*a* in water, where chlorophyll-*a* is a useful proxy for abundance of cyanobacteria/algae commonly associated with eutrophication. Both elements clearly exhibit a positive correlation with chlorophyll-*a*, but an important difference is apparent – growth of chlorophyll-*a* (i.e. bacteria and algae) occurs at P concentrations that are approximately an

order of magnitude lower than what is required of N to produce equal amounts of bacteria and algae. Total P concentrations of 0.1 mg/L (ppm) are likely to produce an aquatic ecosystem of impaired quality, whereas the transition from unimpaired to impaired occurs at ~ 1 mg/L nitrogen.

Although they both tend to be limiting nutrients in most terrestrial and aquatic ecosystems, the differences between N and P are striking. Included among these differences are:

- P has no significant gaseous phase, whereas N is the dominant gas in the atmosphere.
- The primary source of N is the atmosphere, whereas the primary source of P is rock.
- Nitrogen-fixing microbes can introduce reactive N to soils and aquatic ecosystems from air; however, if P is lacking, it cannot be fixed from air and must be introduced to natural systems by rivers, chemical weathering or wind-blown dust.
- P only occurs in the +5 state and is prone to reduction–oxidation controls only if it is sorbed into iron hydroxides or other redox-sensitive phases, whereas N has numerous oxidation states and thus is prone to strong reduction–oxidation controls on speciation.
- Phosphorus tends to be insoluble and partitioned into the particulate fraction (i.e. silt, clay), whereas in surface waters nitrogen typically occurs as the highly soluble nitrate anion.
- While both are limiting nutrients in aquatic ecosystems, P appears to be the truly limiting nutrient in most cases – this has been demonstrated experimentally by Schindler et al. (2008), who found that even in lakes with *excess P and limited N*, eutrophication occurred because the cyanobacteria were able to fix N from the atmosphere in the absence of abundant nitrate in the water. In lakes with *excess N but limited P*, eutrophication did not occur (P cannot be fixed from air via microbially mediated reactions).
- In many estuaries and in the ocean, N may be the true limiting nutrient to eutrophication (Carpenter, 2008).

These differences are important when considering their common environmental impact, eutrophication. Because P is partitioned into the solid phase in terrestrial surface waters, its transport to surface waters is potentially easier to address than that of nitrate, whose high solubility and resistance to sorption (due to its negative charge) make it very mobile. Drainage ditches lined with grasses, hay bales or small retention ponds will tend to foster settling of P-bearing particulates, while the highly soluble nitrate is largely unaffected – it will largely remain in solution and percolate through hay bales or dams and migrate downgradient in solution. Also, whereas the particulate forms of P are effectively immobile in soil and groundwater, nitrate is highly mobile in the subsurface and is capable of contaminating groundwater and of travelling along subsurface flow paths towards surface water bodies. For these reasons, efforts to mitigate nutrient transport to eutrophication-sensitive water bodies are often focused on limiting P transport.

6.4 THE SULFUR CYCLE

Like C and N, sulfur is a highly mobile element with widely ranging oxidation states (from S^{-6} to S^{+2}) and numerous gaseous phases. Its cycling is strongly controlled by microbial processes, and unlike N or P, S often controls fate and transport of heavy metals and arsenic in rocks, soils and waters. Like C, N and P, its cycling has been altered by human activities and one of the main consequences is acid precipitation in the form of sulfuric acid derived from fossil-fuel combustion (especially coal) and in the refining of metals contained in sulfide ores.

6.4.1 Sulfur valence, sulfur species

Given its position in column VI of the periodic table and its valence electron structure of $3s^2 3p^4$, we can predict that S will occur in +6 and –2 valence states, i.e. it can satisfy the octet rule by losing 6 electrons or gaining 2 electrons,

However, in addition to S^{+6} and S^{-2} , it also occurs as S^{+4} , S^{+2} and S^0 . Table 6.2 presents the numerous oxidation states of sulfur and provides examples of many S-bearing compounds.

In most of the compounds shown in Table 6.2, sulfur occurs in a single oxidation state; for example, in galena (lead sulfide), sulfur occurs as the sulfide (S^{-2}) anion; in sulfur dioxide, sulfur is S^{+4} , and in sulfates, sulfur occurs as S^{+6} . This is not the case with pyrite, where the oxidation state of sulfur is best characterized as S_2^{-2} . This is the bisulfide ion, a molecule comprised of two covalently bonded S atoms whose overall charge is –2. The *average* charge on sulfur is –1. The same concept applies to arsenopyrite, where, in a unit cell, an arsenic atom substitutes for a sulfur atom to create As–S pairs whose sum charge is –2.

The sulfide anion and the compounds it forms are generally stable in reducing environments of wetland sediments, low- O_2 crustal environments and deep marine environments. The sulfide anion also occurs in biological molecules, so in living organisms S is mainly found as S^{-2} . The other extreme of S oxidation states is S^{+6} , which in most compounds combines with oxygen to form the sulfate anion. Sulfate-bearing compounds, including sulfate minerals (e.g. gypsum, jarosite) and sulfuric acid, are most stable in oxidized environments such as rivers, the atmosphere, shallow ocean water and oxidized soils. Many of the intermediate oxidation states occur as metastable intermediary phases, notably SO_2 and SO_3 (SO_x), compounds that form during combustion and oxidation of combustion products in air, and also are components of volcanic emissions; these gases are oxidized in the atmosphere to H_2SO_4 . Sulfuric acid typically dissociates completely, or nearly so, to $2H^+ + SO_4^{2-}$; this is because the first dissociation constant for H_2SO_4 is very high (2.4×10^6), and the second dissociation constant for sulfuric acid (the K_a for $HSO_4^- = H^+ + SO_4^{2-}$ is 1×10^{-2}) is high enough that it results in only small amounts of HSO_4^- (the bisulfate anion) in most natural waters (Brimblecombe, 2007).

Table 6.2 Examples of sulfur species and the oxidation states of sulfur in these compounds, presented from most-reduced, S^{-2} -bearing species to the most-oxidized, sulfate-bearing species.

Species	Name	Oxidation State(s) of sulfur
PbS	Galena (lead sulfide)	S^{-2}
ZnS	Sphalerite	S^{-2}
CuFeS ₂	Chalcopyrite	S^{-2}
HgS	Cinnabar	S^{-2}
(CH ₃) ₂ S	Dimethyl sulfide (DMS)	S^{-2}
(CH ₃) ₂ SO	Dimethylsulfoxide (DMSO)	S^{-2}
(CH ₃) ₂ S ⁺ CH ₂ CH ₂ COO ⁻	Dimethylsulfoniopropionate (DMSP)	S^{-2}
C ₇ H ₁₄ N ₂ O ₂ S	Aldicarb ^a	S^{-2}
H ₂ S	Hydrogen sulfide	S^{-2}
COS or OCS	Carbonyl sulfide	S^{-2}
CS ₂	Carbon disulfide	S^{-2}
FeS ₂	Pyrite (iron sulfide)	(S ₂) ⁻²
FeAsS	Arsenopyrite	(AsS) ⁻²
S ₈	Elemental sulfur	S^0
SO	Sulfur monoxide	S^{+2}
H ₂ SO ₃	Sulfurous acid	S^{+4}
SO ₃ ⁻	Sulfite anion	S^{+4}
SO ₂	Sulfur dioxide	S^{+4}
SF ₆	Sulfur hexafluoride ^a	S^{+6}
SO ₃	Sulfur trioxide	S^{+6}
BaSO ₄	Barite (barium sulfate)	S^{+6}
CaSO ₄ · 2H ₂ O	Gypsum (calcium sulfate)	S^{+6}
KFe ³⁺ (OH) ₆ (SO ₄) ₂	Jarosite	S^{+6}
H ₂ SO ₄	Sulfuric acid	S^{+6}
SO ₄ ⁻²	Sulfate anion	S^{+6}

^a These compounds are synthetic.

6.4.2 The global S cycle

A schematic of the sulfur cycle is presented in Fig. 6.13. The dominant reservoir of S is the lithosphere (2×10^{10} TgS), especially pyrite in sediments and in rocks; lesser amounts of S in the lithosphere occur in sulfate minerals (e.g. gypsum) in evaporite deposits. Most of the S in the shallow lithosphere is fixed from sea water into marine sediments as pyrite or other sulfide minerals. Cycling of S from the lithosphere to the next largest reservoir, terrestrial biota and soils, occurs via chemical weathering of sulfides and the associated transformation of S^{-2} to SO_4^{-2} , a form that is

soluble and readily available for plant uptake and transport through soils and surface waters.

S is also transferred from the lithosphere to soils, biota and the atmosphere by volcanism. The main natural terrestrial sources of S to the atmosphere are volcanism, dust and wetland-derived reduced S gases (e.g. hydrogen sulfide); however, anthropogenic activities currently add 3 to 5 times as much S to the terrestrial atmosphere than do natural processes. Rivers transport dissolved sulfate to the oceans, at which point the sulfate may remain in solution, become fixed as a nutrient by marine biota, or undergo reduction to sulfide in anoxic sediments, which can result in

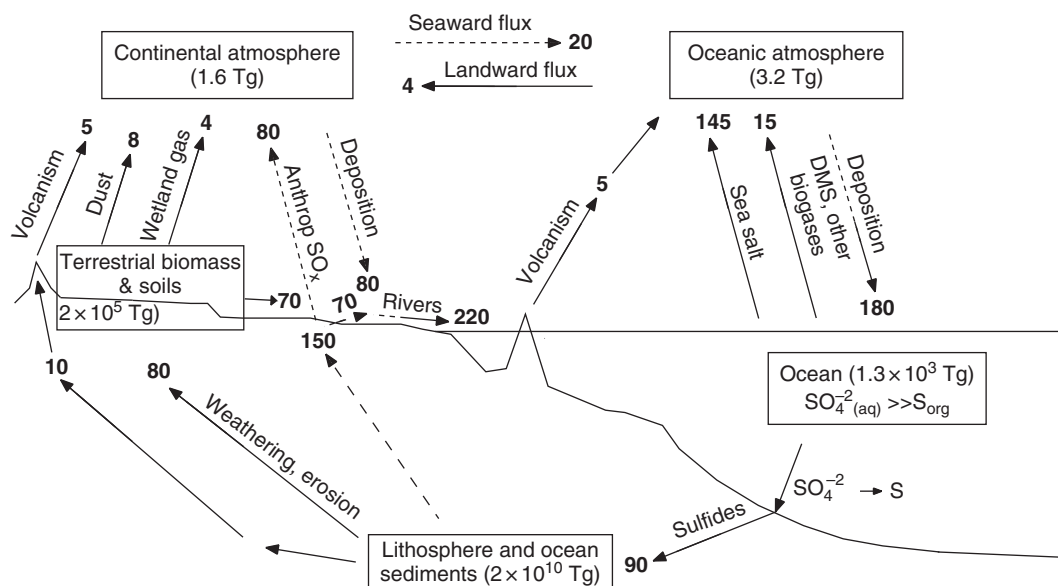


Fig. 6.13 Global sulfur cycle with reservoir masses given in TgS and fluxes in TgS/yr. Fluxes with strong anthropogenic inputs are shown with dashed lines. Of the 150 Tg of sulfur transferred from the crust (fossil fuels, metal mining and refining, fertilizers) to the Earth's surface, 80 TgS/yr are emitted to the atmosphere (mainly by fossil-fuel combustion) and 70 TgS/yr are transferred to soils, lakes, rivers and other terrestrial pools. Data sources are Ivanov (1981), Schlesinger (1997) and Brimblecombe (2007).

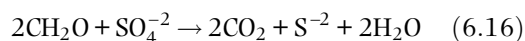
the full-cycle transfer of S back to the lithosphere as pyrite. Within the ocean-oceanic atmosphere system, decaying phytoplankton release dimethyl sulfide (DMS) to the atmosphere, and most of it is rapidly returned to ocean water.

6.4.3 The marine S cycle

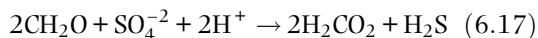
The starting point for any cycle is perhaps somewhat arbitrary, but for sulfur a good starting point is dissolved sulfate in sea water. Sulfate comprises 2712 mg/L of sea water and the main source of this dissolved sulfate is river influx, current estimates of which are 220 teragrams of S per year (TgS/yr). At least one-half of the current river flux is anthropogenic (Brimblecombe, 2007). Atmospheric deposition onto the oceans also delivers a large amount of sulfur to this reservoir, but most

of this is sulfate sea spray that is blown from sea to air and then back to sea again, so it is not a net source of S to ocean water (Fig. 6.13). The oceans also receive a net atmospheric flux of 16 TgS/yr of anthropogenic sulfur (mainly SO_x) carried by prevailing winds from land to sea.

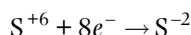
The main net flux of S out of sea water is caused by sulfate reduction that ultimately results in formation and burial of sulfide minerals (mainly pyrite or its disordered diagenetic precursor mackinawite); these reactions typically take place at the sediment/sea water interface in chemically reduced, shallowly buried marine sediments in the presence of organic matter (approximated as CH_2O), or at hydrothermal vents. In the presence of organic matter and anaerobic microorganisms, sulfate is chemically reduced as follows:



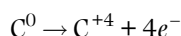
or



The individual sulfur reduction reaction is:



and the paired oxidation reaction (of carbon) is

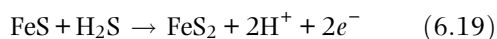


The reactions are balanced in terms of electrons by the fact that two moles of carbon undergo oxidation. Anaerobic bacteria of the genus *Desulfotomaculum* and genus *Desulfovibrio* use sulfur as an electron acceptor, allowing the organisms to obtain energy while also enabling them to act as catalysts for sulfate reduction.

In marine sediments that contain sufficient amounts of Fe^{+2} (i.e. most marine sediments), disordered iron monosulfide will form as a metastable phase before recrystallizing into pyrite in the presence of hydrogen sulfide (Hurtgen et al., 1999):



then



The formation of pyrite and hydrogen sulfide in anoxic marine sediments is shown schematically in Fig. 10.14; also depicted is the fractionation of S isotopes that takes place during this process – isotopically heavy sulfur (^{34}S) is partitioned into pyrite in sediments whereas isotopically light sulfur (^{32}S) is partitioned into hydrogen sulfide, some of which migrates out of sea water and into the atmosphere. Some H_2S may also participate in the FeS to FeS_2 transformation. Pyrite exerts a strong control on cycling of trace metals such as Cu, Pb and Zn because these metals can

substitute for the Fe^{+2} atom; arsenic can substitute for sulfur atoms in pyrite, in some cases reaching thousands of mg/kg As in the pyrite (Smedley and Kinniburgh, 2002). When As and S are in equal abundance, the mineral is termed arsenopyrite.

If the pyrite that forms in marine sediments (and also at hydrothermal vents) is buried and compressed into rock, the sulfur is removed from sea water is sequestered into the crustal S reservoir.

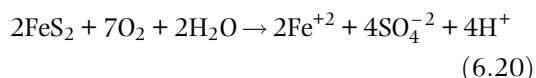
In addition to the fixation of S into marine sediments, another important component of marine S cycling is biological fixation of dissolved sulfate by phytoplankton (e.g. *Dinophyceae* and *Prymnesiophyceae*). The S^{+6} in dissolved sulfate is reduced to S^{-2} and fixed into dimethylsulfoniopropionate (DMSP), a compound that regulates osmosis and other biological functions. When phytoplankton are consumed by higher trophic-level organisms (e.g. zooplankton) or die and undergo decay, DMSP decomposes to dimethyl sulfide (DMS) and methanethiol (CH_3SH). Of these two S-bearing degradation products, methanethiol rapidly disappears, perhaps incorporated into organic matter or degraded by oxidation; most DMS is microbially degraded or fixed into biota but a small amount seeps into the atmosphere where it comprises an important part of marine atmospheric chemistry. The flux of 15 TgS/yr to the atmosphere is the largest single natural source of S to air. For reference, current (early 21st century) emissions of SO_x from fossil-fuel burning and refining and sulfide ore metal refining (and atmospheric oxidation products), are approximately 5 times greater than the DMS flux.

Other trace biogases emitted to the atmosphere from the oceans include carbonyl sulfide and hydrogen sulfide, but their contributions are at least an order of magnitude less than DMS (Schlesinger, 1997). DMS is important in the oceanic atmosphere because (1) in air it is oxidized to sulfate aerosols that act as cloud condensation nuclei, and (2) sulfate aerosols formed

via oxidation of DMS also act to scatter incoming solar radiation, thus contributing a cooling forcing to the climate system. The residence time of DMS in the atmosphere is ~ 2 days, so the majority of DMS flux to the oceanic atmosphere is returned quickly to the oceans, commonly as oxidation products that are re-deposited into the ocean or coastal regions in precipitation ($\sim$ rain).

6.4.4 Soils and biota

Chemical weathering of pyrite-bearing rocks following uplift and erosion introduces sulfur to soils and biota, a process that can be summarized by this weathering reaction (and that is shown in different formulations in equations 1.34, 2.11/2.12, and 4.29):



Note that oxidation of sulfur, not ferrous iron, drives this weathering reaction. The bacteria *Thiobacillus thiooxidans* and *Thiobacillus ferrooxidans* accelerate weathering rates of pyrite but often their effects are most strongly pronounced in acid mine drainage situations. The -2 charge on sulfate (SO_4^{-2}) indicates that it is more-strongly sorbed to anion-exchange sites (particularly sites associated with Fe hydroxides or humus) than are monovalent anions like Cl^- and NO_3^- , but less so the trivalent PO_4^{-3} anion; as a result, sulfate is moderately mobile in soils and available for leaching to surface waters, groundwaters or for plant uptake.

Ultimately, in most soils, sulfur becomes incorporated into biota or is bound to humus and other forms of particulate and dissolved organic matter. When incorporated into plants, most sulfate undergoes assimilatory reduction to S^{-2} so that the sulfur can be bound to carbon in, for example, the amino acids cysteine and methionine. Decomposition of organic matter in

soils commonly results in sequestration of reduced S in humus or, in some cases, can result in oxidation of sulfide to sulfate. Much sulfur is cycled back-and-forth between reduced sulfur in organic matter in soil, sulfate in the soil solution and reduced sulfur in plant tissue.

Sulfur is transferred from soils and terrestrial biota to the atmosphere by two main mechanisms. The first is simply by physical erosion of particles into air by wind. While the annual flux of S from soils and terrestrial biota is relatively small at 8 TgS/yr, it is important climatically because sulfate particles act as cloud condensation nuclei and function as scatterers of incoming solar radiation. The second flux of S from terrestrial ecosystems to atmosphere occurs as reduced gases such as hydrogen sulfide, carbonyl sulfide and DMS that are produced in, and emanate from, wetland sediments and waterlogged soils.

6.4.5 Atmosphere

Due to short residence times of sulfur species in the lower troposphere, most models of the global S cycle treat the terrestrial and oceanic atmospheres as separate reservoirs with relatively small fluxes between the two. 4 TgS/yr migrate from the oceanic atmosphere to the terrestrial atmosphere, and 20 TgS/yr migrate in the opposite direction and most of this flux is anthropogenic S derived from fossil-fuel combustion and refining.

From Fig. 6.13 it is clear that the greatest flux associated with the terrestrial atmosphere is addition of SO_x to the atmosphere from **anthropogenic emissions** – approximately 80% of sulfur emissions to the terrestrial atmosphere are from anthropogenic activities, including:

- 1 coal combustion (typically S comprises 1–3% of coal);
- 2 petroleum refining and combustion; and
- 3 refining of metal-sulfide ores.

The greatest impacts of increased S emissions are negative consequences for human health

(respiratory tract issues like asthma) from sulfur dioxide, and acid rain derived from the formation of sulfuric acid via precursor SO_x and other oxidizable sulfur gases. These topics will be covered in greater detail in Chapters 8 and 9.

The next most abundant net source of S to the atmosphere is volcanic eruptions. Prior to the industrial era, volcanic gases were the main sources of S to the terrestrial troposphere, but now this natural source accounts for less than 10% of total S flux to the terrestrial atmosphere. Volcanic gases include a wide variety of SO_x phases, hydrogen sulfide, and less commonly, carbonyl sulfide and carbon disulfide (Brimblecombe, 2007). These gases generally oxidize to sulfate and dissolve in precipitation although very strong (Plinian) volcanic eruptions can eject sulfates to the stratosphere, where they can persist for 2 to 3 years, exerting a short-term cooling effect on mean global climate – the eruption of Mount Pinatubo in the Philippines in 1991 and the resulting 3 years of lowered mean global temperature is a good example of this effect.

The oceanic atmosphere is less affected by anthropogenic activities than is the terrestrial atmosphere. The majority of SO_x pollution is deposited on land, much as H_2SO_4 acid rain. Within the oceanic atmosphere – ocean system, most biologically derived DMS is rapidly cycled from the atmosphere back into sea water, so little DMS reaches land. Given that DMS production is a function of marine primary productivity (i.e. phytoplankton abundance), and that warming oceans may alter primary productivity, there are concerns about future production rates of DMS.

Possible links between ocean warming and DMS are related to projected changes in ocean dynamics; for example, warmer waters may cause greater ocean stratification, which would decrease upwelling, decreasing nutrient availability and leading to a reduction in marine net primary production, which ultimately would yield a decrease in the DMS production in the ocean and less DMS in air. And, given that DMS

scatters incoming solar radiation and fosters formation of cloud condensation nuclei, less DMS would mean less reflection of incoming radiation and greater warming, potentially initiating a feedback loop where more warming $\rightarrow$ less DMS $\rightarrow$ more warming $\rightarrow$ less DMS, etc. This has become known as the CLAW hypothesis after the four authors who researched the topic (Charlson et al., 1987). Recent modeling by Cameron-Smith et al. (2011) indicates that a warming ocean will cause shifts in DMS productivity, with decreasing production in the tropical ocean and increasing production poleward.

6.4.6 River flux

The greatest flux in the modern global sulfur cycle is the fluvial transport of dissolved and particulate S from land to the oceans, and at least 50% of current river S flux is anthropogenic – pre-industrial riverine S flux was likely ~ 100 TgS/yr (Brimblecombe, 2007), compared to a current global flux of ~ 220 TgS/yr. Natural sources of sulfate in rivers include deposition of volcanic gases and wetland-derived gases plus sulfate derived from pyrite weathering and various organic forms of S including humus, DMS and more. The main anthropogenic species in river water are SO_x and H_2SO_4 derived from fossil-fuel combustion, mining and processing of sulfide ores, and to a lesser extent, organic and inorganic forms of S derived from fertilizers. The S compounds are delivered to streams in many ways, including overland flow, throughflow via soils and groundwater, erosion of stream banks and soils, and direct inputs of SO_4^{2-} from acid mine drainage sites.

In summary, the greatest disturbances to the global sulfur cycle are related to fossil-fuel combustion and mining and processing of metal sulfide ores, and the greatest impacts have been to forested ecosystems and lakes affected by sulfuric acid deposition, a topic to be covered in Chapters 8 and 9. Concerns have also been raised about the

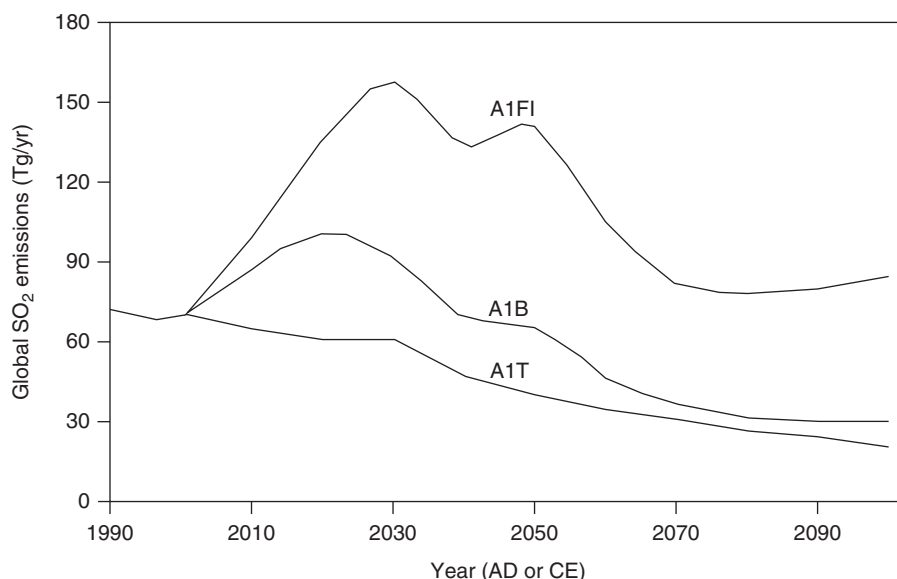


Fig. 6.14 Three scenarios for SO₂ emissions into the 21st century. Each scenario assumes increased population growth to 9 billion by 2050. The difference in the scenarios is that A1FI assumes continued intensive fossil-fuel combustion, especially coal; A1B assumes a balance of fuel sources; and A1T assumes nearly complete introduction of non-fossil-fuel energy sources. Source: IPCC Special Report on Emissions Scenarios. (Nakicenovic and Swart 2000. Reproduced with permission of GRID-Arendal.)

fate of DMS in the oceanic atmosphere given projections of warming oceans and changing oceanic dynamics. On the positive side, anthropogenic sulfur emissions are predicted to decrease into the 21st century (Fig. 6.14) as the result of improved technologies and greater social, economic and political will.

6.5 INTEGRATING THE C, N, P AND S CYCLES

The biogeochemical cycles of these four elements are linked in many ways and understanding these connections is crucial to comprehending and predicting geochemical processes at and near the Earth's surface. Outlined below are some of the main concepts and mechanisms that connect these four critical elements:

1 C, N, P and S are essential elements in living organisms.

2 All four elements are cycled abundantly between biota and organic and inorganic forms.

3 While all four are important components of terrestrial and marine ecosystems, only C, N and S have significant atmospheric fluxes. Very little P is cycled through the atmosphere.

4 P is relatively immobile compared to most species of C, N and S, and as a result it should be easier to control its movement from soils to aquatic systems.

5 All four elements play roles in global climate, from the obvious effects of infrared-active CO₂, CH₄ and N₂O in trapping heat, to solar radiation-scattering SO₄²⁻ in exerting a cooling influence, to the less-obvious role that P plays in controlling primary productivity and hence uptake of CO₂; the topic of global climate is covered in Chapter 7.

6 N and P are limiting nutrients that often control primary productivity in terrestrial and aquatic ecosystems, and hence affect the cycling

of C (through controls on photosynthesis) and S (via plant uptake) as well.

7 Formation of carbonic acid in the atmosphere exerts a strong control on weathering rates and release of P and S to soil solution; formation of sulfuric acid (anthropogenic rate $\gg$ natural rate) can enhance cycling of all elements by speeding up chemical weathering and oxidation of organic matter.

8 Various species of C, N and S are released to air during combustion of fossil fuels, so this one anthropogenic activity greatly affects the cycles of these three elements.

9 Improving environmental quality of terrestrial and aquatic ecosystems and the atmosphere will require better management of all four of these critical elements.

There are clearly other important biogeochemical cycles, including those of oxygen, hydrogen, calcium, arsenic, lead, mercury, various pesticides, fuels and solvents, and many more, but detailed analysis of these topics is not within the scope of this book; however, pertinent concepts related to the behavior and cycling of other elements and compounds are presented in various sections, and a quick glance at the index at the end of the book will direct the reader to specific sections on these components, including their presence and speciation in minerals, their behavior in weathering environments, their occurrence in organic compounds (synthesized or natural), their roles in atmospheric and aqueous chemistry, and more.

QUESTIONS

1. Describe the difference between steady state and dynamic equilibrium and provide an example of each.
2. What is the most-reduced oxidation state of N? Name a compound in which it occurs. Similarly, indicate which is the most-oxidized state of N, and name a compound in which it occurs. Describe why microbes exert such a strong

control on N transformations and use an chemical equation to illustrate your point.

3. What is the most-reduced oxidation state of S? Name a compound in which it occurs. Similarly, indicate which is the most-oxidized state of S, and name a compound in which it occurs.
4. What are the main ways by which humans have altered the S cycle? In particular, indicate the reservoirs from which humans are extracting S, the chemical reactions responsible for its transfer to other reservoirs, and two environmental problems resulting from altered S cycling.
5. Compare the solubility of nitrate and phosphate in soil. Values of the K_{sp} for pertinent species at 25 °C include: $AlPO_4 = 6.3 \times 10^{-19}$, $CaHPO_4 = 1 \times 10^{-7}$, $FePO_4 = 1.3 \times 10^{-22}$, and while not commonly determined for soluble salts, the K_{sp} for KNO_3 is $\sim 10^{-1}$. Write the chemical equation associated with the K_{sp} for each of these 4 reactions. Given these data, in what form(s) in P most likely to occur and in what form(s) is N most likely to occur?
6. How do microbes facilitate nitrification and denitrification?
7. Given that both P and N can cause eutrophication, indicate by what process(es) N and P commonly enter aquatic bodies. Show these pathways with a schematic sketch.
8. Humans are responsible for a greater amount of reactive nitrogen in forests, lakes, farms and streams than is produced by natural processes – what are the main origins of reactive nitrogen (i.e. how is it fixed, both naturally and anthropogenically, into soils and aquatic bodies), and what are its main forms?

REFERENCES

- Brimblecombe, P., 2003. The Global S Cycle. Chapter 14 in *Treatise on Geochemistry*, Volume 8, W.H. Schlesinger, ed, Elsevier B. V., Amsterdam, The Netherlands.

- Cameron-Smith P., Elliott S., Maltrud M., Erickson D., and Wingenter O., 2011. Changes in dimethyl sulfide oceanic distribution due to climate change. *Geophysical Research Letters* 38: L07704, 5 pp. doi:10.1029/2011GL047069.
- Carpenter, S.P., 2008. Phosphorus control is critical to mitigating eutrophication. *Proceedings of the National Academy of Sciences USA* 105: 11039–11040.
- Charlson, R.J., Lovelock, J.E., Andreae, M.O., and Warren, S.G., 1987. Oceanic phytoplankton, atmospheric sulfur, cloud albedo and climate. *Nature* 326: 655–661.
- Dillon, P.J. and Rigler, F.H., 1974. The phosphorus-chlorophyll relationship in lakes. *Limnology and Oceanography* 19: 767–773.
- Duce, R.A., LaRoche, J., Altieri, K. et al. (28 authors), 2008. Impacts of anthropogenic atmospheric nitrogen on the open ocean, *Science* 320: 893–897.
- Einsele, W., 1936. Über die Beziehungen des Eisenkreislaufs zum Phosphorkreislauf im eutrophen See. *Archive for Hydrobiology* 29: 664–686.
- Galloway, J. N., F. J. Dentener, D. G. Capone, E. W. Boyer, R. W. Howarth, S. P. Seitzinger, G. P. Asner, C. Cleveland, P. Green, E. Holland, D. M. Karl, A. F. Michaels, J. H. Porter, A. Townsend, and C. Vörösmarty. 2004. Nitrogen Cycles: Past, Present and Future. *Biogeochemistry* 70: 153–226.
- Graham, J.L., Loftin, K.A., Meyer, M.T., and Ziegler, A.C., 2010. *Cyanotoxin mixtures and taste-and-odor compounds in cyanobacterial blooms from the Midwestern United States*: Environmental Science and Technology, doi:10.1021/es1008938.
- Holland, E. A., Lee-Taylor, J., Nevison, C., and Sulzman, J., 2005. Global N Cycle: Fluxes and N₂O Mixing Ratios Originating from Human Activity. Data set. Available on-line [http://www.daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.
- Howarth, R.W., Jensen, H.S., Marino, R., and Postma, H., 1995. Transport to and processing of P in near-shore and oceanic waters. In H. Tiessen (ed.), *Phosphorus in the Global Environment*, pp. 323–245. Wiley, New York.
- Hurtgen, M.T., Lyons T.W., Ingalls, E.D., and Cruse, A.M., 1999. Anomalous enrichments of iron monosulfide in euxinic marine sediments and the role of H₂S in iron sulfide transformations: Examples from Effingham Inlet, Orca Basin, and the Black Sea. *American Journal of Science*, 299: 556–588.
- Ivanov, M.V., 1981. The Global Biogeochemical Sulfur Cycle. In Likens, G.E. (ed.), *Some Perspectives of the Major Biogeochemical Cycles*, pp. 61–78. SCOPE Report 17. New York, Wiley.
- MacFarling Meure, C., Etheridge, D., Trudinger, C., Steele, P., Langenfelds, R., van Ommen, T., Smith, A., and Elkins, J., 2006. Law Dome CO₂, CH₄ and N₂O ice core records extended to 2000 years BP. *Geophys. Res. Lett.* 33: L14810.
- Michel, F.M., Ehm, L., Antao, S.M., Lee, P.L., Chupas, P.J., Liu, G., Strongin, D.R., Schoonen, M.A.A., Phillips, B.L., and Parise, J.B., 2007. The structure of ferrihydrite, a nanocrystalline material. *Science* 316: 1726–1729. (June 22, 2007). DOI: 10.1126/science.1142525.
- Mortimer, C.H., 1941. The exchange of dissolved substances between mud and water in lakes. *Journal of Ecology* 29: 280–329.
- Nakicenovic, N. and Swart, R. (eds.), 2000. Special Report on Emissions Scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge, UK: Cambridge University Press, 612 pp. Available on-line at: http://www.grida.no/publications/other/ipcc_sr/?src=/climate/ipcc/emission/.
- Paytan, A. and McLaughlin, K., 2007. The oceanic phosphorus cycle. *Chemical Reviews* 107 (2), 563–576. DOI: 10.1021/cr0503613/
- Riskin, M.L., Deacon, J.R., Liebman, M.L., and Robinson, K.W., 2003. Nutrient and Chlorophyll Relations in Selected Streams of the New England Coastal Basins in Massachusetts and New Hampshire, June–September 2001. U.S. Geological Survey Water-Resources Investigations Report 03–4191, 15 pp.
- Ruttenberg, K.C., 2003. The global phosphorus cycle. Chapter 13 in *Treatise on Geochemistry*, Volume 8, W.H. Schlesinger, ed, Elsevier B. V., Amsterdam, The Netherlands.
- Sakamoto, M. 1966. Primary production by the phytoplankton community in some Japanese lakes and its dependence on lake depth. *Archiv fuer Hydrobiologie* 62, 1–28.
- Schlesinger, W.H. 1997. *Biogeochemistry: An Analysis of Global Change*. 2nd edn. Academic Press, San Diego.

- Schlesinger, W. H. 2009. On the fate of anthropogenic nitrogen. *Proceedings of the National Academy of Science USA*, 106 (1): 203–208.
- Smedley, P.L. and Kinniburgh, D.G., 2002. A review of the source, behavior and distribution of arsenic in natural waters. *Applied Geochemistry* 17: 517–568.
- Vitousek, P.M., Aber, J., Howarth, R.W., Likens, G.E., Matson, P.A., Schindler, D.W., Schlesinger, W.H., and Tilman, G.D., 1997. Human alteration of the global nitrogen cycle: causes and consequences. *Issues in Ecology* No. 1, Ecological Society of America, Washington, DC, 17 p.

The Global Atmosphere: Composition, Evolution and Anthropogenic Change

The goal of this chapter is to cover topics that provide background required for understanding the natural structure and composition of the global atmosphere, including some examination of how the atmosphere has varied in composition over time, from billion-year and million-year time scales across geological time to more-recent changes that have occurred in the Pleistocene and Holocene. This chapter also presents some of the fundamental concepts that help us to predict and understand atmospheric circulation, which can be applied to contaminant transport as well as local and regional climate. The chapter concludes by presenting data and concepts related to greenhouse gases (GHGs), including molecular-level properties that impart infrared trapping capability, as well as the role of GHGs in global climate, from pre-industrial time to the 21st century.

7.1 ATMOSPHERIC STRUCTURE, CIRCULATION AND COMPOSITION

The atmosphere differs from the other main components of the Earth system in a few important

ways. First, in terms of mass it is the smallest reservoir at or near the Earth's surface, containing only 5×10^9 Tg of matter as compared to the hydrosphere, which contains 2.4×10^{12} Tg of matter (H_2O), and the crust of the Earth, which contains 2.4×10^{13} Tg of matter (rock). One important result of its small size is that the atmosphere can be very sensitive to small amounts of contamination. A second important distinction is that the atmosphere has a high mixing rate relative to the hydrosphere, soils, sediments and bedrock – what this means is that (1) contaminants can rapidly spread throughout the atmosphere, and (2) the atmosphere has a more homogeneous distribution of constituents than the hydrosphere or the lithosphere. Whereas point-source contaminants in the hydrosphere are often confined to a particular zone within an aquifer or to a specific stream, point source contaminants in the atmosphere spread widely in time frames as short as a few days. For example, sulfuric acid and mercury derived from coal-fired power plants are deposited into terrestrial and aquatic ecosystems hundreds of kilometers downwind of their sources, radioactive

fallout from the 1986 Chernobyl disaster in the Ukraine (part of the USSR at the time) spread to much of northern Europe within hours to days, and sulfate particles derived from very strong Plinian volcanic eruptions can reach the stratosphere and be dispersed globally.

7.1.1 Structure and layering of the atmosphere

The atmosphere is divided into distinct layers characterized by differences in air pressure, temperature and composition. Changes in atmospheric pressure as a function of altitude are explained by this derivation of Dalton's law of partial pressures,

$$p_z = p_0 * e^{(-z/H)} \quad (7.1)$$

where p_z is atmospheric pressure at altitude z , p_0 is atmospheric pressure at sea level, and H is a term known as the "scale height", a measure of the effective thickness of an atmospheric layer that is expressed mathematically as $H = k * T / m * g$ (where k is Boltzmann's constant, T is temperature in Kelvin, m is the molecular mass of the parcel of atmosphere being considered, and g is acceleration due to gravity).

Dalton's law of partial pressures dictates that increasing z causes an exponential drop in atmospheric pressure (Fig. 7.1). Then, from the ideal gas law

$$pv = nRT \quad (7.2)$$

we see that pressure of a gas (or mixture of gases) is proportional to the number of gas molecules (assuming constant volume and temperature). This fact becomes apparent at the elevation of many high-elevation regions and tall mountains (~6 to 9 km), where atmospheric pressure is $\leq 50\%$ of its value at sea level (1 atm = 1013.25 mb), and thus the amount of available oxygen to breathe is $\sim \leq 50\%$ less than at sea level. It is for this reason that hikers become easily exhausted at high elevations and climbers of Mt

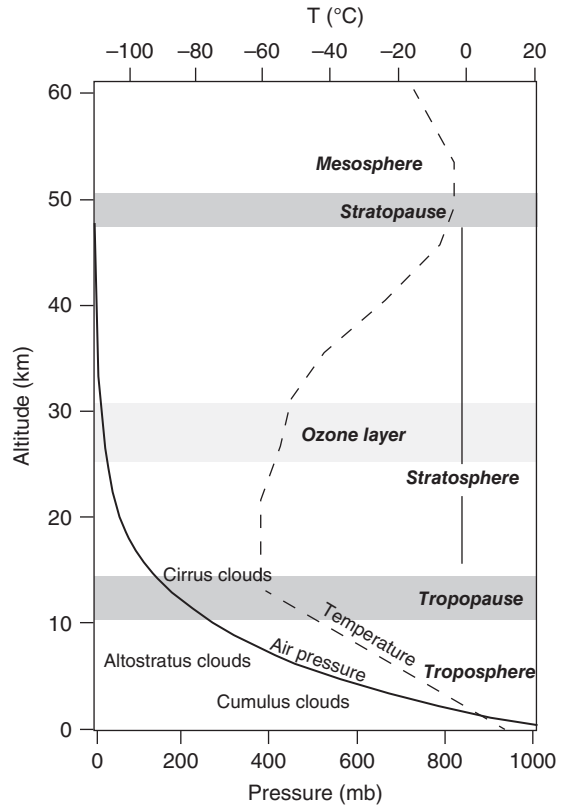


Fig. 7.1 Structure of the atmosphere to an altitude of 60 km. Exponentially decreasing air pressure is indicated by the solid line and air temperature is indicated by the dotted line. Not shown on this figure are the upper mesosphere, the mesopause (temperature minimum at ~80–100 km) and thermosphere (~100–600 km altitude) and exosphere are not shown. The altitude shown for the tropopause is representative of low- to mid-latitudes – see text for explanation.

Everest require oxygen tanks. At an altitude of 50 km, air pressure is only 1 millibar, or 0.1% of air pressure at sea level.

Temperature also decreases markedly with increasing altitude, but only to an altitude of ~15 km, at which point absorption of ultraviolet light (UV) by O_2 molecules produces heat and causes a **temperature inversion** (Fig. 7.1). The lower part of the atmosphere below the temperature inversion is known as the

troposphere – given that the tallest mountains reach elevations of ~ 8.5 to 9 km, and that trans-oceanic flights travel at altitudes of ~ 11 km, nearly all human activities occur within the troposphere. The height of the troposphere varies as a function of latitude. At the equator the top of the troposphere occurs approximately 18 to 20 km above sea level; at 45° , the troposphere extends to an altitude of 8 to 9 km, and at the poles its extent is only 6 to 7 km. Weather-related phenomena and nearly all clouds occur within the troposphere because the inversion at the top of the troposphere (warm stratospheric air resting atop cold upper tropospheric air) prevents air from rising through the **tropopause**, which is the boundary that occurs between the troposphere and the overlying **stratosphere** (Fig. 7.1).

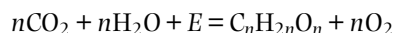
In terms of atmospheric chemistry, the most important part of the stratosphere is the **ozone layer**, where incoming UV radiation causes dissociation of O_2 molecules and formation of O_3 (ozone) molecules. Details on ozone chemistry are presented in Chapter 8 (section 8.9). Above the stratosphere are the **mesosphere**, **thermosphere** and **exosphere** (not shown on Fig. 9.1), layers that combined contain $< 0.1\%$ of the mass of the atmosphere. These layers of the upper atmosphere are characterized by intense UV bombardment and are richer in lighter gases such as H and He; however, because this text is primarily concerned with processes at or near the Earth's surface, details of atmospheric chemistry in layers above the stratosphere are not considered, other than in this chapter with regards to formation of cosmogenic radionuclides.

7.1.2 Geological record of atmospheric composition

Earth's atmosphere initially formed as the result of degassing of volatile compounds (e.g. H_2O , N_2 ,

CO_2) from the mantle and crust and from volatile compounds contained in meteorites, comets and other extraterrestrial impact bodies that collided with Earth early in its history (mainly before 4 Ga). By 3.8 Ga the atmosphere was likely dominated by CO_2 and H_2O with additional constituents such as SO_2 , HCl , Cl_2 , N_2 , H_2 , CH_4 (methane) and NH_3 (ammonia). This early atmosphere likely had little or no O_2 , an idea that is supported by the presence of pyrite cobbles in Archean stream deposits. Pyrite is unstable in the presence of O_2 gas, so these deposits are evidence for the O_2 -poor character of the early atmosphere, and volcanic gases do not include O_2 , so any O_2 in the early atmosphere would have only existed due to photochemical dissociation of small amounts of H_2O into O_2 . However, by ~ 3.5 Ga, photosynthetic bacteria had appeared in the oceans, an evolutionary advance that would eventually result in profound changes to the composition of Earth's atmosphere.

Photosynthesis by cyanobacteria and other early prokaryotes converted CO_2 and H_2O into organic matter and oxygen, and from this point forward the atmosphere began to shift toward one of greater O_2 content at the expense of CO_2 (equation 3.2)



Atmospheric O_2 content has likely varied between 15% and 30% during the ~ 550 Ma of Phanerozoic time (Fig. 7.2) (Berner, 1997). The variation in $\% O_2$ over time has been controlled by changes in photosynthesis related to evolution of new life forms (particularly the appearance of vascular plants at ~ 375 Ma) as well as other aspects of the carbon cycle and the sulfur cycle. For example, the bacterially mediated reduction of sulfate, as illustrated below, is an additional means of producing atmospheric O_2 . In reverse (i.e. right to left), this reaction depicts reaction of O_2 gas with pyrite, an example of an oxidation reaction that would remove O_2 from the atmosphere.

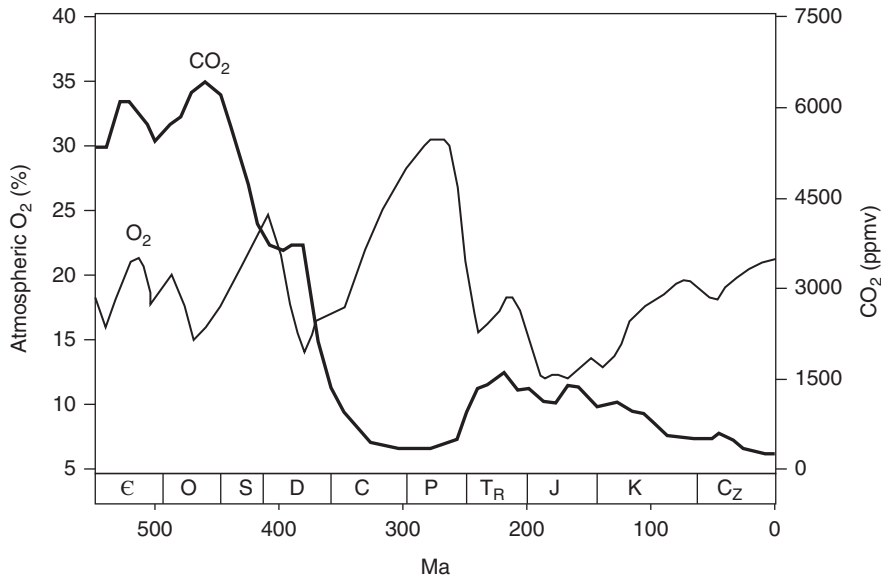
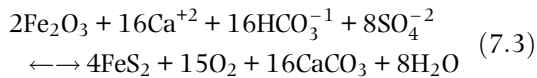


Fig. 7.2 Estimations of atmospheric O_2 and CO_2 during Phanerozoic time. Drawn from models in Berner and Berner (1996) and Berner (2006). Uncertainty in CO_2 values is $\sim \pm 30\%$ of plotted value, meaning that at the Cambrian – Ordovician boundary, CO_2 was $\sim 6000 \pm 1800$ ppmv. From left to right and oldest to youngest, geological periods are Cambrian (C), Ordovician (O), Silurian (S), Devonian (D), Carboniferous (C), Permian (P), Triassic (T_R), Jurassic (J), Cretaceous (K) and Cenozoic (C_z).



Relative to oxygen, carbon dioxide has undergone more profound changes in atmospheric abundance during the Phanerozoic, particularly between 400 Ma and 300 Ma (middle Devonian to middle Carboniferous time), a period during which evolution and appearance of land plants provided a major sink for CO_2 . Since ~ 350 Ma, CO_2 and O_2 have broadly fluctuated in an inverse manner, likely reflecting the influence of photosynthesizing organisms on atmospheric composition (e.g. Lovelock and Margulis, 1974).

The Quaternary (Pleistocene and Holocene) record of atmospheric composition is a useful framework because it is the time period in which humans have evolved and flourished. The onset of the Pleistocene coincides with cooling climate

and expanding ice sheets and over the past 2 Ma humans and their predecessors have witnessed numerous cyclic periods of glacial advance and retreat. Atmospheric CO_2 has decreased through most of the Cenozoic (Fig. 7.2), which is one of the main forcings responsible for Pleistocene ice ages.

7.1.3 Climate proxies

While there are many climate proxies that can be applied to understanding interactions between the composition of the atmosphere and average global temperature, the most commonly applied proxies are those from ice cores. The usefulness of ice cores lies in the fact that they typically contain annual layers of snow that, over time, are compressed into ice. Bubbles of atmosphere become trapped in the

snow and ice, allowing paleoclimatologists to directly measure paleo-atmospheric composition. Variations in stable isotope compositions of hydrogen (δD) and oxygen ($\delta^{18}O$) in H_2O molecules that originally fell as snow provide a mechanism to determine paleo-temperature (details are presented in Chapter 10) – these isotopic values are controlled by global temperature and local temperature, where warmer oceans and air foster heavier (or less-negative) values during warm periods. Dating of ice layers is possible because at shallow depth within the ice, annual layers of snow and ice can be counted backwards in time to determine the age of a given layer; deeper in the core, where individual layers are no longer distinguishable due to compaction, volcanic ash deposits and carbon from large forest fires are common means for determining age.

Well-known ice core records of Pleistocene and Holocene paleotemperature and atmospheric gas concentrations are the Vostok core (Petit et al., 1999) and the EPICA dome core (Augustin et al., 2004), both from Antarctica, and the GISP and GRIP cores from Greenland (e.g. Steffensen et al., 2008 and references therein). Data from the Vostok core are presented in Fig. 7.3a – this core was chosen to illustrate relationships among temperature (and, by definition, the stable isotope signatures δD and $\delta^{18}O$), CO_2 , CH_4 and insolation because the Vostok record extends farther back in time than do the GISP/GRIP records and because pre-420 ka data from the EPICA dome record are of relatively poor resolution compared to pre-420 ka data from Vostok and EPICA.

Atmospheric CO_2 , CH_4 and temperature have shown a very strong coherence for many hundreds of thousands of years (Fig. 7.3a), where cooler temperatures correspond to relatively low amounts of carbon dioxide and methane in air, and warmer periods are correlated to higher amounts of these greenhouse gases. The strong relationships among CO_2 , CH_4 and temperature are best understood in

the framework of one of the main feedback loops operating within the climate system.

7.1.4 Orbital control on C

Variations in the eccentricity of Earth's orbit, changes in the tilt of Earth's axis of rotation (obliquity) and change in the orientation of the rotation axis of Earth (precession) combine to produce periods when mean insolation is greater than average. During these times, increased insolation produces increased average global temperature, which in turn causes changes to the carbon cycle; one consequence of a warming atmosphere is warming of the oceans that results in a net flux of CO_2 out of the oceans. Warming also causes melting of ice caps that, in combination with expansion of warming ocean water, raises sea level, which in turn produces a greater extent of methane-producing coastal swamps; these factors, and many others, combine to increase atmospheric concentrations of CO_2 and CH_4 , which trap more heat in the atmosphere, causing further warming, which releases more CO_2 from oceans and so on. This positive feedback loop causes rapid warming during transitions from ice ages to interglacial periods, a feature that is noticeable in Fig. 7.3a.

During periods when eccentricity, obliquity and precession produce cooler than average temperatures, CO_2 is fixed more readily into oceans, methane-producing swamps shrink, organic matter decomposes more slowly and the concentrations of CO_2 and CH_4 in the atmosphere decrease, trapping less heat, and setting into motion a cooling feedback loop. Global cooling associated with the transition from warm interglacial to ice age occur far more slowly than do glacial to interglacial warming periods, differences that are attributable to the rates and magnitudes of the feedback loops. Also notable is that while the initial driving mechanism behind these

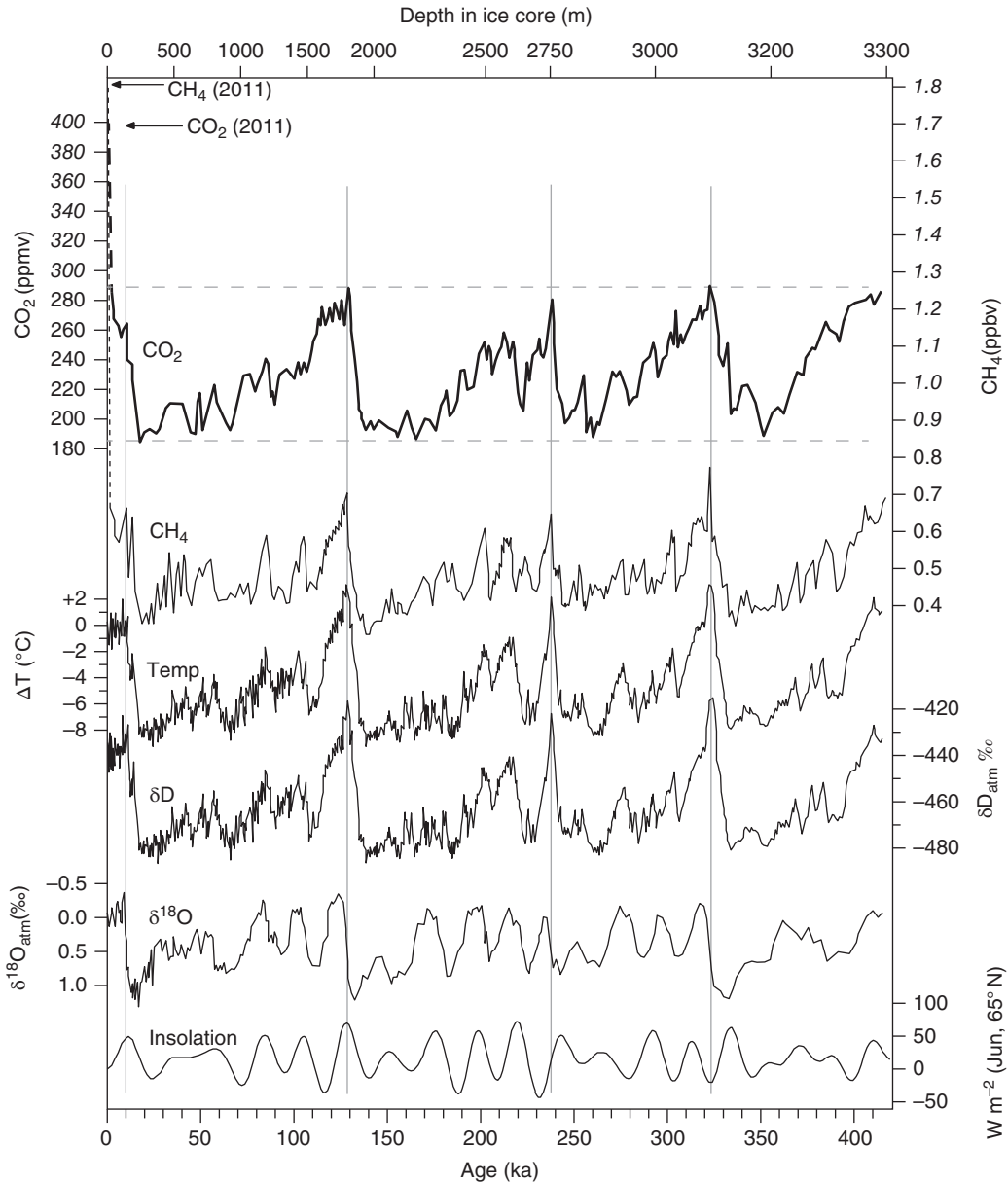


Fig. 7.3a 420 000 year Vostok, Antarctica ice core record of atmospheric CO₂ and CH₄ concentrations (measured from bubbles trapped in ice) as well as changes in atmospheric temperature derived from D/H ratios (expressed as δD) of H₂O molecules. Oxygen isotope data are also shown, as is insolation for the month of June at 65 °S latitude. Note that the pre-industrial range of CO₂ concentrations was ~180 to 290 ppmv, and that CH₄ ranged from ~0.4 to 0.8 ppbv. In June 2011, CO₂ was 395 ppmv and CH₄ was 1.8 ppmv. Values in *italics* in the scales for CO₂ and CH₄ concentrations represent values not seen in the ice cores; rather, these *italicized* concentrations. Vostok data are from Petit et al. (1999) and references cited therein.

shifts in climate appears to be related to the Earth–Sun distance (i.e. changing temperature appears to be the initial trigger), the greenhouse gases CO_2 and CH_4 play important roles in feedback loops that ultimately result in the magnitudes of warming (and cooling) observed in the δD and $\delta^{18}\text{O}$ ice core records.

The orbital mechanisms underlying these climate transitions have occurred cyclically since at least the onset of the Pleistocene epoch, and this phenomenon was discovered by the Serbian mathematician and engineer Milutin Milanković and published in 1914. For that reason, sometimes these orbital variation and their controls on climate on 20 000 to 100 000 year time scales are known as Milankovitch cycles.

The Holocene epoch is often regarded as a time of stable climate and steady greenhouse gas concentrations, yet detailed analysis of bubbles trapped in glacial ice reveals some interesting trends, notably a slow but steady increase in CO_2 from 260 ppmv at 7000 years before the present to 250 years ago, when CO_2 was at its pre-industrial value of 280 ppmv. Methane also increased during the Holocene, although its rise did not begin until 5000 years before the present (Fig. 7.3b). These trends are anomalous compared to the three previous interglacial periods (125 ka, 240 ka, 325 ka) recorded in the Vostok ice core record (Fig. 9.3a), where rapid warming and associated rapid increase in atmospheric CO_2 and CH_4 during deglaciation were followed by

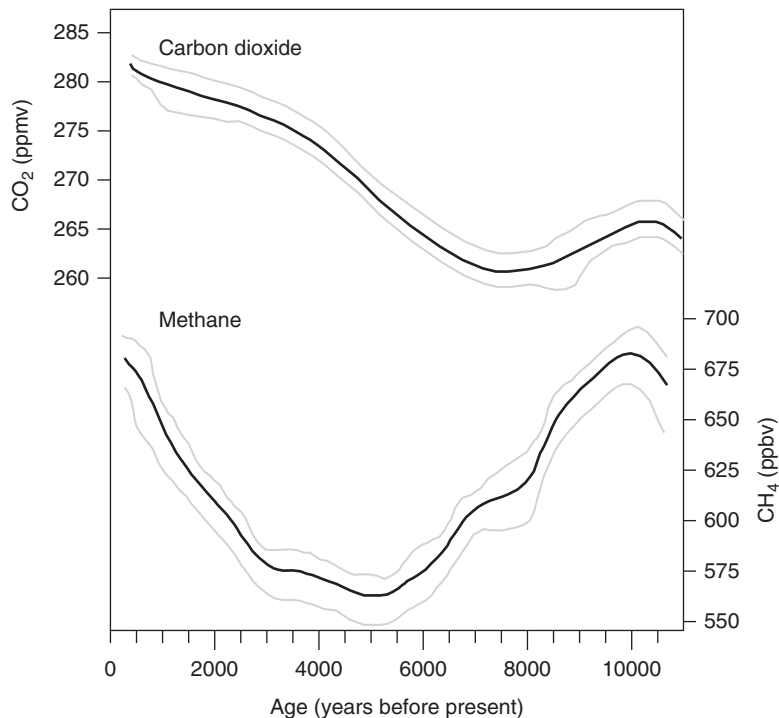


Fig. 7.3b Trends for carbon dioxide and methane during the Holocene epoch (plotted here as ~11 400 to 400 years before the present) derived from air bubbles in glacial ice and firn (from the EPICA Dome C project). Dark black curves are fitted trends and the light gray lines depict estimated uncertainty. Data are available from the USA NOAA National Climate Data Center (<http://www.ncdc.noaa.gov/paleo/data.html>) and were presented in a similar manner by Flückiger et al. (2002).

thousands of years of slow cooling and slowly decreasing CO_2 and CH_4 as Earth headed back towards glacial conditions.

Ruddiman (2003) observed this anomaly and after examining insolation and other potential causes, speculated that the development and spread of agriculture 5000 to 7000 years ago (and its continuation to the onset of the industrial age) was responsible for these trends, mainly due to flooding for rice cultivation (a source of CH_4) and land clearing (for farming) that converted C_{org} sequestered in trees into CO_2 in the atmosphere. In the absence of human activities during the Holocene, atmospheric CO_2 and CH_4 would likely have continued to drop towards glacial maxima values of ~ 190 ppmv rather than recovering after their initial decreases in the earliest Holocene (Ruddiman, 2003). Yet even if humans have been responsible for altering greenhouse gas concentrations for thousands of years, the magnitude of the pre-industrial era increases and the corresponding rates of increase are mere fractions of the trends seen in greenhouse gases over the past two centuries, and this is cause for concern. However, before examining other issues related to greenhouse gases, including mechanisms by which greenhouse gases affect climate and details of post-1800 changes in greenhouse-gas concentrations (topics which are covered in Sections 7.3 and 7.4), the text now explores topics related to atmospheric composition, circulation and precipitation.

7.1.5 Composition of the current atmosphere

The **composition of the modern atmosphere** (to an altitude of 25 km) is presented in Table 7.1. The concentrations of many of these gases, including nitrogen, oxygen, argon, neon, helium and xenon, are relatively constant over time scales on the order of 10^3 to 10^4 years or longer. These gases do not vary appreciably over

Table 7.1 Concentrations of selected fixed gases in the atmosphere (sea level). Data are from Spiro and Stigliani (2006) and the Global Monitoring Division of the National Oceanic and Atmospheric Administration (USA) data on greenhouse gases. Data from Earth System Research Laboratory of US National Oceanic and Atmospheric Administration (<http://www.esrl.noaa.gov/gmd/ccgg/trends/>).

Gas	Concentration
Nitrogen (N_2)	78.09%
Oxygen (O_2)	20.94%
Argon (Ar)	0.93%
Carbon Dioxide (CO_2)	395 ppmv ^a
Neon (Ne)	18.2 ppmv
Helium (He)	5.2 ppmv
Methane (CH_4)	1.8 ppmv ^a
Krypton (Kr)	1.1 ppmv
Hydrogen (H_2)	0.55 ppmv
Nitrous oxide (N_2O)	0.32 ppmv ^a

^a The concentrations of these gases are increasing due to anthropogenic emissions.

glacial–interglacial cycles, as do carbon dioxide and methane, and they do not vary over time scales of hours to days, as does water vapor.

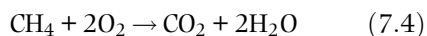
The concentration of water vapor ranges from $< 1\%$ in arid air to 4% in very hot humid (tropical) air, so when water vapor content is at its lowest extreme, the actual concentrations of the constant gases (N_2 , O_2 , Ar, etc.) are effectively the same as the values given in Table 9.1, and the sum of the main constant components (i.e. the fixed gases) $\text{N}_2 + \text{O}_2 + \text{Ar}$ is 99.96% ; however, when water vapor is 4% on a humid, sticky day in the tropics, the actual concentrations of the fixed gases are proportionally lower. In this case, the amount of N_2 gas in humid air with $4\% \text{H}_2\text{O}_g$ is $78.08\% \times 96/100 = 74.96\%$. Similarly, the amount of $\text{O}_2 = 20.95\% \times 96/100 = 20.11\%$, $\text{Ar} = 0.89\%$, and the sum of the fixed gases $\text{N}_2 + \text{O}_2 + \text{Ar}$ is 95.96% .

The abundance of O_2 gas makes the atmosphere a strongly oxidizing environment and this helps to explain the presence of CO_2 and H_2O , gases that form by oxidation of, for example, organic matter

The common units of measure for concentrations of atmospheric gases are based on abundance of a gas on a volumetric basis, reflecting the fact that one mole of each gas occupies approximately equal volumes (~22.4 liters per mole). In tangible terms, if a trace gas is present in the atmosphere at a concentration of 1 part per million volumetrically (1 ppmv), there will be 1 cm³ of that gas in 10⁶ cm³ of air. It also implies that 1 ppmv of a given gas corresponds to 1 mol of that gas per 10⁶ moles of air, and similarly that this gas will effectively comprise 1 molecule out of every 10⁶ molecules of air.

and methane. These oxygen-bearing gases are relatively stable in the atmosphere and stand in contrast to gases that are more stable in reducing environments such as CH₄ and H₂S; however, these gases (and others) do exist in the atmosphere at relatively constant concentrations, at least during the ~12 000 years of the Holocene prior to the industrial age (e.g. CH₄ concentration was ~0.8 ppmv throughout the Holocene; its current concentration is 1.8 ppmv).

If we consider the fate in the atmosphere of chemically reduced gases like CH₄, we would predict that they will react with O₂, as follows:



(Chapter 8 explores the complexities of oxidation reactions that typically involve other species such as the hydroxyl and peroxy radicals, but this equation serves to emphasize the point being illustrated here.)

From tables of geochemical constants, the equilibrium constant for this reaction can be written as:

$$K_{\text{eq}} = [\text{CO}_2] * [\text{H}_2\text{O}]^2 / [\text{CH}_4] * [\text{O}_2]^2 = 10^{140} \quad (7.5)$$

and from this equation the concentrations of the gases can also be represented as partial pressures:.

$$K_{\text{eq}} = p_{\text{CO}_2} * p_{\text{H}_2\text{O}} / p_{\text{CH}_4} * p_{\text{O}_2}^2 = 10^{140} \quad (7.6)$$

The K_{eq} of 10^{140} indicates that an equilibrium assemblage of CH₄, O₂, CO₂ and H₂O will be very strongly dominated by the products, which in this case are carbon dioxide and water. Using partial pressure values for O₂, CO₂ and H₂O of 0.21 atm, 0.004 atm and 0.01 atm, respectively (if N₂ were part of the calculation its partial pressure value would be 0.78 atm, where the sum of all partial pressures = 1.0 atm), in the K_{eq} equation, as follows:

$$K_{\text{eq}} = 0.0004 \text{ atm} * 0.01 \text{ atm} / p_{\text{CH}_4} * 0.21 \text{ atm} = 10^{140}$$

and rearranging to solve for p_{CH_4} :

$$p_{\text{CH}_4} = 0.0004 \text{ atm} * 0.01 \text{ atm} / 10^{140} * 0.21 \text{ atm} = 1.9 * 10^{-145} \text{ atm}$$

it is apparent that the concentration of methane in the atmosphere based on equilibrium considerations should be 10^{-145} atm. This is an incredibly small amount. When compared to the actual concentration of 2 ppmv ($2 * 10^{-6}$ atm), it becomes very clear that the amount of methane in air is not controlled by chemical equilibrium; instead, methane and many other gases in air are out of equilibrium and their concentrations are mainly controlled by kinetic factors. What are these factors? Given that the fate of most atmospheric CH₄ is oxidation to CO₂ via equation 7.4, one important control on CH₄ abundance is the probability of collision or interaction of O₂ and CH₄ molecules in air (or on the surfaces of particulate substrates); another is the rate at which CH₄ and O₂ dissociate in the presence of each other (once molecules of the gases have collided) to form the products CO₂ and H₂O. These rate-limiting steps lend a strong kinetic control to the concentration of methane in Earth's atmosphere.

7.1.6 Air circulation

Air circulation plays an important role in transport of natural and anthropogenically derived

compounds in the atmosphere, and much of the air circulation is driven by differences in density. For an air mass at sea level at constant temperature (e.g. 20 °C), the main control on the **density of air** is the amount of water vapor. This is because the other main gases in the atmosphere (N_2 , O_2 , Ar) are present at effectively fixed values—dry air (0% H_2O) contains 78.08% N_2 , 20.95% O_2 and 0.93% Ar. A very humid end-member will contain 4% water vapor and proportionally less N_2 , O_2 and Ar than what is shown in Table 7.1. The density of these end-member air mass can be determined by carrying out a mass-balance calculation as follows:

For dry air with no H_2O :

$$[0.7808 * 28.01 \text{ g/mol (N}_2\text{)}] + [0.2095 * 32.00 \text{ g/mol (O}_2\text{)}] + [0.0093 * 39.95 \text{ g/mol (Ar)}] = \mathbf{28.95 \text{ g/mol.}}$$

For humid air with 4% H_2O :

$$[0.7496 * 28.01 \text{ g/mol (N}_2\text{)}] + [0.2011 * 32.00 \text{ g/mol (O}_2\text{)}] + [0.0089 * 39.95 \text{ g/mol (Ar)}] + [0.004 * 18.01 \text{ g/mol (H}_2\text{O)}] = \mathbf{27.86 \text{ g/mol.}}$$

While perhaps counterintuitive, humid air is less dense than dry air, and this density difference drives circulation. Lower density air (generally

humid and warm) is transported convectively upward in the atmosphere (i.e. it rises and carries heat and water vapor upward), and when humid air masses collide with dry air masses, the dry air dives under the humid air and this forces the humid air to higher altitudes in the atmosphere. This results in condensation of water vapor that can lead to rain, snow or other forms of precipitation.

Air temperature also drives circulation – hot air has greater kinetic energy and greater average distance between air molecules compared to cool (and often dry) air, so its density is lower and as a consequence, it rises; cool air sinks. Heating of the Earth's surface and its atmosphere is mainly caused by **insolation** (incident solar radiation), and while radioactive decay within the Earth and volcanic and hydrothermal systems also can heat the Earth's surface and its atmosphere, the main control on atmospheric circulation is insolation and the absorption or reflectance of solar rays. Insolation is greatest near the equator (Fig. 7.4), and with increasing latitude poleward, the sun strikes the Earth at progressively more oblique angles where more of it is scattered, absorbed, attenuated or reflected. The result is that the equatorial belt receives the greatest amount of insolation (i.e. incoming heat) of any region at the Earth's surface. At latitudes greater

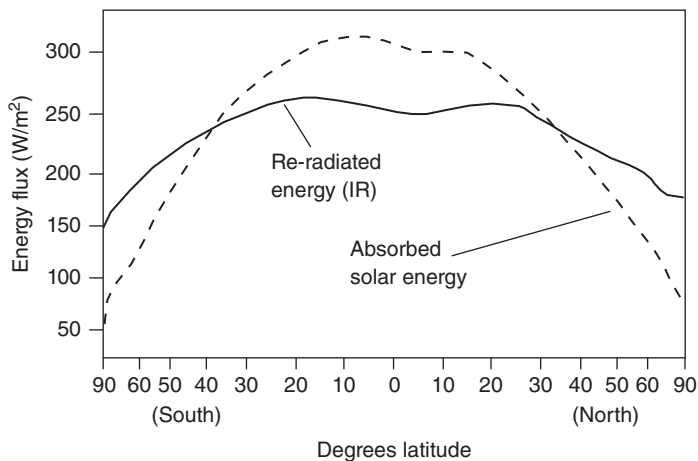


Fig. 7.4 Variation in insolation at Earth's surface (visible, UV, IR; dashed line) and reradiated infrared radiation from Earth's surface (solid line) plotted as a function of latitude. Note net influx of energy in tropics and subtropics and net loss of energy at latitudes greater than $\sim 35^\circ$. Non-linearity of the x-axis reflects scaling by cosine of the latitude. (Modified from Vonder Haar and Suomi 1971.)

than 35°N and 40°S , more energy is lost as infrared radiation than is gained as incoming solar radiation (Fig. 7.4).

The occurrence of warm air (and soils, lakes and oceans) in the tropics leads to high evaporation rates and warm, low-pressure, humid air within ~ 10 to 20 degrees of the equator (highly productive tropical forest ecosystems also contribute to high water vapor content through transpiration). As this warm, moist air rises through the atmosphere, it cools and eventually loses its moisture through condensation and precipitation as

tropical rains. Cold, dry, high-pressure air from higher latitudes rushes towards the equator in response to the pressure gradient – the cooler, higher-pressure air dives beneath the convectively uplifting tropical low-pressure air, setting into motion what, in the simplest sense, is known as a **Hadley cell** (named after 18th century meteorologist George Hadley), a sketch of which is depicted in Fig. 7.5.

Global air circulation is described by the **three-cell model** (Fig. 7.5). As warm, humid tropical air masses migrate upward into higher levels of

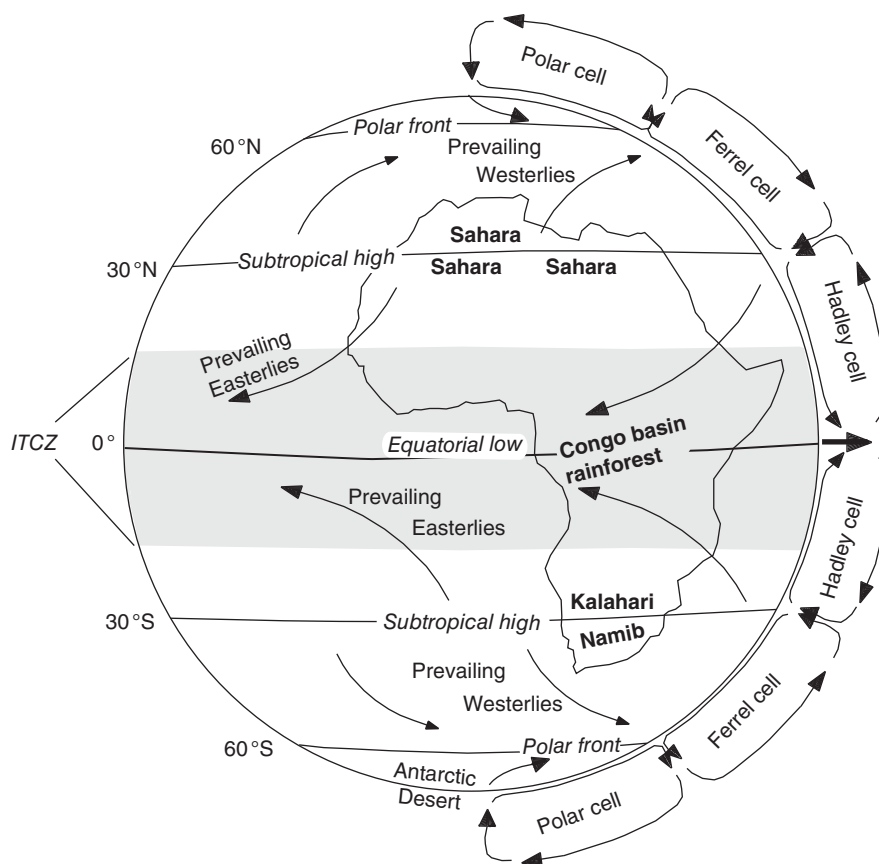


Fig. 7.5 Atmospheric circulation patterns on a global scale, with Africa presented here as an example of a continent with climate zones that are strongly controlled by global air circulation. Note the occurrence of tropical rainforests (e.g. Congo Basin) within the ITCZ, locations of deserts (e.g. Sahara, Kalahari, Namib) at 30 degrees north and south latitude, as well as prevailing wind patterns that transport contaminants at many scales, from regional (acid rain, tropospheric ozone) to global (notably, transport of persistent organic pollutants from the northern continents to the Arctic in prevailing westerlies of the northern hemisphere Ferrel cell).

the troposphere, they cool through radiative heat loss and water vapor is lost from the air mass as rain that falls in the tropics. Radiative cooling of tropic-derived, poleward-migrating air in the upper troposphere increases its density and causes it to sink back to the Earth's surface at approximately 30°N and 30°S of the equator. This sinking air (now dry) heats adiabatically in the lower levels of the atmosphere, and the rate of heating is more rapid than the rate at which the rising tropical air cooled because the air mass is not thermally buffered by the high heat capacity of H₂O. Fully saturated tropical air cools and heats at a rate of 5 °C per 1000 m; conversely, dry air cools and heats at a rate of approximately 10 °C per 1000 m.

The rapid heating of sinking, dry air at 30 degrees north and south of the equator produces two belts of **subtropical high** pressure regions that contain many of the great deserts of the world, including the Sahara Desert of northern Africa, the Mojave Desert of southwestern North America, the Arabian Desert of the Arabian Peninsula and the Namib and Kalahari deserts of southwestern Africa (Fig. 7.5). Some of this descending air returns toward the equator, setting up a convection cell still known as a **Hadley cell**. Air masses that return to the tropics from 30° north and south latitude are initially hot and dry, so they gather moisture as they migrate back towards the equator. The region where these two equator-bound air masses collide is known as the **intertropical convergence zone (ITCZ)**. The collision of equator-bound air masses combined with high moisture content and strong insolation causes air to rise convectively in the ITCZ – the latitudinal extent of the ITCZ is ~ ± 20 degrees north / south of the equator, and this is where tropical rainforests occur.

Some of the air that descends back to the Earth's surface at 30° migrates poleward at or near the Earth's surface, and because this air is hot and dry, water is evaporated from the Earth's surface as air in this cell migrates poleward. Collision with cold, dry polar air

produces precipitation in a belt known as the **Polar Front**. Some of the uplifted poleward-migrating air continues poleward where it ultimately sinks to the surface at the pole, from where it migrates back towards the equator. Convection cells that deliver air from the mid-latitudes towards the poles contribute to the transport of **Pb, DDT, PCBs and other persistent pollutants** from industrialized areas to Arctic and Antarctic regions (e.g. Murozumi et al., 1969; Eckhardt et al., 2007).

While the three-cell model might imply that air travels in north–south directions, in reality air masses are deflected from these orthogonal pathways by rotation of the Earth; in the northern hemisphere, migrating air masses are deflected to the right (clockwise), while in the southern hemisphere, air masses are deflected to the left (counter-clockwise). This force, known as the **Coriolis Effect**, results in prevailing winds such as the tropical **easterlies** (also known as the northeasterly and southeasterly trade winds) between the equator and 30°N or S; the prevailing **westerlies** between 30 and 60° latitude; and prevailing easterlies between 60° and the poles.

Two consequences the Coriolis Effect on contaminant transport and climate are: (1) sulfur dioxide (SO₂), nitrogen oxides (NO_x), mercury (Hg) and other contaminants from coal-fired power plants in the central United States and England are transported downwind in a northeasterly direction entrained in air carried in Ferrel cells towards pristine forest ecosystems in northeastern North America and Scandinavia, respectively, where they ultimately fall as components of acid rain and as various species of Hg; and (2) tropical easterlies migrating across the Andes Mountains of northwestern South America produce orographic precipitation in the mountains (see below) and the very pronounced rainshadow of the Atacama Desert. In a similar manner, prevailing westerlies migrating across the Himalaya are responsible for the Gobi Desert rainshadow.

7.2 EVAPORATION, DISTILLATION, CO₂ DISSOLUTION AND THE COMPOSITION OF NATURAL PRECIPITATION

The purest water in the hydrologic cycle occurs in the atmosphere, where total dissolved solids typically are < 5 mg/L (ppm). By comparison, river water typically contains 50–500 mg/L TDS, groundwater typically contains 100–1000 mg/L TDS and ocean water contains 35 000 mg/L TDS (Table 4.1).

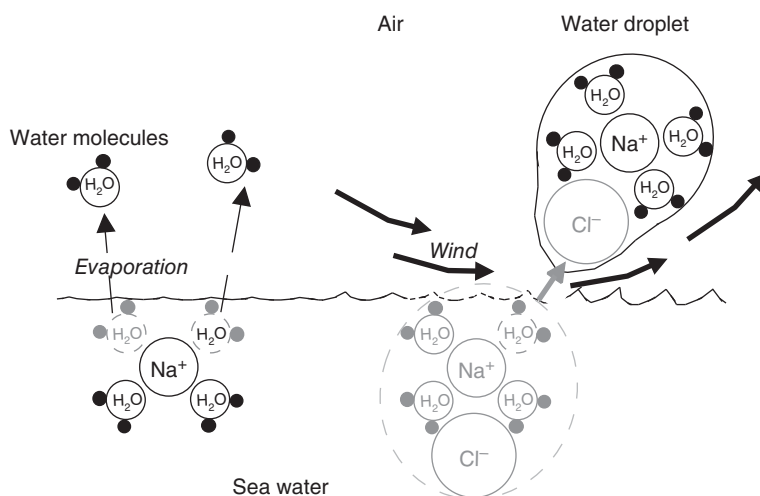
Most atmospheric H₂O is derived via evaporation from ocean water, and the fate of dissolved solids during evaporation helps to explain the occurrence of both very low TDS values of atmospheric H₂O and the high TDS values of ocean water (as well as evaporation-dominated water bodies such as the Dead Sea and the Great Salt Lake). Consider a surface water body (e.g. lake, ocean or river) with dissolved ions such as Na⁺, Ca⁺², HCO₃⁻ and Cl⁻. In order for a water molecule to evaporate, it must achieve a kinetic energy that enables it to overcome the intermolecular bonds between the H₂O molecule in question and adjacent H₂O molecules as well as bonds between the H₂O molecule and dissolved ions. In terms of surface and

atmospheric composition, the difference in the mass of a water molecule (18 Daltons) as compared to the combined mass of a water molecule bonded to a calcium ion (H₂O + Ca⁺² = 58 Daltons) or to a chloride ion (H₂O + Cl⁻ = 53 Daltons) dictates that when H₂O molecules evaporate from surface water bodies they tend to leave behind dissolved ions (Fig. 7.6). This process is a form of **distillation** and it is the main reason for the low amounts of dissolved ions in atmospheric water (relative to surface water and subsurface water).

The solids left behind in a stove-top pan or tea kettle are visual examples of dissolved solids left behind when water evaporates (or boils, although the greater kinetic energy associated with boiling will cause greater amounts of dissolved salts to enter air). In addition to boiling, two additional processes that introduce dissolved salts and particulates to water in the atmosphere are:

1 winds that whip up water droplets from the ocean (Fig. 7.6) – these droplets introduce Na⁺, Cl⁻ and other dissolved ions into the atmosphere; the residence time of these salts in air is 2 to 3 days, so they often “rain out” in coastal regions (note from Table 4.1 that coastal rain typically contains higher TDS than continental interior rain);

Fig. 7.6 Schematic sketch showing evaporation of isolated water molecules from sea water on the left, which depicts evaporation as distillation that leaves dissolved ions in the surface water body. On the right, wind is shown physically eroding a droplet of water containing dissolved salts into air. In reality, many molecules of NaCl may be contained in the water droplet.



2 winds in arid regions, where dust storms introduce fine-grained particulates from desert soils into the atmosphere; some of these particles (termed aerosols) are soluble salts (e.g. halite or gypsum), whereas some others are clay minerals with adsorbed ions that can desorb from exchange sites and become dissolved in water droplets.

Given that molecules of carbon dioxide in air dissolve in water droplets to form carbonic acid, rainfall is typically at least slightly acidic, and when in equilibrium with atmospheric CO_2 , the pH of rainfall is 5.6. The calculations required to determine pH of rain in equilibrium with varying CO_2 concentrations are presented in Chapter 5. As we will see below, reactions of SO_2 and NO_x with water and other constituents of the atmosphere can produce rainfall that is much more acidic than what is produced by CO_2 - H_2O equilibrium. In some cases, naturally sourced S and N gases (e.g. volcanic eruptions) can produce rain with $\text{pH} < 5$, but in most cases, acid rain is related to combustion of fossil fuels.

7.3 THE ELECTROMAGNETIC SPECTRUM, GREENHOUSE GASES AND CLIMATE

The interaction of infrared radiation with trace gases that possess symmetrical molecular structures with a central atom (e.g. CO_2 , CH_4 , H_2O , O_3 and N_2O) produces an effect known as the greenhouse effect. In the absence of this phenomenon, Earth's average temperature would be approximately -18°C ; in the presence of the greenhouse effect, Earth's current average temperature is approximately 15°C (and soon to be 16°C).

In order to understand the greenhouse effect and trapping of heat by trace atmospheric gases, it is first important to understand the electromagnetic radiation emitted by the sun, for this is the main source of energy to heat Earth's atmosphere. Much of this radiation is visible – the sun lights our days and illuminates the moon

at night – but some also occurs at wavelengths (λ) slightly greater and less than those of visible solar radiation (Fig. 7.7).

7.3.1 Electromagnetic spectrum

The visible part of the solar spectrum occurs at wavelengths in the range of 380 to 780 nm. The shortest-wavelength visible light is violet ($\lambda = 380$ to 420 nm), while the longest wavelength light is red (620 to 780 nm). Adjacent to violet light at shorter wavelengths (and thus, higher frequencies) is ultraviolet light ($\lambda = 1$ to 400 nm). The high-frequency (high-energy/short-wavelength) UV radiation can burn skin and cause mutations in cells of living organisms. At the other end of the visible spectrum, radiation with wavelengths greater than those of red light is known as infrared (IR) radiation, with wavelengths ranging from 780 nm to 10^6 nm (1 mm). The amounts of each type of radiation incident upon Earth's outermost atmosphere and upon Earth's surface at sea level are shown in Fig. 7.7, and it is apparent from these spectra that the greatest amount of radiation arrives as visible light followed in abundance by IR and UV radiation.

The effects of atmospheric gases (e.g. H_2O , CO_2 and O_3) as absorbers of incoming radiation are indicated by the differences between the incident radiation and the radiation that ultimately reaches Earth's surface. Notable examples are absorption of UV by ozone and absorption of IR by water and carbon dioxide. Visible light is not particularly affected by these gases; rather, visible radiation that does not reach Earth's surface is most likely scattered and directed back towards outer space by aerosols (especially sulfates) and by water droplets.

7.3.2 Re-radiation from earth surface

The combined visible, IR and UV radiation that reaches rocks, soil and water at the Earth's surface heats these substances and they in turn

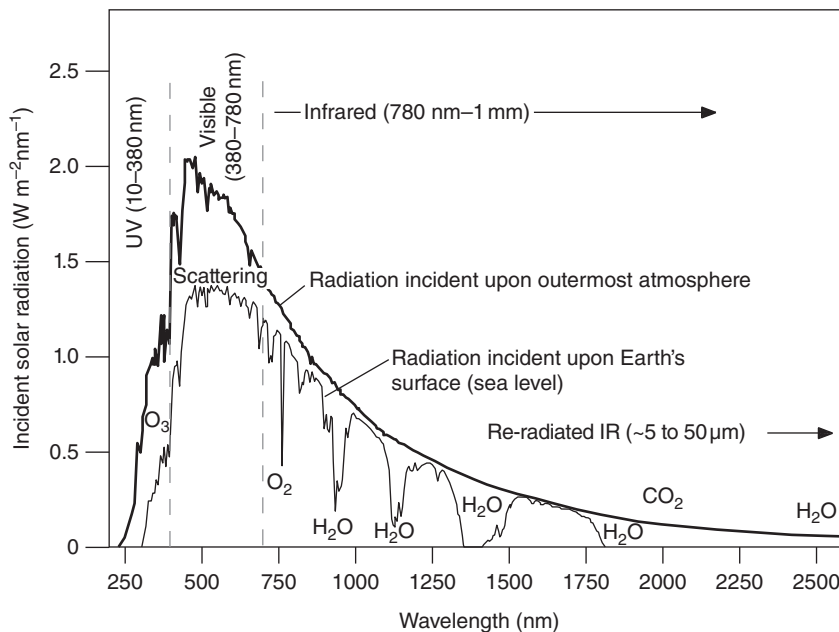


Fig. 7.7 Spectrum of incoming solar radiation showing (a) radiation incident upon outermost atmosphere, and (b) radiation incident upon Earth's surface. It is clear that the atmosphere absorbs and scatters appreciable amounts of incoming ultraviolet (10–380 nm) and visible radiation (380–780 nm). Sulfate aerosols and water droplets are common scatterers of incoming radiation. Absorption of incident radiation by specific gases in specific ranges (e.g. by water between ~1300 and 1500 nm and by ozone in the UV range of ~240 to 380 nm) is indicated on the figure. The image source is the ASTM 1.5 solar output spectrum, a standard reference solar spectrum that is available on-line at: <http://rredc.nrel.gov/solar/spectra/am1.5/>.

re-radiate the heat back towards outer space as IR. At the molecular level, the high frequency of incoming UV radiation increases the frequency of molecular vibrations in objects at the Earth's surface, which causes the objects to increase in temperature (you can feel this on a sunny day); in turn, heat is carried away from the objects at the Earth's surface as infrared radiation with a peak wavelength of 10 μm and in the range of ~3 to 50 μm . These values are consistent with the expected black body radiation from a planet at ~15 °C (the mean temperature of Earth) based on the following Wien's law calculation, as follows:

$$\lambda_{\text{peak}} = b/T \quad (7.8)$$

The term b is Wien's displacement constant, $2.897721 \times 10^{-3} \text{ K m}$, T is temperature in Kelvins, and λ_{peak} is the wavelength at which the emitted radiation is at its maximum intensity (i.e. the modal λ or the λ at the top of the peak). Given the units of b and T , the result is in units of meters. Using 288 K as the temperature of Earth, the maximum wavelength of radiation emitted is $1.01 \times 10^{-5} \text{ m}$ or 10.1 μm (10 100 nm). This value represents the highest intensity of the peak shown in Fig. 7.8.

By comparison, doing the same calculation for the sun using a surface temperature of 6000 K, peak λ is $4.83 \times 10^{-7} \text{ m}$ or 483 nm, which corresponds to the actual solar peak as is visible in the spectrum shown in Fig. 7.7.

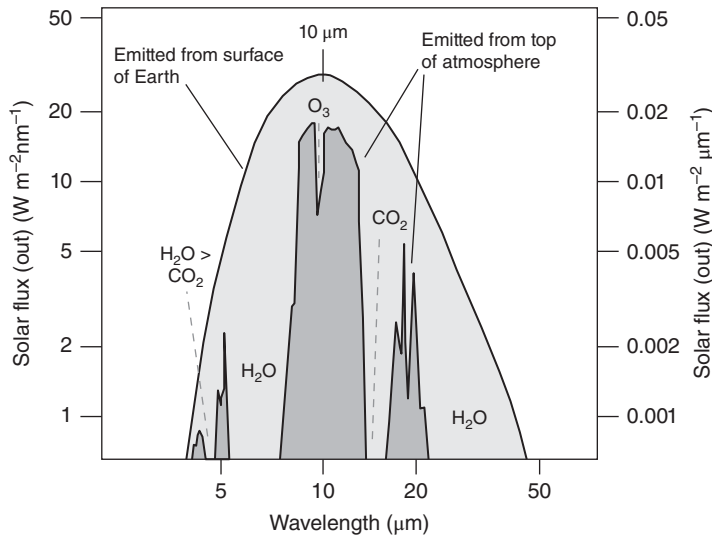


Fig. 7.8 Spectrum of infrared radiation that leaves the atmosphere compared to the ideal spectrum predicted for Earth emanating black body radiation calculated using Wien's law. The difference is mainly due to absorption of IR radiation by greenhouse gases. Note that the x -axis scale is logarithmic and that the units shown here (μm) differ by a factor of 1000 relative to the x -axis in Fig. 7.7.

The role of H_2O and CO_2 as the greenhouse gases with the greatest ability to absorb outgoing IR radiation is evident in Fig. 7.8. H_2O absorbs strongly in the ranges of ~ 3 to $4 \mu\text{m}$, 5 to $7 \mu\text{m}$ and $> 22 \mu\text{m}$, whereas CO_2 absorbs strongly at ~ 13 to $15 \mu\text{m}$. Details on absorption of IR radiation at the molecular level by H_2O , CO_2 and other gases are presented in Section 7.4.

The surface temperature of Earth (or other planets and moons) can be calculated if we can quantify incoming solar radiation (i.e. the amount of energy Earth receives from the sun) and outgoing radiation (i.e. energy emitted by Earth as re-radiated IR radiation). The basis of this calculation is the Stefan–Boltzmann Law, which states that energy radiated by a “**black body**” in space (e.g. Earth) is proportional to the fourth power of the absolute temperature of the body:

$$E = k * T^4 \quad (7.9)$$

where k is the Stefan-Boltzmann constant (see below) and T is absolute temperature in Kelvin.

In the absence of greenhouse gases, the balance of incoming solar energy and re-radiated heat can be determined by the following equation:

$$(E_s/4) * (1 - \alpha) = \sigma * T_E^4 \quad (7.10)$$

Components of this equation include:

- E_s is quantified in terms of W m^{-2} (this is the solar constant, a value that represents the average amount of energy received by Earth from the sun; estimates vary, but it is safe to say that within the past 30 years it has not varied outside the range of 1360 to 1370 W m^{-2});
- E_s multiplied by $1/4$, which is the ratio of the cross-sectional area of Earth as a disc in space to its total spherical surface area – this is necessary because solar radiation is intercepted by the area of the cross-sectional disc of the Earth (πr^2) but is radiated outward as a function of the area of the entire Earth sphere ($4\pi r^2$);
- $\alpha = 0.3$ (this is the estimated albedo of Earth, a unitless parameter – it quantifies the fraction of solar energy reflected back into outer space. What this value indicates is that the other 70% of incoming solar radiation is absorbed at the surface and re-radiated as IR radiation);
- $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$ (this is the Stefan–Boltzmann constant, where K is Kelvin);
- T_E represents temperature (in K) of Earth surface.

Rearranging to solve for T_E :

$$T_E = [(1365/4 \text{ W m}^{-2}) * (1 - 0.3) / (5.67 * 10^{-8} \text{ W m}^{-2} \text{ K}^{-4})]^{1/4}$$

$$T_E = 255 \text{ K} = -18^\circ \text{C}$$

The average temperature of Earth is clearly not -18°C . The difference in this “black body” temperature calculation and Earth’s actual surface temperature lies in the effect exerted by atmospheric gas molecules with symmetrical structures, mainly H_2O , CO_2 , and CH_4 as well as O_3 , N_2O and chlorofluorocarbons (CFCs).

7.3.3 Greenhouse effect and heat trapping

The **greenhouse effect** can be incorporated into the Stefan-Boltzmann Law as **effective infrared transmission** (g_h), a term that expresses heat trapping by the greenhouse effect.

$$(E_S/4) * (1.0 - \alpha) = g_h * \sigma * T_E^4 \quad (7.11)$$

At an atmospheric CO_2 concentration of 280 ppmv, the value of g_h is 0.61. What this means is that 61% of re-radiated IR radiation escapes the atmosphere headed for outer space; the other 39% interacts with CO_2 , H_2O and other greenhouse gases and is reflected back toward Earth’s surface. The mechanism of this interaction is presented in Section 7.4. At a g_h value of 0.61 (i.e. indicative of pre-industrial CO_2 levels), the global average surface temperature is determined as follows:

$$T_E = [(1367 \text{ W m}^{-2})/4 * (1.0 - 0.30) / (0.61 * 5.67 * 10^{-8} \text{ W m}^{-2} \text{ K}^{-4})]^{1/4}$$

$$T_E = 288.28 \text{ K} = \sim 15^\circ \text{C}$$

The greenhouse effect raises Earth’s average temperature by about 34°C ; however, too much of a

good thing can be problematic, and we have entered an era where concentrations of greenhouse gases have reached levels not seen in at least hundreds of thousands of years (Section 7.4). One of the ways to forecast how increasing greenhouse gas concentrations will affect Earth’s surface temperature is to perform calculations based on the first principles described above. For example, considering an atmospheric CO_2 concentration of 550 ppmv, the value of g_h becomes 0.60 (i.e. only 60% of outgoing infrared radiation escapes to outer space) and the calculation of global average surface temperature is as follows:

$$T_E = [(1367 \text{ W m}^{-2})/4 * (1.0 - 0.3) / (0.60 * 5.67 * 10^{-8} \text{ W m}^{-2} \text{ K}^{-4})]^{1/4}$$

$$T_E = 290.47 \text{ K} = 16.47^\circ \text{C} = \sim 16^\circ \text{C}$$

This simple calculation only takes into account radiative forcing from CO_2 , but not feedback loops, other greenhouse gases (including water), changes to albedo and myriad other elements of the climate system, so it underestimates the warming associated with a $\sim$ doubling of CO_2 . For example, accounting for a slight decrease in the albedo of Earth from 0.3 to 0.28, we obtain the following result:

$$T_E = [(1367 \text{ W m}^{-2})/4 * (1.0 - 0.28) / (0.60 * 5.67 * 10^{-8} \text{ W m}^{-2} \text{ K}^{-4})]^{1/4}$$

$$T_E = 291.51 \text{ K} = 18.51^\circ \text{C} = \sim 19^\circ \text{C}$$

Thus, in a scenario involving a $\sim$ doubling of CO_2 and $\sim 7\%$ decrease in albedo (e.g. from 0.3 to 0.28 due to melting ice caps and other changes), average global temperature would increase by $\sim 4^\circ \text{C}$. While these calculations are clearly not as comprehensive as global climate models, they are the types of considerations and calculations that are used in modeling climate.

7.4 GREENHOUSE GASES: STRUCTURES, SOURCES, SINKS AND EFFECTS ON CLIMATE

As mentioned above, what differentiates greenhouse gases from non-greenhouse gases like N_2 (78.1% of dry air), O_2 (20.9% of dry air) and Ar (0.93% of dry air) is the symmetrical character of their molecular structures.

7.4.1 Molecular structures and vibrations of greenhouse gases

Examples of many greenhouse gas molecules are presented in Fig. 7.9.

For each of these gases, you will note a center of symmetry about which these molecules can stretch and bend when struck by IR rays (note: only one center of symmetry is shown for methane, but many centers of symmetry can be drawn). Ozone is symmetrical because bonds in both molecules shown on the left in Fig. 7.9 are resonance hybrids of the O_3 molecule – in other words, ozone contains two end-member structures that contain a single bond on one side and a double bond on the other, but these bonds resonate between the two sides of the central O atoms, producing an overall bond order of 1.5 for each side (depicted on the right in parentheses). The story is similar with N_2O , which is not truly symmetrical, but one of its resonant modes contains double bonds on each side of the central N atoms, making its structure similar to that of CO_2 , where the presence of a central atom permits stretching and bending of bonds on either side in the presence of IR radiation. Figure 7.10 illustrates the ways in which CO_2 molecules vibrate in the presence of IR radiation of different wavelengths, keeping in mind that all greenhouse gases will behave in similar ways in the presence of IR radiation of a range

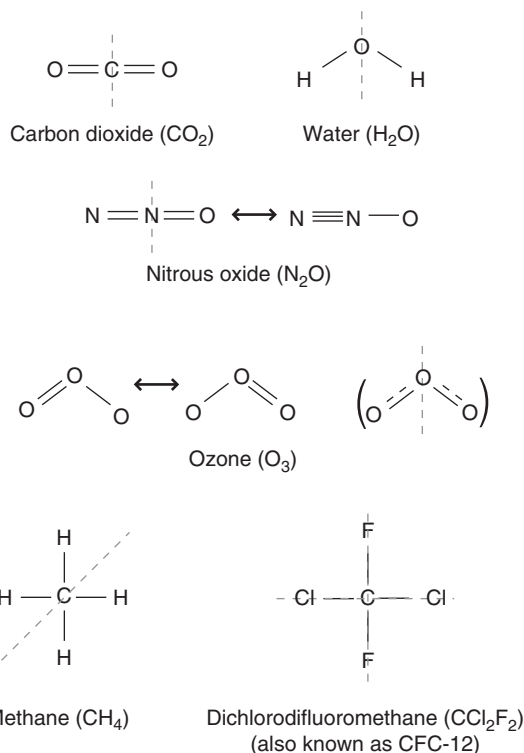


Fig. 7.9 The molecular structure of important greenhouse gases. In each case, dashed lines are drawn to emphasize the center(s) of symmetry in the molecules that allows them to stretch, flex and bend in the presence of infrared radiation. (Note: for CH_4 and CFC-12, four centers of symmetry can be drawn from the 2-dimensional representations shown – for clarity, not all of these centers of symmetry are shown).

of wavelengths. The wavelengths of IR radiation absorbed by CO_2 molecules are shown for each of the vibrational modes.

The vibrations shown in Fig. 7.10 cause IR radiation to be re-radiated in all directions, so while some of the IR energy will be directed towards the outer limits of the atmosphere, some will also be sent back towards the Earth surface, and this is how greenhouse gases “trap” heat (Fig. 7.11).

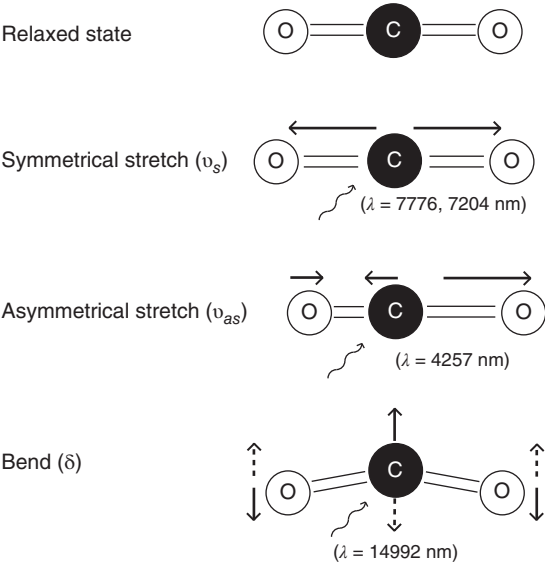


Fig. 7.10 Stretching and bending of the symmetrical carbon dioxide molecule in the presence of infrared radiation of varying wavelength.

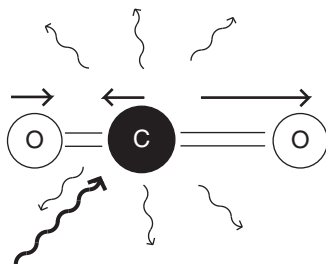


Fig. 7.11 Re-radiation of infrared energy by a CO_2 molecule undergoing asymmetrical stretching. The incident IR ray is the bold wavy line and the thin squiggly lines are IR radiation re-emitted by molecular vibrations. Note that the net effect is re-direction of outgoing IR radiation back towards Earth's surface.

In addition to the vibrations shown, there are also rotations that broaden the range of wavelengths over which CO_2 (and other greenhouse gases) can absorb IR. In contrast, the lack of a central atom in atmospheric gases like N_2 , O_2 , Ar, H_2 prevents them from vibrating (via stretching and bending) in the presence of IR radiation, and thus they are not capable of “trapping” outgoing IR radiation.

7.4.2 Greenhouse gases, radiative forcing, GWPs

Our understanding of the ability of CO_2 and other greenhouse gases to trap heat in the atmosphere is nothing new – it was recognized over one hundred years ago by the great Swedish chemist and Nobel prize winner Svante Arrhenius, who in the late 1800s became interested in the causes of ice ages. In 1896 he predicted that if CO_2 levels were to drop by 50%, Earth's surface temperature would fall by 4 to 5 °C; conversely, a doubling of CO_2 levels from 280 ppmv to 560 ppmv would produce an increase in surface

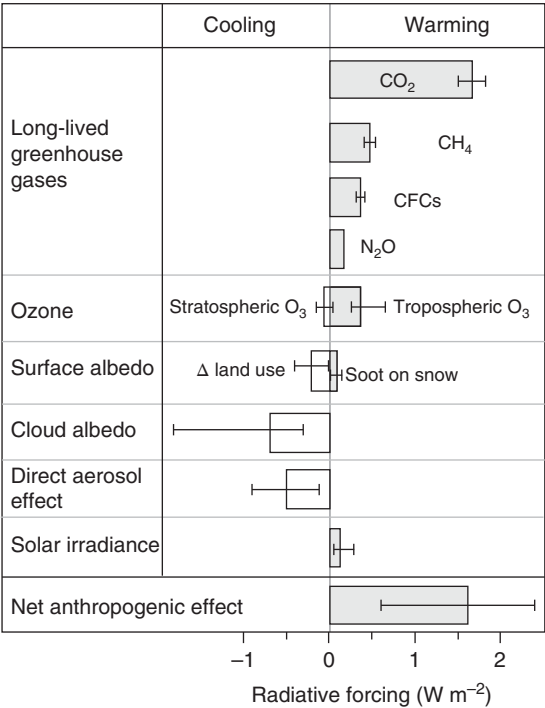


Fig. 7.12 Changes in radiative forcings from 1750 to 2005 including individual gases, changes to surface and cloud albedo, the direct effects of aerosols (e.g. sulfate particles), slight increase in energy emitted from the sun, and the net impact of human activities on radiative forcing. The net effect of changing atmospheric composition is an increase of 1.55 W m^{-2} in radiative forcing. (Modified from Forster et al. 2007.)

Table 7.2 Selected properties of greenhouse gases. Data are from Oak Ridge National lab (http://cdiac.ornl.gov/pns/current_ghg.html) and Forster et al. (2007).

	CO ₂ ppmv	CH ₄ ppmv	N ₂ O ppbv	O ₃ pptv	CFC-12 pptv
Atmospheric concentration					
Pre-1750	280	0.8	288	25	0
1990	355	1.7	310		484
Current (2011)	395	1.8	322	34	535
	Years	Years	Years	Years	Years
Atmospheric lifetime	100 ± 50*	12	114	<1	100
	100-yr	100-yr	100-yr	100-yr	100-yr
Global warming potential (GWP) (per-molecule basis)	1	25	298	na*	10,900
	W m ⁻²	W m ⁻²	W m ⁻²	W m ⁻²	W m ⁻²
Increased radiative forcing* (Integrated for whole atmosphere)	1.66	0.48	0.16	0.35	0.17

* The residence time of CO₂ in the atmosphere is difficult to precisely quantify due to the myriad removal mechanisms. A 100-year time horizon cannot be applied to ozone (O₃) given its short residence time. Increased radiative forcing is expressed as the change since 1750.

temperature of about 5 to 6 °C. Although current models predict that a doubling of CO₂ will yield a temperature increase more in the range of 3 ± 1 °C, Arrhenius' first-principles calculations were on the right track.

The net influence of all greenhouse gases and other climate-forcing agents (e.g. aerosol particles or net insolation) on the average net IR radiation leaving the atmosphere is known as **radiative forcing**. The combined radiative forcing of all greenhouse gases in the atmosphere produces the greenhouse effect; however, not all greenhouse gases contribute equally to this effect. Relatively abundant gases such as H₂O and CO₂ exert strong forcings based on their sheer abundance relative to trace gases. However, some of the trace gases, particularly CH₄, N₂O and CFCs, exert strong forcing due to their relatively high per-molecule **global warming potentials** (GWP), a parameter that is controlled in part by molecular structure (i.e. ability to vibrate in the presence of re-radiated IR), in part by their long residence times in the atmosphere, and also in part by where in the IR spectrum these gases are capable of absorbing IR radiation.

Figure 7.12 and Table 7.2 present data for some representative greenhouse gases (IPCC Fourth Assessment Report, 2007), where all values are normalized to CO₂ = 1. Recall that data on carbon cycling, including fluxes and reservoirs, are also presented in Chapter 5, notably in Fig. 5.5 and Table 5.4.

Note that many of the trace gases have GWPs tens to thousands of times greater than CO₂, indicating that small increases can have disproportionately large effects. In addition to per-molecule differences in absorption potential, a very important factor that contributes to the high GWP of chlorofluorocarbons like CFC-12 is that they are capable of absorbing IR radiation at frequencies outside of the ranges of absorption of the main greenhouse gases H₂O and CO₂. Figure 7.13 displays the IR spectrum from 1000 to 15 000 nm, a range over which H₂O and CO₂ both display strong absorption bands (note: absorption bands are defined as regions within the IR spectrum where these gases absorb high amounts of re-radiated IR radiation). Two points are worth noting here.

1 In some regions of the spectrum, all IR radiation is absorbed, so no increases in concentrations

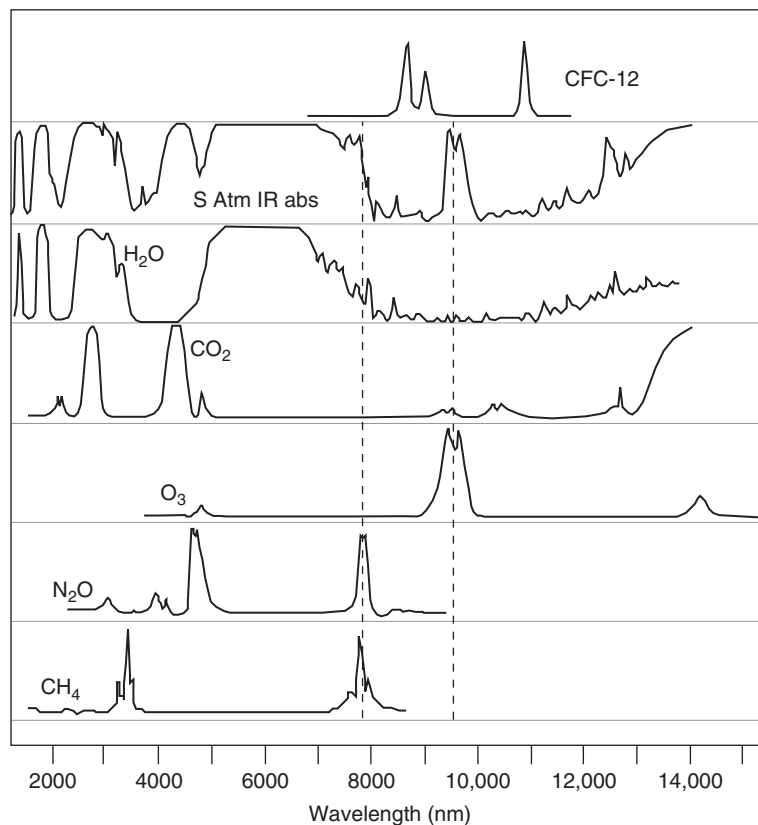


Fig. 7.13 Infrared spectrum of selected greenhouse gases showing the wavelengths in which each gas is infrared active, i.e. the regions in the spectrum where a given gas absorbs IR radiation. “S Atm IR Abs” is the spectrum produced by summing spectra of all gases with natural sources. Note that CFC-12 absorbs IR in two regions where naturally produced and anthropogenically enhanced gases (H_2O , CO_2 , O_3 , N_2O , CH_4) do not absorb (i.e. 8000–9000 and ~11 000 nm). Thus, any small additions of CFC-12 will have a large impact on IR absorption. Also note that IR absorption capacity is effectively saturated between ~5000–7000 nm, so any additional water (or other molecule that absorbs in this region) will not lead to increased absorption in this region.

of H_2O or CO_2 will have an impact on heat trapping, at least not in these IR ranges. Ranges in which the atmosphere is already saturated include ~4250 nm, where the asymmetrical stretching vibration of CO_2 (Fig. 7.11) absorbs 100% of re-radiated IR; another is between ~5000 to 7000 nm, where H_2O vibrations absorb 100% of re-radiated IR. However, there are many regions where the atmosphere is undersaturated, e.g. with respect to CO_2 at 10 200 – 10 500 nm and with respect to CH_4 and N_2O at 7500–8000 nm.

2 There are regions of the spectrum (e.g. between ~8000–11 000 nm) where neither H_2O nor CO_2 absorb IR radiation. It is these regions where additions of trace gases, especially ozone, nitrous oxide and chlorofluorocarbons, have great potential to absorb IR where little or none is currently being absorbed (or wasn't being absorbed prior to anthropogenic emissions). You will note that absorption lines for CFC-12 occur in the regions where neither H_2O nor CO_2 absorb. The same is true for N_2O and CH_4 , although these are less obvious than CFC-12 because the absorption

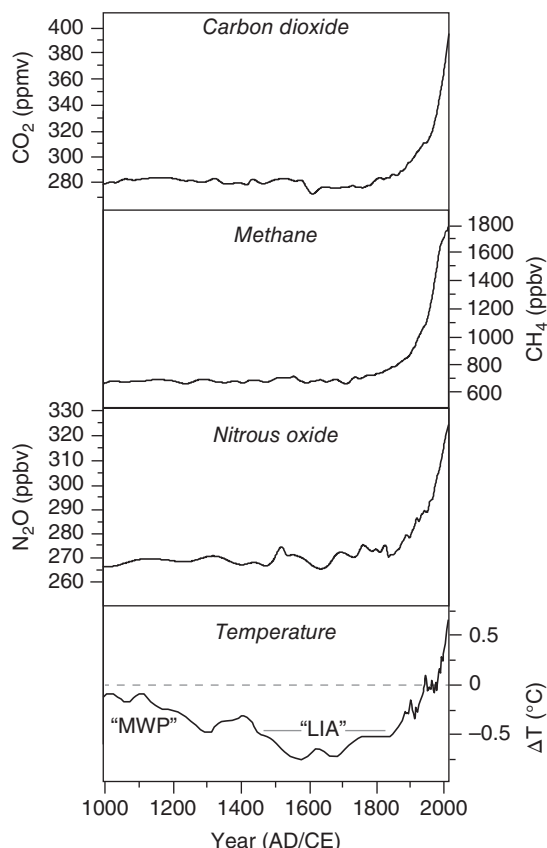


Fig. 7.14 Greenhouse-gas trends and temperature anomaly (relative to average temperature for the period 1951–1980) for the time period 1000 AD to 2010 AD. CH_4 , CO_2 and N_2O data for the period 1000 AD to 2004 AD are from MacFarling-Meure et al. (2006), with data from 2005 through July 2011 obtained from the NOAA Annual Greenhouse Gas Index (AGGI) web page accessible at <http://www.esrl.noaa.gov/gmd/aggi/>. Temperature data for the period 1000 AD–1880 are derived from ice core and firn data published by Moberg et al. (2005), and 1880 through 2010 data are from the Goddard Institute for Space Studies Surface Temperature Analysis web page (<http://data.giss.nasa.gov/gistemp/graphs/>), accessed 1 July 2011. Data are global mean land-ocean temperature index referenced to the base period 1951–1980. The temperature curve is a five-year running average. Two climate anomalies often referred to in the literature, the Medieval Warm Period (MWP) and the Little Ice Age (LIA), are shown here.

bands for these gases partially overlap or are close to bands for H_2O and CO_2 .

As human population has grown, very short-term anthropogenic activities (e.g. transformation of hydrocarbons stored deep within Earth's crust into energy and atmospheric CO_2) have accelerated the pace of CO_2 flux to the atmosphere. Whereas atmospheric CO_2 appears to have been present at an atmospheric concentration of 260 to 280 ppmv during the Holocene (12 000 years ago to ~250 years ago; Fig. 9.3b), the rate of addition of CO_2 to the atmosphere now greatly exceeds the rate at which it is removed from the atmosphere by geological, hydrological and biological processes; as a result, CO_2 has progressively increased in concentration since the onset of the industrial revolution. Current concentration (2014) is ~400 ppmv. Figures 7.14 presents millennial-scale trends in greenhouse-gas concentrations.

7.4.3 Global warming

As we learned in Chapter 5, *inputs* of anthropogenic carbon to the atmosphere are relatively easy to quantify, and while natural processes annually fix approximately one-half of these emissions into marine and terrestrial reservoirs, we are still left with the inescapable conclusion that human activities are adding greenhouse gases to the atmosphere; *net* flux of carbon to the atmosphere is currently on the order of 4 PgC/yr (4×10^3 TgC/yr), and the build-up of CO_2 , CH_4 , N_2O and other greenhouse gases leaves us with the inescapable conclusion that Earth's surface temperature is warming – to date, warming in the **troposphere** has reached a global average of 0.7 °C (Fig. 7.14). In response, other systems in contact with the atmosphere such as soils, shallow ocean waters and lakes also, on average, are warming.

The impacts and scope of this warming, including rising sea level, changing weather

and climate (cooling in some places, warming in others, desertification in some areas, increased rainfall in others), changes to fluxes in the hydrologic cycle (including increased flood intensity and melting glaciers), increased storm intensity, loss of CH₄ from deep-sea sediments, peat bogs and permafrost regions, and ecosystem shifts and losses, are too vast to cover within the scope of this chapter. In order to learn more of the science of climate change the reader is referred to Hansen et al. (2010) and in-depth books by Ruddiman (2007), Fletcher (2012) and many others.

QUESTIONS

- At current 400 ppm CO₂ in the atmosphere, the term for effective infrared transmission (g_h) = 0.61, and current albedo is ~0.30. Using the Stefan–Boltzmann Law, determine mean global temperature for the current atmosphere, and also for a potential future world with atmospheric CO₂ concentration of 550 ppmv (g_h = 0.60) and an albedo diminished by 10% (i.e. from 0.30 to 0.27).
- Describe the relationship between CO_{2(atm)} and mean global temperature over the past 400 000 years? Is it a direct correlation? Explain why or why not. Also examine the relationship between CO_{2(atm)} and mean global temperature over the past 400 million years. Is it a direct correlation? Explain why or why not.
- Calculate the density of relatively dry air (0.5% H₂O vapor) relative to humid air with 2.8% H₂O vapor. Express the answer in g/mol and in terms of g/m³. State any assumptions you might make to simplify the calculation.
- Diagram 5 greenhouse gases and explain what molecular properties cause them to “trap” infrared radiation. Describe how these compounds interact with IR at the molecular level. For comparison, also diagram two non-greenhouse gases.
- Critically assess this comment: adding more CO₂ to the atmosphere will not cause warming because the IR absorbing capacity of CO₂ in the atmosphere is already saturated in the 4250–4260 nm range. Is this statement true? False? Partially true? Justify.
- Using 288 K as the temperature of Earth, determine the maximum wavelength of radiation emitted from Earth’s surface.

REFERENCES

- Augustin, L., et al., 2004. Eight glacial cycles from an Antarctic ice core. *Nature*, 429(6992): 623–628. doi:10.1038/nature02599.
- Berner, R.A., 1997. The rise of plants and their effect on weathering and atmospheric CO₂. *Science* 276: 544–546.
- Berner R.A., 2006. GEOCARBSULF: a combined model for Phanerozoic atmospheric O₂ and CO₂. *Geochimica et Cosmochimica Acta* 70: 5653–5664. (doi:10.1016/j.gca.2005.11.032).
- Berner, E.K. and Berner, R.A. 1996. *Global Environment: Water, Air and Geochemical Cycles*. Prentice-Hall, Upper Saddle River, New Jersey, USA.
- Eckhardt, S., Breivik, K., Manø, S., Stohl, A., 2007. Record high peaks in PCB concentrations in the Arctic atmosphere due to long-range transport of biomass burning emissions. *Atmospheric Chemistry and Physics*, 7: 4527–4536.
- Fletcher C., 2012. *Climate Change: What the Science Tells Us*. Wiley, New York.
- Flückiger, J., Monnin, E., Stauffer, B., Schwander, J., Stocker, T.F., Chappellaz, J., Raynaud, D., and Barnola, J.-M., 2002. High resolution Holocene N₂O ice core record and its relationship with CH₄ and CO₂, *Global Biogeochemical Cycles* 16 (1). 10.1029/2001GB001417.
- Forster, P., V. Ramaswamy, P. Artaxo, T. Berntsen, R. Betts, D.W. Fahey, J. Haywood, J. Lean, D.C. Lowe, G. Myhre, J. Nganga, R. Prinn, G. Raga, M. Schulz and R. Van Dorland, 2007. Changes in Atmospheric

- Constituents and in Radiative Forcing. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. Accessed on-line: <http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter2.pdf>.
- Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010. "Global Surface Temperature Change". *Reviews of Geophysics* 48 (4): doi:10.1029/2010RG000345.
- Lovelock, J.E. and Margulis, L., 1974. Atmospheric homeostasis by and for the biosphere: the Gaia Hypothesis. *Tellus* 14: 1–10.
- MacFarling Meure, C., Etheridge, D., Trudinger, C., Steele, P., Langenfelds, R., van Ommen, T., Smith, A., and Elkins, J., 2006. Law dome CO₂, CH₄ and N₂O ice core records extended to 2000 years BP, *Geophysical Research Letters* 33: L14810, doi:10.1029/2006GL026152. <ftp://ftp.ncdc.noaa.gov/pub/data/paleo/icecore/antarctica/law/law2006.txt>.
- Moberg, A, Sonechkin, D.M. Holmgren, K. Datsenko, N. M., and Karlén, W., 2006. Highly variable Northern Hemisphere Temperatures Reconstructed from Low- and High-Resolution Proxy Data. *Nature* 433: 613–617.
- Murozumi M.T., Chow J., Patterson C.C., 1969. Chemical concentration of pollutant lead aerosols, terrestrial dusts and sea salts in Greenland and Antarctic snow strata. *Geochimica et Cosmochimica Acta*, 33: 1247–1294.
- Petit, J.R., Jouzel, J., Raynaud, D., Barkov, N.I., Barnola, J.M., Basile, I., Bender, M., Chappellaz, J., Davis, J., Delaygue, G., Delmotte, M., Kotlyakov, V.M., Legrand, M., Lipenkov, V., Lorius, C., Pepin, L., Ritz, C., Saltzman, E., and Stievenard, M., 1999. Climate and Atmospheric History of the Past 420,000 years from the Vostok Ice Core, Antarctica. *Nature* 399: 429–436.
- Ruddiman, W.F., 2003. The anthropogenic greenhouse era began thousands of years ago. *Climatic Change* 61: 261–293.
- Ruddiman, W.F., 2007. *Earth's Climate: Past and Future* (2nd edn). W.H Freeman and Co., New York.
- Spiro, T.G. and Stigliani, W.M., 2006. *Chemistry of the Environment*, 2nd edn Prentice Hall, Upper Saddle River, NJ, USA, 489 p.
- Steffensen, J.P., Andersen, K. K., Bigler, M., Clausen, H. B., Dahl-Jensen, D., Fischer, H., Goto-Azuma, K., Hansson, M., Johnsen, S. J., Jouzel, J., Masson-Delmotte, V., Popp, T., Rasmussen, S. O., Röthlisberger, R., Ruth, U., Stauffer, B., Siggaard-Andersen, M., Sveinbjörnsdóttir, A. E., Svensson, A., White, J.W. C., 2008. High-resolution Greenland ice core data show abrupt climate change happens in few years. *Science* 321: 680–684.
- Vonder Haar, T.H. and Suomi, V.E., 1971. Measurements of the Earth's radiation budget from satellites during a five-year period. Part 1: Extended time and space means. *Journal of Atmospheric Science* 28: 305–314.

Urban and Regional Air Pollution

The changes to the atmosphere discussed in Chapter 7 mostly pertain to gases that are not directly toxic to humans; after all, animals exhale CO_2 , and CH_4 and N_2O are emitted by natural processes. These gases are only harmful to humans in very high concentrations. Furthermore, CO_2 is emitted from such a wide range of sources (e.g. from point-source coal emissions to non-point-source automobile and household wood stoves), and this presents barriers to regulation that are more difficult to approach than solutions to commonly monitored and regulated “air pollutants”. As a result, emissions of greenhouse gases are not regulated as air pollutants and their emissions go largely unchecked.

The focus of regulatory policies has been on atmospheric constituents that are directly harmful to humans (or contribute to the formation of smog) and often are emitted from point sources (e.g. coal plants, sewage-treatment facilities) that are feasible to regulate. Well-known examples include lead (Pb), carbon monoxide (CO), fine-grained particulate matter (PM-2.5), nitrogen oxides (NO_x), ground-level ozone (O_3), sulfur

dioxide (SO_2), benzene and other volatile organic compounds, and arsenic (As) and cadmium (Cd). The first six of these are the criteria pollutants regulated by the US EPA in outdoor air. The European Commission regulates additional pollutants, including As, Cd and Ni, polycyclic aromatic hydrocarbons (PAHs) and benzene.

The health effects of air pollutants are wide ranging, from respiratory illnesses, especially for asthmatics (NO_x , SO_2 , PM-2.5 and O_3), to diminished delivery of oxygen to vital organs (CO), to impaired cognitive and neurological capabilities (Pb). Many volatile organic compounds (e.g. benzene and PAHs) are carcinogenic.

The following sections on air pollution and atmospheric chemistry will be organized as follows. First to be covered will be the important role that oxygen and water play in atmospheric reactions; decomposition of molecules of O_2 and H_2O in the presence of UV radiation create highly reactive free radicals, examples of which include oxygen, hydroxyl and peroxy radicals. Once these basics of atmospheric chemistry are covered, the text presents information related to

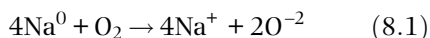
sources of air pollutants, recent trends in atmospheric abundance of pollutants and the chemical transformations responsible for decomposition of pollutants (e.g. of volatile organic compounds, NO_x , O_3 , CO , SO_2) or mechanisms by which less-chemically reactive constituents (e.g. particulates, Pb and other metals) are removed from air.

Additional topics in Chapter 8 include NO_x and SO_2 as agents of acid deposition ($\sim$ acid rain) formation, the roles of NO_x and O_3 in photochemical smog formation and the differences between and impacts of ground-level ozone and stratospheric ozone.

8.1 OXYGEN AND ITS IMPACT ON ATMOSPHERIC CHEMISTRY

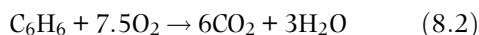
The abundance and reactivity of O_2 exerts a very strong control on the composition of the atmosphere. It creates a highly oxidizing environment that makes reduced or not-completely oxidized gases (e.g. volatile organic compounds, N_2O , CH_4 , CO , NO_x , SO_2) unstable, and reaction with oxygen is a common means of contaminant transformation in the atmosphere.

O_2 has a very high electronegativity (only fluorine's is higher) and hence will have a strong tendency to react with electron donors, a good example of which is sodium metal (Na^0), an atom that will lose a valence electron to O_2 gas. The reaction stoichiometry is:



Sodium metal is not expected to occur in the atmosphere because it is so unstable in the presence of free O_2 , so early in Earth's history virtually all sodium would have oxidized to Na^+ and either formed compounds (e.g. albite, $\text{NaAlSi}_3\text{O}_8$) or dissolved in water.

Organic matter (benzene in this case) is unstable in the atmosphere and undergoes oxidation to carbon dioxide:



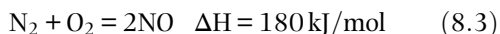
In this case, carbon atoms are converted from an average charge of -1 in benzene to $+4$ in carbon dioxide. However, as is apparent for the reaction of methane to carbon dioxide in Chapter 7 (equation 7.4), mere thermodynamic instability does not necessarily guarantee that organic matter or other chemically reduced trace constituents of the atmosphere will be oxidized. A typical benzene molecule will spend hours to days in the atmosphere before undergoing oxidation. Methane residence time in the atmosphere is 12 years and its concentration is very many orders of magnitude greater than would be predicted based solely on thermodynamic considerations – the strong kinetic control on these reactions was partially described in Section 9.1.

Factors controlling reactions of O_2 and other gases in the atmosphere include:

- 1 The probability of a collision involving oxygen and the reduced gas – trace gases may persist for longer than thermodynamically predicted due to relatively low amounts of collisions with O_2 .
- 2 The fleeting opportunity for O_2 and a gas like CH_4 to react once they collide – these are gases moving at high speeds so they commonly collide and immediately bounce off each other in a Brownian motion style. This creates very different reaction conditions from those at the surface – a good example of the opposite condition would be the drawn-out time frame involved with O_2 dissolved in slow-moving soil water in contact with pyrite.
- 3 Water is a critical medium for oxidation reactions. Owners of old airplanes store their planes in arid deserts. This is interesting because the H_2O -poor desert atmosphere contains more oxygen than humid air, yet planes will rust if stored in a rainforest. The obvious key here is the role of water, functioning as a medium for the

chemical reaction; thus, thinking about atmospheric reactions, if O_2 and an oxidizable gas both dissolve in water (e.g. a rain drop), the probability that reaction will occur will increase.

4 The O_2 molecule must overcome an energy barrier before it can react. A good example of an oxidation reaction that is not thermodynamically stable in the atmosphere at ambient temperatures (as compared to those in combustion chambers) is this reaction depicting oxidation of nitrogen gas to nitrogen monoxide.



The higher electronegativity of O_2 would have led us to predict that N_2 would lose electrons to O_2 , but the reaction is not spontaneous. The positive enthalpy of formation indicates that this is an endothermic reaction; with added heat (e.g. in internal combustion engines, coal power plant furnaces and hot forest fires), the reaction is more likely to occur. The cause of oxygen's inability to oxidize nitrogen gas in the atmosphere lies in the valence-electron structure of the O_2 molecule.

Oxygen atoms in O_2 molecules in the ground state have an electronic configuration of $1s^2 2s^2 2p^4$, indicating that O_2 contain four electrons in the valence shell. The four electrons can fill six possible locations distributed among three paired p orbitals. There are two possibilities for filling these sites, one involving an electron pair in the $2p_x$ orbital and two unpaired parallel electrons in the $2p_y$ and $2p_z$ orbitals (a configuration known as **triplet oxygen**; Fig. 8.1A). This is the stable, or ground-state, form of O_2 .

The other main configuration of O_2 is termed **singlet oxygen**, and consists of oxygen atoms with electron pairs in $2p_x$ and $2p_y$ orbitals but no electrons in $2p_z$ (Fig. 8.1B) – this hybridization of oxygen yields a molecule whose energy level is 94 kJ/mol above the energy level of triplet oxygen. It is far less stable than triplet oxygen and

exceedingly more likely to react with reduced compounds (e.g. organic molecules).

The differing electron arrangements can be depicted schematically as is shown in Fig. 8.1.

Because of its valence-electron structure, singlet oxygen can react with electron-pair donors and thus is far more reactive than triplet O_2 . One of the ways that singlet oxygen forms in the atmosphere is via reaction of triplet O_2 with organic compounds that contain *unpaired* electron and act as electron donors. Given that UV radiation can transform molecules to elevated energy states (i.e. with respect to valence electrons), these organic compounds with unpaired electrons often have been altered by UV radiation, making them more likely to react with triplet O_2 to yield the highly reactive singlet O_2 .

8.2 FREE RADICALS

While singlet oxygen is responsible for some of the oxidation reactions in the atmosphere, most of the airborne chemical species controlling oxidation

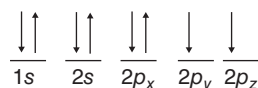


Fig. 8.1A Electron orbitals and spins (spin-up, spin-down) in triplet oxygen. This is the electron configuration of stable, ground-state oxygen.

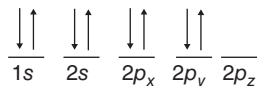
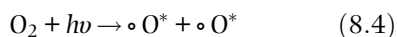


Fig. 8.1B Electron orbitals and spins (spin-up, spin-down) in singlet oxygen. Note vacant $2p_z$ orbital and paired electrons of opposite spin in $2p_y$. Singlet hybridization of O_2 is far more reactive than triplet oxygen.

reactions atmosphere are various types of **free radicals**, a class of atoms and molecules that possess unpaired electrons in valence orbitals. Most free radicals are electrically neutral, but some are ionic; in most cases, the unpaired electrons make free radicals extremely reactive. Free radicals typically are fragments of compounds that have been split apart by UV radiation or by chemical reactions between compounds and other free radicals, and the unpaired electrons that cause these atoms or molecules to be classified as free radicals are the remnants of paired electrons that once occupied a single (σ) chemical bond in a molecule.

What makes free radicals so unstable and reactive is the high-energy state of the unpaired electron. This electron will seek to find another electron with which to pair up and create a stable chemical bond. Often, in order to pair up the high-energy electron, the free radical will react with a molecule containing paired electrons (these types of molecules are far more common than those with unpaired electrons); in doing so, the reaction will create a new free radical. In some cases, two free radicals will react and the result is molecules that are no longer free radicals, but usually the result is formation of a different type of free radical. This can set off a “free-radical chain reaction” where a reaction of a free radical with a molecule containing paired electrons produces a new free radical and new molecule containing paired valence electrons, and so on. An example of such a chain reaction is the catalytic destruction of stratospheric ozone presented in Section 8.9.

A simple and illustrative example of free-radical formation is the **photodissociation** of oxygen in the presence of UV radiation:



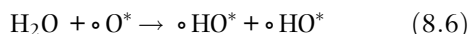
The symbol “ $\bullet$ ” denotes that the atom (or molecule) is a free radical, the superscript “ $*$ ” indicates the occurrence of a high-energy valence electron,

and the resulting oxygen *atoms* are high-energy free radicals.

One of the most important free radicals in atmospheric chemistry is the hydroxyl radical, $\bullet\text{HO}^*$, which forms in many ways, one of which is photodissociation of a water molecule:



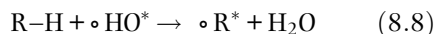
Hydroxyl radicals can also form when a high-energy free-radical oxygen atom reacts with a water molecule:



Hydroxyl radicals react with many compounds in the atmosphere, including nitrogen oxides, sulfur oxides and organic compounds. Hydroxyl radicals commonly react by either adding themselves to molecules or by abstracting a hydrogen atom from a molecule. An example of the latter (in this case, octane + hydroxyl radical) is shown here:

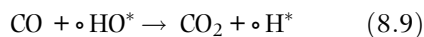


Often, hydroxyl radical reactions with organic compounds are generically represented as:

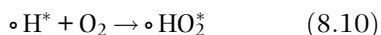


In this case, “R” represents some fragment of an organic compound and “ $\bullet\text{R}^*$ ” is the free radical formed when the hydroxyl radical stripped a hydrogen atom from the organic compound. The organic compound free radical is now very unstable and will likely react rapidly, and in this way, the hydroxyl radical contributes to decomposition of volatile organic compounds in the atmosphere.

Hydroxyl radicals also oxidize carbon monoxide to carbon dioxide, as follows:



Note that, in addition to CO_2 , a hydrogen radical is also produced. This neutral hydrogen atom, with its unpaired single electron, is very reactive and often combines with oxygen molecules to form the **hydroperoxyl radical**:



The hydroperoxyl radical, $\bullet \text{HO}_2^*$, also commonly forms as an intermediate product in the oxidation of most hydrocarbon combustion products and by-products.

This introduction to free radicals will provide a useful framework as we begin to analyze sources, residence times and reaction pathways of various air pollutants. The following sections discuss concepts related to pollutants in ambient air (i.e. outdoor air). Table 8.1 summarizes US and EU standards for ambient air and Plate 10 shows an urban air reporting system in Granada Spain.

8.3 SULFUR DIOXIDE

The majority of anthropogenic sulfur dioxide emissions globally are derived from coal combustion, which contributes approximately double the amount of SO_2 annually ($\sim 30 \text{ Tg SO}_2\text{-S}$) to the global atmosphere as does petroleum combustion, and as much atmospheric SO_2 as is contributed by all other anthropogenic non-fossil-fuel SO_2 sources combined (e.g. metal refining, international shipping, biomass fuels, waste incineration, and more).

Global anthropogenic emissions of SO_2 (Fig. 8.2) peaked in the early 1970s and generally declined until 2000, but began to rise again in the early 2000s as China and other industrializing countries have increasingly turned to coal as an energy source. Anthropogenic SO_2 emissions in 2010 were $\sim 60 \text{ Tg SO}_2$ (Smith et al., 2011), an amount that far outpaces natural SO_2 sources, which are on the order of $\sim 10 \text{ Tg SO}_2$ per year, predominantly derived from volcanic emissions (Stoiber et al.,

Table 8.1 Ambient Air Quality Standards According to the US Environmental Protection Agency (EPA) and the EU European Commission (EC).

Pollutant	Level (EPA)	Avg Time	Level (EC)	Avg Time
SO_2	30 ppb ($78 \mu\text{g}/\text{m}^3$)	1 yr	–	–
SO_2	–	–	$125 \mu\text{g}/\text{m}^3$	24 h
SO_2	75 ppb ($196 \mu\text{g}/\text{m}^3$)	1 h	$350 \mu\text{g}/\text{m}^3$	1 h
NO_2	53 ppb ($100 \mu\text{g}/\text{m}^3$)	1 yr	$140 \mu\text{g}/\text{m}^3$	1 yr
NO_2	100 ppb ($188 \mu\text{g}/\text{m}^3$)	1 h	$200 \mu\text{g}/\text{m}^3$	1 h
CO	9 ppm ($10 \mu\text{g}/\text{m}^3$)	8 h	$10 \mu\text{g}/\text{m}^3$	8 h
CO	35 ppm ($40 \mu\text{g}/\text{m}^3$)	1 h	–	–
$\text{PM}_{2.5}$	$15.0 \mu\text{g}/\text{m}^3$	1 yr	$25 \mu\text{g}/\text{m}^3$	1 yr
$\text{PM}_{2.5}$	$35.0 \mu\text{g}/\text{m}^3$	24 h	–	–
PM_{10}	$150 \mu\text{g}/\text{m}^3$	24 h	$50 \mu\text{g}/\text{m}^3$	24 h
O_3	0.075 ppm ($150 \mu\text{g}/\text{m}^3$)	8 h	$120 \mu\text{g}/\text{m}^3$	8 h
Pb	$0.15 \mu\text{g}/\text{m}^3$	3 month	$0.5 \mu\text{g}/\text{m}^3$	1 yr
As	–	–	$6 \text{ ng}/\text{m}^3$	1 yr
Cd	–	–	$5 \text{ ng}/\text{m}^3$	1 yr
Ni	–	–	$20 \text{ ng}/\text{m}^3$	1 yr
Benzene	–	–	$5 \mu\text{g}/\text{m}^3$	1 yr
PAHs (quantified by measuring Benzo[a]pyrene)	–	–	$1 \text{ ng}/\text{m}^3$	1 yr

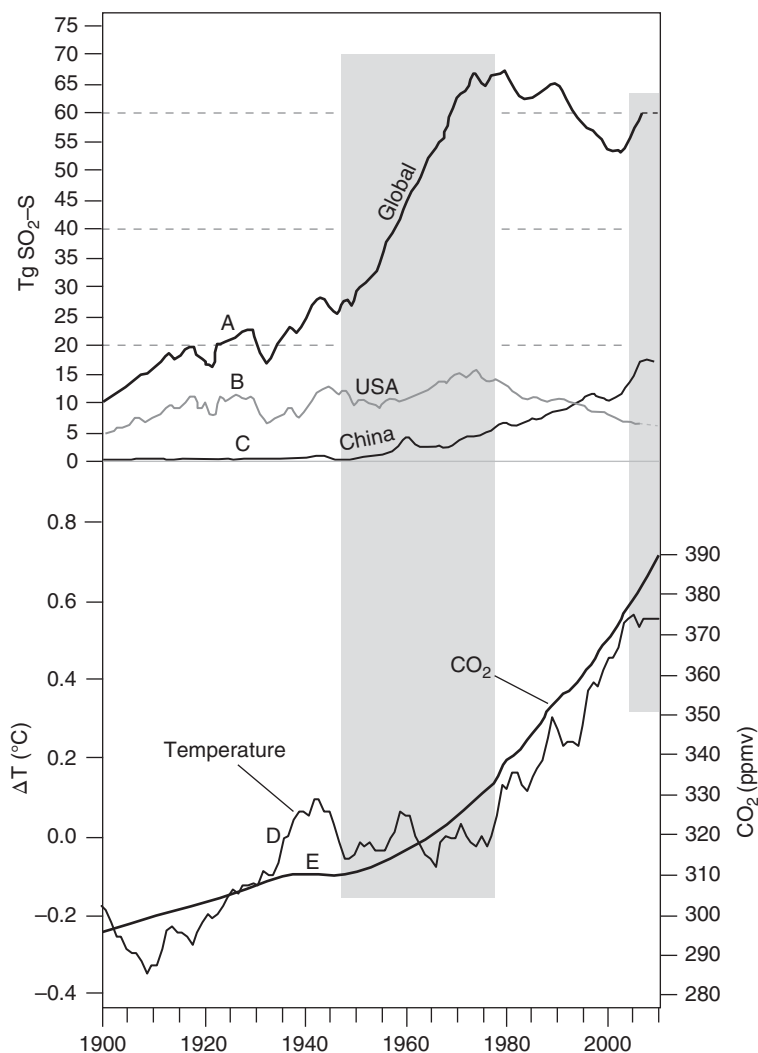


Fig. 8.2 SO₂ data from 1900–2010 showing (A) global emissions, (B) United States emissions, (C) China emissions, as well as global mean surface temperature (D) and atmospheric CO₂ concentration (E). Gray bars indicate time period where elevated SO₂ emissions (and resulting sulfate particle formation) may have mitigated greenhouse-related warming (e.g. Kaufmann et al., 2011). Sources for A, B and C are Smith et al. (2011), with C updated to 2008 by data from Lu et al. (2010); D is from **US Global Climate Change Report** (<http://www.globalchange.gov/publications/reports/scientific-assessments/us-impacts/full-report>) and E is from USA NASA Goddard Institute for Space Studies Surface Temperature Analysis web page accessible at <http://data.giss.nasa.gov/gistemp/graphs/> (note: data are accessible in tabular form as well as graphical form).

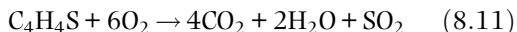
1987; refer to sulfur cycle section in Chapter 6). It is important to note that natural emissions of SO₂ constitute only ~10% of total natural S emissions; this is because the majority of natural S emissions to the atmosphere occur as low residence time sea salts (as sulfate) and dimethylsulfide (DMS, the gas derived from marine microorganisms). Viewed in this way, anthropogenic SO₂ emissions in 2010 were approximately 6 times greater than SO₂ from natural sources.

Coal typically contains 1 to 3% sulfur, which is contained mainly in pyrite (comprising 25 to 50%

of the sulfur in coal) and in thiophene (comprising 30 to 75% of S in coal) (Huffman et al., 1991). Pyrite (FeS₂) is a common mineral found in chemically reduced deposits such as coal, but it does not occur in hydrocarbons, helping to explain why coal typically contains double the sulfur content of liquid hydrocarbons.

Thiophene is an organic molecule with a pentagonal cyclic structure and the molecular formula C₄H₄S. The average oxidation state of sulfur in pyrite is S⁻¹; in thiophene, it is S⁰. Combustion of coal or any other fossil fuel causes

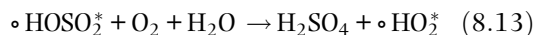
oxidation of reduced sulfur (S^{-1} , S^0) to S^{+4} and results in formation of SO_2 , which, for thiophene, can be expressed as:



The highest oxidation state of sulfur is S^{+6} and this is the form most stable in the oxidizing atmosphere, so once emitted from smokestack or tailpipe, SO_2 is eventually oxidized from $S^{+4} \rightarrow S^{+6}$ according to a wide array of reactions. Given the instability of the hydroxyl radical and the tendency of S^{+4} in SO_2 to lose two additional electrons, this is a common reaction:

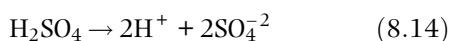


The compound $HOSO_2^*$ (which is also written as HSO_3^*) is an intermediate in the trajectory of sulfur dioxide to sulfuric acid; the oxidation state of sulfur is S^{+5} , so in order to reach the +6 oxidation state, $HOSO_2^*$ reacts with water and oxygen to produce sulfuric acid and the hydroperoxyl radical:



The hydroperoxyl radical will go on to rapidly react with other gases and the sulfuric acid will dissolve in water droplets and rain out of the atmosphere.

The first dissociation constant of sulfuric acid is very high ($K_a = 2.4 \times 10^6$) and the second (K_a for $HSO_4^- = H^+ + SO_4^{2-}$ is 1×10^{-2}) is sufficiently high so that nearly all H_2SO_4 completely dissociates (Brimblecombe, 2003) as follows:



The products of this reaction dissolve in water droplets and then are rained out of the atmosphere. The kinetics of the stages of transformation

from SO_2 to dissolved H^+ and SO_4^{2-} result in typical atmospheric residence times of SO_2 -derived sulfur of 2 to 8 days (Lelieveld, 1993).

While these reactions demonstrate that SO_2 will eventually react to form dissolved H^+ and SO_4^{2-} , it is important to recognize two important consequences of SO_2 reaction kinetics:

1 An SO_2 molecule typically remains in air for 1 to 2 days (Lee et al., 2011), during which time it can exert negative effects on the human respiratory system.

2 SO_2 and its derived reaction products (e.g. $\bullet HOSO_2^*$, H^+ , SO_4^{2-}) are commonly carried hundreds of kilometers in prevailing winds, meaning that SO_2 -derived sulfuric acid is often finally delivered in precipitation far downwind of the original SO_2 emissions. More on this will be presented in Section 8.10.

Also important to recognize is the role of sulfates as scatterers of incoming solar radiation, so when combustion adds SO_2 to air, one of the climate considerations will be the potential cooling effect exerted by sulfate particles. Cooling from the 1950s to the 1970s and mitigated warming between 2000 and 2010 (Fig. 8.2) may be due to sulfate emissions that are capable of counteracting radiative forcing of increased greenhouse-gas concentrations (e.g. Kauffmann et al., 2011). The main sources of SO_2 (and therefore, sulfates) in the 1950s–1970s were Europe and the United States, but stricter emissions standards (e.g. The Clean Air Act and subsequent amendments in the USA) have resulted in decreased SO_2 emissions in spite of increased fossil fuel use. Mitigated warming from 2000–2010 may be related to increased SO_2 emissions from southeast Asia, although emissions controls in China are already beginning to diminish SO_2 emissions from this region (Fig. 8.2). If Kaufmann et al. (2011) are correct in their assessment of the role of sulfates as warming mitigators, the high rates of temperature increase seen from ~1980 to 2000 are likely again appear.

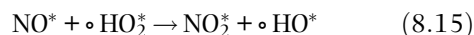
8.4 NITROGEN OXIDES

The abundant atmospheric molecules N_2 and O_2 do not react with each other at ambient atmospheric temperatures due to the strength of covalent bonds in these molecules and triplet O_2 's electron configuration (Fig. 8.1a), but in the elevated temperature of internal combustion engines and coal furnaces (typically $\geq 2000^\circ\text{C}$), N_2 and O_2 react to form NO_x – and the higher the temperature of combustion, the greater the amount of NO_x produced. (The heat of combustion provides the energy needed to facilitate this endothermic reaction). Of course, low combustion temperatures are undesirable because they create sooty emission products, so higher temperatures (~ 2000 to 2200°C) are necessary for efficient combustion and cleaner emissions in furnaces and engines, and one of the unavoidable results is the dissociation of N_2 and O_2 ; some dissociated molecules re-form as N_2 and O_2 , but others inevitably are reformulated into NO (“nitric oxide” or nitrogen monoxide) and NO_2 (nitrogen dioxide). There are also natural processes that foster formation of NO_x at elevated temperatures, particularly lightning and forest fires (thus, an additional anthropogenic source of NO_x to the atmosphere is biomass burning, whether in the field or in power-generating plants).

Regarding terminology, NO_x consists of (NO and NO_2); when other oxidized forms of nitrogen are also included (e.g. HNO_3 , NO_3 radical, N_2O_5), the blanket term is NO_y .

The nitrogen in NO has a +2 charge and in NO_2 it is N^{+4} . The highest oxidation state of nitrogen is +5, so in the atmosphere NO_x will oxidize to ultimately form N^{+5} (its form in the NO_3^- anion). The majority of NO_x produced by combustion of fossil fuels is NO, a stoichiometry dictated by the greater abundance of N_2 compared to O_2 in air pulled into the engine. NO emitted from stacks or tailpipes typically reacts relatively rapidly with free radicals; the example here involves reaction with

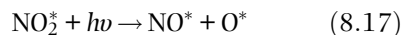
the hydroperoxyl radical to produce nitrogen dioxide and a hydroxyl radical:



NO can also react with O_2 to form NO_2 as follows:



The resulting NO_2 can follow different pathways. One is that it may decompose during the daytime in the presence of incoming UV radiation into nitrogen monoxide and oxygen atom radicals:

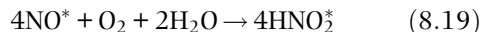


The oxygen radical commonly will bond with O_2 to form ozone, and this is one of the key reactions in the photochemical smog system. More on that topic will be presented in Section 8.8.

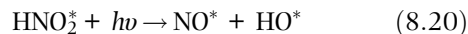
The NO^* radical may also react with O_2 (e.g. singlet O_2 or O_2 that has been turned into an electronically excited O_2 radical by UV):



NO can also react with oxygen and water to form nitrous acid:

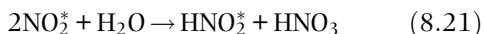


This is a **homogeneous reaction** because both NO^* and O_2 are in the gaseous state. Note that during this reaction the oxidation state of nitrogen has increased from +2 to +3. Nitrous acid is prone to photodissociation in the presence of UV to yield nitrogen monoxide and hydroxyl radicals, unstable and highly reactive products that then will immediately participate in other reactions (some of which are shown above and below).



Nitric acid can form by the **heterogeneous reaction** of nitrogen dioxide gas and water

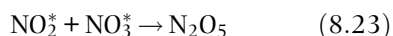
droplets. This is a heterogeneous reaction because the reactants occur in two different states of matter, gas (NO_2) and liquid (H_2O):



Nitric acid can also form in the homogeneous gas phase reaction of nitrogen dioxide and a hydroxyl radical:



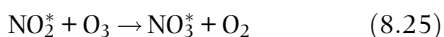
and by a pair of reactions that occur in the absence of solar radiation (i.e. at night):



and



The second of these two reactions takes place on the surface of an aerosol, and the nitrate radical (NO_3^*) shown also forms at night as follows:



HNO_3 is a soluble acid ($K_a = 2.4 \times 10^{-1}$) that dissolves in water to yield H^+ and NO_3^- , soluble ions that are removed from the atmosphere in precipitation. Along with H_2SO_4 , HNO_3 is a component of acid deposition (Section 8.10).

The residence time of NO and NO_2 are dependent on the composition of the atmosphere (i.e. presence/absence of HO^* , O^* , etc.) and time of day (presence/absence of solar radiation). If ambient air contains ozone (and thus O^*), NO will react very rapidly (minutes to hours); if not, reaction with the hydroxyl radical typically takes place within 1 to 2 days. In the presence of the hydroperoxyl radical, NO_2 will react in time spans of ~ 1 to 3 h; in the presence of intense solar radiation, NO_2 will decompose even more rapidly.

In the absence of HO_2^* or UV, NO_2 will react with HO^* on times spans of ~ 1 to 3 days.

NO_x is problematic for many reasons:

- 1 It is a biologically available form of nitrogen that contributes to eutrophication.
- 2 It ultimately results in formation of nitric acid, one of the two main components of acid deposition (along with sulfuric acid).
- 3 NO_2 is problematic for the human respiratory system, particularly for lungs affected by asthma or other respiratory disorders.

Globally, NO_x derived from fossil fuel-combustion (Fig. 8.3) shows a similar trend to global SO_2 emissions (Fig. 8.2), although it is evident that emissions controls have been somewhat more effective regarding SO_2 compared to NO_2 (Fig. 8.4).

To summarize NO_x atmospheric chemistry very succinctly, the NO emitted during fossil-fuel combustion is transformed into NO_2 that is then ultimately transformed to HNO_3 and rained out of the atmosphere as a component of acid deposition and as reactive nitrogen that contributes to eutrophication. However, as we have seen, the potential reactions of NO_x are abundant and include many species and various homogeneous and heterogeneous reactions that proceed in the direction $\text{NO} \rightarrow \text{NO}_2$, but also $\text{NO}_2 \rightarrow \text{NO}$. These reactive gases are extremely important components of the troposphere and stratosphere, and given their importance in photochemical smog, acid deposition and stratospheric ozone, NO_x will appear again in subsequent sections in this chapter.

8.5 CARBON MONOXIDE

Carbon monoxide is an odorless, tasteless gas that is toxic to humans because it has a much stronger affinity for hemoglobin sites than does oxygen (the difference is ~ 300 -fold), and for this reason, excess CO in air will diminish delivery of oxygen to cells in the body. At ambient concentrations (e.g. 100–200 ppbv in most air), CO will occupy

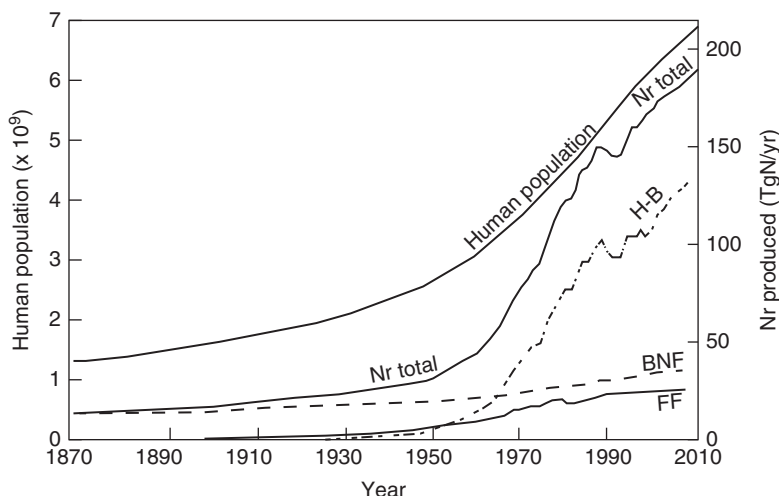


Fig. 8.3 Global human population from 1870 to 2010 (left axis is population in billions) plotted along with production of reactive nitrogen from different sources (the right axis is TgN/yr). Nitrogen sources are total reactive nitrogen; reactive nitrogen produced by the Haber–Bosch process (H–B); biologically fixed nitrogen (including natural sources plus microbial fixation in agricultural plants); and fossil-fuel emissions of Nr (FF). (Galloway 2007 and Erisman et al. 2008.)

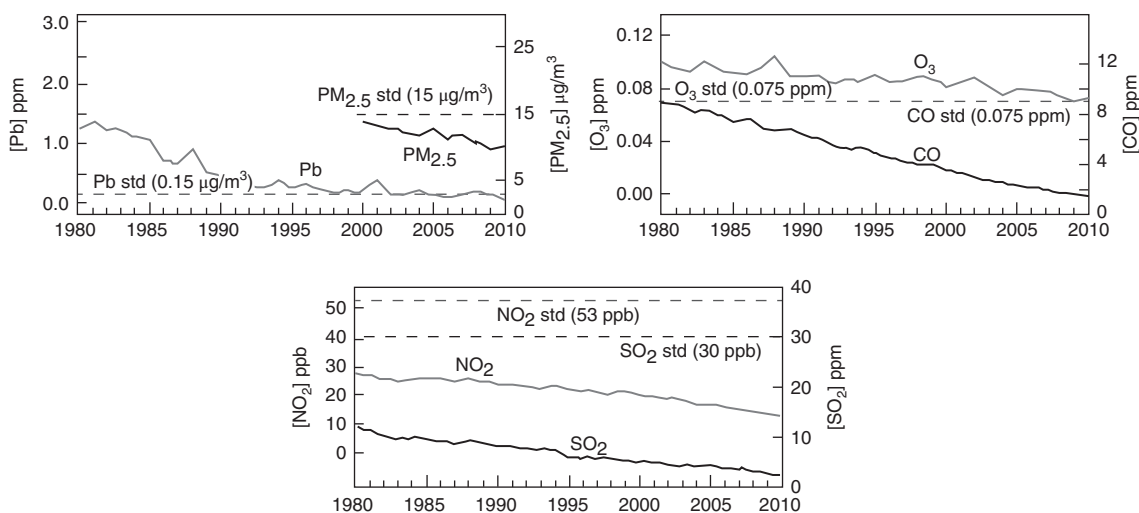


Fig. 8.4 Trends in ambient air contaminants in the US for the period 1980–2010, with the exception of $\text{PM}_{2.5}$, for which the standard was developed in the late 1990s. For details on standards, refer to Table 8.1.

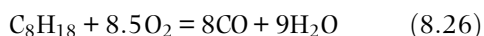
1% of hemoglobin binding sites; however, in contaminated urban air with concentrations of 10–50 ppmv (the EPA standards are 9 and 35 ppmv; Table 8.1), up to ~10% of hemoglobin

sites will be occupied by CO, causing headaches and difficulty breathing. In tunnels, parking garages and other enclosed areas where fossil-fuel emissions are not ventilated, concentrations can

reach 100 ppmv, and in indoor air affected by improperly venting furnaces and stoves, humans can lose consciousness when values reach 250 ppmv. Death is nearly certain at concentrations ≥ 750 ppmv. Even at lower levels (e.g. ~ 5 – 10 ppmv), humans with heart problems can experience heart attacks when exposed for short periods. In most cases, the effects of CO poisoning can be reversed by breathing uncontaminated air or by O_2 medical therapy.

Annual emissions of CO to the atmosphere are ~ 2800 Tg CO per year, and the three main sources are (1) oxidation of methane by the hydroxyl radical, which accounts for 800 Tg CO per year, (2) biomass burning, at 700 Tg CO per year, and (3) combustion of fossil fuels, which produces 650 Tg CO per year. These three sources comprise $> 75\%$ of the total CO flux to the atmosphere. Other sources include partial oxidation of organic compounds in air and emissions from vegetation and the oceans (Seinfeld and Pandis, 2006).

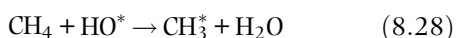
Incomplete combustion, whether during fossil-fuel combustion in a low air: fuel mixture, or during biomass burning, can be illustrated by this reaction of octane with oxygen:



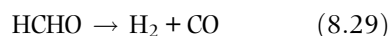
Increased airflow (i.e. increased O_2) will foster combustion that produces carbon dioxide:



In reality, the kinetics of combustion dictate that at least some CO will form during combustion, and the high concentrations of vehicle emissions and other sources of volatile hydrocarbons in urban and industrial areas can combine to produce CO-contaminated air. For example, derivation of CO from the oxidation of methane is a multi-step process, beginning with reaction of methane with the hydroxyl radical:



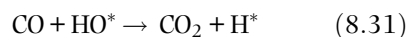
After additional steps, the reaction series culminates with photodissociation of formaldehyde ($HCHO$ or H_2CO) and, in some cases, subsequent rapid reaction with O_2



or



The details of this reaction series are beyond the scope of this book but are covered in Seinfeld and Pandis (2006) and other similar sources. The mean residence time of CO in the atmosphere is in the range of 30–100 days, and the main mechanism by which carbon monoxide is removed from the atmosphere is via reaction with the hydroxyl radical to produce carbon dioxide and a hydrogen radical:



Emissions controls and increased fuel efficiency have helped to reduce CO emissions in most countries since the early 1970s (see Fig. 8.4 for data on CO in the US).

8.6 PARTICULATE MATTER

Particulates in the atmosphere (commonly termed aerosols) are derived from many natural processes as well as human activities (which in most cases amplify natural processes). The main natural processes include

- wind erosion of soil, especially from arid regions, in which case the main aerosols are minerals such as calcite, gypsum, illite, iron hydroxides and much more;
- emissions of volcanic ash, which range from tiny bits of exploded volcanic rock to sulfate particles;
- wind-derived sea salts (sea spray, mainly sulfates); and
- soot from forest fires and emission of organic particulates from trees.

The size range of aerosols is ~ 1 nm ($\sim$ the size of a single unit cell of a clay mineral) to ~ 20 μm (fine silt); in reality though, most particulates are in the range of 0.1 μm ($10\,000$ nm) to 10 μm . Particles finer than 0.1 μm are simply less common, and at sizes greater than 10 μm , particles either are not easily entrained into air or, if they are eroded into air by wind, settle relatively rapidly.

In terms of chemical and physical characteristics, aerosols include:

1 Relatively insoluble inorganic particles, i.e. minerals (e.g. illite, quartz, kaolinite).

2 Relatively insoluble organic particles derived from incomplete combustion of biomass and hydrocarbons as well as from natural plant-generated organic compounds. These compounds range from soot, which is dominated by clusters of hexagonally arranged carbon atoms (like in graphite), and organic molecules such as benzo(a)pyrene (see below) and terpenes emitted by coniferous trees.

3 Soluble ionic salts, mostly chlorides (e.g. NaCl) as well as sulfates such as H_2SO_4 , NH_4SO_4 and NaSO_4 and others. These compounds function as particulates when they act as condensation nuclei. Ultimately, the salt will likely dissolve in the droplet once it coalesces around the salt nucleus.

4 Toxic metals (e.g. Cd, Cr, Ni), metalloids (e.g. As, Se), and organic compounds (e.g. polycyclic aromatic hydrocarbons or PAHs) are often sorbed onto particulates (mainly of classes 1 and 2 listed above), and thus can increase the toxicity of the particulate.

Aerosols can also be classified based on how they arrive in the atmosphere. Those that are introduced to the atmosphere in the solid state (e.g. sea salts, mineral particles, soot) are often referred to as **primary particulates**. Aerosols that form in the atmosphere from gases are sometimes referred to as **secondary particulates**, a good example of which is SO_4^{2-} derived from SO_2 or DMS.

The amount of human-derived particulate matter in a given area varies widely, with populated

areas containing the greatest amount of anthropogenic emissions; for example, Weijers et al. (2011) indicate that human activities account for roughly 75–85% of total particulates in the Netherlands. Globally, anthropogenic sources may account for $\sim 20\%$ of total particulates in the atmosphere (Wallace and Hobbs, 2006), a lower value that reflects the dominance of natural sources in rural areas (e.g. windblown dust in deserts, sea salts in coastal areas, terpenes in pine forests, etc.). The main anthropogenic sources are (1) combustion of fossil-fuel and biomass burning that results in soot and other organic aerosols, and (2) enhanced erosion of mineral and organic particles, mainly from roads and cleared lands (in some cases this can lead to desertification).

The functions of particulates in the atmosphere are covered in various detail in other sections of the text, so will only be presented in summary form here. Their main roles are as:

1 condensation nuclei;

2 scatterers of incoming solar radiation, especially sulfates;

3 absorbers in incoming solar radiation, especially black soot;

4 transporters of soil nutrients via wind-blown dust (e.g. Sahara to the tropics).

Countering their important roles in the atmosphere is the fact that particulates can be quite harmful to human health. In general, any particulates finer than 10 μm , particularly those finer than 2.5 μm ($\text{PM}_{2.5}$), are considered potentially harmful because they are difficult to expel from lungs, and thus they can accumulate in alveoli. So while quartz is not a notoriously toxic mineral, it can cause silicosis of the lungs when large amounts of $\text{PM}_{2.5}$ quartz are inhaled over years to decades. Of particular concern are asbestiform mineral grains, especially from the amphibole group (e.g. crocidolite, see Chapter 2), and compounds known to be carcinogenic such as benzo(a)pyrene, a polycyclic aromatic hydrocarbon (PAH) whose structure is shown in Chapter 3 (Fig. 3.10) and whose concentration

in air is regulated by the European Commission (Table 8.1). The only $PM_{2.5}$ class that is likely not harmful in elevated amounts are soluble, non-toxic molecules like sea salts. Fortunately, as is the case with many other ambient air pollutants, concentrations have shown notable decreases as a result of enforceable air standards (Fig. 8.4).

8.7 LEAD (PB)

The improvement in air quality with respect to lead since the 1970s is one of the great success stories of environmental regulation. Lead is a toxic metal that attacks the central nervous, cardiovascular, reproductive, and immune systems; it also impairs kidney function and can decrease the oxygen-carrying capacity of blood. Some of the most notable effects of lead exposure are manifested in children, where diminished development results in lowered IQs. In adults, lead can result in high blood pressure and heart disease as well as the other effects listed above. It may persist for extended periods in the human body because it can accumulate in bones (Pb^{+2} behaves somewhat like Ca^{+2} in the human body), then decades later be released from bone back into blood during periods of bone turnover, e.g. menopause and osteoporosis (Silbergeld et al., 1988).

Prior to the successes of restrictions imposed by law, automobiles were the main source of lead emissions to the atmosphere. The reason is that tetraethyl lead, $(C_2H_5)_4Pb$, was added to gasoline so as to prevent knocking (i.e. to improve combustion and engine output). However, it is clear from Fig. 8.4 that emissions have dropped drastically since the 1970s; the main sources of lead to the atmosphere in the 21st century are point sources such as smelter and metal refineries, waste incinerators and lead-acid battery manufacturers.

8.8 HYDROCARBONS AND AIR QUALITY: TROPOSPHERIC OZONE AND PHOTOCHEMICAL SMOG

Ozone (O_3) in the troposphere is problematic because, when inhaled, it can impair lung function. Common symptoms include throat irritation, inflammation of lung tissue, chest pain, coughing and congestion. Long-term exposure may permanently scar lung tissue, and it is particularly harmful to people with asthma, bronchitis and emphysema. Furthermore, ozone is a stronger oxidizer than O_2 so when it is elevated above background in the lower troposphere, it is also harmful to leaves of plants (and thus reduces crop yields and affects forest health) and can oxidize anything else prone to oxidation, notably rubber. These negative impacts can appear at concentrations as low as 100 ppbv (refer to Table 8.1 for current standards).

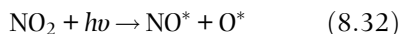
Photochemical smog and tropospheric ozone concentrations are closely related, so will be presented together in this section. In contrast to unhealthy ozone in the troposphere, ozone in the stratosphere is beneficial to life because it reacts with incoming UV radiation and thus limits the amount of UV that reaches Earth's surface – this topic is examined in Section 8.9.

There are two main types of smog (smoke + fog). One consists mainly of smoke, soot, and polycyclic aromatic hydrocarbons (PAHs) in air produced by partial combustion of fossil fuels in urban and industrial areas. The type of smoky, sooty air that plagued London for decades, most notably the 1952 smog event that resulted in thousands of premature deaths, is emblematic of this type of smog, where the contaminants are **primary pollutants** derived directly from tailpipes and smokestacks. This type of smog also resulted in 20 immediate deaths and thousands of cases of respiratory ailments in the small mill town of Donora, Pennsylvania, USA in 1948.

Photochemical smog is a different type of air pollution because it involves a series of reactions

resulting in a wide range of **secondary air pollutants** that are derived from primary pollutants, mainly NO_x and volatile organic compounds (VOCs) emitted by combustion of fossil fuels. The other key ingredient in photochemical smog is sunlight, hence the “photo” in photochemical. Photochemical smog is harmful to people because it consists of elevated concentrations of NO_2 , ozone and other substances that cause respiratory problems.

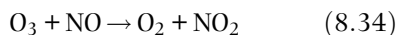
The series of photochemical smog reactions begins with the reaction of $\text{NO} \rightarrow \text{NO}_2$ as shown above in Section 8.4. Nitrogen dioxide is a brown gas which contributes to the characteristic color of photochemical smog. In the presence of sunlight, NO_2 breaks down, releasing an isolated, reactive oxygen atom:



The oxygen atom is then free to react with O_2 to form ozone:



Ozone is unstable so quickly reacts with nitric oxide (nitrogen monoxide):



Continued production of NO by combustion and resulting formation of NO_2 followed by photodissociation into NO^* and O^* results in continued production of ozone, and NO_x and ozone would reach a steady-state concentration controlled by NO production (from combustion) if not for the added impact of volatile organic compounds (VOCs) emitted by incomplete combustion.

Combustion by-products play an important role because they contribute to the formation of hydroperoxyl radicals that react with NO to produce NO_2 , which dissociates in the presence of sunlight to yield NO and O that then produce ozone according to the reactions shown above.

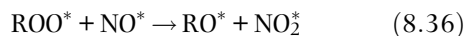
Thus, VOCs raise ozone concentrations above what would be produced by NO_x alone. Hydroperoxyl radical formation from VOCs can be described with the series of reactions shown below, understanding that these examples are representative of a very large family of reactions.

In these examples, R is used as a variable to represent a fragment of an organic molecule, commonly a broken (or partially combusted) hydrocarbon chain. For example, R could represent C_7H_{15} , a heptane molecule minus one H, in which case RH would be the complete heptane molecule C_7H_{16} ; alternatively, R could designate a fragment of an aromatic molecule (e.g. C_6H_5 , benzene minus one H atom), or any other potential by-product of incomplete combustion.

Given that hydrocarbon fuels are rich in alkanes, the “R” involved in photochemical smog is typically derived from an alkane. When emitted from a tailpipe or smokestack, the hydrocarbon can react with oxygen atom radicals and electronically excited oxygen gas to produce a hydrocarbon radical and a hydroxyl radical:



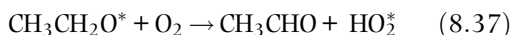
If R is an alkyl chain (e.g. the heptane fragment suggested above), the term for ROO^* (or RO_2^*) is **alkylperoxyl radical**; it is highly reactive and commonly will react with NO to produce an alkoxy radical (RO^*) plus nitrogen dioxide:



In this way, VOCs (e.g. uncombusted alkanes) contribute to smog formation by producing NO_2 , which yields O^* in the presence of sunlight, resulting in formation of ozone. Thus, VOCs and NO produced by combustion both yield NO_2 and are therefore both triggers for photochemical smog reactions.

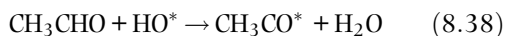
Some by-products of combustion can react to form additional harmful secondary pollutants. For example, the alkoxy radical may react

with O_2 to produce an aldehyde and a hydroperoxyl radical. In this case, to illustrate the behavior of H, an ethyl group is shown as the basis of the alkoxy radical (although keep in mind that alkoxy radicals typically are much larger molecules):



CH_3CHO is acetaldehyde (also known as ethanal in IUPAC system or as ethyl aldehyde), a probable carcinogen. If the radical above had been CH_3O^* rather than $CH_3CH_2O^*$, the product would have been $HCHO$ (formaldehyde), an irritant to eye, nose, throat, and skin and also a probable carcinogen. Then, to make matters worse, the hydroperoxyl radical will contribute to the formation of NO_2 and, through the reaction series shown above, ozone as well.

Acetaldehyde will react with the hydroxyl radical to produce a new alkoxy radical plus water:



This radical will rapidly react with oxygen to yield the acetylperoxy radical $CH_3COO_2^*$:



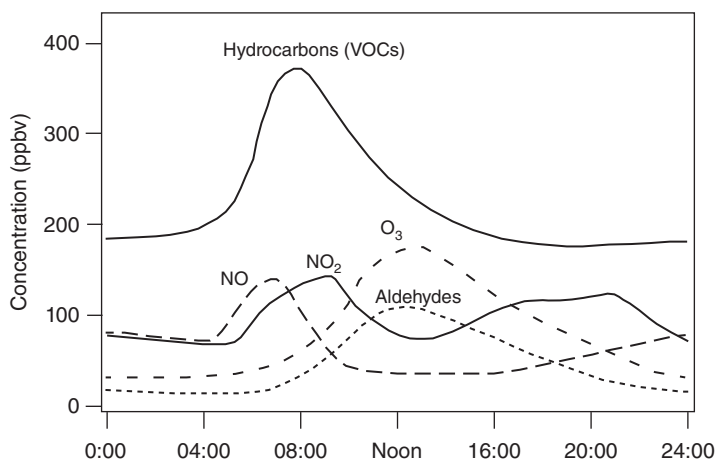
The acetylperoxy radical rapidly reacts with NO_2 to form the secondary pollutant **peroxyacetyl nitrate (PAN)**, a relatively stable gas and eye irritant commonly associated with photochemical smog.



The stability of PAN in air is very temperature-dependent, ranging from ~ 30 min at $25^\circ C$ (Bridier et al., 1991) to 5 yr at $-25^\circ C$ (Kleindienst, 1994); it often persists long enough to serve as a means of transporting reactive N in the atmosphere.

Figure 8.5 depicts a typical photochemical smog sequence instigated by morning automobile traffic showing important aspects of photochemical smog chemistry. First, as commuters begin their way to work, concentrations of VOCs and NO rapidly increase. Shortly thereafter we see evidence for the $NO \rightarrow NO_2$ reactions that occur in the presence of gases such as HO_2^* and O_2 . Then, dissociation of NO_2 in the presence of sunlight results in production of radical oxygen atoms (O^*), which react rapidly with O_2 to produce ozone, which commonly reaches its peak concentration in the afternoon. The drawn-out peak of ozone reflects production from NO_x as well as from VOCs and their by-products.

Fig. 8.5 Representative example showing concentrations of smog constituents instigated by morning vehicle traffic. (Adapted from Finlayson-Pitts & Pitts 2000 and Spiro & Stigliani 2006.)



Thus, in summary, Fig. 8.5 serves to illustrate (A) the effect of sunlight, particularly on dissociation of NO_2 and formation of O_3 , (B) the sequence of reactions, (C) the approximate overall rates of reactions, and (D) common concentrations of smog components, from background levels to peak levels. Also apparent is that NO_2 reaches its lowest concentration around noon, when maximum incoming solar radiation strongly drives $\text{NO}_2 + h\nu \rightarrow \text{NO}^* + \text{O}^*$. Concentrations of nitrogen dioxide then begin to rise in the early evening hours as photochemical dissociation of NO_2 slows down.

Most large cities are affected by photochemical smog to at least some extent. Notable examples are Los Angeles, Mexico City, Beijing and Rome. In many cases topography and climate can exacerbate problems, especially in cases where cities are situated in basins prone to air inversions that prevent circulation and allow primary and secondary contaminants to accumulate.

8.9 STRATOSPHERIC OZONE CHEMISTRY

At an altitude of approximately 25 to 30 km in the middle of the stratosphere, ozone concentrations are approximately $\sim 15\,000$ ppbv, or roughly 100 times greater than the amount of ozone present on a smoggy urban day. In the stratosphere, O_2 atoms are split apart into radical oxygen atoms by ultraviolet radiation (effectively, high-energy photons) with wavelengths < 242 nm. If you refer back to Fig. 7.7, you will see a graphical representation of the range of UV absorption by ozone. The pair of reactions is simple:



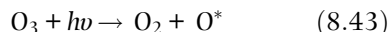
Then (equation 8.33):



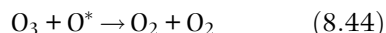
The second of these reactions occurs in the presence of a third molecule (e.g. N_2) which absorbs

some of the energy released by this reaction, explaining why there is no radical on the product (O_3) side.

Ozone concentrations are limited by dissociation in the presence of UV with $\lambda < 325$ nm:



and by reaction of O_3 with radical oxygen atoms (produced by dissociation of O_2 or O_3 by UV):



While both ozone and radical oxygen atoms are inherently unstable, the rate of this reaction is limited by the relatively low probability of a productive interaction of both gas-phase reactants in the diffuse stratosphere.

The UV-driven dissociation of O_2 and O_3 is what is often referred to as absorption of UV by the ozone layer – high-energy UV rays (with λ between ~ 230 and 320 nm) are consumed when bonds are broken, and this is crucial to life on Earth given the ability of UV to break bonds in organic molecules of living organisms, causing cancer and other problems.

The altitude of the ozone layer is controlled by the intensity of incoming UV radiation (it is much higher in the upper atmosphere compared to the lower atmosphere) coincident with the fact that there still are sufficient amounts of O_2 at that altitude to produce O_3 . Above 30 km, oxygen is too dispersed to result in significant ozone formation. The “thickness” of the ozone layer, or the abundance of stratospheric ozone, varies with latitude, being greatest in the tropics where insolation is greatest, and lowest at the poles, where insolation is lowest. Ozone production is also greater during summer months compared to winter months.

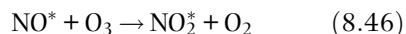
The long-term, natural steady-state concentration of ozone in the stratosphere is ~ 250 – 300 Dobson units, where one Dobson unit represents the hypothetical thickness of the ozone layer if it were pure ozone compressed under conditions of

0 °C and 1 atm. One Dobson unit (DU) = 10^{-2} mm or 10 μm ; thus, if the thickness of the ozone layer above some part of the Earth is 300 DU, the thickness of pure ozone in the stratosphere under hypothetical conditions of 0 °C and 1 atm would be 3 mm. One DU also equals 4.46×10^{-4} moles of O_3 per m^2 cross section of stratosphere (i.e. the amount of ozone in a vertical column 1 m $\times$ 1 m). The DU is named after G.M.B. Dobson of Oxford University, one of the early scientists to investigate atmospheric ozone and designer of the Dobson Spectrometer, the standard instrument used for ground-based measurements of stratospheric ozone.

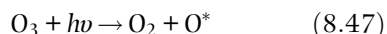
In the early 1980s, scientists working for the British Antarctic Survey published a paper demonstrating the presence of a “hole” in the ozone layer over Antarctica (Farman et al., 1985). The BAS had been measuring Antarctic stratospheric ozone concentrations since 1956 and the precipitous drop in ozone in the 1970s and 1980s (Fig. 8.6) was so extreme that it originally had been rejected as instrumental error. Previous research by the Dutch chemist Paul Crutzen in the late 1960s had determined that emissions of nitrous oxide, a stable gas in the troposphere, could reach the stratosphere and catalytically destroy ozone as follows:



Once produced, NO^* reacts with ozone to produce nitrogen dioxide and oxygen:



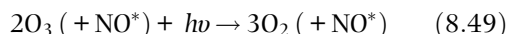
Given that UV causes dissociation of O_3 to produce radical oxygen atoms:



Radical oxygen atoms are then available to react with nitrogen dioxide:



The sum of reactions 8.46 through 8.48 is:



NO^* is ultimately neither produced nor consumed during the series of reactions, so it acts as a catalyst in the destruction of ozone. However, calculations indicated that nitrogen oxides were not present in sufficient quantities to produce the observed decrease in ozone indicated in Fig. 8.6; around the same time, a paper published in 1974 by University of California at Irvine researchers M.J. Molina and F.S. Rowland had demonstrated that chlorofluorocarbons, which like N_2O are effectively stable in the troposphere, could migrate up to the stratosphere and catalytically destroy ozone in much the same way that nitrogen oxides can.

In the presence of intense UV radiation of the stratosphere, CFCs (using CFC-11, or trichlorofluoromethane, as the example) dissociate into a chlorine atom radical and a CFC radical:

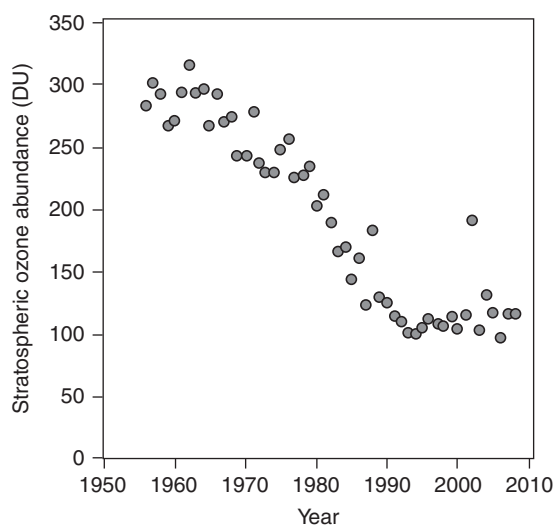
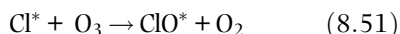
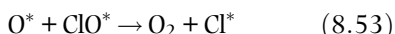
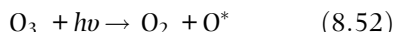


Fig. 8.6 Average stratospheric ozone abundance, in Dobson units (DU), for the month of October above Halley Station, Antarctica. Data are from the British Antarctic Survey (www.antarctica.ac.uk).

The chlorine radical then proceeds to behave like NO*:



Radical oxygen atoms produced by ozone dissociation can then regenerate radical Cl atoms by these two reactions:



The sum of reactions 8.51 through 8.53 is:



In this way, CFCs act to catalytically destroy ozone, and for their research, Crutzen, Molina and Rowland were awarded the Nobel Prize in Chemistry in 1995.

Interestingly, the stability of CFCs in the troposphere is part of what made them appealing as refrigerator fluids and propellents (in aerosol cans). Since the Montreal Protocol of 1987 and subsequent amendments, CFCs have largely been phased out and replaced by hydrochlorofluorocarbons (HCFCs, e.g. CHClF_2), molecules which have shorter residence times in the troposphere ($\sim 1\text{--}10$ yr) and are less likely to reach the stratosphere. HCFCs, however, are greenhouse gases, so research and development continues to seek compounds that are not environmentally harmful. Also, the long residence times of CFCs ($\sim 10^2$ yr) indicate that they will continue to wreak havoc on the ozone layer well into the 21st century (Fig. 8.7).

The interaction of the unique Antarctic climate and halogen chemistry plays an important role in ozone destruction. Atmospheric chemist Susan Solomon of the US National Oceanic and Atmospheric Administration (NOAA) attributed the strong seasonal component of ozone destruction (Fig. 8.8) to the presence of polar

stratospheric clouds (PSCs), wispy high elevation clouds that form only at very low temperatures (-80 to -85°C) during Antarctic winter. Observations had indicated a pronounced annual decrease in the stratospheric ozone above Antarctica during the early southern hemisphere spring (September and October), and Solomon determined that PSCs provide a template for chemical reactions that result in a large increase in concentrations of reactive chlorine in late winter (i.e. August).

The solid surfaces (ice crystals) in PSCs provide a template upon which less-reactive secondary chlorine gases (e.g. ClONO_2) can be converted to chlorine gas (Cl_2). This is a reaction that does not readily occur in the gas state. Then, as air begins to warm in late winter (August), the reservoir of Cl_2 accumulated by reactions on PSCs escapes from the ice crystal surfaces into the air, is converted to reactive Cl^* upon the return of sunlight and incoming UV radiation during early Antarctic spring, and then rapidly catalyzes ozone destruction – note the precipitous decline in ozone in September and October shown in Fig. 8.8. Due to differences in topography (the presence of large mountains in Antarctica and none at the North Pole) that result in differences in atmospheric circulation, these conditions do not occur in the Arctic.

Methyl bromide (CH_3Br) is another source of an ozone-depleting halogen, Br in this case. It has long been used as a fumigant in agricultural applications and it also has natural sources, mainly biomass combustion and marine plankton metabolism. Like CFCs, CH_3Br (MeBr) dissociates in the presence of UV and bromine radicals participate in ozone reactions in a manner similar to chlorine radicals. Its use is now being phased out in many countries.

Given that the “ozone hole” is located above the South Pole, the greatest impacts of ozone depletion are felt in the southern hemisphere in spite of the fact that the majority of CFC emissions have

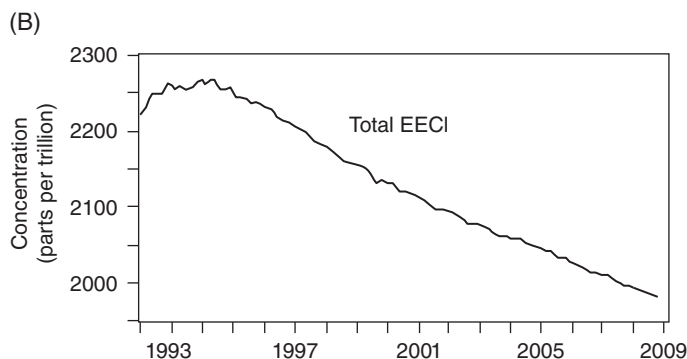
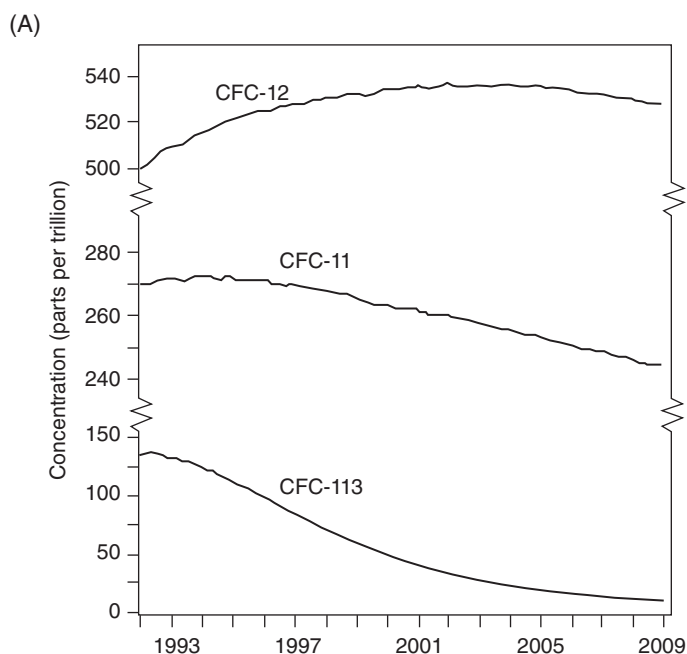


Fig. 8.7 CFC trends. (A) Trends in selected CFCs; (B) overall decrease in chlorine in the atmosphere expressed as total effective equivalent chlorine (EECl) for the time period 1992–2010. Data are from the Global Monitoring Division of the US NOAA Earth System Research laboratory (<http://www.esrl.noaa.gov/gmd/hats/>).

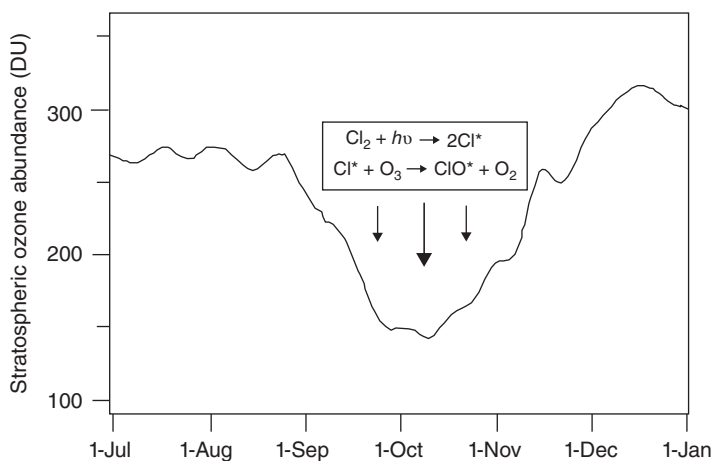


Fig. 8.8 Temporal evolution of ozone abundance above the South Pole showing average daily values from July 1 through January 1 for the years 1986 to 1995. Annual decrease in ozone abundance from September through November is due to the return of sunlight (and UV) and reactions of ozone with Cl^* in southern hemisphere springtime. See text for discussion of the role of polar stratospheric clouds (PSCs). (Data from Angell et al. 1995.)

been derived from industrialized countries of the northern hemisphere.

8.10 SULFUR AND NITROGEN GASES AND ACID DEPOSITION

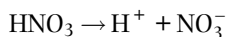
The pH of rainwater in equilibrium with carbon dioxide is ~ 5.6 , as can be determined from calculations as presented in Section 5.1. In some cases, unpolluted rainfall may have a pH below 5.6, especially in cases where nitrogen or sulfur compounds are emitted from forest fires, wetlands, oceans or volcanoes. In fact, rains that fall downwind of some volcanoes (e.g. Volcan Poas in Costa Rica, Plate 9) can be very acidic, with pH values as low as 2. However, such low values are generally constrained to within a few km of the volcanic source, and as a convention, “acid rain” is defined as having a $\text{pH} < 5$. Because water is delivered to the surface of Earth in many forms, including rain, snow, fog and dry deposition, the term acid precipitation is often used, as is the even more-comprehensive term is **acid deposition**, which is the term mainly used in this text. This latter term includes acids delivered to soils and water bodies as dry deposition, e.g. acids adsorbed onto particles, as well the acids dissolved in precipitation.

Fossil-fuel emissions are the primary sources of gases NO_x and SO_2 that transform in the atmosphere to nitric acid and sulfuric acid.

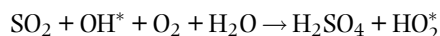
Formation of nitric acid can be represented as this gas-phase reaction (equation 8.22: note that the hydroxyl radical is written by some as OH^* and by others as HO^*):



and the dissociation reaction:



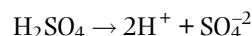
Formation of sulfuric acid can be visualized as a gas-phase reaction (equation 8.13):



Or as an aqueous-phase reaction (i.e. in water droplets):



And the dissociation reaction:



As we have seen, the chemical reactions involving nitrogen and sulfur species in the atmosphere are more complex than this, and the details can be found earlier in this chapter.

One important aspect of the reactions that ultimately result in formation of HNO_3 and H_2SO_4 are the effects of reaction kinetics. The gas-phase reaction of NO_2 with OH^* occurs approximately ten times faster than the reaction of SO_2 with OH^* (Seinfeld and Pandis, 2006); reaction of NO_2 with OH^* directly produces HNO_3 whereas reaction of SO_2 with OH^* yields the intermediate phase HSO_3^* , which then must undergo at least one more reaction stage before sulfuric acid is formed (e.g. dissociation to SO_3^* then reaction with H_2O , among many possibilities). The point here is that the reaction of NO_2 to HNO_3 occurs more rapidly than reaction of SO_2 to H_2SO_4 , and this can have implications for transport distance of N-derived acid deposition compared to S-derived acid deposition. In particular, in urban air with abundant OH^* and NO_2 , acid rain and acid fog may be deposited very close to the source of emissions, particularly in cities situated in enclosed basins like Los Angeles, where acid fogs have been observed to contain more than twice as much nitrate as sulfate (Waldman et al., 1982).

In general, acid deposition dominated by HNO_3 would be expected in areas with concentrated automobile use (hydrocarbons produce twice as much NO_x as SO_2), whereas acid deposition affected mainly by coal combustion will produce more than twice as much H_2SO_4 as HNO_3 (not accounting for measures to reduce S or N

emissions on site). Given that atmospheric residence time of NO_x , SO_2 and their derivatives are approximately 1–3 days (before being deposited in rain, snow or as dry deposition), and that air masses tend to travel 300–500 km per day, acid deposition commonly is delivered to the land or ocean surface 300–1500 km from its source.

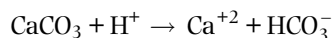
Prevailing winds shown in Fig. 7.5 illustrate the general directions in which acid deposition will be transported from source to deposition site. In the industrialized northern hemisphere between 30 and 60 degrees north, prevailing transport of HNO_3 and H_2SO_4 is to the north and east. Two of the more notable examples occur in the eastern United States and southeastern Canada, where coal power plant and industrial and automobile emissions from the southeast, Midwest and mid-Atlantic regions are carried downwind on prevailing westerlies to rural areas and forested ecosystems; and in northwestern Europe, where emissions from the UK and northern Europe affect downwind forested ecosystems, notably those of Scandinavia. Industrialization has led to acid deposition problems in eastern China, and to a lesser extent, in northern India, and at numerous sites globally where base metal and oil refining and heavy automobile traffic produce acids in the atmosphere.

Impacts of acid deposition are caused both by increased H^+ in precipitation water as well as increased concentrations of the nitrate and sulfate anions, and given that most rain and melted snow are cycled through soils before being incorporated into plants or delivered to streams and lakes, soil geochemical and biochemical reactions largely dictate the effects of acid deposition on a given area. Among changes to soil chemistry include acidification, loss of calcium and other nutrients, increased solubility of aluminum (toxic to most plants) and other metals, build-up of nitrate and sulfate, increased decomposition rates of natural soil organic matter, and leaching of H^+ and NO_3^- into surface water bodies; in turn, these changes produce further alterations to

surface water chemistry, producing negative consequences such as acidification, eutrophication (due to increased NO_3^-), and decreased aquatic ecosystem health.

Given that acid deposition consequences are strongly controlled by soil mineralogy, chemistry and biota, these topics are covered in Chapter 9. However, some of the negative consequences take effect before acid deposition even reaches the soil. These include:

- Damage to leaves and needles of trees that are directly exposed to acidic water droplets in forests of high mountain regions bathed in clouds and fog. In the presence of acids in air, essential nutrients can be stripped away from leaves and needles (e.g. exchange of H^+ for Ca^{+2}), making trees more susceptible to damage by other environmental factors such as disease, pests and cold winter weather. That said, most of the negative impacts of acid deposition on forested ecosystems are caused by changes to soil chemistry (Chapter 9).
- Chemical disintegration of buildings, bridges, statues, memorials and other objects made of limestone, marble or certain metals (e.g. bronze, copper). An illustrative example of the effect of acid deposition is the way in which rain water of pH = 3 to 5 attacks the calcite in limestone and marble:



In this way, acid deposition effectively accelerates the natural chemical weathering of carbonate rocks, converting solid calcite to dissolved calcium and bicarbonate ions.

- Because the sulfates derived from dissociation of sulfuric acid are light-scattering particles, acid deposition can make air hazy and impair views, which can negatively impact tourism in otherwise pristine regions.

Environmental policy designed to reduce SO_2 and NO_x emissions is beginning to result in decreases in acid deposition in many regions where impacts have been the most severe, notably in eastern North America and northwest and eastern

Europe. Indications that emissions of SO_2 in China are beginning to decrease (Fig. 8.2) suggests that acid deposition may begin to decline there as well. Plates 14–16 presents three series of maps showing increasing pH and decreasing nitrate and sulfate concentrations of deposition in the United States in 1994, 2002 and 2009. There are a few important considerations when looking at these maps.

1 The trajectories of H^+ , NO_3^- and SO_4^{2-} show the potential effectiveness of legislation in improving air quality.

2 Amendments to the 1970 US Clean Air Act were initially successful at addressing sulfuric acid deposition, but not nitric acid deposition. Between 1994 and 2002, notable improvements were made in sulfate emissions, but nitrate emissions increased during this period, and it was not until the 2002 to 2009 period that nitrate emissions began to improve. The reason for the difference between sulfate and nitrate is that the majority of sulfuric acid is derived from point source coal-fired power plants emissions, which are easier to control at the point source than are the non-point-source automobile and agricultural emissions responsible for much of NO_x emissions.

3 Rather than the broad regional pattern depicted by sulfate emissions in the 1990s, nitrate in precipitation in some areas exhibits localized bulls-eye patterns, often surrounding cities such as Salt Lake City, Utah (SLC) and the Colorado Springs – Denver – Fort Collins corridor (the Front Range, FR) in Colorado, much of it related to automobile emissions. Decreases from 2002 to 2009 are at least partly due to US Clean Air Act amendments requiring that, by 2004, sport utility vehicles (SUVs) and other light-duty trucks become subject to the same national pollution standards as cars.

4 Nitrogen is broadly elevated across the agriculturally intensive High Plains (HP) states, and in spite of reductions between 2002 and 2009, this region continues to produce some of the highest

N concentrations in the US. In much of the central High Plains, ammonium comprises 25–75% of total nitrogen deposition (Schmeltz, 2009). One certain source of NH_4^+ is animal feedlots (concentrated animal feeding operations or CAFOs; Todd et al., 2008), where ammonia derived from urea and other waste products is transformed to ammonium ($\text{NH}_3 + \text{H}^+ \rightarrow \text{NH}_4^+$), and the abundance of ammonium in air in this region explains why the high N deposition in the Midwest is not correlated to acid deposition. Rather than contributing to acid deposition as does NO_x , ammonia emissions tend to neutralize acidic precipitation, yet are still problematic because they contribute to N pollution of surface and groundwaters, including the Gulf of Mexico “dead zone” at the mouth of the Mississippi River.

5 While atmospheric deposition of H^+ , NO_3^- and SO_4^{2-} are decreasing, it will likely require decades for forested and aquatic ecosystems to return to pre-acid deposition conditions. Unfortunately, the compositions of sandy, inherently nutrient-poor soils typical of many forested alpine regions have already been significantly altered, as have the pH and dissolved ion chemistry of lakes in these areas. Accumulated acidity, sulfate and nitrate in soils will continue to leach into surface water bodies for many years or decades, so ecosystem recovery, especially of high alpine forests and lakes, will take time.

It is important to keep in mind that the maps in Plates 14–16 show *concentrations* of H^+ , NO_3^- and SO_4^{2-} , but they do not show total *quantities* deposited on the landscape (where quantities are expressed in units such as kg/hectare or mol/hectare, rather than mg/L or mmol/L). This becomes an important consideration when precipitation is not uniform; for example, a mountain range that receives twice as much annual precipitation as an adjacent valley will receive twice as much H^+ , NO_3^- and SO_4^{2-} if the concentrations of these ions are equal in rainfall in the two areas. In the northern Appalachians of North America, total

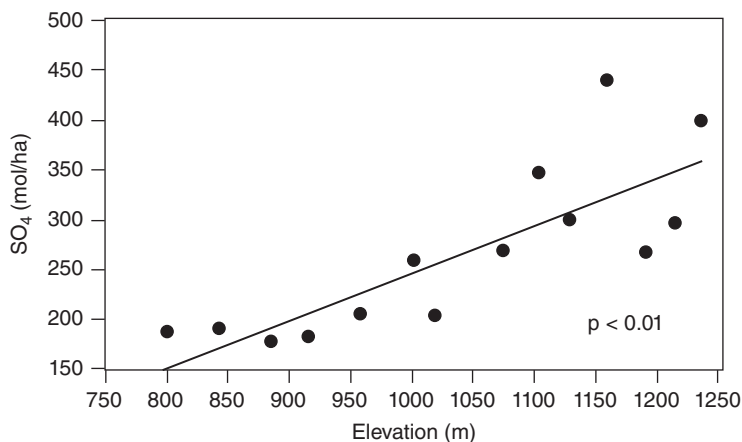


Fig. 8.9 Correlation between elevation and sulfate deposition in the Catskill Mountains of New York State. Data and statistical analysis are from Lawrence et al. (1999).

deposition of H^+ , NO_3^- and SO_4^{2-} in high-elevation areas is often 3 to 7 times greater than deposition to nearby low-elevation sites (Lovett and Kinsman, 1990). Figure 8.9 depicts the strong orographic control on the quantity of sulfate deposited in a mountainous region of northeast North America affected by acid deposition, demonstrating the added acid deposition load that falls on alpine regions compared to valleys and plains.

The *concentration* of sulfate in eastern New York State does not appreciably vary (~ 1.5 – 2.1 mg/L in 1994); if anything, the sulfate concentration in precipitation falling on the Catskill Mountains of southeastern New York State appears to have been slightly lower (1.8 mg/L) in 1994 than it was at lower elevation near New York City (2.1 mg/L) (Plate 15). However, the orographic precipitation effect yields significantly greater sulfuric acid *deposition rates* at higher elevation (~ 400 – 450 mol/ha) than at lower elevation (~ 175 – 200 mol/ha) (Fig. 8.9). This combination of orographic effect and inherently nutrient-poor soils in alpine regions combine to explain why the greatest impacts of acid deposition are typically seen at high elevations. More on this topic is presented in the framework of soil biogeochemical and mineralogical reactions in Chapter 9.

8.11 TRACE ELEMENTS IN ATMOSPHERIC DEPOSITION: ORGANOCHLORINE PESTICIDES, MERCURY AND OTHER TRACE ELEMENTS

Numerous toxic trace elements and compounds, both organic and inorganic, are emitted to the atmosphere from a wide range of sources, from energy generation to transportation, agriculture, residential activities and industry. These include the six pollutants presented above (SO_2 , NO_x , particulates, Pb, CO, O_3), but many others as well, including a wide range of organic compounds, including fuel derivatives, industrial by-products, solvents and pesticides, as well as trace metals and metalloids.

8.11.1 Pesticides in air

Common organic compound contaminants found in air include VOCs derived from incomplete combustion (as discussed above), pesticides applied to farmlands, golf courses, residential lawns, parks, etc., and dioxins and furans released from paper-pulp bleaching, waste incineration and manufacture of organochlorine chemicals

used as pesticides and solvents. Usually, harmful effects from organic compounds in air are chronic and express themselves in the form of cancer, disruption of endocrine and nervous system function, or other diseases after years of low-level exposure, but there have also been cases of acute poisoning associated with accidents, perhaps the most tragic being the explosion at a pesticide manufacture plant in Bhopal, India in 1984 that released methyl isocyanate, CH_3CNO , into the air of Bhopal, resulting in thousands of poisoning deaths. Methyl isocyanate is an intermediate compound produced during the manufacture of pesticides, carbaryl in the Bhopal case. Carbaryl, a likely carcinogen according to the US EPA, is marketed as Sevin and is a common insecticide used in residential and agricultural applications. Like many pesticides it is an acetylcholinesterase inhibitor (also referred to as an anti-cholinesterase chemical); compounds in this class inhibit the breakdown of acetylcholine by the cholinesterase enzyme, and thus impair central nervous system function.

Pesticides enter air in two main ways. First, given that pesticides are often applied by aerial spraying, some enter air immediately as aerosols or volatiles. Second, the high vapor pressure of most synthetic pesticides means that at least some portion will vaporize into air from the soil or leaf surfaces, and volatilization can continue for days after application. In some cases, 80–90% of applied pesticide can volatilize within two to three weeks of application (Unsworth et al., 1999), upon which it can be transported on prevailing winds to down-gradient areas. Decomposition of pesticides in air occurs by hydrolysis, reaction with OH^* radicals and by deposition, commonly in rain. Pesticides that resist hydrolysis and reaction with OH^* can be transported hundreds to thousands of kilometers, and many organochlorine insecticides have been detected in polar air, waters, and terrestrial ecosystems (Bidleman et al., 1989). While typical concentrations of pesticides in air are in the parts per trillion and parts

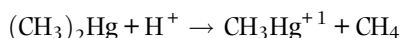
per billion range, their propensity to undergo bioaccumulation in organisms and biomagnification in ecosystems means that trace amounts in air can have noticeable effects on ecosystems and humans too.

8.11.2 Hg in air

Mercury (Hg) is a heavy metal that causes impaired neurological development in children and the developing unborn resulting in impaired brain and nervous system development. Common symptoms include diminished cognitive ability, memory deficits, and incomplete development of skills in the areas of language, fine motor movement and visual-spatial relationships. These effects are most pronounced when exposure occurs during fetal development. Effects on adults are generally not observed except in unusually high concentrations associated with point source contamination or occupational exposure, and include tremors, mood swings, irritability, neuromuscular changes (e.g. muscle atrophy and twitching), headaches and diminished cognitive abilities. The most extreme case of mercury poisoning likely occurred in the 1950s in Minimata, Japan, where a polyvinylchloride (PVC) manufacturing plant was using Hg^{+2} as a catalyst. Discharges to Minimata Bay contaminated sediments and the inorganic Hg^{+2} was transformed to the far more toxic dimethyl mercury, which worked its way up trophic levels via bioaccumulation and biomagnification, ultimately resulting in hundreds of deaths and developmental disabilities in infants exposed during fetal development.

The main pathways by which Hg is released to air in the 21st century are coal combustion and waste incineration. According to the UN Environment Programme, coal combustion is the largest single source of Hg to air and the US EPA estimates that 87% of anthropogenic Hg introduced to air is derived from point-source combustion, mainly

from coal-fired power plants, waste incineration (e.g. mercury thermometers, fluorescent light bulbs and batteries) and industrial energy generation. Incineration of fuels and wastes volatilizes mercury and it escapes to the atmosphere mainly as elemental mercury (Hg^0) which then is oxidized to Hg^{+1} and Hg^{+2} . These inorganic forms of Hg (Hg^{+1} , Hg^{+2}) are less toxic than Hg^0 and organo-mercury (e.g. methylmercury), but when deposited on soils or in lake and marine sediments, Hg^{+1} and Hg^{+2} can be methylated by microbially mediated reactions to produce methylmercury (CH_3Hg) or dimethyl mercury ($[(\text{CH}_3)_2\text{Hg}]$). $(\text{CH}_3)_2\text{Hg}$ can volatilize from soils and sediments but when it occurs in slightly acidic to acidic forest soils, the dimethyl mercury can react with H^+ to form a water-soluble methyl mercury cation:



Because $\text{CH}_3\text{Hg}^{+1}$ is soluble it tends to be leached into surface waters where it can be incorporated into organisms, and once ingested, methyl mercury crosses the intestinal wall. Then, because it binds to S atoms in proteins, it accumulates in the tissues of organisms and tends to biomagnify, causing upper-trophic-level organisms (e.g. predatory fish and birds) to contain the highest concentrations of methylmercury in the food chain. When humans eat trout, salmon and other high trophic level fish, they too can accumulate Hg, and this is most problematic for young children and developing fetuses. Emissions controls are beginning to have positive impacts in some areas, notably Canada, where emissions have decreased $\sim 80\%$ since the early 1970s, in the EU, where emissions are down $\sim 60\%$ since the late 1970s, and in the US, the EPA is in the process (2011) of implementing new rules that should appreciably decrease Hg emissions.

There are far too many toxic trace metals and metalloids emitted to the atmosphere each year to cover in detail given the scope of this text. We will

limit the scope of this final segment of this chapter by referring to Table 8.1, which indicates that, in addition to the 6 criteria pollutants regulated by the US Clean Air Act, the European Commission also regulates arsenic, cadmium, nickel, benzene and polycyclic aromatic hydrocarbons. The main sources of the latter two are fossil-fuel combustion and this topic was covered in Section 9.5, mainly under the topic of photochemical smog. Both benzene and PAHs are carcinogenic volatile organic compounds (VOCs), and their impacts, sources, fates and atmospheric reactions are similar to those for the hydrocarbons presented in Section 8.8.

8.11.3 As, Cd and Ni

The metalloid **arsenic** and the trace metals **cadmium** and **nickel** all share common sources; for example, they all occur in sulfide metal ores and thus can be released from ore smelters. They also occur in trace amounts in coal and are released upon combustion in the absence of strong emissions controls. All three largely occur in air in the particulate phase, often sorbed onto soot or other aerosols (e.g. oxides and hydroxides). Because nickel is more abundant in the crust of the Earth, natural and anthropogenic processes release approximately 10 times as much Ni to air compared to As and Cd, so its lower toxicity is counteracted by its greater abundance. Cd and Ni both mainly occur as divalent cations (whether in compounds or sorbed) and neither has a gas phase; arsenic, however, has multiple oxidation states and can also occur in the gaseous state. Given that the sources and atmospheric fates of Cd and Ni are dictated by the behavior of particulates (Section 8.6 above), their speciation and behavior is simpler than As. For example, arsenic in air can occur as oxides, hydroxides or weak acids comprised of As^{+3} or As^{+5} ; compounds containing As^{+3} such as As_2O_3 , H_3AsO_3 and $\text{H}_2\text{AsO}_3^{-1}$ are known as arsenites (the As^{+3}

ion is also known as “arsenite”), whereas compounds with As^{+5} , e.g. As_2O_5 , H_3AsO_4 and $\text{H}_2\text{AsO}_4^{-1}$, are known as arsenates (and the As^{+5} ion is also known as “arsenate”). Arsenite compounds are more toxic than arsenate, and compounds consisting of arsenic in lower oxidation states (As^0 , As^{-3}) are much less common in the atmosphere than the more-oxidized arsenite and arsenate forms.

In addition to particulate forms (mainly sorbed As) and dissolved forms (e.g. H_2AsO_4^-), arsenic also can occur as As^{-3} in gases, mainly arsine (AsH_3) and organo-As gases such as methylarsine (CH_3AsH_2), dimethyl arsine, $[(\text{CH}_3)_2\text{AsH}]$ and trimethyl arsine, $[(\text{CH}_3)_3\text{As}]$. These gases are produced in chemically reducing environments, especially water-logged soils. In rice paddy soils of the Bengal Fan of Bangladesh and southeast India, groundwater used for irrigation contains elevated arsenic from the natural aquifer materials, and formation of trimethyl arsine is fostered by the presence of soil organic matter (reactive methyl groups), warm temperatures and microbial activity in the reducing environment of the rice paddy. Volatilization of arsines from soils may account for ~1 to 5% of total arsenic in the atmosphere globally, although this amount will vary depending on local factors (e.g. mainly redox conditions and temperature). Residence times of arsines and organo-arsines in air is on the order of hours to days, so they have the potential to be transported downwind. Their decomposition is driven by reaction with hydroxyl radicals (HO^*), a process that ultimately results in formation of arsenite or arsenate that may sorb to particulates or dissolve in water droplets and be removed from the atmosphere (Mestrot et al., 2011a, b). Arsenic was also formerly an ingredient in insecticides such as lead arsenate (PbHAsO_4), which was widely applied to fruit trees and other crops to control moths and other pests. By the 1950s, most arsenical insecticide applications had given way to the new generation of organochlorine pesticides, particularly DDT

(dichlorodiphenyltrichloroethane); however, decades of application of arsenical insecticides led to build-up in soils (particularly the organic-rich topsoil [O and A horizons]) in many agricultural regions, and As can enter the atmosphere via wind erosion of particulates. Once applied as lead arsenate, As has a strong affinity for soil organic matter, so often is transported in air sorbed to organic matter aerosols.

QUESTIONS

1. What is a free radical? Provide two examples and explain their origin in the atmosphere. Explain why free radicals are so important in atmospheric chemistry, providing 2 chemical equations to support your response.
2. The reaction of N_2 and O_2 to form NO has a positive enthalpy change: $\Delta H = 180 \text{ kJ/mol}$. Based on ΔH , is this a spontaneous reaction? If not, why does NO exist in the atmosphere?
3. What is the difference between singlet oxygen and triplet oxygen, and why is this significant for atmospheric chemistry?
4. Discuss the origin of methylmercury in aquatic ecosystems. In particular, what is the initial source of Hg to the atmosphere, and what transformations take place in with organic matter in soils and sediments to enhance Hg mobility and toxicity? Show with a schematic sketch, annotated notes and chemical equations.
5. Compare reaction pathways and kinetics of formation of H_2SO_4 vs. HNO_3 in air. Which reaction occurs faster, and why? What implications does this have for length of travel in air before H_2SO_4 vs. HNO_3 are removed from air via precipitation or dry deposition?
6. What are the major sources of N and S pollution in air?
7. By what molecular-scale mechanism does the ozone layer protect the Earth's surface from UV radiation?

8. Show the processes and compounds responsible for the catalytic destruction of ozone. Discuss origins of ozone-destroying compounds, their stabilities in air, and show with chemical equations their catalytic effects.
9. Describe, with appropriate chemical reactions, the role of VOCs in photochemical smog. Consider also the role of air temperature in the process.

REFERENCES

- Angell, J.K., Gelman, M.E., Hofmann, D., Long, C.S., Miller, A.J., Nagatani, R.M., Oltmans, S., Planet, W.G., 1995. Southern Hemisphere Winter Survey. National Oceanic and Atmospheric Administration (United States). Accessed on-line: http://www.cpc.ncep.noaa.gov/products/stratosphere/winter_bulletins/sh_95/.
- Bidleman, T.F., Patton, G.W., Walla, M.D., Hargrave, B.T., Vass, W.P., Erickson, P., Fowler, B., Scott, V. and Gregor, D.J. 1989. Toxaphene and other organochlorines in Arctic ocean fauna: evidence for atmospheric delivery. *Arctic* 42: 307–313.
- Brimblecombe, P., 2003. The Global S Cycle. Chapter 14 in *Treatise on Geochemistry*, Volume 8, W.H. Schlesinger, ed, Elsevier B. V., Amsterdam, The Netherlands.
- Bridier, I., Caralp, F., Loirat, H., Lesclaux, R., Veyret, B., Becker, K. H., Reimer, A., and Zabel, F., 1991. Kinetic and Theoretical-Studies of the Reactions $\text{CH}_3\text{C}(\text{O})\text{O}_2 + \text{NO}_2 + \text{M} \leftarrow \rightarrow \text{CH}_2\text{C}(\text{O})\text{O}_2\text{NO}_2 + \text{M}$ between 248-K and 393-K and between 30-Torr and 760-Torr. *Journal of Physical Chemistry* 95: 3594–3600.
- Erismann, J.W., Sutton, M.A., Galloway, J., Klimont, Z., Winiwarter, W., 2008. How a century of ammonia synthesis changed the world. *Nature Geoscience* 1: 636–639.
- Farman, J.C., Gardiner, B.G. and Shanklin, J.D., 1985. Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction. *Nature* 315: 207–210.
- Finlayson-Pitts, B.J., Pitts J.N., Jr, 2000. *Chemistry of the Upper and Lower Atmosphere—Theory, Experiments, and Applications*. Academic Press, San Diego, 969 p.
- Galloway, J.N., 2003. The Global Nitrogen Cycle. Chapter 12 in *Treatise on Geochemistry*, Volume 8, W.H. Schlesinger, ed., Elsevier, Amsterdam, The Netherlands.
- Huffman, G.P., Mitra, S., Huggins, F.E., Shah, N., Vaidya, S., Lu, F., 1991. Quantitative analysis of all major forms of sulfur in coal by X-ray absorption fine structure spectroscopy. *Energy and Fuels* 5: 574–581.
- Kaufmann, R.K., Kauppi, H., Mann, M.L., and Stock, J.H., 2011. Reconciling anthropogenic climate change with observed temperature 1998–2008. *Proceedings of the National Academy of Sciences (USA)*, doi:10.1073/pnas.1102467108.
- Kleindienst, T.E., 1994. Recent developments in the chemistry and biology of peroxyacetyl nitrate. *Research on Chemical Intermediates* 20: 335–384.
- Lawrence, G.B., David, M.B., Lovett, G.M., Murdoch, P. S., Burns, D.A., Stoddard, J.L., Baldigo, B.P., Porter, J.H., Thompson, A.W., 1999. Soil calcium status and the response of stream chemistry to changing acidic deposition rates. *Ecological Applications* 9: 1059–1072.
- Lee, C., Martin, R.V., van Donkelaar, A., Lee, H., Dickerson, R.R., Hains, J.C., Krotkov, N., Richter, A., Vinnikov, K., and Schwab, J.J., 2011. SO_2 emissions and lifetimes: Estimates from inverse modeling using in situ and global, space-based (SCIAMACHY and OMI) observations. *Journal of Geophysical Research* 116, D06304, doi:10.1029/2010JD014758.
- Lelieveld, J., 1993. Modelling of heterogeneous chemistry in the global troposphere. In: J.T. Kiehl (ed.) *Clouds and Climate*, pp. 423–446, National Center for Atmospheric Research, Boulder, Colorado, USA.
- Lovett, G.M., and Kinsman, J.D., 1990. Atmospheric pollutant deposition to high-elevation ecosystems. *Atmospheric Environment* 24A (11): 2767–2786.
- Lu, Z., Streets, D.G., Zhang, Q., Wang, S., Carmichael, G.R., Cheng, Y.F., Wie, C., Chin, M., Diehl, T., and Tan, Q., 2010. Sulfur dioxide emissions in China and sulfur trends in East Asia since 2000. *Atmospheric Chemistry and Physics* 10: 6311–6331. doi:10.5194/acp-10-6311-2010.
- Mestrot, A., Feldmann, J., Krupp, E.M., Hossain, M.S., Roman-Ross, G., and Meharg, A.A., 2011a. Field Fluxes and Speciation of Arsines Emanating from

- Soils. *Environmental Science & Technology* 45 (5): 1798–1804.
- Mestrot, A., Merle, J.K., Broglia, A., Feldmann, J., and Krupp, E.M., 2011b. Atmospheric Stability of Arsine and Methylarsines. *Environmental Science & Technology* 45 (9): 4010–4015.
- Schmeltz, D., 2009. Spatial and Temporal Trends in Sulfur and Nitrogen Deposition for the United States. Regional Scientific Workshop on Acid Deposition in East Asia 2009 12–13 October 2009, Tsukuba, Japan. Accessed on-line: <http://www.eanet.cc/meeting/RWS/presentation/5.pdf>.
- Seinfeld, J.H. and Pandis, S.N., 2006. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change* (2nd edn). John Wiley & Sons, Inc., Hoboken, New Jersey, USA. 1203 pp.
- Smith, S.J., van Aardenne, J., Klimont, Z., Andres, R.J., Volke1, A., and Delgado Arias, S., 2011. Anthropogenic sulfur dioxide emissions: 1850–2005. *Atmospheric Chemistry and Physics* 11: 1101–1116.
- Spiro, T.G. and Stigliani, W.M., 2006. *Chemistry of the Environment*, 2e. Prentice Hall, Upper Saddle River, NJ, USA.
- Stoiber, R.E., Williams, S.N. and Heubert, B., 1987. Annual contribution of sulfur dioxide to the atmosphere by volcanoes. *Journal of Volcanology and Geothermal Research* 33: 1–8.
- Todd, R.W., Cole, N.A., Clark, R.N., Flesch, T.K., Harper, L.A., Baek, B., 2008. Ammonia emissions from a beef cattle feedyard on the Southern High Plains. *Atmospheric Environment* 42: 6797–6805.
- Unsworth, J.B., Wauchope, R.D., Klein, A.W., Dorn, E., Zehe, B., Yeh, S.M., Akerblom, M., Racke, K.D., and Rubin, B., 1999. Significance of the long range transport of pesticides in the atmosphere. *International Union of Pure and Applied Chemistry (IUPAC) pesticide report* 41, *Pure and Applied Chemistry* 71: 1359–1383.
- Waldman, J.M., Munger, J.W., Jacob, D.J., Flagan, R.C., Morgan, J.J. and Hoffman, M.R., 1982. The chemistry of acid fog. *Science* 218: 677–680.
- Wallace, J.M., Hobbs, P., 2006, *Atmospheric Science: An Introductory Survey*. Elsevier, Amsterdam, The Netherlands, 504 pp.
- Weijers, E. P., Schaap, M., Nguyen, L., Matthijsen, J., Denier van der Gon, H. A. C, ten Brink, H. M., and Hoogerbrugge, R., 2011. Anthropogenic and natural constituents in particulate matter in the Netherlands. *Atmospheric Chemistry and Physics* 11: 2281–2294.

Chemical Weathering and Soils

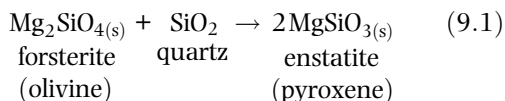
9.1 PRIMARY MINERALS, MINERAL INSTABILITY, CHEMICAL WEATHERING MECHANISMS AND REACTIONS, SOIL-FORMING FACTORS, AND PRODUCTS OF CHEMICAL WEATHERING

Soils play a critical role in the composition of Earth's surface and near-surface. They are the interface between the atmosphere and fresh water systems (streams, lakes, groundwater); they influence cycling of atmospheric CO₂ and other gases; sulfuric and nitric acid and other air pollutants are deposited onto soil so the ultimate fate of these and many other air pollutants is decided by soil biogeochemical reactions. Plant nutrients are derived from mineral weathering and organic matter reactions in soil, and most fresh surface water on Earth has passed through soil at some point (i.e. water in lakes, streams and aquifers typically passes through soil first), so soil reactions often impart a strong signature to surface water. The formation, evolution and composition of soil

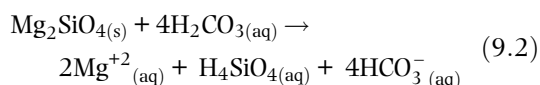
is a complex process that is controlled by geological parent materials (rock and sediment), plants, microbes and biochemical reactions, climate, topography, tectonically driven uplift and erosion, and time, and this chapter explores a series of topics in environmental and low-temperature soil geochemistry.

To some extent soils can be understood from a thermodynamic point of view (e.g. in general, high-temperature igneous minerals are unstable at Earth surface temperatures, and hydrous low-temperature minerals are stable in soils), but soil mineral assemblages are nearly always in a state of disequilibrium, commonly consisting of minerals that, given more time, would undergo reactions to decrease the energy of the system. An illustrative example is the presence of quartz and forsterite (olivine) in glacial till soils whose mineral assemblage is a function of glacial erosion of parent rock and subsequent deposition, and thus is mostly controlled by physical processes rather than by chemical equilibrium. Under most conditions in middle to deep crustal levels and in

magmas, this reaction will occur in the presence of metamorphic or igneous fluids (water or magma) in the direction of increasing stability:



However, the low temperatures (typically 0–25 °C) of soils do not facilitate this reaction, and instead what will happen eventually (perhaps over thousands or tens of thousands of years) is that unstable forsterite will decompose via chemical weathering in the presence of weak acids, typically carbonic acid delivered by precipitation, to yield dissolved magnesium, silica and bicarbonate:



Quartz is far more resistant to chemical weathering than is forsterite and tends to be relatively stable in soils, so we would predict that, over time, the quartz–forsterite soil will trend towards one dominated by quartz, and the soil solution (the water in the pores of the soil) will contain dissolved magnesium, silica and bicarbonate (assuming a simple forsterite weathering scenario with no inputs of Na^+ , Ca^{+2} , Cl^- , SO_4^{-2} or other constituents).

9.1.1 Goldich stability sequence

The early 20th century geochemist Samuel Goldich examined the changing abundances of primary minerals in crystalline igneous and metamorphic rocks undergoing chemical weathering and realized that he could model chemical weathering as the inverse of Bowen's reaction series (Goldich, 1938). Figure 9.1 depicts Bowen's reaction series, which shows the order of crystallization of silicate minerals from igneous melts as they cool; when examined through the lens of chemical weathering, it is referred to as

the **Goldich stability sequence** or **Goldich dissolution series**.

On the discontinuous, ferromagnesian side of Bowen's reaction series, the first mineral to crystalline is olivine (forsterite) at ~1200–1400 °C, followed by pyroxene, hornblende and biotite. Eventually, non-ferromagnesian minerals will crystallize, starting with K-feldspar followed by muscovite and then, finally, at ~600–800 °C, quartz. On the plagioclase feldspar side of the reaction series, anorthite (Ca-plagioclase) crystallizes first from magma, and then through a continuous reaction series, evolves to Na-Ca-plagioclase and then eventually albite (Na-plagioclase) with decreasing temperature.

How does Bowen's reaction series relate to mineral stability in soils and chemical weathering? The answer lies in the fact that this igneous crystallization sequence also represents, from top to bottom, increasing polymerization of SiO_4 tetrahedra, from isolated SiO_4 tetrahedra in olivine to progressively increased polymerization of SiO_4 tetrahedra in pyroxene through the micas and K-feldspar to completely polymerized SiO_4 tetrahedra in quartz, where all O atoms are covalently bonded to Si atoms of adjacent silica tetrahedra (details of SiO_4 polymerization are presented in Chapter 2). The increasing polymerization corresponds to an increasing abundance of covalent bonds that are far less susceptible to decomposing in the presence of water (and H^+ ions derived from weak acids involved in chemical weathering) than are the ionic bonds that dominate halide salts such as NaCl.

In reality, most minerals consist of a mix of covalent and ionic bonds – a good example is forsterite olivine used in the examples above, which contains dominantly ionic bonds between Mg^{+2} and SiO_4^{-4} , and covalent bonds between Si and O atoms within SiO_4^{-4} . Thus, dissolution of olivine will result in dissolved Mg^{+2} and silica (as H_4SiO_4) because the ionic bonds are susceptible to breaking in the presence of water, but the covalent bonds in silica are not (“silica” commonly refers to SiO_2 , H_4SiO_4 or SiO_4^{-4}). The opposite extreme

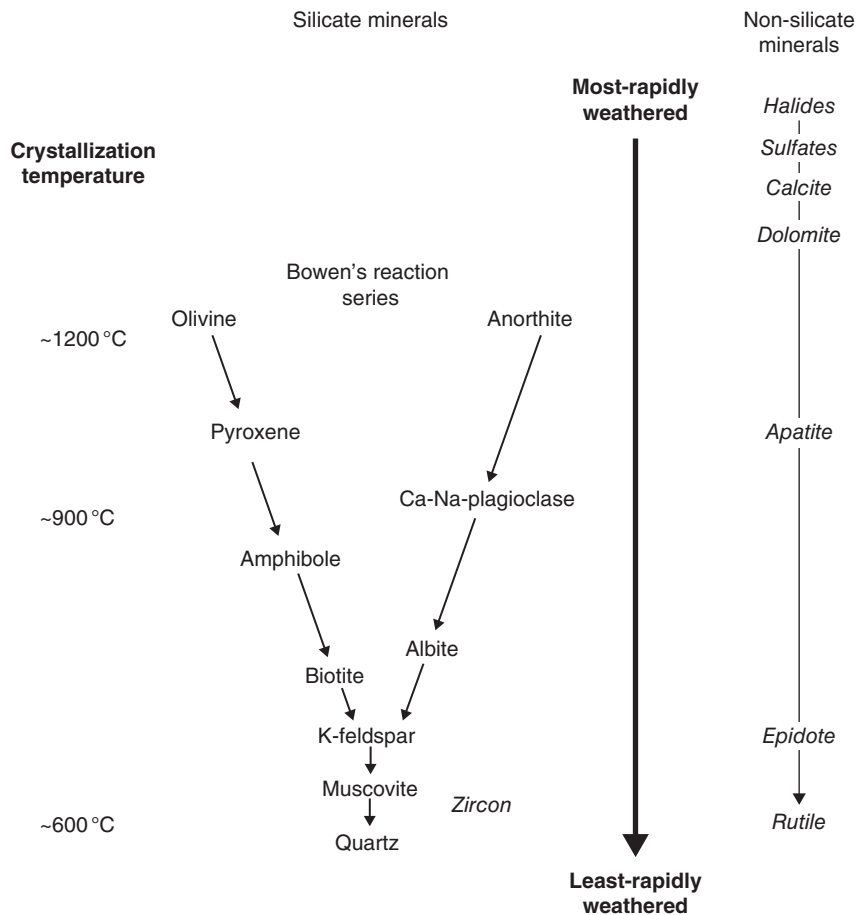


Fig. 9.1 Bowen's reaction series / Goldich dissolution series showing susceptibility to chemical weathering, from least stable at the top (halides, e.g. NaCl) to most stable at the bottom (quartz); after Goldich (1938), Lasaga et al. (1994) and Drever (1997). Minerals which are not part of the original Bowen and Goldich series are in *italics*. Igneous crystallization temperatures are approximate and do not apply to the non-silicate minerals on the right.

in Bowen's reaction series / Goldich's stability sequence is quartz, where the absence of ionic bonds results in a mineral that strongly resists chemical weathering.

9.1.2 Weathering rates

The differences in weathering rates between minerals depicted in Fig. 9.1 span many orders of magnitude (Table 9.1). One way to understand

relative weathering rates is to quantify the amount of time required to dissolve a 1-mm diameter sphere of a given mineral in, for example, a dilute solution at a pH of 5. Quantified in this way, a 1-mm sphere of calcite will dissolve in approximately 1 month, whereas an equal-size crystal of forsterite olivine will require 2000 years, and a 1-mm sphere of quartz will persist for approximately 34 million years before completely dissolving (Lasaga et al., 1994). Comparisons of field-determined weathering rates and

Table 9.1 Weathering rates of common rock-forming minerals assuming hypothetical 1-mm sphere of each mineral in a dilute solution of pH = 5. Data compiled from Chou et al. (1989), Lasaga et al. (1994), Drever (1997), Brantley (2005), Buss et al. (2008) and numerous sources cited therein. Variation in rates between time (yr) versus dissolution rate and log dissolution rate reflects different approaches to measuring rates. See Brantley (2005) for details. The composition of hornblende is $(\text{K},\text{Na})_{0-1}(\text{Ca},\text{Na},\text{Fe},\text{Mg})_2(\text{Mg},\text{Fe},\text{Al})_5(\text{Si},\text{Al})_8\text{O}_{22}(\text{OH})_2$.

Mineral	Composition	Time (yr)	Dissol. rate (mol/m ² /s)	Log dissol. rate (mol/m ² /s)
Quartz	SiO ₂	34 × 10 ⁶	4.1 × 10 ⁻¹⁴	-13.4
Kaolinite	Al ₂ Si ₂ O ₅ (OH) ₄	6 × 10 ⁶	1 × 10 ⁻¹³	-13
K-feldspar	KAlSi ₃ O ₈	740 × 10 ³	5.0 × 10 ⁻¹³	-12.4
Muscovite	KAl ₂ Si ₃ AlO ₁₀ (OH) ₂	720 × 10 ³	3.2 × 10 ⁻¹⁴	-12.5
Phlogopite	KMg ₃ Si ₃ AlO ₁₀ (OH) ₂	670 × 10 ³	3.2 × 10 ⁻¹⁴	-12.5
Albite	NaAlSi ₃ O ₈	500 × 10 ³	6.3 × 10 ⁻¹³	-12.2
Hornblende	see caption	500 × 10 ³	6.3 × 10 ⁻¹³	-12.2
Diopside	CaMgSi ₂ O ₆	140 × 10 ³	3.6 × 10 ⁻¹²	-11.4
Anorthite	CaAl ₂ Si ₂ O ₈	80 × 10 ³	4.0 × 10 ⁻¹²	-11.4
Enstatite	MgSiO ₃	16 × 10 ³	3.2 × 10 ⁻¹¹	-10.5
Tremolite	Ca ₂ Mg ₅ Si ₈ O ₂₂ (OH) ₂	10 × 10 ³	1.1 × 10 ⁻¹¹	-11
Forsterite	Mg ₂ SiO ₄	2 × 10 ³	3.6 × 10 ⁻¹⁰	-9.4
Dolomite	CaMg(CO ₃) ₂	1.6	3 × 10 ⁻⁷	-6.5
Calcite	CaCO ₃	0.1	1 × 10 ⁻⁶	-6

laboratory-determined rates often differ by as much as an order of magnitude, but given that the differences in rates among minerals differ by many orders of magnitude, the values given in Table 9.1 are useful indicators of relative weathering rates of common minerals. Ultimately, understanding weathering rates is important for assessing the potential of an ecosystem to neutralize effects of acid rain or acid mine drainage, for predicting rates of release of plant nutrients (e.g. Ca, K, P) from minerals, for understanding rates of CO₂ uptake over human to geological time scales, and for determining potential for sorption or release of trace-metal contaminants.

9.1.3 Chemical weathering

The changes to the structure and composition of minerals that take place in rocks or sediments exposed at the surface of the Earth or in soils, and that result in release of ions to solution and/or formation of new minerals, are known as

chemical weathering. **Primary minerals** are inherited from igneous, sedimentary or metamorphic rocks, or sediments derived from these rocks, and they are often unstable at the Earth's surface because they commonly form under conditions quite different than Earth surface conditions. They therefore are prone to decomposition by a variety of mechanisms, including dissolution, hydration, hydrolysis, and oxidation. The minerals that form as a result of chemical weathering are known as **secondary minerals** and commonly consist of clay minerals (e.g. smectite, vermiculite, kaolinite), hydroxides (e.g. gibbsite, goethite), oxides (e.g. hematite), carbonates (e.g. calcite) and sulfates (e.g. gypsum). Secondary mineral assemblages are strongly controlled by climate, parent material, soil age and, perhaps to a lesser extent, the two other soil-forming factors (biotic effects and topography).

One of the main driving forces behind chemical weathering is the action of carbonic acid dissolved in precipitation (rain and snow) on minerals. (Refer to Chapter 5 for details on

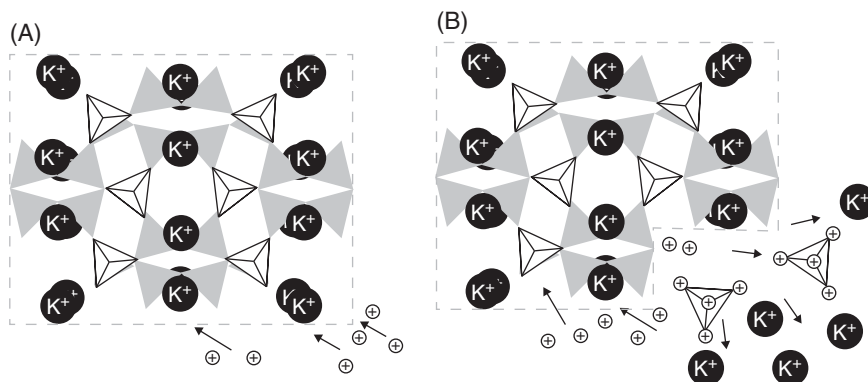
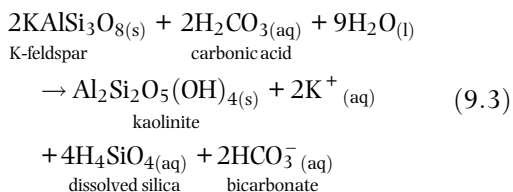


Fig. 9.2 Schematic sketch of K-feldspar hydrolysis. Diagram A is unweathered feldspar in a soil solution with hydrogen ions (small circles with “+” sign) attacking an edge of the crystal. Gray triangles represent silica tetrahedra whose apices face into the page and the dashed line represents the edge of the crystal. Diagram B depicts the same crystal after hydrolysis has attacked an edge, showing two silica tetrahedra which are no longer bonded to K^+ , but rather are bonded to H^+ , resulting in dissolved silica (H_4SiO_4) in solution. Potassium ions are also shown in solution, but in order to maintain clarity in the diagrams, not all K^+ can be shown. Note also that anions in solution (e.g. HCO_3^-) are not shown.

formation of carbonic acid). The reaction shown below, the dissolution of potassium feldspar (microcline, orthoclase or sanidine) in the presence of carbonic acid and water, is a **hydrolysis reaction** that produces a solid clay mineral (kaolinite) as well as dissolved potassium, silica and bicarbonate.

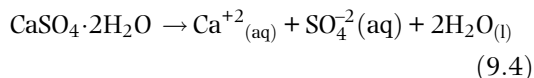


In a hydrolysis reaction, H^+ ions (in this case, derived from H_2CO_3) replace cations (in this case, K^+) in the mineral structure, resulting either in complete dissolution of the primary mineral, or in formation of a secondary mineral.

The kaolinite that forms by the hydrolysis of feldspars inherits SiO_4 tetrahedra from the primary feldspar and incorporates Al from the feldspar and H^+ and H_2O from soil water into the

clay mineral octahedral sheet. This reaction also indicates that some silica ends up dissolved in solution as H_4SiO_4 , and depending on rainfall and soil permeability, may wind up being leached out of the soil and into surface water or groundwater. The chemical weathering of olivine presented earlier in this section is also a hydrolysis reaction; in that case, four H^+ ions replace two Mg^{+2} ions, resulting in the transformation of $Mg_2SiO_4(s)$ to $H_4SiO_{4(aq)}$ (and Mg^{+2} in solution). Hydrolysis of potassium feldspar is presented schematically in Fig. 9.2.

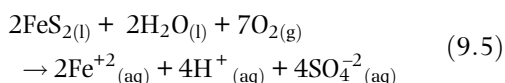
Minerals dominated by ionic bonds, particularly salts, are prone to simple **dissolution** in the presence of water. The strengths of bonds between hydrogen atoms in water and anions in the salt, or between oxygen atoms in water and cations in the salt, are greater than the bonds between cations and anions in the salt, so the mineral dissolves generically into X^+ and Y^- . Dissolution of the hydrated sulfate gypsum is a good example:



Note that the covalent bonds in sulfate prevent decomposition of S–O bonds in water, so the resulting products are ions of calcium and sulfate plus molecules of water released from the structure upon dissolution.

The presence of chemically reduced atoms in minerals that form in reducing environments in the crust makes certain minerals prone to oxidation when in contact with or close proximity to the atmosphere. Soil water typically contains enough dissolved oxygen to facilitate **oxidation reactions**, and of the ions in minerals that are prone to oxidation in soils, Fe^{+2} , S^{-2} are the most important. Others are far less abundant (e.g. As, Mn) or, as is the case with carbon, generally occur in their most-oxidized state in minerals; for example, in carbonate minerals, carbon occurs as C^{+4} and so will not decompose by oxidation. Organic matter, on the other hand, with carbon ranging from C^{-4} to C^0 , is prone to oxidation in soils.

The oxidation of pyrite is quite common in soils, and it occurs via the oxidation of sulfur from S^{-2} to S^{+6} (also presented in equations 1.34, 2.11/2.12, 4.29):



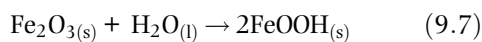
The ferrous iron released by sulfur oxidation is then oxidized to ferric iron and incorporated into iron hydroxide:



The goethite produced will impart an orange color typical of iron-bearing soils, and the electron shown as a product will be consumed by an electron receptor such as a soil microbe or perhaps soil nitrate undergoing reduction to N_2 or N_2O . In fact, microbial activity often catalyzes this reaction because it is an energy source. Ferrous iron originally present in silicate minerals also typically undergoes oxidation in a similar manner after being released by hydrolysis.

Both of these oxidation reactions (9.5 and 9.6) are anomalous as far as chemical weathering is concerned because they produce acid, rather than consuming acid. In rocks and soil rich in pyrite or other sulfide minerals, oxidative weathering can result in **acid rock drainage**, the natural analog to anthropogenic acid mine drainage. That said, it is important to realize that the net effect of chemical weathering is to **buffer** the acidity inherent to rainfall, mainly through hydrolysis reactions which consume H^{+} and sequester it in mineral structures (e.g. in the octahedral sheet of clays) or in dissolved products (e.g. H_4SiO_4).

Hydration reactions involve incorporation of water molecule(s) into the structure of an existing mineral, resulting in a new structure in which the water molecule is re-arranged into H^{+} and OH^{-} in the new hydrated mineral structure. The hydration of hematite is a common example of a hydration reaction:



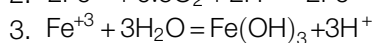
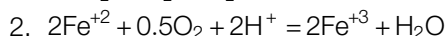
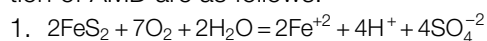
The product of this reaction, the mineral goethite, is sometimes written as $\text{Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$, which helps to illustrate the hydration of the primary mineral hematite (Fe_2O_3), but in reality this is an oversimplification that does not reflect dissociation of H_2O into H and O upon incorporation into the goethite structure. The hematite to goethite transformation results in decreasing size of the unit cell, incorporation of hydrogen atoms between oxygen atoms of iron octahedra, disappearance of Fe from one-half of octahedral sites (Gualtieri and Venturelli, 1999), producing vacant octahedra not present in hematite, and associated decrease in density of the mineral. Also, to clarify, hydration differs from hydrolysis in that hydration involves full incorporation of H_2O into the mineral structure, whereas hydrolysis involves replacement of cations in minerals with H^{+} .

Chemical weathering is augmented by **physical weathering**, which is the mechanical disintegration of rocks that increases exposed surface area, allowing greater contact of water

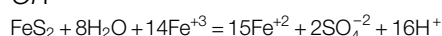
Acid Mine Drainage (AMD) and Sulfide Weathering

Weathering of sulfide minerals is distinct in that it *produces* acid rather than consuming acid like most minerals. In rocks and soils with abundant pyrite (likely 5% or more), naturally acidic water may form. This natural occurrence is sometimes referred to as acid rock drainage. When sulfide-rich rock is mined, the high surface area:volume ratio of the crushed ore rock increases the reaction rate, and when mined rock is exposed to atmosphere or oxidized water, it will generate more sulfuric acid than the soil or rock can

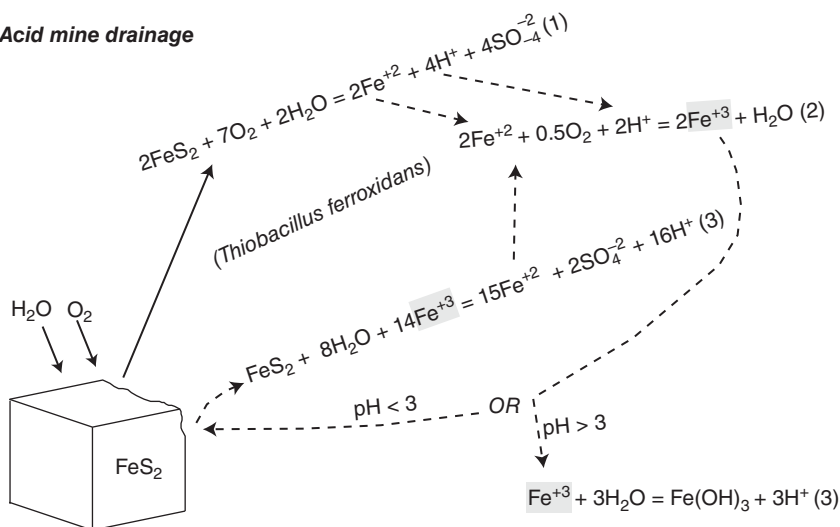
buffer (the main buffer would be silicate or carbonate weathering). The reaction is effectively catalyzed by bacteria such as *Thiobacillus ferrooxidans* who take advantage of sulfide anions as an electron source. The series of reactions associated with generation of AMD are as follows:



OR



Acid mine drainage



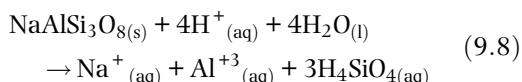
The schematic diagram depicts reactions involved in acid drainage, including the capacity of ferric iron to oxidize pyrite.

and air with rock and mineral surfaces. Common physical weathering processes include wind and water erosion, mass wasting (e.g. landslides), expansion of fractures by growing plant roots, freeze-thaw cycles of ice and liquid water that increase lengths and widths of fractures, and thermal expansion and contraction (especially in deserts) that leads to exfoliation.

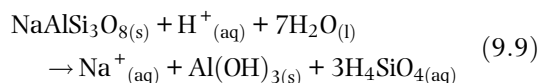
9.1.4 Consequences of chemical weathering: dissolved species and secondary minerals

Products of chemical weathering range from secondary mineral grains to dissolved ions. Reactions that result in complete dissolution of a primary mineral are known as **congruent**

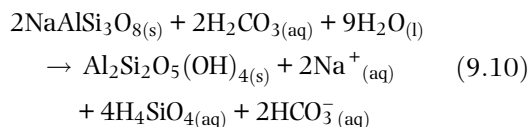
reactions, an example of which is hydrolysis of albite under acidic conditions that cause aluminum to remain in solution (rather than forming a secondary mineral):



Under slightly acidic conditions in well-leached soils (e.g. sandy forest soils or some tropical soils), pH may be too close to neutral for aluminum to remain in solution (see Section 4.5) and the solid aluminum hydroxide mineral known as gibbsite $[\text{Al}(\text{OH})_3]$ will form:



Weathering reactions that produce dissolved ions and solid minerals, like equation 9.9, are known as **incongruent reactions**. From these two reactions we can see that, for a given primary mineral, chemical weathering could result in either congruent weathering or incongruent weathering. In fact, the congruent reaction that results in production of gibbsite is not the only possible congruent reaction of albite; with different weathering stoichiometry, kaolinite can also form:



Whether or not weathering will proceed congruently or incongruently, and the ultimate fate (solubility, mobility) of atoms and molecules released from primary minerals during chemical weathering, depends on many factors, most of which fall within the scope of soil-forming factors, but before we begin to explore details of soil-forming factors, we will first examine a dataset

that will allow us to understand the inherent mobilities and solubilities of the ions and molecules released to the soil solution by chemical weathering.

9.1.5 Geochemical quantification of elemental mobility in soil

Table 9.2 presents geochemical data from a soil sequence in Costa Rica showing the compositional changes that take place during chemical weathering. The parent materials for the soils in this particular region are river sediments, and the two soils, one 10 000 years old, the other 125 000 years old, are evolving on river terraces originally comprised of the parent material indicated in the table. Thus, the main process responsible for the changes seen in this soil sequence is the effect of time, and the data allows us to examine which elements are relatively immobile over time, and which tend to be soluble and leached out of the soil over time.

Calculations in Table 9.2 are presented in two ways. The first quantifies the change in concentration of each element or mineral, as follows:

$$\% \text{ change} = 100 * (\text{wt\% soil} - \text{wt\% PM}) / \text{wt\% PM} \quad (9.11)$$

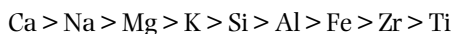
Using this measure we see that the concentrations of some elements have decreased markedly – Mg, Ca and Na concentrations in 125 ka soil are 86 to 96% less than concentrations in the parent material. Others have increased markedly – Ti, Fe and Zr are present in concentrations from 69 to 99% greater than what they are in the parent material. K decreases by 64%, Si exhibits a small decrease (12%) and Al increases by ~29%. Using % change as a measure of compositional change, soluble, mobile element concentrations may decrease by as much as 100% (i.e. be completely removed by leaching). Concentrations of

Table 9.2 Geochemical changes in a soil sequence where 10 ka and 125 ka terrace soils are derived from parent material (PM) comprised of river sediments dominated by basaltic and andesitic volcanic detritus intermixed with soil minerals derived from erosion further upstream. Data are obtained from B horizons of soils along the central Pacific coast of Costa Rica (Fisher and Ryan, 2006; Ryan and Huertas, 2009). All values for PM, 10 ka and 125 ka (except for Zr [ppm]) are presented as wt.% element or mineral. The two columns on the right are explained in the text. The mass-transfer coefficient does not apply to minerals.

	PM	10 ka	125 ka	% change (total)	Mass-Transfer Coeff. (τ_j)
Si	26.5	25.3	23.2	-12.4	-0.48
Ti	0.88	1.37	1.77	+99.7	+0.18
Al	5.24	5.99	6.75	+28.9	-0.24
Fe	4.26	6.42	7.20	+68.9	0
Mg	2.35	0.97	0.32	-86.3	-0.92
Ca	2.37	0.18	0.09	-96.2	-0.98
Na	0.52	0.12	0.05	-90.7	-0.95
K	0.47	0.18	0.17	-64.3	-0.79
Zr	129	164	219	+69.8	0
Quartz	13.1	11.5	18.8	+43.5	NA
Plagioclase	12.8	2.8	<0.5	-100	NA
Augite	4.1	<0.5	<0.5	-100	NA
Smectite	42.2	40.6	<2	-100	NA
Kaolin	15.9	27.6	57.8	+263	NA
Goethite	6.2	6.9	3.2	-48.4	NA
Hematite	<0.5	<0.5	7.8	+inf	NA

immobile elements will increase as soluble elements are removed, potentially to values exceeding 100% (this can occur when base cation concentrations in parent material are high).

Using percent change in concentration, element mobility can be ranked from most mobile to least mobile:



Note that high-charge cations Al^{+3} , Fe^{+3} , Zr^{+4} and Ti^{+4} are the least mobile, consistent with data presented in Fig. 4.9. Recall that Si^{+4} forms moderately soluble H_4SiO_4 .

Another approach to determining element mobility is to calculate the mass-transfer coefficient, τ_j , for each element (Brimhall and Dietrich, 1987). This is done by determining two ratios, one of the concentration of a given element (component j) in soil to its concentration

in parent material, and the other of the concentration of a presumed immobile element (commonly Zr or Ti; component i) in parent material versus its concentration in soil. The equation is as follows:

$$\tau_j = [(C_{j,w} * C_{i,p}) \div (C_{j,p} * C_{i,w})] - 1 \quad (9.12)$$

where w indicates weathered material and p indicates parent material and C indicates “concentration of”; for example, $C_{j,w}$ represents the concentration of element j in weathered material and $C_{i,p}$ represents the concentration of the immobile element in parent material.

Using this approach, a value of -1 indicates complete depletion, a value of zero indicates no change (this is the case for Fe in Table 9.2 and also, by definition, Zr), and any value >0 implies addition from an outside source (e.g. aeolian deposition). We can see that both rankings

give the same relative mobilities, with the low-charge base cations (Ca^{+2} , Na^+ , Mg^{+2} , K^+) showing greatest mobilities (i.e. greatest potential to be leached out of soil), and the high-charge Al^{+3} , Fe^{+3} , Zr^{+4} and Ti^{+4} showing the least mobility. Thus, we can predict that base cations released by chemical weathering are the most likely to remain in solution following weathering of a mineral and, given sufficient water leaching downward through soil, that the base cations will be depleted with time, leaving a soil enriched in the relatively insoluble Al^{+3} , Fe^{+3} , Zr^{+4} and Ti^{+4} .

9.1.6 Quantifying chemical weathering: CIA

The extent of chemical weathering can be assessed by quantifying the changes to soil chemical composition or mineralogy that have taken place since onset of weathering. The **chemical index of alteration** (CIA; Nesbitt and Young, 1982) is a widely applied weathering index that includes relatively immobile aluminum in the numerator and aluminum as well as mobile calcium, sodium and potassium in the denominator. Values used in this equation are bulk soil concentrations (rather than, for example, the composition of the exchangeable fraction) as oxide molar concentrations, and the CaO used in the calculation must only apply to CaO derived from silicate minerals (i.e. calcium in carbonates is to be excluded).

$$\text{CIA} = 100 * \text{Al}_2\text{O}_3 \div (\text{Al}_2\text{O}_3 + \text{CaO} + \text{Na}_2\text{O} + \text{K}_2\text{O}) \quad (9.13)$$

Given that Al accumulates in soils as they are leached and that Ca, Na and K are depleted by leaching, CIA will increase with increased chemical weathering and leaching. If we apply this formula to the 10 ka and 125 ka soils shown in Table 9.2

(and convert wt% element to molar oxide values), we determine that:

$$\begin{aligned} \text{CAI (PM)} &= 100 * 0.19 \div (0.19 \\ &\quad + 0.083 + 0.023 + 0.012) = 62.3 \\ \text{CAI (10 ka)} &= 100 * 0.22 \div (0.22 + 0.0063 \\ &\quad + 0.0052 + 0.0046) = 93.2 \\ \text{CAI (125 ka)} &= 100 * 0.25 \div (0.25 + 0.0031 \\ &\quad + 0.0022 + 0.0043) = 96.3 \end{aligned}$$

As we would predict, the CAI value increases with increasing leaching associated with increased soil age. The upper limit of CAI, i.e. the value corresponding to complete leaching of soluble base cations, is 100. The lower limit is dictated by concentrations of Al, Ca, Na and K in parent material. For example, fresh, unweathered basalt has a CAI of ~30–40, whereas granite has CAI values of ~45–50. River sediments, like the parent material shown in Table 9.2, typically are in the CAI range of 60–70.

Many other indices using molar concentrations of soil chemical components have been developed to assess the extent of chemical weathering, including the sum of base cations divided by the sum of immobile cations ($\Sigma \text{ base cations} \div \Sigma [\text{Al}_2\text{O}_3 + \text{Fe}_2\text{O}_3 + \text{TiO}_2]$), the ratio of Fe_2O_3 to FeO (which quantifies oxidation of ferrous to ferric iron), the sum of base cations divided by Al_2O_3 , and various others summarized in Birkeland (1999). All take advantage of the fact that certain elements (mainly high-charge cations) are inherently immobile in soil and tend to accumulate with time, whereas others (mainly low-charge base cations) are inherently mobile or unstable (e.g. Fe^{+2}) and tend to decrease in concentration with time.

Soils exposed to pronounced leaching (mainly in the humid tropics) over hundreds of thousands to millions of years eventually evolve to a mineralogy dominated by gibbsite ($\text{Al}[\text{OH}]_3$) with additional kaolinite and iron oxides and/or hydroxides (e.g. hematite and/or goethite) – if rich enough in Al, these soils are known as **bauxite**

deposits and they form an important ore of aluminum. In addition to Ti and Zr, other trace elements such as the rare-earth elements Hf, Nb and Ta, and the lanthanides Th (and sometimes U) are also very immobile in weathering environments, so they too tend to be concentrated in bauxites and thus are sometimes mined from these deposits.

9.1.7 Soil profile

An Alfisol **soil profile** formed via weathering of basalt on the Nicoya Peninsula of Costa Rica is shown in **Plate 13**. The combined processes of addition of leaves, decaying tree trunks, decaying roots and other organic matter to the upper part of a soil, and chemical weathering below the top of the soil, produce soil profiles with relatively organic-rich upper horizons and mineral-rich lower horizons. The upper horizons, termed O horizons and A horizons, contain decomposing organic matter ranging from identifiable plant remains to black, fine-grained humus. These horizons are commonly referred to as topsoil, and the presence of organic matter imparts a dark brown to black color to these horizons, which typically contain 0.5 to 5% organic matter.

Underlying the upper horizons are mineral horizons with relatively little organic matter and which are dominated by colors of either the parent minerals or secondary minerals (e.g. yellows, oranges and reds of iron oxides and hydroxides, or whites and grays of carbonates). Because it is closer to the surface, the B horizon is more evolved than the underlying C horizon. B horizons are often dominated by secondary clays and hydroxides (or carbonates in arid climates), whereas C horizons more closely represent parent material. In cases where soils have formed via chemical weathering of bedrock, C horizons are underlain by saprolite, which, while altered by weathering, will still contain evidence of sedimentary layering, metamorphic foliation, igneous textures or other

properties inherent to the parent rock. E horizons (also known as albic horizons) occur in highly leached forest soils where the combination of high precipitation and the added effect of organic acids results in leaching of even iron, producing layers that are white to gray and often dominated by quartz.

In addition to the inherent mobilities of elements and weatherability/stability of minerals in soil, the main soil-forming factors, as first described by the Swiss soil scientist Hans Jenny in the 1930s and 1940s, ultimately exert a very strong control on soil geochemistry. These soil-forming factors are climate (cl), organisms/biotic effects (o), relief/topography (r), parent material (p), and time (t); given that anthropogenic (a) activities also influence soils, they are often included as well, and the factors are often referred to by the acronyms **clorpt** or **clorpt(a)**. The following section will briefly present these factors as they relate to soil geochemistry.

9.1.8 Soil-forming factors

Climate

Climate is very influential in determining soil type because it strongly controls the intensity of mineral decomposition; it also influences whether or not products of primary mineral dissolution remain in soil or are leached out of soil towards surface water or groundwater. The two main factors associated with climate are precipitation and temperature, which are commonly quantified as mean (or median) annual precipitation (MAP) and mean (or median) annual temperature (MAT). Increased MAP provides the medium in which chemical weathering reactions occur (i.e. water) as well as protons (H^+) that drive hydrolysis reactions. A particularly illustrative example from Hawaii, where other soil-forming factors (e.g. basaltic parent material, temperature, soil age) are relatively constant, demonstrates the

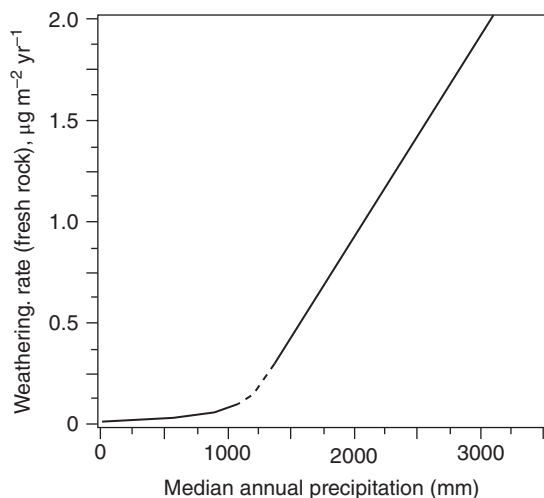


Fig. 9.3 Weathering rate as a function of precipitation in Hawaii where parent material is freshly exposed basalt. Diagram redrawn from data presented by Stewart et al. (2001). The sharp transition in rate between ~1100 and 1400 mm MAP may represent a threshold across which increased leaching results in decreased concentrations of base cations to the point that reactions are driven strongly to the right (i.e. weathering products).

effect of precipitation on chemical weathering rates (Fig. 9.3).

But while MAP certainly strongly affects rate of chemical weathering, it is often useful to determine effective precipitation (precipitation minus evapotranspiration) because soils exposed to the greatest effective precipitation will experience the greatest amount of **leaching** (downward movement of water through soil towards the water table), and hence, the greatest loss of soluble ions (e.g. Ca^{+2} , Na^{+2} , Mg^{+2} , SO_4^{-2} ; HCO_3^-). Conversely, in areas with low effective precipitation, and especially where evaporation is greater than precipitation (i.e. deserts), carbonates (e.g. calcite) or sulfates (e.g. gypsum) may crystallize in soil because soluble ions are less likely to be leached out, and thus reach saturation.

Temperature also strongly controls weathering. An old adage in biogeochemistry says that

for every increase of 10°C , rates of biogeochemical reactions double, and at ambient temperatures of soils, rates may even more than double for each 10°C increase (Davidson and Janssen, 2006). This concept is embedded in the Arrhenius equation (equation 1.76), which indicates that increasing temperature will increase k (the rate constant), increasing the rate of chemical reaction:

$$k = A * e^{-E_a/RT}$$

The combined effect of high MAP and MAT make the humid tropics the region with the highest weathering rates, and conversely, the cold deserts of Antarctica are characterized by the lowest weathering rates on Earth. Hot desert soils may experience rapid chemical weathering following rainfall events, but evaporation tends to inhibit leaching of soluble base cations out of soil, leading to formation of soil layers rich in calcite or gypsum and fostering formation of sodium smectites.

Organisms

From microbial activity that enhances reaction rate (especially of redox reactions) to uptake of nutrients by roots that decreases concentrations of base cations and other nutrients in soil solution, the activity of organisms plays an important role in chemical weathering and soil formation. The rhizosphere, the zone that extends ~2 mm out from roots, is compositionally different than the rest of the soil; roots commonly release organic acids to enhance solubility of nutrients, and concentration gradients caused by nutrient uptake cause nutrients to migrate towards the root, enhancing reactions even outside of the rhizosphere. Decomposition of leaves, woody material and other forms of organic matter produces humic and fulvic acids that contribute to hydrolysis; burrowing worms enhance aeration and

permeability; and decay of organic matter produces solid humus that retains nutrients, contributes to retention of soil moisture, and binds soil particles together, improving soil structure and permeability.

Lichens are an illustrative example of how organisms affect chemical weathering and soil formation. They consist of a fungus and a photosynthesizer (e.g. green algae) in a symbiotic relationship that enhances weathering of bare rock surfaces. Their nature and effect can be summarized as:

- Lichens: exude organic acids (e.g. oxalic acid) that accelerate rate of mineral dissolution (both by H^+ attack and by chelation) – this provides mineral nutrients (e.g. Ca, K, P) to the photosynthesizer.
- Photosynthesizer: produces H_2CO_3 (via respiration, enhancing mineral dissolution) and sugars (energy source for itself and fungus)

Both members of a lichen will produce nutrient-rich organic matter as they decay, thus beginning the process of soil formation on exposed rock surfaces.

Relief (topography)

The position in the landscape of a soil can either enhance or impede chemical weathering and soil formation. Given that chemical weathering is enhanced by leaching, and that water tables tend to be deeper (farther below the surface) in summits or higher elevation areas, hilltops and shoulders of hills tend to be well-drained and see greater chemical weathering rates than do soils in footslope or toeslope areas (the bases of hills), where shallow water tables inhibit downward leaching (Fig. 9.4). Often, base cations and soluble anions leached out of higher elevations are carried in solution along groundwater flow paths toward lower elevations, then deposited in toeslope regions, a process that can inhibit chemical weathering in toeslopes by increasing concentrations of reaction products (e.g. Na^+ , Ca^{+2} , H_4SiO_4 , HCO_3^-) in the soil solution. The low (deep) water table in summit positions also tends to facilitate oxidation reactions, whereas the high (shallow) toeslope water table creates more-reducing conditions that not

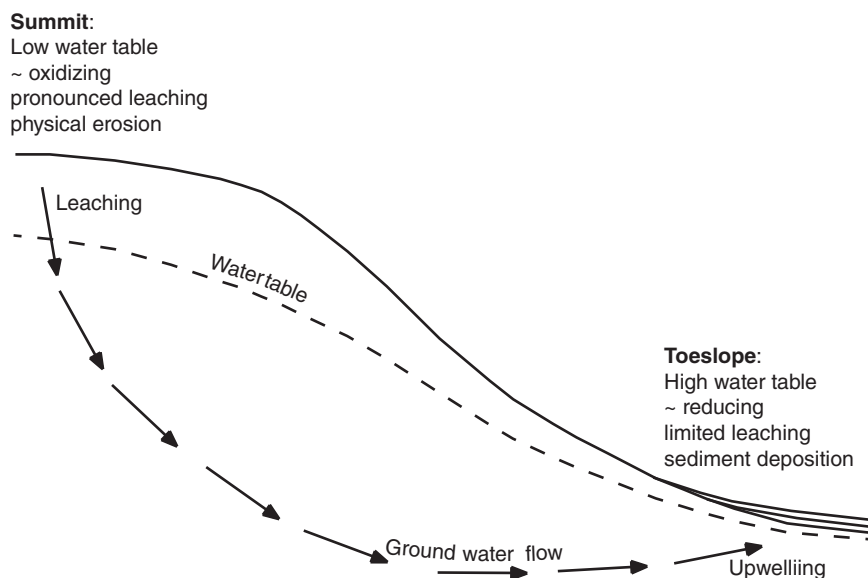


Fig. 9.4 Sketch of a topographic transect indicating factors that tend to enhance chemical weathering rates in summit positions, and those that tend to inhibit chemical weathering in toeslope positions.

only inhibit oxidation reactions but can foster reduction reactions.

Soils in toeslopes with shallow water tables often exhibit mottling or gley texture consisting of iron hydroxide splotches or spheres (typically 1 mm to 1 cm in diameter) in a matrix of gray, chemically reduced soil – in these soils redox reactions often exert strong control on element mobility. Orientation with respect to the sun is important too – in arid and semi-arid regions, north-facing slopes retain greater moisture than do south-facing slopes; therefore, leaching tends to be greater in soils on north-facing slopes, whereas on south-facing slopes, soils are more evaporation-dominated and soil solutions will be richer in base cations and other soluble soil constituents.

Another factor related to topography is tectonic uplift, which can enhance chemical weathering rates through exposure of fresh, unweathered rock. For example, chemical weathering rates measured in terms of dissolved solutes transported from basins by streams are much greater in the Andes Mountains than they are in the Amazon Basin (Stallard and Edmond, 1983), but slow physical erosion rates in the ~topographically flat Amazon Basin coupled with hot, moist conditions allow very thick, highly leached, Al–Fe-rich soils to form. Continual uplift and erosion in the Andes exposes fresh rock that weathers rapidly, but thick soils generally do not form due to steep slopes and high physical erosion rates.

Parent material

Based on information provided in Fig. 9.1 and Table 9.1, it is clear that the mineralogical content of rock or sediment will control the rate of chemical weathering, soil formation and the composition of the soil solution. Basaltic parent material dominated by olivine, pyroxene and Ca-plagioclase will weather far more rapidly than granitic parent material dominated by K-feldspar, muscovite and quartz. Moreover, the cationic composition of soil water with basaltic parent

material will contain a strong Ca–Mg signature whereas granitic parent materials will produce waters enriched in K and Na. Limestones, dolostones and marbles are prone to rapid chemical weathering in humid climates that far outpaces weathering rates of basalts, although in arid climates carbonates rocks are relatively resistant to weathering. Physical attributes are important too – greater amounts of exposed surface area in fractured rocks compared to non-fractured rocks will facilitate influx of water and enhance weathering and soil formation; similarly, permeable sediments or rock will likely weather more rapidly than impermeable sediment or rock of the same mineralogical composition.

Time

Even though most rock-forming minerals (calcite, olivine, pyroxene etc) are unstable in soil (to varying degrees), their decomposition in weathering regimes has a kinetic component, much like the ice cube that does not immediately melt on a hot summer day in spite of the obvious instability of the situation. Figure 9.1 and Table 9.1 both contain information pertaining to the time factor; so too does Table 9.2, which presents data on the evolution of a tropical soil sequence with time. Plagioclase, which comprises ~13% of parent material (river sediment), decreases to ~3% in 10 000 year-old soils, and is undetectable in 125 000 year-old soils. We also observe increasing abundance of the secondary mineral kaolin, from 16% in parent material, to 28% in 10 000 year-old soil, and ultimately (in this sequence), to 58% in 125 000 year-old soil. The trend toward increasing kaolin is expected because kaolinite and halloysite (pedogenic members of the kaolin group) are abundant in tropical soils, and form via the incongruent dissolution of feldspars and other silicates (more on soil minerals is presented in Section 9.2).

Trends of chemical weathering rate, soil formation rate, leaching of base cations, build-up of

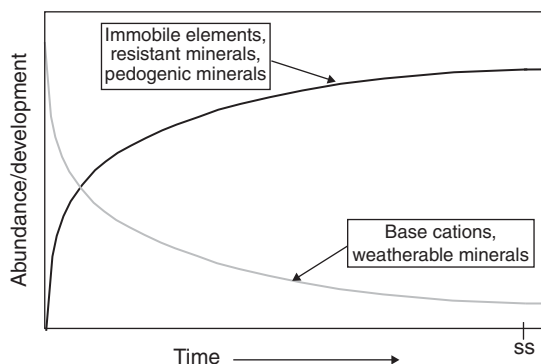


Fig. 9.5 Curves depicting typical evolution of soil characteristics with time. The black curve represents constituents that increase in concentration or degree of development with time (e.g. $[\text{Al}^{+3}]$, carbonate horizon thickness), whereas the gray curve depicts the trajectory of constituents that tend to decrease over time (e.g. $[\text{Na}^+]$, [parent materials]). See text for full explanation. The location on the x-axis labeled as “ss” indicates the point in time at which steady-state conditions have developed.

immobile elements, and other pedogenic processes tend to follow parabolic curves; in other words, rates are rapid at the onset and then taper with time. Figure 9.5 depicts typical parabolic (logarithmic) curves where the black curve would represent increasing concentrations of pedogenic clays (e.g. kaolinite), pedogenic hydroxides and oxides (e.g. goethite, hematite, gibbsite), cations that are relatively immobile and tend to accumulate in soil with time (e.g. Fe^{+3} , Ti^{+4} , Zr^{+4}), and minerals that are resistant to chemical weathering (e.g. quartz, zircon, rutile). The black curve could also represent thickness of weathering rinds, depth of the soil profile or, in arid climates, thickness of calcite or gypsum horizons. The gray curve is representative of soil components that decrease over time; examples include leaching of base cations and weathering of primary minerals, both of which initially occur rapidly, but then decrease in rate with time. Examples are in Fisher and Ryan (2006).

This is typical behavior for systems that are out of equilibrium; at first, rates of change are rapid in the direction of equilibrium, but eventually

decrease as the system approaches equilibrium. In some cases, soils appear to reach a state of equilibrium, typically after tens of thousands to millions of years, where the time frame depends on the other soil-forming factors (e.g. tropical soils will evolve towards the low-slope part of the parabolic curve more rapidly than will semi-arid soils).

Given that chemical weathering reactions are not reversible in soil, a true state of dynamic equilibrium cannot exist in soil (e.g. the reaction of Na-feldspar to kaolinite is not reversible at ambient soil temperatures), but many soils do appear to achieve a steady state, or quasi-steady state with respect to soil development, where curves like those shown in Fig. 9.5 eventually reach a slope of zero, or nearly so. One scenario in which a steady state might develop involves initial rapid chemical weathering of plagioclase and other Ca-bearing minerals accompanied by rapid leaching of calcium; then, eventually, the low, long-term rate of leaching may be balanced by aeolian additions of Ca, and the flux relationship is $\text{Ca}_{\text{in}} = \text{Ca}_{\text{out}}$. Birkeland (1999) presents a more-detailed account of arguments for and against the existence of steady-state conditions in soil, and while there is disagreement of whether steady states exist in soils, from a practical point of view soils often do reach a point where so little change occurs in soils greater than a certain age. For this reason it may be practical to model landscape-scale soil properties as if the older soils have reached steady state. In the humid tropics, for example, soils older than ~100 000 years old are nearly uniformly characterized by a kaolinite-rich, nutrient-poor status regardless of whether the soil is 100 ka, 500 ka or 1000 ka (1 Ma).

Anthropogenic factors

While the overall chemical and mineralogical character of soils in most regions is dominated by the natural processes described above, human activities have certainly altered soil compositions in many areas. For example, acid rain accelerates

leaching of base cations and decomposition of natural organic matter, leading to changes in the composition of the soil solution and of soil particles; emissions of trace metals from smelters and coal-fired power plants directly alter the composition of soil via deposition; additions of nutrients (e.g. Ca, K, N, P) and pesticides to enhance crop growth alters soil chemistry, as does improper disposal or storage of mine wastes, fuels and solvents. These topics are presented in greater detail in sections towards the end of this chapter.

9.1.9 Soil classification – soil orders and geochemical controls

The US Department of Agriculture has developed a systematic nomenclature for soils known as Soil Taxonomy that has gained broad acceptance globally. It is presented here in part because of

its widespread use, but also because of its systematic approach in which many of the criteria used to classify soils are based on soil mineralogy and/or geochemistry. This section also includes terms that are not part of the USA Soil Taxonomy system (e.g. laterite, chernozem, etc.) but may be encountered in studies of soil geochemistry.

All soil orders have the root “sol” preceded by a root that carries some relationship to the soil itself. The reader should also be aware that soil orders are further classified into suborders, of which there are 55, great groups, of which there are at least 200, and other criteria that ultimately result in complex, yet informative terms such as *halloysitic*, *isohyperthermic Tropeptic Haplorthox*. Although useful information to know, the scope of this text does not permit this level of detail, and the reader is referred to Birkeland (1999), USDA on-line information on soil taxonomy, and numerous other sources.

Brief summary of USDA Soil Orders. The properties of the first seven are strongly controlled by mean annual precipitation and thus are arranged more or less in order of decreasing leaching (i.e. from humid tropical to temperate to semi-arid to arid); the properties of the last five (Andisols, Histosols, Gelisols, Inceptisols and Entisols) are less-easily organized by precipitation and so are presented separately.

Oxisols: highly leached, intensely weathered tropical soils that form in tropical and subtropical environments. Little or no remaining primary minerals; rich in oxides and hydroxides of Al and Fe as well as kaolinite. Nutrient-poor (low Ca, K and Mg), acidic (pH ~ 4–5), limited ability to retain fertilizer-added nutrients, low base saturation.

Ultisols: strongly leached, clay-rich soils in forested ecosystems of temperate subtropics, e.g. subtropical forests of India and the southeastern United States. Low base saturation (<35%), usually low inherent fertility, rich in

kaolinite, less leached than oxisols. The terms **laterite** and **latosol** apply to Ultisols in some cases.

Spodosols (also known as **podsoils** or **podzols**): cool, humid temperate coniferous forest soils. Leached E and B horizons, relatively nutrient poor. Organic acids produced by coniferous needle decomposition result acidic (pH < 4) E/A/B horizons. Very distinctive layering of dark brown/black A horizon over white-gray E horizon over bright red/orange B horizon over tan C horizon.

Alfisols: moderately leached soils in temperate, deciduous forest ecosystems. % BS > 35%. Clay-rich B horizons (vermiculite, illite or smectite with higher CEC than the kaolinites of oxisols or ultisols). Relatively high nutrient level, good agricultural soils. Abundant in the agriculturally productive US Midwest, in temperate deciduous forests of Europe, central Asia, Africa and elsewhere.

Vertisols: rich in smectite, form in areas with pronounced dry and rainy seasons (e.g.

monsoonal climates of western India, eastern Sudan and southeastern Texas, USA). The shrink – swell nature of smectites lends distinctive characteristics to Vertisols; during the dry season, smectites dehydrate (water is lost from the interlayer) and deep, wide desiccation cracks form. During rainy periods, smectites rehydrate and the cracks disappear. Shrink–swell creates engineering problems (e.g. cracked foundations, heaved roads), yet nutrient rich.

Mollisols (also known as **Chernozems**): natural grassland ecosystem soils between ~35 to 55 degrees latitude (e.g. American Great Plains, the Pampa of Argentina, the Russian Steppe, parts of Mongolia, and other grassland ecosystems). The combination of relatively high primary productivity (of grasses) and low decomposition and leaching rates produces thick, soft, humus-rich A horizons (mollic epipedons) that are outstanding for agriculture.

Aridisols (also known as **Sierozems**): Desert soils with calcite- or gypsum-bearing B horizons. Limited leaching and pronounced evaporation of soil water result in crystallization of otherwise soluble minerals at depths typically ~10 to 100 cm. In well-developed aridisols, 15–50-cm thick layers of well-cemented calcite or gypsum may form. High nutrient concentrations, so with irrigation have potential to be productive agricultural soils. Major concern is sodification (Section 9.6).

Andisols: young soils in volcanic ash deposits. Low density ($< 0.90 \text{ g/cm}^3$), low organic matter content, abundance of disordered volcanic glass and metastable weathering products such as allophone and imogolite. High water-retention capacity, high cation- and anion-exchange capacities, nutrient-rich, tend to retain fertilizer nutrients well. Occur mainly in moist forested ecosystems on the flanks or bases of volcanoes (e.g. Japan, Philippines).

Histosols: organic rich peat-bog/wetland soils. Typically contain $\geq 20 \text{ wt\%}$ organic carbon, total soil thickness is at least 40 cm. High organic matter content means very low bulk density, often $< 0.3 \text{ g cm}^{-3}$. Typically form where anaerobic conditions and restricted drainage inhibits the decomposition of plant matter; thus, Histosols are a reservoir of carbon that can be lost to the atmosphere when wetlands are drained. Histosols are also referred to as peats and mucks.

Gelisols: soils where permafrost occurs within 2 m of the surface. Geographically limited to polar regions and some high-elevation regions. Low temperature inhibits decomposition of organic matter, so like Histosols, Gelisols store large quantities of organic carbon. Warming of polar and alpine regions will likely result in release of some of this stored organic carbon to the atmosphere. **Tundra soils** is a term sometimes used to refer to Gelisols.

Entisols: soils that exhibit almost no pedogenic development. Typically are recently deposited sand (e.g. river floodplains, beaches) or virtually unweathered sediments exposed by mass wasting. No distinctive horizon (with the possible exception of a thin A horizon), properties are determined by the properties of the unconsolidated parent material. With time, Entisols evolve into one of the other soil orders; e.g., recently deposited glacial till in cool temperate forest will likely evolve into a Spodosol given thousands or tens of thousands of years of pedogenesis.

Inceptisols: like Entisols but more evolved; intermediate to Entisols and other soil order (mainly Oxisol, Ultisol, Spodosol, Alfisol, Mollisol, Aridisol). Given their transitional state, Inceptisols are widely distributed across climate zones, parent materials, and the other soil-forming factors. In the humid tropics, Inceptisols (e.g. tropepts) will eventually evolve into Oxisols; in grassland ecosystems, Inceptisols will evolve with time into Mollisols; and so on.

9.2 SECONDARY MINERALS, CONTROLS ON THEIR FORMATION, AND MINERAL STABILITY DIAGRAMS

The abundance of primary minerals in soil is controlled by a combination of geological processes (e.g. intrusion and exposure of granite, or deposition of glacial till) and the resistance of chemical bonds in primary minerals to chemical weathering. The abundance of secondary minerals, on the other hand, is controlled by the soil-forming factors, particularly climate, parent material and time, although the others are locally important too. The intense leaching environment of the humid tropics fosters soils rich in Al_2O_3 , Fe_2O_3 , and SiO_2 and H_2O that combine to form minerals such as kaolinite ($\text{Al}_2\text{Si}_2\text{O}_5[\text{OH}]_4$), hematite (Fe_2O_3), and gibbsite ($\text{Al}[\text{OH}]_3$) regardless of parent material; however, young tropical soils, typically those < 10 ka, still retain base cation and Si concentrations that are high enough to result in formation of metastable pedogenic smectite (e.g., $[\text{Ca}_{0.1}] [\text{Al}_{1.7} \text{Fe}_{0.3}] [\text{Si}_{3.8} \text{Al}_{0.2}] \text{O}_{10} [\text{OH}]_2$) rather than kaolinite, so while climate is ultimately the controlling factor on tropical soil mineralogy and geochemistry, kinetic factors clearly play an important role as well.

9.2.1 Factors controlling soil mineralogy

Ultimately, pedogenic mineralogy is controlled by the composition of the soil solution from which the minerals crystallize, and there are a few basic principles of soil formation that help to understand and predict soil mineralogy:

- 1** Effective precipitation is correlated to leaching of base cations, so highly leached soils are likely to evolve to a mineralogy dominated by kaolinite, hematite and/or goethite and gibbsite.
- 2** Arid and semi-arid region soils characterized by low effective precipitation and limited leaching of base cations will foster formation of pedogenic

minerals that contain base cations, including smectite, illite, calcite, and gypsum. Abundance of parent minerals is likely to be greater in arid region soils as compared to humid tropical soils of the same age.

- 3** Soil physical properties are important; for example, permeable sandy soils will foster more extensive leaching of base cations or other substances than will impermeable soils.

- 4** Poorly drained areas (e.g. wetlands, toeslopes) tend to accumulate base cations and have higher Si:Al ratios than well-leached soils. This favors formation of smectite over kaolinite.

- 5** Compared to primary minerals, secondary (pedogenic) minerals tend to be fine-grained and typically are concentrated in the clay size fraction (< 2 μm); thus, they are more reactive and have higher cation exchange capacities than do primary minerals.

Figures 9.6 and 9.7 depict the variation in soil pedogenic mineralogy as a function of climate. Figure 9.6 depicts some of the main aspects of soil mineralogy as controlled by precipitation, notably the occurrence (and stability) of the Ca-minerals gypsum and calcite in arid region soils (MAP < 500 mm), the dominance of smectite in arid and semi-arid regions with MAP ~200–1000 mm, the prevalence of kaolinite (and/or halloysite) in humid temperate to tropical climates with MAP ~1000–3000 mm, and ultimately the increasing abundance of hydroxides and oxides of Al and, to a lesser extent, Fe in extremely leached, highly evolved tropical soils (bauxites).

Figure 9.7 presents information on the latitudinal variation in soil mineralogy and thickness of the soil profile, from the tropics poleward through the desert belt, temperate and boreal regions and finally up into tundra and polar regions.

Tropical regions foster deep soils rich in kaolinite and gibbsite, although in deeper, less-weathered parts of the soil profile, smectites and mixed-layer kaolinite–smectites occur as well. These less-evolved minerals also occur at the surface in the humid tropics where soil age is young

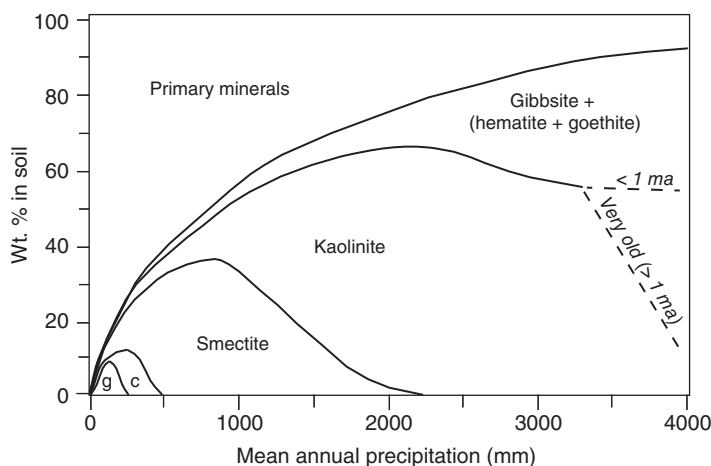


Fig. 9.6 Pedogenic mineralogy plotted as a function of mean annual precipitation. Given the complex interactions of the five soil-forming factors, this sketch is to be viewed as a simplified model intended to show dominant trends. In addition to pedogenic mineral trends, also note progressive decrease in % primary minerals (from ~100% to ~10%) with increasing precipitation. Modified from Sherman (1952), data in Birkeland (1999), and many other sources. Symbols: “g” = gypsum, “c” = calcite. Dashed lines indicate variation in soil mineralogy of differing ages in the humid tropics.

(e.g. soils on a 2000-year-old lava flow or uplifted river or beach deposit). Soil mineralogy in desert regions trends towards smectite and calcite and the thin soils reflect the lower intensity of chemical weathering; one exception to this general rule is higher-elevation soils that receive orographic precipitation that enhances leaching.

Soil mineralogy in temperate regions is even more difficult to generalize due to stronger effects of parent material and topography (e.g. elevation) compared to the tropics and arid regions. Rainfall and leaching are intermediate to the humid tropics and arid regions, so temperate region soil mineralogy (and soil depths) fall somewhere between end-member deep kaolinite–gibbsite soils (tropics) and shallow smectite–calcite soils (arid). Temperate soil minerals include vermiculite, interstratified biotite–vermiculite (also known as hydrobiotite), illite, interstratified mica–smectite (which can form via hydrolysis and leaching of K^+ in mica interlayers), as well as intermixed kaolinite and smectite. Polar regions generally experience the least-intense

chemical weathering of global climate zones, so are likely to remain rich in primary minerals.

9.2.2 Mineral stability diagrams

While soils typically do not achieve thermodynamic equilibrium, **mineral stability diagrams** based in thermodynamic principles are useful in examining likely trajectories of secondary mineral growth and transitions in soil. The main rationale is that, for a soil with a given chemistry of soil water (soil solution), certain mineral(s) will be stable, and the system will trend in that direction. To view one extreme, strongly leached soils rich in dissolved aluminum and lacking Si and base cations will foster crystallization of gibbsite ($Al(OH)_3$), but not calcite or smectite. Mineral stability diagrams can be constructed from congruent reactions between primary minerals and secondary minerals, and also congruent reactions between two different secondary minerals.

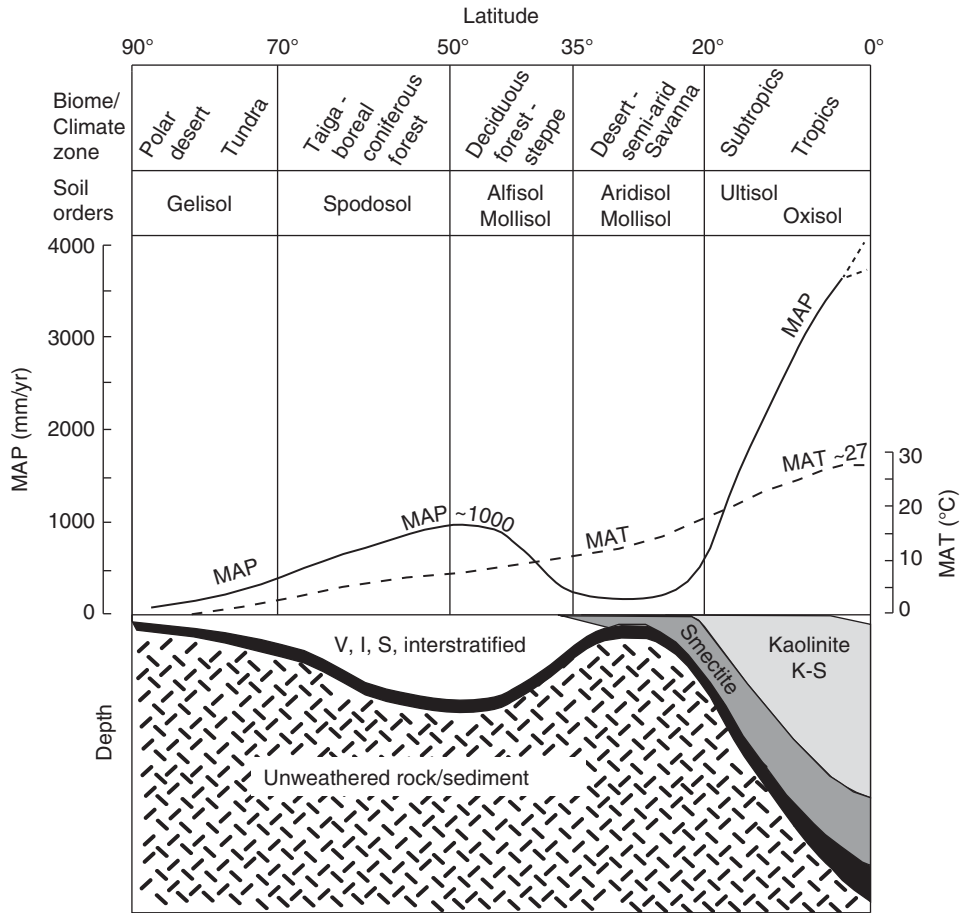
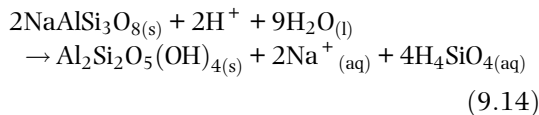


Fig. 9.7 Schematic sketch of soil minerals and relative depths of typical soils across a latitudinal transect, from pole to equator. Soil depth in the tropics may be tens of meters, whereas in other regions is typically < 2 m. Common soil orders for each latitudinal zone are provided. MAP = mean annual precipitation, MAT = mean annual temperature, V = vermiculite, I = illite, S = smectite, K-S = interstratified kaolin-smectite. (Modified from Strakhov 1967 and Birkeland 1999.)

If we return to the weathering of albite (Na-plagioclase feldspar), we can begin to construct a mineral stability diagram:



The equilibrium constant for this reaction is:

$$K_{\text{eq}} = [\text{Na}^+]^2 * [\text{H}_4\text{SiO}_4]^4 / [\text{H}^+]^2 \quad (9.15)$$

And for the purposes of formulating the equation such that variables are arranged in a useful way, arranged into:

$$K_{\text{eq}} = [\text{H}_4\text{SiO}_4]^4 * ([\text{Na}^+]^2 / [\text{H}^+]^2) \quad (9.16)$$

Using rules of logarithms, this equation can be rearranged into:

$$\text{Log}(K_{\text{eq}}) = 4 * \text{log}[\text{H}_4\text{SiO}_4] + 2 * \text{log}([\text{Na}^+] / [\text{H}^+]) \quad (9.17)$$

then

$$\frac{1}{2} \log (K_{\text{eq}}) = 2 * \log [\text{H}_4\text{SiO}_4] + \log ([\text{Na}^+]/[\text{H}^+]) \quad (9.18)$$

We can now rearrange these terms into an equation of a line:

$$\log ([\text{Na}^+]/[\text{H}^+]) = -2 * \log [\text{H}_4\text{SiO}_4] + \frac{1}{2} \log (K_{\text{eq}}) \quad (9.19)$$

In this case, $\log ([\text{Na}^+]/[\text{H}^+]) = y$; $-2 = m$ (slope); $\log [\text{H}_4\text{SiO}_4] = x$; and $\frac{1}{2} \log (K_{\text{eq}}) = b$. We know that, plotted on log x - and y -axes, the slope will be -2 and the y intercept will be one-half of the log of the equilibrium constant for the reaction.

All that remains now to plot this equation is the equilibrium constant, which can be calculated from standard free energies of formation of the species and phases involved in the reaction (values used in these calculations are given in Table 9.3). The principles of this topic are covered in Chapter 1, and tell us that, in a general case, the Gibbs free energy of formation ($\Delta G^\circ_{\text{R}}$) for

the reaction is computed according to this equation (1.52):

$$\Delta G^\circ_{\text{R}} = \sum n_x * G^\circ_{\text{fx}}(\text{products}) - \sum n_x * G^\circ_{\text{fx}}(\text{reactants})$$

Pertaining to the congruent weathering of albite to kaolinite, the equation is:

$$\begin{aligned} \Delta G^\circ_{\text{R}} = & [(G^\circ_{\text{f}}(\text{kaol})) + 2(G^\circ_{\text{f}}(\text{Na}^+)) \\ & + 4(G^\circ_{\text{f}}(\text{H}_4\text{SiO}_4))] - [2(G^\circ_{\text{f}}(\text{albite})) \\ & + 2(G^\circ_{\text{f}}(\text{H}^+)) + 9(G^\circ_{\text{f}}(\text{H}_2\text{O}_{(\text{l})}))] \end{aligned} \quad (9.20)$$

and

$$\begin{aligned} \Delta G^\circ_{\text{R}} = & [(-3778) + 2(-261.7) + 4(-1316)] \\ & - [2(-3712) + 2(0.0) + 9(-237.2)] \\ = & -6.6 \text{ kJ} \end{aligned}$$

From Chapter 1 (equation 1.61) we learned that the Gibbs free energy of formation and the equilibrium constant for a reaction are related (for units of kJ/mol) as follows:

$$\Delta G^\circ_{\text{R}} = -5.708 * \log(K_{\text{eq}})$$

Applying this to the weathering of albite to kaolinite:

$$-6.6 \text{ kJ} = -5.708 * \log(K_{\text{eq}})$$

Then

$$\log(K_{\text{eq}}) = -6.6 / -5.708$$

and

$$K_{\text{eq}} = 10^{(-6.6 / -5.708)} = 10^{1.2}$$

We can now go back to the equation of the line and add in the log of this K_{eq} value ($\log [10^{1.2}] = 1.2$) for “b”, the y -intercept:

Table 9.3 Thermodynamic data used in development of the mineral stability diagram in Fig. 9.9. Data are from Dean (1979) and Robie et al. (1979). Given uncertainty in Gibbs free energy values, precision is shown to only 1 decimal place. See Appendix III for a more-extensive table of thermodynamic data.

Compound or species	Name	G°_{f} (kJ/mol)
$\text{NaAlSi}_3\text{O}_8$	Albite (low albite)	-3712
$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$	Kaolinite	-3778
$\text{Na}_{0.33}\text{Al}_2(\text{Si}_{3.67}\text{Al}_{0.33})\text{O}_{10}(\text{OH})_2$	Na-beidellite	-5354
$\text{Al}(\text{OH})_3$	Gibbsite	-1144
H^+	Hydrogen ion	0
$\text{H}_2\text{O}_{(\text{l})}$	Water	-237.2
$\text{Na}^+_{(\text{aq})}$	Sodium ion	-261.7
H_4SiO_4	Dissolved silica	-1316

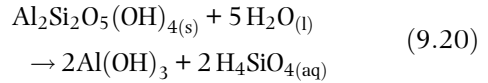
$$\log ([\text{Na}^+]/[\text{H}^+]) = -2 * \log [\text{H}_4\text{SiO}_4] + \frac{1}{2} (1.2) \quad (9.21)$$

$$\log ([\text{Na}^+]/[\text{H}^+]) = -2 * \log [\text{H}_4\text{SiO}_4] + 0.6 \quad (9.22)$$

This gives us the equation for a straight line with a slope of -2 and y -intercept of $+0.6$ (line labeled “1” on Fig. 9.8. The variable on the x -axis is $\log$

$[\text{H}_4\text{SiO}_4]$ and the variable on the y -axis is $\log ([\text{Na}^+]/[\text{H}^+])$.

Using this same approach, we can plot the line between gibbsite and kaolinite using this reaction as the basis:



The equilibrium constant is:

$$K_{eq} = [\text{H}_4\text{SiO}_4]^2 \quad (9.21)$$

Using rules of logarithms, this equation can be rearranged into:

$$\text{Log} (K_{eq}) = 2 * \log [\text{H}_4\text{SiO}_4] \quad (9.22)$$

then

$$\frac{1}{2} \log (K_{eq}) = \log [\text{H}_4\text{SiO}_4] \quad (9.23)$$

We can now rearrange these terms into an equation of a line, as above (equations 9.21 and 9.22) but understanding that $\log ([\text{Na}^+]/[\text{H}^+]) = 0$ for this reaction:

$$0 = -\log [\text{H}_4\text{SiO}_4] + \frac{1}{2} \log (K_{eq}) \quad (9.24)$$

The equilibrium constant for this reaction, derived using the same approach used for the example of albite $\rightarrow$ kaolinite, is $10^{-7.7}$, resulting in this equation:

$$0 = -\log [\text{H}_4\text{SiO}_4] - 3.9 \quad (9.25)$$

Or

$$\log [\text{H}_4\text{SiO}_4] = -3.9 \quad (9.26)$$

Thus, the line separating kaolinite and gibbsite is vertical (labeled “2” on Fig. 9.8 and crosses the x -axis at -3.9 (because, by definition, $\log [\text{H}_4\text{SiO}_4]$ is the variable on the x -axis). The lack of involvement of Na^+ and H^+ in this reaction intuitively indicates that changes to $[\text{H}_4\text{SiO}_4]$ in

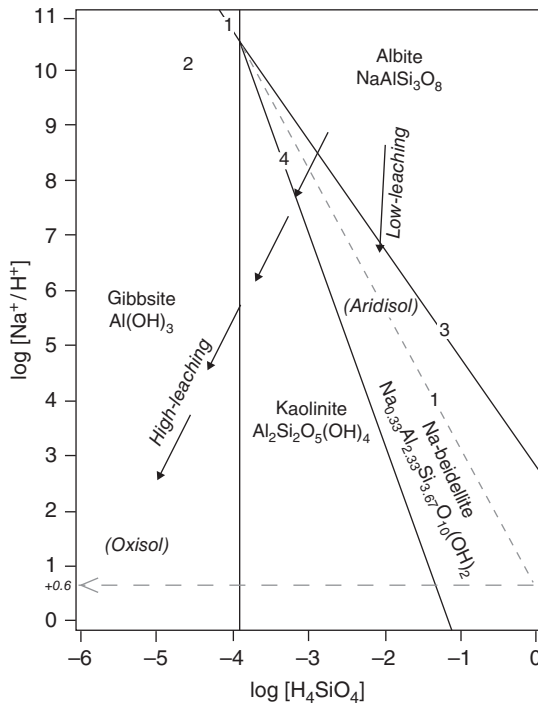
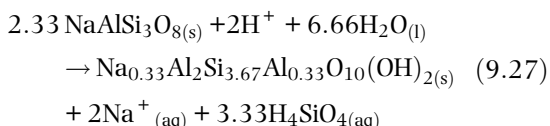


Fig. 9.8 Mineral stability diagram showing stability fields for gibbsite, kaolinite, Na-beidellite (a smectite) and albite, at 25 °C and 1 atm. Note that both axes are logarithmic. The x -axis represents dissolved Si (as H_4SiO_4) in soil water; the y -axis represents increasing $[\text{Na}^+]$ towards the top of the diagram, and increasing $[\text{H}^+]$ (decreasing pH) towards the bottom. Arrows depict representative trends of soil water under a leaching dominated scenario (high-leach) such as might be found in well-leached soils such as Oxisols, Ultisols, and Spodosols, and a limited leaching scenario (low-leach) like would be expected to occur in an Aridisol, Mollisol or Vertisol. Approximate compositions of soil waters in Oxisols and Aridisols are shown.

the soil solution are all that is required to drive this reaction.

Another common mineral involved in weathering of albite is sodium smectite, which will be represented here as sodium beidellite, a common clay mineral in soils that are less leached than those that produce abundant kaolinite:



The equilibrium constant for this reaction is:

$$K_{\text{eq}} = [\text{Na}^+]^2 * [\text{H}_4\text{SiO}_4]^{3.33} / [\text{H}^+]^2 \quad (9.28)$$

Then re-arranging and applying rules of logarithms, this equation can be rearranged into:

$$\log(K_{\text{eq}}) = 3.33 * \log[\text{H}_4\text{SiO}_4] + 2 * \log([\text{Na}^+]/[\text{H}^+]) \quad (9.29)$$

then

$$\frac{1}{2} \log(K_{\text{eq}}) = 1.67 * \log[\text{H}_4\text{SiO}_4] + \log([\text{Na}^+]/[\text{H}^+]) \quad (9.30)$$

We can now rearrange these terms into an equation of a line:

$$\log([\text{Na}^+]/[\text{H}^+]) = -1.67 * \log[\text{H}_4\text{SiO}_4] \\ + \frac{1}{2} \log(K_{\text{eq}}) \quad (9.31)$$

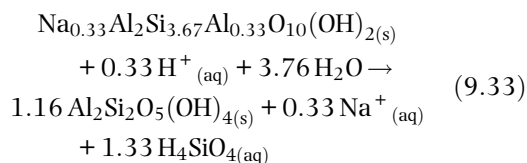
Pertaining to the congruent weathering of albite to Na-beidellite, the equation is:

$$\Delta G^\circ_{\text{R}} = [(-5354) + 2(-261.7) + 3.33(-1316) \\ - 2.33(-3778) + 2(0.00) + 6.66(-237.2)] \\ = -30.9 \text{ kJ} \quad (9.32)$$

The slope of the albite–Na-beidellite line is -1.67 (equation 9.31) and dividing -30.9 by -5.708 and putting the result into log form gives a K_{eq}

of $10^{5.4}$, and because the term on the right side of the equation is $\frac{1}{2} \log(K_{\text{eq}})$, the y-intercept is 2.7.

In order to complete this four-component mineral stability diagram, it is necessary to calculate the equation of the line for the reaction of Na-beidellite to kaolinite, a transition that occurs in soils over time as Na^+ and H_4SiO_4 are leached away.



For this reaction, the equilibrium constant is:

$$K_{\text{eq}} = [\text{Na}^+]^{0.33} * [\text{H}_4\text{SiO}_4]^{1.33} / [\text{H}^+]^{0.33} \quad (9.34)$$

Skipping a couple of steps, we arrive at:

$$\log(K_{\text{eq}}) = 1.33 * \log[\text{H}_4\text{SiO}_4] \\ + 0.33 * \log([\text{Na}^+]/[\text{H}^+]) \quad (9.35)$$

Then (see above for examples including all steps):

$$\log([\text{Na}^+]/[\text{H}^+]) = -4 * \log[\text{H}_4\text{SiO}_4] + 3 * \log(K_{\text{eq}}) \quad (9.36)$$

The slope of this line (line “4” on Fig. 9.8) is -4 and the y-intercept (from the Gibbs free energy calculation, as in the examples above), is -4.7 .

From this diagram we see that higher concentrations of H_4SiO_4 favor albite, as does increasing the ratio of $[\text{Na}^+]/[\text{H}^+]$. These observations are reasonable in a qualitative sense – albite has a higher molar concentration of Si than does kaolinite (e.g. Si to Al in albite is 3:1, whereas in kaolinite it is 1:1), and 1 mol of albite contains 1 mol of Na; kaolinite contains no Na. Other considerations related to this diagram include:

1 In cases where leaching is intense enough to simultaneously decrease dissolved Na and Si in

the soil solution, albite may directly transform to kaolinite. Numerous studies have documented direct transformation of feldspars to kaolinite. In cases where some Na is leached but the less-mobile Si is not, Na-beidellite will form. Numerous studies have also documented transformation of feldspars to smectites.

2 Starting at $\log [\text{H}_4\text{SiO}_4] = -2$ and $\log [\text{Na}^+/\text{H}^+] = 2$, shifting the system composition vertically upward (i.e. increasing dissolved sodium or raising pH) will (or could) cause kaolinite to transform to Na-beidellite. From this same point $(-2, 2)$, shifting the system composition horizontally to the right (i.e. increasing dissolved Si) will (or could) also cause reaction of kaolinite to Na-beidellite. So too will combinations of increasing Na and Si (however, in reality kaolinite may persist in the Na-beidellite field for thousands or millions of years due to kinetic factors).

3 In the framework of chemical weathering, leaching of Na^+ and dissolved silica drives the reaction of albite towards kaolinite (and ultimately to gibbsite), but in cases where either net leaching is limited to non-existent (e.g. arid regions), or where kinetic factors have not yet permitted sufficient leaching for kaolinite to form (i.e. young soils), Na-beidellite and similar smectite clays occur in soils (even in some cases in young humid tropical soils). At the other end of the weathering spectrum, intense leaching and sufficient time leach Na^+ (and other base cations) and silica to such an extreme that gibbsite becomes a dominant soil mineral.

As we recognize the usefulness of mineral stability diagrams, we must also consider uncertainty in measurements of thermodynamic data used in mineral stability diagrams, and the applicability of these diagrams and thermodynamic principles to the analysis of chemical weathering. One concern is that under the low-temperature conditions in soils, there can be no true thermodynamic equilibrium between primary minerals like albite and secondary minerals like kaolinite (or Na-beidellite, etc.). While the weathering reactions

that result in formation of kaolinite and Na-beidellite clearly do occur, these reactions are not reversible under soil pressure-temperature conditions, so even as concentrations of Na and Si increase into the albite stability field, kaolinite (or Na-beidellite) will not react to form albite at 25 °C.

Sources of uncertainty also arise from the fact that the Na, Ca and K smectites shown in Table 9.3 are idealized end-member compositions that lack Fe and Mg, yet soil smectites typically contain Fe and Mg, so the commonly used Gibbs free energy values do not account for compositional complexity. This will affect the slopes and *y*-intercepts of lines in mineral stability diagrams (note the difference in Figure 9.9 diagrams which involve similar minerals constructed from different thermodynamic data). Similarly, variation in grain sizes and abundance of cleavage planes and other crystallographic discontinuities can result in different mineral solubilities, which in turn will affect Gibbs free energies – clearly coarse-grained, hexagonal kaolinite will have a different Gibbs free energy than will a poorly crystalline, fine-grained, Fe-bearing kaolinite.

In spite of limitations, the types of diagrams shown in Figures 9.8 and 9.9 are commonly used in environmental geochemistry. As long as we recognize that slopes and locations of phase boundaries (lines) are limited by some degree of uncertainty, the diagrams do show approximate stability fields of primary minerals and secondary minerals, allowing us to understand relative stabilities of minerals in the presence of solutions and likely directions of chemical reactions in the presence of fluids of known composition.

9.3 SOILS AND THE GEOCHEMISTRY OF PALEOCLIMATE ANALYSIS

Given the effect of climate on soil formation, numerous studies have extrapolated knowledge

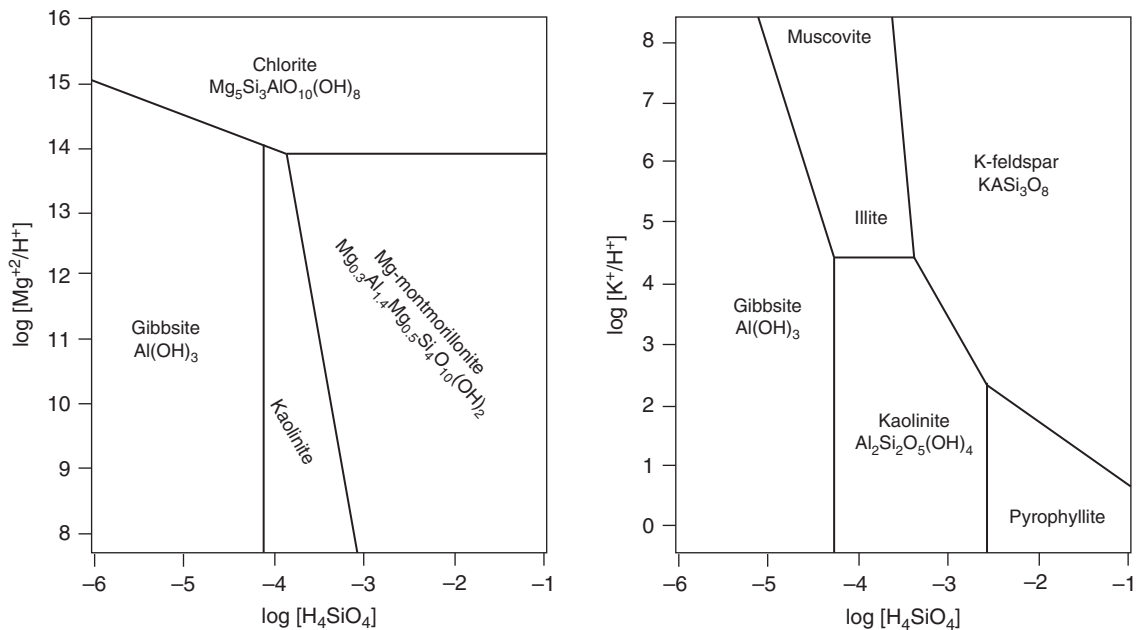


Fig. 9.9 Mineral stability diagram for systems Al-Mg-Si-O-H (left, modified from Lee et al., 2003) and K-Al-Si-O-H, 25 °C, 1 atm.

of climate and pedogenesis to the interpretation of paleoclimate signatures preserved in paleosols, ancient soils typically preserved in sediments or sedimentary rock (Retallack, 2001; Sheldon and Tabor, 2009). This is particularly useful in terrestrial sedimentary sequences where paleosols are often preserved in fluvial floodplain deposits, alluvial fans and other aggrading sequences. As you might predict, paleosols from humid tropical paleoenvironments should be highly leached, Al-Fe-rich and dominated by kaolinite, whereas paleosols formed in arid environments should be rich in smectite, gypsum or calcite. Of course time is also an important factor in soil formation, and given that some young tropical soils are rich in pedogenic smectite, caution must be exercised and other approaches incorporated (e.g. paleobotany of seeds or root casts). While it is beyond the scope of this chapter to delve into detailed examples, it is worth drawing attention to this field and the thorough body

of literature based in paleosol analysis and paleo-environmental interpretation.

9.4 EFFECTS OF ACID DEPOSITION ON SOILS AND AQUATIC ECOSYSTEMS

Dissolution of CO_2 into water produces carbonic acid (H_2CO_3), and this is the origin of naturally acidic rainfall (or snow) with a pH = 5.6 (Chapter 5). In areas where natural sources of S or N (Chapter 6) cause the formation of H_2SO_4 or HNO_3 in air, the pH of precipitation may drop below pH of 5.6, becoming extreme in cases where volcanic gases result in pH < 3, but these occurrences tend to be insignificant globally. Anthropogenic acid rain with pH < 5 is the focus of the following section, particularly the response of soils to anthropogenic acid deposition

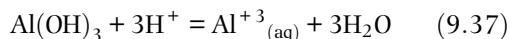
from sulfuric and nitric acids derived from combustion of fossil fuels or, less commonly, from industrial emissions. Chapter 8 (Section 8.11) presents the origin and atmospheric chemistry as well as definitions of acid deposition relative to acid precipitation and acid rain. The following section delves into one of the major impacts of fossil fuel emissions – acid deposition – on soils and the aquatic ecosystems to which they are linked.

The main constituents added to soil in acid precipitation (in excess of their abundance in natural precipitation) are H^+ , SO_4^{2-} and NO_3^- , so understanding the way in which these ions interact with soil minerals and organic matter helps to explain most of the negative impacts of acid deposition. It is also important to understand that some soils orders (e.g. Spodosols, Oxisols, Ultisols) are naturally acidic, in large part due to decomposition of organic matter that produces humic and fulvic acids. In these soils, however, microbially facilitated decomposition of organic acids (to CO_2 , H_2O and other compounds) tends to minimize their long-term negative impacts compared to the more-stable, stronger minerals acids H_2SO_4 and HNO_3 . Due to their inherent natural acidity, Spodosols and related coniferous forest soils are more prone to negative effects of acid deposition than would be naturally alkaline soils (e.g. Aridisols, Mollisols).

9.4.1 Increased solubility of Al in acidic soil solution

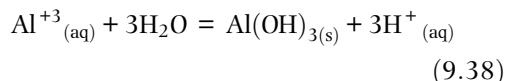
Aluminum is nearly insoluble at neutral pH (Fig. 4.10), but as pH drops below ~6, Al is released to solution as the polyatomic cations $Al(OH)_2^+$ and $Al(OH)^+$ or the ion Al^{+3} (equations 4.8–4.12), depending on pH (Fig. 4.10). The main cause of increased Al solubility in acidic soils is the ability of H^+ to dissolve Al-bearing minerals, including the Al-hydroxide gibbsite ($Al[OH]_3$) (Fig. 9.10) as well as

aluminosilicates. The sum of equations 4.8, 4.9 and 4.10 helps to visualize the net effect of H^+ of Al solubility:



This is effectively an acid–base reaction, but unfortunately it is one that buffers solutions at a pH ~3–4 (compare that to buffering by carbonate species to pH ~5.5 to 8.0). Al is toxic to most plants, and one example of how elevated Al may inhibit plant growth in acidic soil environments is via incorporation of Al into cell walls. Ca^{+2} enhances cell-wall stability, but in acidic soils where plant roots incorporate excess Al in place of Ca, cell-wall stability is impaired and plants become less resistant to cold, disease and other stresses.

Aluminum that leaches from acidic soils into acidic lakes or is released from sediments in acidic lakes is also potentially harmful to fish. When Al solubility in soils is increased as described above, soluble Al can be leached out of soils or sediments and into surface water bodies; also, Al in sediments may be leached into lake water if lake pH drops below pH ~5.5. Dissolved Al in lakes or streams is problematic because fish breathe by passing water through their gills and extracting dissolved oxygen from the water. The pH inside a fish gill is typically between 6 and 7, so if the water brought into the gill contains elevated amounts of dissolved Al, here is an example of what can happen:



Al is insoluble in circum-neutral fish gills so solid $Al(OH)_3$ precipitates in or on the gills. This then causes the gill to excrete mucous in an attempt to deal with the irritation, and the ultimate result is that the light gray mucous + $Al(OH)_3$ coating causes fish to suffer from respiratory distress and cell necrosis that can lead to death.

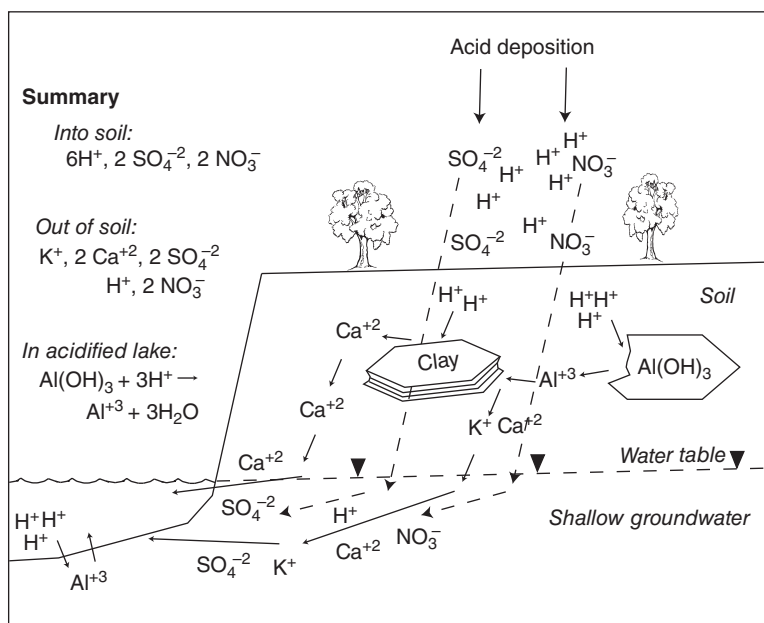


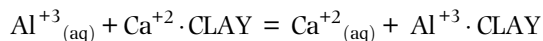
Fig. 9.10 Schematic diagram of effects of acid deposition on soil and adjacent lake. With the exception of the reaction in the lake sediments, the diagram attempts to balance ionic charges; for example, -6 mol of charge from sulfate and nitrate and $+6$ mol of charge from H^+ enter the soil, and -6 mol of charge from sulfate and nitrate enter the lake, balanced by $+6$ mol of charge (2Ca^{+2} , 1K^+ and 1H^+).

9.4.2 Displacement of adsorbed nutrient cations

Addition of excess H^+ to soil alters the composition of exchangeable cations adsorbed onto surfaces of clays and humus (Fig. 9.10). If we view H^+ (or H_3O^+) as a cation, we can begin to understand impacts to the adsorbed plant-available cation composition of particle surfaces. Two primary factors enhance sorption of H^+ onto particle surfaces and result in desorption of cations such as Ca^{+2} and K^+ : (1) the high activity (or concentration) of H^+ in acidic soils, and (2) the relatively low ratio of hydration radius to charge (Table 4.6), where the relatively small radius of H_3O^+ contributes to its attraction to exchange sites and allows it to compete with large radius Ca^{+2} and K^+ ions, especially in H^+ -rich (acidic) solutions. Thus, the “ H^+ cation” alone may result

in loss of nutrient cations from particle surfaces, and then once Ca^{+2} and other cation nutrients are free to migrate in the soil solution, they are prone to leaching and loss from the soil.

Increased solubility of Al can also have a profound effect on the exchangeable cation pool. The high charge and relatively small hydration radius of the Al^{+3} cation mean that it is more strongly attracted to particle surfaces than are larger, lower-charge nutrient cations K^+ and Ca^{+2} (Table 4.6). The exchange reaction below (similar to equation 2.6) depicts adsorption of Al and desorption of Ca with increasing activity of Al in solution.



Al for Ca exchange is depicted schematically in Fig. 9.10, and with time, soils become Ca-depleted

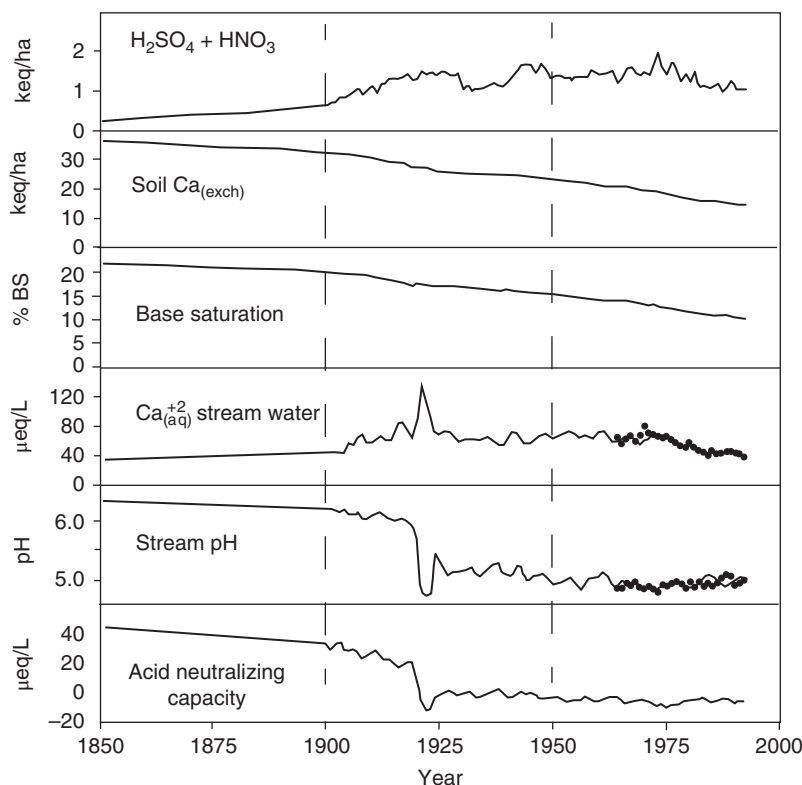


Fig. 9.11 Study of the Hubbard Brook Experimental Forest (HBEF), New Hampshire, northeast USA showing changes in atmospheric deposition of mineral acids, soil exchangeable Ca, base saturation, dissolved Ca in stream water, stream pH and ANC. Data consist of measured and modeled data, and data points are shown for late 20th century stream measurements to show comparison of measured and modeled. Early 1920s spikes in Ca, pH and ANC are attributed to short-term forest disturbance (tree harvesting) and release of Ca^{+2} and H^{+} from the forest floor. (Adapted from Gbondo-Tugbawa and Driscoll 2003.)

and Al-rich, an imbalance that is harmful to native ecosystems and to agriculture. Initially, dissolved Ca often increases in lake and stream water as it is flushed out of soil, an effect visible in Fig. 9.11 (especially early 1920s). Later, as soils become saturated with Al, it too begins to leach into surface water bodies (Fig. 9.12).

9.4.3 Leaching of base cations enhanced by increased NO_3 and SO_4

Soil solutions must maintain electrical neutrality, and the high flux of soluble anions (NO_3^- , SO_4^{2-})

through and out of soils requires that cations also be leached to balance the change of the leaching nitrate and sulfate. Given that Al^{+3} and H^{+} tend to displace Ca^{+2} (and K^{+} to a lesser extent) from exchange sites, the pronounced leaching of anions through soil and into groundwater results in depletion of Ca^{+2} , a critical nutrient cation in many forest ecosystems.

Two other considerations related to loss of base cations from soils, which are usually less important than the mechanisms described above, are (1) the physical removal of base cations when trees are harvested (Fig. 9.11), and (2) emissions controls that have reduced the amounts of

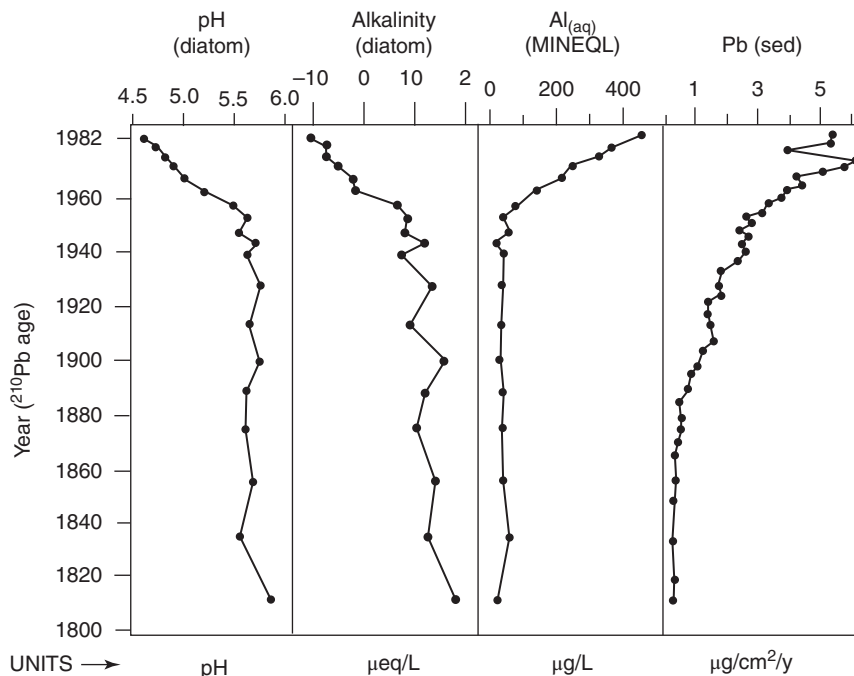


Fig. 9.12 Data from a sediment core from Big Moose Lake in the southwestern Adirondack region, New York State, USA. Age of sediment layers was determined by ^{210}Pb and ^{137}Cs dating (Chapter 11). Alkalinity and pH were determined from diatoms in the lake sediment and $\text{Al}_{(\text{aq})}$ by modeling using MINEQL. Original data and analytical details are in Charles et al. (1987).

particulates from smokestacks, resulting in a return to pre-industrial (i.e. $\sim$ low) deposition of base cations onto soils.

9.4.4 Decrease of soil buffering capacity and base saturation

Soils are capable of buffering acid deposition in four main ways:

1 by release of Ca and other base cations adsorbed onto particle surfaces in exchange for H^+ (shown schematically in Fig. 9.10).

2 via chemical weathering, especially of carbonates and other reactive minerals that consume H^+ as they decompose (e.g. $\text{CaCO}_3 + \text{H}^+ = \text{Ca}^{+2} + \text{HCO}_3^-$), and

3 reaction of dissolved bicarbonate with H^+ ($\text{HCO}_3^- + \text{H}^+ \rightarrow \text{H}_2\text{CO}_3$).

4 reaction of organic matter (e.g. humus) with H^+ . In some cases humus may adsorb H^+ onto exchange sites but in other cases H^+ may result in breakdown of organic matter.

The combined effect of mechanisms 1 through 4 is the **acid-neutralizing capacity** (ANC) of the soil, i.e. ANC is the capacity of particulates plus dissolved species in soil (or a water body) to neutralize acid. It is effectively synonymous with alkalinity. Soils rich in calcite or with abundant clays or humus that are capable of exchanging base cations for H^+ will withstand the effects of acid deposition better than will soils rich in quartz, feldspars and other minerals that are largely non-reactive and have very low CEC. Soils or water bodies with $\text{ANC} < 50 \mu\text{eq/L}$ are typically considered sensitive to acidic deposition (they have a low capacity to buffer acids), whereas soils or waters with ANC values $> 500 \mu\text{eq/L}$ are relatively resistant to

acidification. For definition of alkalinity and how it is determined, refer to Chapter 5 (Section 5.1, Figures 5.1 and 5.2, and equations 5.15 and 5.16). ANC can become negative if the concentration of acids exceeds bases – an example is an acidic stream or lake lacking bicarbonate and with minerals acids H_2SO_4 and HNO_3 and a pH of 4.

Base saturation is a measure of the abundance of exchangeable base cations (in cmol_c/kg ; refer to Chapter 2, Section 2.5) relative to the total cation exchange capacity (CEC), also expressed in units of cmol_c/kg . It can be quantified as follows:

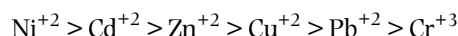
$$\% \text{BS} = \left[\frac{(\text{Na}^+ + \text{Mg}^{+2} + \text{K}^+ + \text{Ca}^{+2})}{(\text{total CEC})} \right] * 100 \quad (9.39)$$

When exchange sites are dominated by the base cations (i.e. Na^+ , Mg^{+2} , Ca^{+2} and K^+), the percent BS will be high; for example, Mollisols (semi-arid grassland soils) or Aridisols typically have percent BS values $> 50\%$ and often $> 80\%$. Conversely, when exchange sites are dominated by Al^{+3} , Fe^{+3} and H_3O^+ , as they often are in acidic soils, base saturation values are typically $< 20\%$. Note that in Fig. 9.11, % BS values of the Spodosols in the watershed of the Hubbard Brook Experimental Forest were $\sim 20\%$ prior to anthropogenic acid deposition; after a century of acid deposition, the percent BS is approximately 10% . In other words, 50% of the base cations formerly occupying exchange sites on soil particles have been leached out of soils and replaced by Al^{+3} , H_3O^+ and perhaps Fe^{+3} . Soils with low percent BS tend to be nutrient-poor and incapable of buffering acid deposition.

9.4.5 Acid deposition and heavy metals

A general rule of thumb in environmental geochemistry is that, as is the case with Al and Fe, the mobility of heavy metals increases with decreasing pH. The Eh–pH diagrams of Cu and Pb shown in Fig. 4.13 indicate that these metals

are most soluble under acidic conditions, and the adsorption edge curves in Fig. 4.18 reveal that heavy metals tend to desorb from particle surfaces as pH decreases. The order of heavy metal desorption with decreasing pH (below pH = 7) is:



In other words, as soil pH decreases, Ni^{+2} will become mobile more readily than will the other metals listed, and Pb^{+2} and Cr^{+3} are least likely to be mobilized (although in reality the plot for Cr^{+6} in the polyatomic anion CrO_4^{-2} is more apt for oxidized soil environments than the reduced Cr^{+3} form; refer to Fig. 4.18). Field studies and laboratory experimentation summarized by Wilson and Bell (1996) indicate that Ni^{+2} and Zn^{+2} tend to be most readily mobilized and leached from soil as pH decreases due to acid deposition. Cd^{+2} forms strong complexes with particulate organic matter and this enhances its retention in soil even as it is released from mineral (e.g. Fe-hydroxide) surfaces. Pb^{+2} forms similarly strong complexes with organic matter, a major factor in its low mobility in soils. The cause of the > 5 -fold increase in Pb concentration in sediments of Big Moose Lake (Fig. 9.12) is not leaching from soils, but rather atmospheric deposition of anthropogenic Pb into the lake and onto the surrounding landscape, where it is able to be physically eroded into the lake (Chapter 8).

9.5 SOILS AND PLANT NUTRIENTS

A quick examination of the chemical composition of plants gives us insight into nutrients that plants must obtain from soil. The carbon and oxygen in plants is derived from atmospheric CO_2 via photosynthesis, and the hydrogen used by plants to form carbohydrates and other organic molecules can be derived from air or soil water. Nearly all other components of plants, i.e. the mineral nutrients (Table 9.4), must be obtained from the soil,

Table 9.4 Examples of nutrient concentrations in grasses (e.g. alfalfa, clover, ryegrass), apple leaves, pine needles and tomato leaves. Data obtained from Brady and Weil (2002) and the University of Arkansas Cooperative Extension Service (http://www.uaex.edu/Other_Areas/publications/pdf/FSA-3043.pdf), and various other sources. Typical values for carbon, hydrogen and oxygen in bulk plant tissue are: C = 40–50%, H = ~5% and O = 40–50%.

	Grasses	Apple Leaves	Pine Needles	Tomato Leaves
N (wt%)	1.0–5.0	1.5–2.5	1.0–1.5	3.0–5.0
K (wt%)	1.0–4.0	1.0–2.0	0.30–0.45	2.5–4.5
Ca (wt%)	0.20–2.0	1.0–1.5	0.10–0.20	1.5–4.0
Mg (wt%)	0.10–0.50	0.25–0.50	0.05–0.10	0.45–0.70
P (wt%)	0.20–0.50	0.15–0.30	0.10–0.20	0.30–0.50
S (wt%)	0.10–0.50	0.10–0.30	0.05–0.15	0.50–1.0
B (ppm)	10–50	20–50	3–10	30–60
Cl (ppm)				
Cu (ppm)	5–75	5–20	1–10	5–15
Fe (ppm)	25–250	50–250	20–100	100–200
Mn (ppm)	25–150	25–100	50–600	50–200
Mo (ppm)	0.1–0.5			
Zn (ppm)	20–80	20–50	20–50	30–50

and the majority of the mineral nutrients are, in fact, derived from chemical weathering of minerals, at least initially – the “initially” caveat refers to the fact that decomposing organic matter provides nutrients (originally derived from minerals) to plants.

Plant nutrients are commonly divided into the macronutrients and micronutrients, where **macronutrients** occur in concentrations greater than 0.1% (by wt.) and **micronutrients** typically comprise < 200 ppm of bulk plant tissue. Because they are not derived from the soil solution, carbon, hydrogen and oxygen typically are not classified as nutrients, *strictu sensu*. The macronutrients are N, P, K, Ca, Mg, and S and the micronutrients include boron, chloride (Cl⁻, not Cl₂ gas), copper, iron, manganese, molybdenum, nickel and zinc, and the only one of these that is generally not derived from rock weathering is nitrogen, which is fixed into soils from the atmosphere by symbiotic relationships of plants and nitrogen-fixing bacteria. (Anthropogenic NO_x emissions and fertilizer additions are now responsible for the

majority of N fixation in many soils; refer to Chapter 6 for details on the nitrogen cycle.) The remainder of the macronutrients and micronutrients all are mineral-derived and are liberated to the soil by chemical weathering (Table 9.5). Later, after mineral-derived nutrients are incorporated into plants, decaying plant matter (e.g. humus) also becomes a source of all nutrients, as do exchange sites on surfaces of clay minerals, hydroxides and humus.

The roles of nutrients in plants include nitrogen as a component of proteins; phosphorus as a component of adenosine di- and tri-phosphate and, thus, a critical element in photosynthesis; calcium as a constituent of cell walls and magnesium as a component of chlorophyll; boron in production of carbohydrates; and copper and chloride in regulating plant metabolism. Note that of the main elements in the continental crust and soils, neither Al nor Na are plant nutrients and when they are abundant in the exchangeable/plant available fraction, they may impair plant growth.

Table 9.5 Plant nutrients, their common forms and sources. For each of the macronutrients, decomposing organic matter is also a source.

Nutrient	Source	Plant-available Form
C	Atmosphere (photosynthesis)	CO ₂
O	Atmosphere (photosynthesis)	CO ₂
H	Atmosphere, soil water	H ₂ O
<i>Macronutrients</i>		
N	Fixation from N ₂	NO ₃ ⁻ > NH ₄ ⁺
P	Apatite (Ca ₅ [PO ₄] ₃ OH)	HPO ₄ ⁻² , H ₂ PO ₄ ⁻ , etc.
K	Micas, K-feldspars	K ⁺
Ca	Calcite, plagioclase feldspars	Ca ⁺²
Mg	Dolomite, chlorite & many silicates	Mg ⁺²
S	Sulfides, sulfates	SO ₄ ⁻²
<i>Micronutrients</i>		
B	Borates, micas	H ₂ BO ₃ ⁻¹
Cl	Salts, micas, precipitation	Cl ⁻
Cu	Sulfides (e.g. chalcopyrite)	Cu ⁺²
Fe	Pyrite, Fe-oxides/hydroxides, silicates	Fe ⁺² or Fe ⁺³
Mn	Iron oxides	Mn ⁺²
Mo	Sulfides, oxides	MoO ₄ ⁻² , HMoO ₄ ⁻
Ni	Millerite, trace in some silicates	Ni ⁺²
Zn	Sphalerite, trace in some silicates	Zn ⁺²

9.6 SALINE AND SODIC SOILS

In strongly leached soils like those of the tropics or humid temperate forests, weathering of primary minerals and leaching of base cations tends to result in naturally acidic soils (e.g. Spodosols, Oxisols). On the contrary, in hot arid climates, chemical weathering is often intense during short periods of time following rainfall, but when evaporation exceeds infiltration, as is typically the case in arid region soils, base cations liberated from primary minerals tend not to be leached from soil. In some cases, strong evaporative forces draw soil water upward toward the surface, carrying dissolved base cations upward, resulting in accumulation of base cations in soil A and B horizons.

Chemical weathering of primary minerals such as silicates (e.g. equations 9.2 and 9.8 through 9.10) and carbonates (equations 2.7 and 4.2) are often the main source of base cations in soil, although in some cases coastal areas may receive

appreciable Na in rainfall (Table 4.1), and weathering of marine shales may produce soils with exchangeable Na⁺ higher than normal. A very important anthropogenic source of Na is irrigation water. Whereas rainwater typically contains ~1 ppm Na⁺, river water or groundwater in arid regions often contains > 100 ppm Na⁺; thus, when crops are irrigated, Na⁺ is added to soil, often in high concentrations, and if evaporation exceeds infiltration rate, Na⁺ will not be leached out of soil, but rather accumulate.

At this point a couple of definitions are necessary. **Saline** soils have high concentrations of base cations balanced by anions Cl⁻, SO₄⁻², HCO₃⁻/CO₃⁻² and/or NO₃⁻, and are defined by the UN FAO (Food and Agriculture Organization) as having electrical conductivity > 4 dS/m (TDS > 2500 mg/L). They are typically alkaline, with pH in the range of 8 to 10. In terms of plant growth, excess soil salinity makes less water available to plants than is actually in the rooting

zone. This is because as the salt concentration increases, so too does the osmotic pressure of the soil solution – the pressure difference is caused by high TDS in soil water and low TDS in root water. The osmotic pressure gradient would be for water to seep out of root to dilute the soil solution. This not only makes it difficult for plants to draw water into roots, but in fact in some cases roots may lose water to the salty solution according to the osmotic pressure gradient. Another problem with saline soil water is that, in addition to the osmotic effect, excessively high concentrations of Ca, Na and Mg may limit the uptake of other essential plant nutrients.

Sodic soils are a special subgroup of saline soils that contain very high concentrations of exchangeable Na; according to the UN FAO, **sodic soils** possess exchangeable sodium percentage (ESP) > 15%. The ESP formula is much like the one for percent base saturation, with variables expressed in $\text{cmol}_\text{c}/\text{kg}$, and in this case the numerator is solely exchangeable Na^+ (mmol/kg).

$$\% \text{ESP} = [\text{Na}^+ \div (\text{total CEC})] * 100 \quad (9.40)$$

In typical saline or sodic soils, the exchange sites are overwhelmingly dominated by Na, Ca and Mg, so in some cases the denominator of ESP is presented as the sum of those three exchangeable cations. The sodium absorption ratio (SAR) is also used to define sodic soils, where the threshold is 12 and variables are expressed in $\text{mmol}_\text{c}/\text{L}$:

$$\text{SAR} = \frac{[\text{Na}^+]}{\sqrt{([\text{Na}^+] + [\text{Ca}^{+2}])}} / 2 \quad (9.41)$$

Other than the fact that Na is not a plant nutrient and its abundance in sodic soil solutions can poison plants or impede uptake of nutrient cations, abundant exchangeable Na^+ also negatively impacts soil structure, especially in clay-rich soils. Because of its low ionic charge and relatively high

hydration radius (Table 4.6), the charge on hydrated Na^+ is weak and Na^+ does not strongly attract 2:1 layers of smectite clays; in effect, Na^+ causes dispersion of clay particles, which results in destruction of the aggregate structure of soils. This leads to poor drainage (low permeability) that impairs movement of water throughout the soil and limits migration of nutrients towards roots. In contrast, the relatively high charge and low hydration radius of Ca^{+2} (Table 4.6) means that soils with abundant exchangeable Ca^{+2} will contain smectite 2:1 layers that are collapsed around the hydrated Ca^{+2} , an attribute that fosters formation of soil aggregates and enhances permeability (Fig. 9.13).

Soil sodification is an important environmental and social issue globally because 15–20% of the world's agricultural soils are saline soils under irrigation. Continued addition of Na in irrigation water brings these soils closer to a sodic state, and once reached, sodic soils may end up abandoned or in need of remediation – two common approaches to remediation are (1) to try to flush Na out of the soil (often limited by lack of water in arid regions, or made difficult by poor soil structure; Fig. 9.13), or (2) to add Ca (e.g. powdered gypsum) to increase the Ca:Na ratio and decrease activity of Na. Yet often the low permeability of sodic soils inhibits flushing/leaching of Na, and additions of Ca to improve Ca:Na ratio eventually cannot overcome increasing salinity.

9.7 TOXIC METALS AND METALLOIDS

Certain factors must be considered when trying to predict or understand the behavior of a given element or compound in soil (or surface water or groundwater). Typically only a fraction of the total (bulk) amount of an element or compound in soil or sediment is mobile (i.e. capable of being transported in solution or as extremely

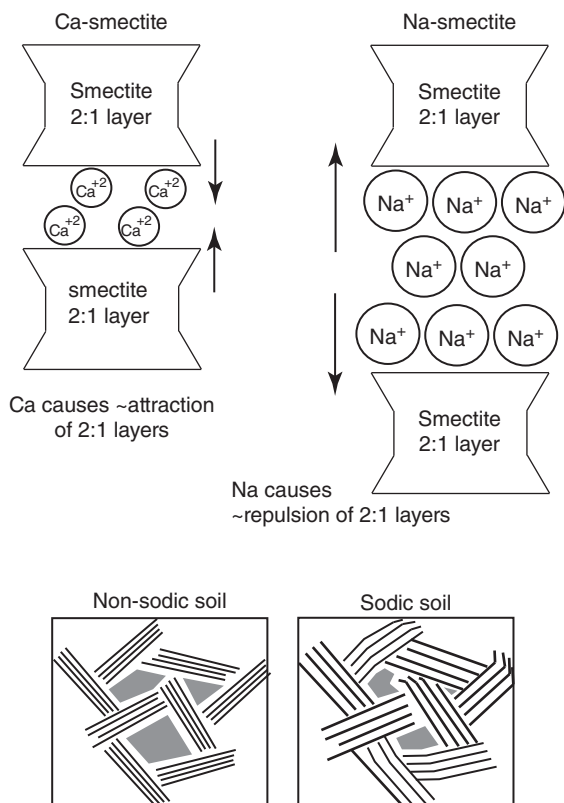


Fig. 9.13 Schematic sketches showing the effects of exchangeable Ca^{+2} vs Na^{+} in soils. The top sketches diagram the capacity of Ca^{+2} to collapse smectite 2:1 layers around the interlayer vs. the way that Na^{+} causes expansion and dispersion of smectite 2:1 layers. In soil (bottom), crystals of clay are shown. Ca^{+2} leads to formation of aggregates that enhance soil permeability, whereas clay expansion and dispersion caused by Na^{+} reduces pore size and decreases permeability (shown by gray polygons). (Modified from McBride 1994.)

fine-grained colloidal suspensions, or taken up by plants or other organisms); the remainder will be immobile, and mainly bound in (or onto) minerals, organic matter or other solids or organisms that prevent its release into solution.

Questions that can help to determine mobility of metal or metalloid (or most elements or compounds in environmental geochemistry) include:

- Does the element or compound tend to form complexes with organic matter? If so, does this limit its mobility, as would be the case of complexation with particulate organic carbon (POC, e.g. humus), or does organic matter enhance mobility, as would be the case of complexation of certain metals with dissolved organic carbon (DOC, e.g. humic acid).
- Is the element or compound likely to be adsorbed to particle surfaces? (Figures 4.17 and 4.18).
- How do solution pH and redox state (i.e. Eh or pe) affect mobility of the element or compound (Figures 4.14–4.16)?
- Is it likely that the element or compound will co-precipitate in minerals (as is the case with many trace metals and arsenic in pyrite or Fe-hydroxides; Chapter 4, Section 4.8)? In this case, the mobility of the trace element will be strongly controlled by the stability of the host mineral (a good example of arsenic in Fe-hydroxides is presented in Appendix I).
- Will the element or compound combine with other elements to form a mineral in which the element (e.g. Pb, Mn, Zn) is a dominant component (e.g. PbCO_3 , MnO , ZnS)?

These questions seek to determine **speciation** (Chapter 4, Section 4.4) and address issues such as: what% of substance X is plant available or likely to be leached from soil into surface water? Will substance X be released to solution if the water table is lowered by excessive pumping of water wells and sediments become aerated (i.e. are the toxic elements co-precipitated in sulfides that might be decomposed by oxidation)? How much of substance X is complexed to particulate organic matter?

The various reservoirs (or “pools”) of a given element in soil, as well as mechanisms responsible for transfer of an element from one reservoir to another, are shown in Fig. 9.14. To some extent, the speciation of any element will be controlled by inherent characteristics such as ratio of ionic charge to radius (“ionic potential”; Table 4.6), but the environment also exerts strong controls.

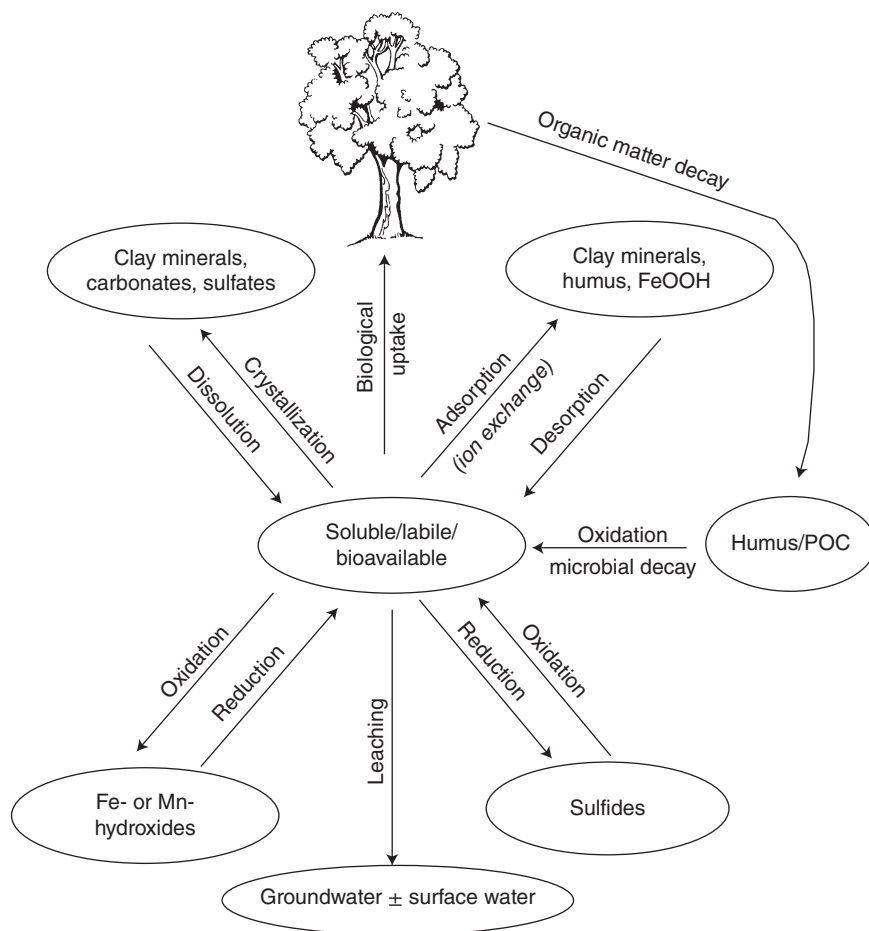


Fig. 9.14 Sketch of the main reservoirs, pathways and mechanisms associated with speciation of trace metals and other substances in soil. Many of these processes may be strongly influenced by microbial activity, including oxidation of sulfides, organic matter decay and mineral crystallization. (Modified from McBride 1994.)

Two examples of environmental control on potentially toxic trace elements are provided to give an insight into trace element behavior:

In reducing waters (e.g. low- O_2 groundwater), manganese tends to be soluble as Mn^{+2} but when exposed to an oxidizing environment (e.g. the casing inside a well where air circulates), Mn tends to precipitate as hydrated Mn oxide, a black gooey (when wet) or powdery (when dry) precipitate. Unpublished analysis by the author of a viscous black coating on a water filter attached to a well (producing from a black shale aquifer) was identified by X-ray diffraction as partially hydrated

birnessite, a manganese oxide mineral with the chemical formula $Na_{0.7}Ca_{0.3}Mn_7O_{14} \cdot nH_2O$. Manganese is a potential neurotoxin that, with chronic long-term exposure, may cause symptoms similar to those associated with Parkinson's disease; accordingly, the redox state of well water will influence the amount of dissolved Mn in drinking-water supplies. Speciation of Mn in a shallow oxidized aquifer would likely be as birnessite or similar Mn oxide or hydroxide, whereas in anoxic groundwater Mn is more likely to be dissolved as Mn^{+2} .

Lead arsenate ($PbHAsO_4$) was widely used as a pesticide in fruit orchards from the late 1800s

until the 1940s, when cheaper and (in some ways) safer DDT was adopted. The decades of lead arsenate application provides the opportunity to compare the behavior of Pb and As under effectively identical environmental factors – for a given soil, the mineralogy, organic matter composition and content, temperature, moisture content, biological community, permeability, pH, redox conditions and any other potential factors available to affect one of the elements is also there to affect the other, and results reveal important differences in their inherent behaviors. Pb forms very strong complexes with particulate organic carbon (e.g. humus) and sorbs to iron hydroxides, and thus is nearly immobile (unless soil erosion occurs, in

which case Pb may be physically eroded and transported along with particulates); As also tends to be largely immobile as it is similarly sorbed to POC and hydroxides. However, As can become mobilized when soils are fertilized with phosphate. According to Fig. 6.8, phosphate will exist as a mix of H_2PO_4^- and HPO_4^{2-} at soil pH of 6–8. Similarly, in this pH range arsenic will occur as a mix of H_2AsO_4^- and HAsO_4^{2-} . Given the smaller radius of P^{+5} relative to As^{+5} , the radius of phosphate anions will be smaller than of arsenate anions – Carron et al. (1958) report that the radius of $\text{PO}_4^{3-} = 1.75 \text{ \AA}$ compared to $1.86 \text{ \AA}$ for AsO_4^{3-} . You can probably predict the effect – phosphate anions will be attracted to adsorption

One common approach to determining speciation of elements or compounds in soil is sequential chemical extraction (SCE). A frequently used SCE approach was developed

by Tessier et al. (1979) and has been slightly modified by subsequent researchers (e.g. Rauret et al., 1999; Young et al, 2006). A representative example of SCE is:

Target Phase/Pool	Reagent
1 Exchangeable ions, soluble in weak acids (e.g. carbonates)	0.11 M acetic acid (pH = 3)
2 Reducible (e.g. Fe-, Mn-hydroxides)	0.5 M hydroxylamine hydrochloride (pH = 1)
3 Oxidizable (sulfides, organic matter)	8.8 M hydrogen peroxide (pH = 2.5)
4 Residual (some silicates, well-crystallized oxides)	Aqua regia solution (e.g. 50% HNO_3 + 20% HCl)

In each of the 4 stages, a known mass of dried soil or sediment (e.g. 3.0 g) is reacted with a known volume (e.g. 25 or 100 mL) of the reagent. Often the same initial powder is exposed sequentially, so that first the weakly held (bio-available, ~soluble) exchangeable and carbonate-bound ions are released to solution, then after washing to remove extracted ions and leftover previous reagent, the same powder can be reacted in the presence of hydroxylamine hydrochloride to extract ions bound to (or in) Fe or Mn hydroxides, and so on, up

through stage 4. One important caveat is that the reagents are not always selective; for example, Ryan et al. (2008) found that certain Fe–Mg smectite and chlorite clays dissolve during stages 1–3. Given that trace metals substitute into the octahedral sheet, interpretation without mineralogical analysis might assign those metals to hydroxides or sulfides rather than clays. For this reason many geochemists refrain from using specific phases, but rather indicate that a certain metal is, e.g. associated with the oxidizable fraction.

or exchange sites more strongly than will arsenate, so as the amount of phosphate increases by fertilizer application, arsenate mobility will increase (Peryea and Kammereck, 1997). Thus, the inherent difference between cationic Pb^{+2} and anionic H_2AsO_4^- and HAsO_4^{2-} leads to different behavior when environmental factors (in this case, addition of phosphate) change.

9.8 ORGANIC SOIL POLLUTANTS AND REMEDIATION (FUELS, INSECTICIDES, SOLVENTS)

The main concepts and principles related to behavior of organic pollutants in the environment are presented in Chapters 3 (Section 3.3) and 8 (Section 8.8), and Appendix I presents a case study involving remediation of groundwater contaminated with organic solvents. The following section, therefore, draws attention to pertinent issues at the intersection of organic compounds and soil. The reactions that take place in soil control potential vaporization of VOCs into buildings, potential uptake of organic pollutants into organisms (including crops), potential leaching into aquifer drinking-water supplies, and potential for transport of organic contaminants into surface waters.

In soil, organic compounds nearly always occur in one of five possible states: sorbed (e.g. to soil organic carbon or minerals), gaseous (especially in the case of VOCs; Chapter 3, Section 3.3, Table 3.4), liquid (i.e. NAPLs), aqueous/dissolved (e.g. ~soluble benzene, TCE), or incorporated into organisms (i.e. plants, animals); therefore, quantifying the factors that control uptake into these reservoirs is vital for predicting and understanding behavior of organic compounds in the environment. Figure 9.15 is a schematic of the main reservoirs and reactions that occur involving organic contaminants in soils. It is worth emphasizing an important difference between the behavior of organic contaminants as shown

in Fig. 9.15, compared to potential pathways of inorganic metals and metalloids shown in Figure 9.14 – the potential for organic compounds to decompose into metabolites (Chapter 3 Section 3.3). An inorganic element such as As or Hg will always be As or Hg, and although its speciation may change, it will not (with the exception of radioactive elements) decompose into non-toxic end-products such as CO_2 and H_2O (which is the ultimate step in decomposition of hydrocarbons, for example). Another important difference between organic compounds and inorganic elements or molecules is that organics may occur in soil as insoluble liquids (NAPLs, Chapter 3, Section 3.3), whereas inorganic elements and substances in the environment do not.

Two factors strongly control sorption of organic compounds – the abundance of soil organic carbon (POC, humus) and, to usually to a much lower extent, the mineralogy of the soil. Non-polar organic compounds such as BTEX and most VOCs have very high K_{OW} values (Equation 3.4), indicating they are strongly attracted to soil organic carbon, and numerous studies have indicated that the extent of sorption of solvents, fuels and herbicides in soil is proportional to the quantity of soil organic matter. What this means is that soils with high organic matter content (e.g. Mollisols, Alfisols, Spodosols) will tend to sorb and limit migration of organic compounds; conversely, in soils with low organic matter content, as is often the case with Aridisols, Entisols and Oxisols, clay minerals and other inorganic solids may be an important variable exerting control on sorption. Within soil minerals there is great variability in sorption – smectites and vermiculites have high cation exchange capacity so they may adsorb cationic species (e.g. the herbicides diquat and paraquat), but in cases of non-polar organic compounds, the exposed basal tetrahedral sheet of 2:1 layers may be capable of adsorption. Two examples illustrate the complexity of adsorption to soil clays: Na-smectite has a greater capacity to adsorb atrazine

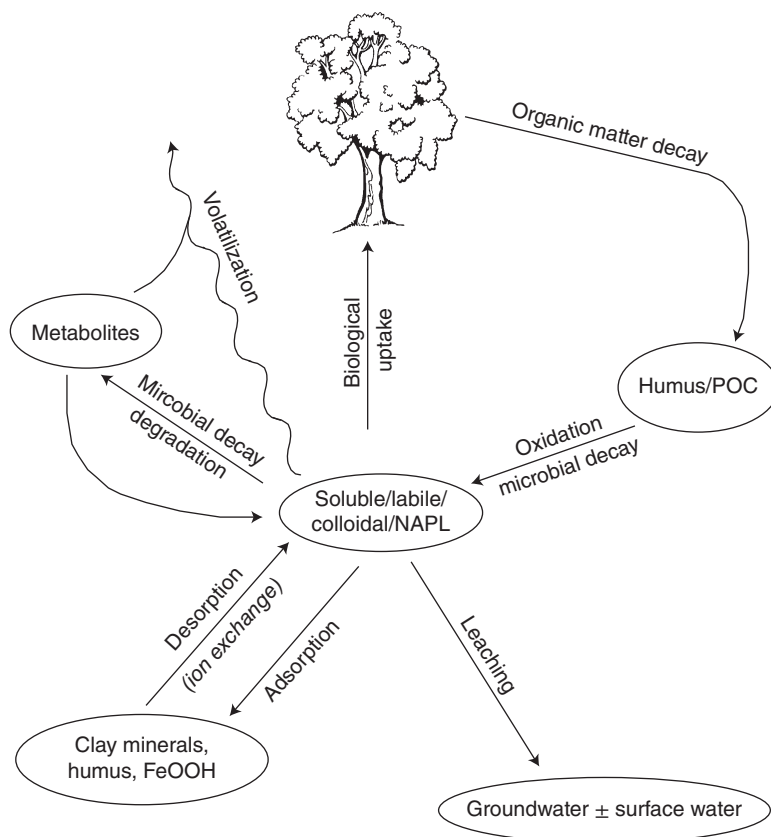


Fig. 9.15 Schematic diagram of potential reservoirs and reactions of organic contaminants once they have been introduced into soil. For most organic compounds the soluble/labile/colloidal/NAPL reservoir (i.e. likely to be mobile) is small given low solubilities (Table 3.4) and tendency to sorb to organic matter.

(a weakly polar herbicide) than does Ca-smectite (Prima et al., 2002), yet smectite was found to have no capacity to adsorb either of the non-polar compounds toluene or benzene (López-Cortés et al., 2008).

Understanding the behavior of organic pollutants in soils, surface water and groundwater is clearly a challenging proposition. Some of the basic principles are provided above and in Section 3.3 of Chapter 3, and further analysis of this topic also provided in Appendix I in the form of a case study that provides insight into strategies for remediation of organic contaminants. This case study reveals the complex relationships among inorganic minerals and elements, organic

compounds, microbial communities and other environmental controls such as redox and pH.

QUESTIONS

1. The data in Table 9.2 indicate that goethite (FeOOH) decreases with time in tropical soils, and at the same time, hematite (Fe_2O_3) increases. From these data, one might hypothesize that goethite transforms to hematite in soils at $\sim 25^\circ\text{C}$ over time spans of 10^4 to 10^5 years. Assess this hypothesis using thermodynamic data; in other words, is the transformation of goethite to hematite

spontaneous? Is goethite stable in these soils? Is it metastable? Is hematite stable relative to goethite? Or vice versa?

2. At what soil solution conditions do kaolinite and gibbsite exist in equilibrium?
3. Determine the exchangeable sodium percentage (ESP) for a soil that is 40% smectite, 30% calcite, 20% quartz and 10% illite with exchangeable Na^+ of 1210 mg/kg. Is this a sodic soil? Explain the negative effects that high amounts of exchangeable Na^+ have on soils (consider chemical and physical effects). In what type of climate does this problem occur, and why?
4. Explain why Goldich found Bowen's reaction series useful in predicting weathering rates of igneous minerals in soil. Compare relative rates derived from this approach to the rates presented in Table 9.1 and indicate a factor that might explain any differences.
5. Citing mineral formulas, explain why smectite and calcite are common in Aridisols, whereas kaolinite and gibbsite are common in Oxisols. Comment on how this difference could affect plant nutrient retention in Aridisols vs. Oxisols.
6. Explain why kinetic factors are often more influential in determining soil mineralogy than thermodynamic factors, citing three examples.
7. Determine relative gain and loss of elements associated with weathering of basalt to a tropical Alfisol B horizon (Plate 13). Use the data provided to rank elements in terms of percentage change as well as by calculating the mass-transfer coefficient, τ_j , for each element. Major elements are presented as wt% oxides and trace elements as mg/kg (ppm).

For both methods, rank elements in terms of mobility, and also compare results produced by the two methods.

8. Construct a mineral stability diagram like the one in Fig. 9.8, replacing albite with K-feldspar and Na-beidellite with K-montmorillonite.
9. What are common fates of H^+ delivered to soils by acid deposition? What effect does added H^+ have on Ca^{+2} and Al^{+3} ? Explain the problem associated with high amounts of NO_3^- and SO_4^{2-} fluxing through soils acid affected by acid deposition.
10. Provide two chemical equations that depict the oxidation of pyrite in acid mine (or rock) drainage systems (include one involving ferrous iron as reactant). Also describe the role of bacteria in these reactions.
11. List and briefly describe three important ways by which the behavior of organic pollutants (e.g. VOCs) in soil differs from the behavior of inorganic elements (e.g. Cr, Pb).

REFERENCES

- Birkeland, P.W., 1999. *Soils and Geomorphology*. Oxford University Press, New York, 430 pp.
- Bonan, G., 2010. Nitrogen and Climate. American Meteorological Society Climate Briefing Series. The Role of Nitrogen in Global Change. Russell Senate Office Building, United States Senate Washington, DC. Accessed on-line 5-December 2013. http://www.cgd.ucar.edu/staff/bonan/seminars/Bonan_Nitrogen_AMS.pdf.
- Brady, N.C. and Weil, R.R., 2002. *The Nature and Properties of Soils*, 14th edn, Prentice Hall. Upper Saddle River, New Jersey.
- Brantley, S.L., 2005. Reaction kinetics of primary rock-forming minerals under ambient conditions. In J.I.

	SiO_2	TiO_2	Al_2O_3	$\text{Fe}_2\text{O}_{3(t)}$	MnO	MgO	CaO	Na_2O	K_2O	P_2O_5	Cr	Ni	Sr	Zr
B-horizon	44.6	1.78	27.7	21.3	0.29	1.75	0.79	0.04	0.06	0.64	436	326	69	104
Basalt	52.3	1.32	19.5	16.1	0.16	5.53	1.81	0.26	2.18	0.24	442	281	240	71

- Drever (ed.), *Surface and Ground Water, Weathering and Soils. Treatise on Geochemistry*, H.D. Holland and K.K. Turekian (executive editors), Elsevier. pp. 73–118.
- Brimhall, G.H. and Dietrich, W.E., 1987. Constitutive mass balance relations between chemical composition, volume, density, porosity, and strain in metasomatic hydrochemical systems: results on weathering and pedogenesis. *Geochimica et Cosmochimica Acta* 51: 567–587.
- Buss, H.L., Sak, P.B., Webb, S.M., and Brantley, S.L., 2008. Weathering of the Rio Blanco quartz diorite, Luquillo Mountains, Puerto Rico : Coupling oxidation, dissolution, and fracturing. *Geochimica et Cosmochimica Acta* 72: 4488–4507.
- Charles, D.F., Whitehead, D.R., Engstrom, D.R., Fry, B.D., Hites, R.A., Norton, S.A., Owen, J., Roll, L.A., Schindler, S., Smol, J.P., Uutala, A.J., White, J.R., and Wise, R.J. 1987. Paleolimnological evidence for recent acidification of Big Moose Lake, Adirondack Mountains, N.Y. (U.S.A.). *Biogeochemistry* 3: 267–296.
- Chou, L., Garrels, R.M., and Wollast, R., 1989. Comparative study of the kinetics and mechanisms of dissolution of carbonate minerals. *Chemical Geology* 78: 269–282.
- Davidson, E.A. and Janssens, I.A., 2006. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature* 440: 165–173. doi:10.1038/nature04514.
- Dean, J.A., 1979. *Lange's Handbook of Chemistry*, 12th edn, McGraw-Hill, New York.
- Drever, J.I., 1997. *The Geochemistry of Natural Waters: Surface and Groundwater Environments*, 3rd edn. Simon and Schuster, Upper Saddle River, New Jersey, USA.
- Fisher G.B. and Ryan P.C., 2006. The smectite to disordered kaolinite transition in a tropical soil chronosequence, Pacific Coast, Costa Rica. *Clays and Clay Minerals* 54: 571–586.
- Gbondo-Tugbawa, S.S. and Driscoll, C.T., 2003. Factors controlling long-term changes in soil pools of exchangeable basic cations and stream acid neutralizing capacity in a northern hardwood forest ecosystem. *Biogeochemistry* 63: 161–185.
- Goldich, S.S., 1938. A study in rock-weathering. *Journal of Geology* 46: 17–58.
- Gualtieri, A.F. and Venturelli, P., 1999. In situ study of the goethite-hematite phase transformation by real time synchrotron powder diffraction. *American Mineralogist* 84: 895–904.
- Hendriks, I.E., Duarte, C.M., and Alvarez, M., 2010. Vulnerability of marine biodiversity to ocean acidification: a meta-analysis. *Estuarine, Coastal and Shelf Science* 86: 157–164. doi:10.1016/j.ecss.2009.11.022.
- Lasaga, A.C., Soler, J.M., Ganor, J., Burch, T.E., and Nagy, K.L., 1994. Chemical weathering rate laws and global geochemical cycles. *Geochimica et Cosmochimica Acta* 58: 2361–2386.
- Lee, B.D., Sears, S.K., Graham, R.C., Amrhein, C. and Vali, H., 2003. Secondary Mineral Genesis from Chlorite and Serpentine in an Ultramafic Soil Toposequence. *Soil Science Society of America Journal* 67: 1309–1317.
- López-Cortés, C., Osorio-Revilla, G., Gallardo-Velázquez, T., and Arellano-Cárdenas, S., 2008. Adsorption of vapor-phase VOCs (benzene and toluene) on modified clays and its relation with surface properties. *Canadian Journal of Chemistry* 86: 305–311.
- McBride, M.B., 1994. *Environmental Chemistry of Soils*. Oxford University Press, New York. 406 p.
- Nesbitt, H.W. and Young G.M., 1982. Early Proterozoic climates and plate motions inferred from major element chemistry of lutites. *Nature* 299: 715–717.
- Prima, S., Evangelou, V.P., and McDonald, L.M., 2002. Surface exchange phase composition and nonionic surfactant effects on the nonequilibrium transport of atrazine. *Soil Science* 167: 260–268.
- Rauret, G., López-Sánchez, J.F., Sahuquillo, A., Rubio, R., Ure, A.M., Davidson, C.M., and Quevauviller, Ph., 1999. Improvement of the BCR three step sequential extraction procedure prior to the certification of new sediment and soil reference materials J. Environ. Monit. 1: 57–61.
- Retallack, G.J. 2001. *Soils of the Past: an Introduction to Paleopedology*. 2nd edn, Blackwell, Oxford, 600 p.
- Robie, R.A., Hemingway, B.S., and Fisher, J.R., 1979. Thermodynamic properties of minerals and related substances at 298.15 K and 1 bar 105 pascals pressure and at higher temperatures. United States Geological Survey Bulletin 1452. Accessed on-line 5-December 2013: <http://pubs.usgs.gov/bul/1452/report.pdf>.
- Ryan P.C. and Huertas, F.J., 2009. The temporal evolution of pedogenic Fe-smectite to Fe-kaolin via

- interstratified kaolin-smectite in a moist tropical soil chronosequence. *Geoderma* 151: 1–15.
- Ryan, P.C., Hillier, S., and Wall, A.J., 2008. Stepwise effects of the BCR sequential chemical extraction procedure on dissolution and metal release from common ferromagnesian clay minerals: a combined solution chemistry and X-ray powder diffraction study. *Science of the Total Environment* 407: 603–614.
- Sheldon, N.D. and Tabor, N.J., 2009. Quantitative paleoenvironmental and paleoclimatic reconstruction using paleosols. *Earth-Science Reviews* 95: 1–52.
- Spiro, T.G. and Stigliani, W.M., 2006. *Chemistry of the Environment*, 2nd edn, Prentice Hall, Upper Saddle River, NJ, USA.
- Stallard, R.F. and Edmond, J.M., 1983. Geochemistry of the Amazon: 2. The influence of geology and weathering environment on the dissolved load. *Journal of Geophysical Research, Oceans and Atmospheres* 88: 9671–9688.
- Stewart, B.W., Capo, R.C. and Chadwick, O.A., 2001. Effects of rainfall on weathering rate, base cation provenance, and Sr isotope composition of Hawaiian soils. *Geochimica et Cosmochimica Acta* 65: 1087–1099.
- Strakhov, N.M., 1967. *Principles of Lithogenesis*. Oliver and Boyd, Edinburgh.
- Tessier, A., Campbell, P.G.C., and Bisson, M., 1979. Sequential extraction procedure for the speciation of particulate trace metals. *Analytical Chem.* 51: 844–851.
- Wallace, J.M. and Hobbs, P., 2006, *Atmospheric Science: An Introductory Survey*. Elsevier, Amsterdam, The Netherlands.
- Wilson, M.J. and Bell, N., 1996: Acid deposition and heavy metal mobilization. *Appl. Geochem.* 11: 133–137.
- Young, S.D., Zhang, H., Tye, A.M., Maxted, A., Thums, C., and Thornton, I., 2006. Characterizing the availability of metals in contaminated soils. I. The solid phase: sequential extraction and isotopic dilution. *Soil Use Management* 21: 450–458.

Stable Isotope Geochemistry

Numerous isotopes of major and trace elements are useful in a wide variety of applications in the field of environmental geochemistry, particularly as tracers of geochemical processes and for determining ages of sediments, minerals (e.g. in rocks, faults and paleosols), groundwater, and organic matter. They have been used to determine origin of atmospheric contaminants (e.g. $^{202}\text{Hg}/^{198}\text{Hg}$), sources of nitrates in groundwater ($^{15}\text{N}/^{14}\text{N}$), global paleotemperature ($^{18}\text{O}/^{16}\text{O}$), the role of microbes in mineral reactions ($^{56}\text{Fe}/^{54}\text{Fe}$), and much more. Most elements have two or more naturally occurring isotopes, but interestingly, not all do. Notable examples of environmentally significant elements that have do not have multiple naturally occurring isotopes include arsenic (^{75}As), fluorine (^{19}F), manganese (^{55}Mn), and phosphorus (^{31}P) and needless to say this prevents analysis of their behavior in the environment by stable isotope geochemistry.

10.1 STABLE ISOTOPES – MASS DIFFERENCES AND THE CONCEPT OF FRACTIONATION

Stable isotopes are non-radioactive isotopes of an element. Common examples include ^{16}O – ^{17}O – ^{18}O , ^{14}N – ^{15}N and ^{32}S – ^{34}S . Examining the **chart of the nuclides** (Fig. 10.1), it is evident that the horizontal rows contain various isotopes of a given element. Using a familiar example, oxygen ($Z = 8$) contains three naturally occurring isotopes, ^{16}O , ^{17}O and ^{18}O , the main difference being the number of neutrons (8, 9 and 10, respectively). The base cation magnesium also has three stable isotopes (^{24}Mg , ^{25}Mg , ^{26}Mg) with 12, 13 and 14 neutrons, respectively, and we can also see that carbon has two stable isotopes (^{12}C and ^{13}C) and one radioactive isotope (^{14}C) (as well as the very rare ^{11}C).

Because isotopes of the same element have the same Z and the same electron structure (allowing

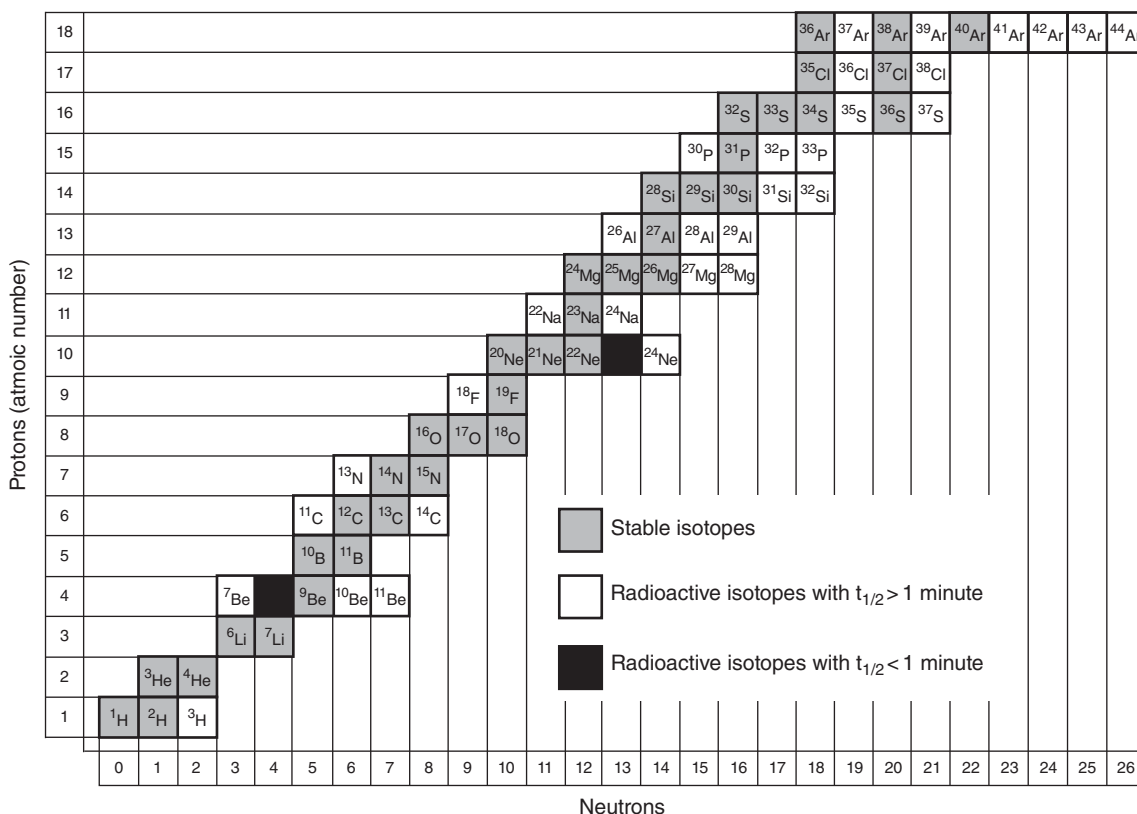


Fig. 10.1 Chart of the low atomic mass nuclides. Note that most elements have two or more stable isotopes (e.g. sulfur has four, ^{32}S , ^{33}S , ^{34}S and ^{36}S), but some have only one stable isotope (e.g. Be, F, Na, Al, P and As [not shown]). Radioactive isotopes included in this chart that occur in sufficient quantities to be useful in environmental geochemistry include ^3H (tritium), ^{10}Be , ^{14}C , ^{26}Al and ^{36}Cl . This chart also includes many isotopes with very short half-lives that effectively do not occur in nature (e.g. ^{11}C , ^{37}S , ^{18}F), and it also does not show numerous others with very short half-lives.

for changes to valence state that occur equally for the different isotopes of a given element), they behave very similarly, but not identically, in geochemical systems. Each of the three isotopes of oxygen will bond with silicon to form SiO_2 . Given that each of the oxygen atoms could be any of the three oxygen isotopes, possible SiO_2 molecules include Si^{16}O_2 , $\text{Si}^{16}\text{O}^{17}\text{O}$, $\text{Si}^{16}\text{O}^{18}\text{O}$, and so on. Silicon also occurs as three stable isotopes, ^{28}Si , ^{29}Si and ^{30}Si , so the combinations of silicon and oxygen isotopes further expands the array of SiO_2 molecular mass configurations to include molecules such as $^{28}\text{Si}^{16}\text{O}_2$, $^{28}\text{Si}^{16}\text{O}^{17}\text{O}$, $^{28}\text{Si}^{16}\text{O}^{18}\text{O}$ and so on.

In biochemical reactions, both ^{14}N and ^{15}N can be incorporated into proteins, and both can bond with oxygen to form nitrate (NO_3^-). The same can be said for stable carbon isotopes (^{12}C and ^{13}C) and the radioactive carbon isotope, ^{14}C – all three isotopes can be incorporated into CO_2 , organic matter or carbonate minerals. The subtle variations in behavior caused by the difference in mass between isotopes of the same element (or the compounds that the different isotopes form, e.g. H_2^{16}O vs. H_2^{18}O), is known as **mass-dependent fractionation (MDF)**, and this is the crux of stable isotope analysis.

Examination of the periodic table of the elements reveals that nearly all elements have **non-integer atomic masses**. For example, iron has an atomic mass of 55.85 g/mol, which can also be stated as 55.85 Daltons (Da), or 55.85 atomic mass units (where 1 a.m.u. = 1 Da). To be certain, there are no Fe atoms with 29.85 neutrons, which would be the number required to produce an atomic mass of 55.85 Da when added to 26 ($Z = 26$ for Fe). Rather, the value of 55.85 is produced by a mass balance of Fe isotopes with the following abundances (http://ie.lbl.gov/education/parent/Fe_iso.htm):

Fe ⁵⁴	5.58%
Fe ⁵⁶	91.95%
Fe ⁵⁷	2.18%
Fe ⁵⁸	0.30%

The mass balance calculation is expressed as

$$\text{Atomic mass (Da)} = 54 * (0.0558) + 56 * (0.9195) + 57 * (0.0218) + 58 * (0.0030) = 55.92$$

The slight discrepancy between what is seen on the periodic table of the elements (55.85 Da) compared to the value produced by the mass balance computed above (55.92 Da) is caused by the fact that all atomic masses are normalized to ¹²C, which is defined as having a mass of exactly 12 Da. Similarly, the same calculation for the Mg isotopes ²⁴Mg (78.99%), ²⁵Mg (10.00%) and ²⁶Mg (11.01%) produces an atomic mass of 24.320 Da, as compared to the ¹²C scaled value of 24.305 Da.

Water serves as an illustrative example of the differences that stable isotopes impart to molecules. Both of the naturally occurring hydrogen isotopes, ¹H and ²H (the heavier hydrogen isotope, ²H, is typically referred to as “D”, for deuterium) will combine with oxygen to form H₂O – most water molecules are ¹H₂O, but a small amount are ¹HDO or D₂O. When considering the oxygen isotopes in water, the possible types of water molecules expand

to ¹H₂¹⁶O, ¹H₂¹⁷O, ¹H₂¹⁸O, D₂¹⁶O, and so on. In all, there are nine stable isotopic combinations of water when considering all combinations of H, D and the three stable isotopes of oxygen. There are ten stable isotope combinations in carbon dioxide produced by ¹²C, ¹³C and the three stable isotopes of oxygen (Table 10.1).

The two main properties that make stable isotopes particularly useful in geochemical studies are:

- 1 the difference in mass of molecules with different isotopes, such as nitrate (NO₃⁻) with ¹⁵N vs. ¹⁴N, and;

- 2 the difference in covalent bond strength of the compounds formed by the different isotopes – nitrate formed by the heavier isotope ¹⁵N (i.e. ¹⁵NO₃⁻) will be more stable than nitrate formed by the lighter ¹⁴N isotope (i.e. ¹⁴NO₃⁻).

Perhaps the most straightforward example of how stable isotopes behave in natural systems is illustrated by evaporation of water molecules from a body of water such as the ocean, a wetland or a lake. The mass of H₂¹⁶O is 18 g/mol, whereas H₂¹⁸O is 20 g/mol. Based on its lower molecular mass, the probability of an H₂¹⁶O molecule evaporating is greater than the probability of an H₂¹⁸O molecule evaporating. Fractionation caused by evaporation produces atmospheric water with a higher ratio of ¹⁶O/¹⁸O than the original water body, and as evaporation proceeds, the surface water body becomes progressively enriched in water with ¹⁸O. This causes the

Table 10.1 Examples of stable isotope combinations in water and carbon dioxide.

Water:

¹H₂¹⁶O, ¹H₂¹⁷O, ¹H₂¹⁸O
 DH¹⁶O, DH¹⁷O, DH¹⁸O
 D₂¹⁶O, D₂¹⁷O, D₂¹⁸O

Carbon dioxide:

¹²C¹⁶O₂, ¹²C¹⁶O¹⁷O, ¹²C¹⁶O¹⁸O, ¹²C¹⁷O¹⁸O,
¹²C¹⁷O₂, ¹²C¹⁸O₂,
¹³C¹⁶O₂, ¹³C¹⁶O¹⁷O, ¹³C¹⁶O¹⁸O, ¹³C¹⁷O¹⁸O,
¹³C¹⁷O₂, ¹³C¹⁸O₂

$^{16}\text{O}/^{18}\text{O}$ ratio of the water body to decrease as the amount of ^{18}O increases with progressive evaporation. Similarly, when water vapor in the atmosphere condenses, the heavier H_2^{18}O and ($^1\text{H}^{18}\text{O}$, D_2^{18}O , D_2^{16}O , etc.) molecules will be more likely to condense due to their greater mass (and hence lower vibrational energy that stabilizes intermolecular bonds between water molecules) as compared to H_2^{16}O . To emphasize this point a little further, as an air mass migrates and cools, water molecules in the first precipitation to fall will be enriched in D and ^{18}O (relative their lighter counterparts) as compared to precipitation that falls farther down gradient.

Stable isotopes of the light elements such as hydrogen, carbon, nitrogen, oxygen and sulfur have been applied for longer and with greater frequency than isotopes of heavier elements (e.g. Fe, Cu). Part of the answer lies in the percent difference in mass between the isotopes under consideration (and the compounds they form). The difference in mass between H_2^{18}O and H_2^{16}O is $(20-18)/18 = 11\%$. In other words, a molecule of H_2^{18}O is 11% heavier than a molecule of H_2^{16}O . On an atomic level, ^{18}O is 12.5% heavier than ^{16}O , and the mass of D is twice that of ^1H . Conversely, the difference in mass between the two most abundant stable isotopes of iron, ^{54}Fe and ^{56}Fe , is only 3.7%.

The relatively small difference in the mass of stable isotopes of heavy elements means that fractionation will be much less pronounced and more difficult to detect, so stable isotopes of heavy elements are much less commonly applied to geochemical and biochemical studies than are the lighter isotope stable systems. However, with the development of new analytical techniques (particularly MC-ICP-MS) that offer higher precision and greater sensitivity, stable isotope geochemists are examining with increasing frequency stable isotope fractionation in heavier elements such as Fe, Zn, Cu, Mo, Ti, Hg and many others.

10.2 DELTA (δ) NOTATION

Whether considering H, C, N, O or S, there is always one isotope whose abundance is far greater than the other stable isotope(s) of that element. For example, ^1H comprises 99.985% of all hydrogen; D is only 0.015% of all hydrogen atoms. Similarly, ^{14}N is 99.63% of all nitrogen, and ^{15}N comprises only 0.37% of stable nitrogen atoms (Table 10.2).

To present isotope abundances as simple ratios of heavier to lighter isotope would produce very small values. In the case of hydrogen, the average ratio of D : ^1H is ~ 0.00015 . Thus, in order to produce values that are more convenient for comparison and for analysis of fractionation, a simple equation is applied to stable isotope pairs. The result is known as the **delta notation**, where the ratio of heavy isotope to light isotope in a sample (of water, mineral, gas, leaf tissue, etc.) is compared to the ratio of heavy to light isotope in a standard. Using hydrogen as an example (and remembering that D stands for the heavier isotope ^2H):

$$\delta\text{D}(\text{‰}) = \left(\frac{(\text{D}/\text{H})_{\text{sml}} - (\text{D}/\text{H})_{\text{std}}}{(\text{D}/\text{H})_{\text{std}}} \right) * 1000 \quad (10.1)$$

Using the delta notation, if the ratio of D/H in the sample is lower than the D/H ratio in the standard, the resulting “delta D” value will be negative. Conversely, samples with D/H ratios that are higher than the standard D/H value will produce positive values of δD . To produce values in the approximate range of -100 to $+100$, the ratio in large brackets is multiplied by one thousand. It is common to multiply ratios by one hundred to get per cent values (for interest rates or winning percentages of sports teams); in this case, multiplying by 1000 results in values expressed in “per mil”, i.e. per thousand. Per cent is represented by the symbol %, per mil is represented by the symbol ‰.

The standards used in reference to stable isotope pairs vary depending on the composition of the standard and the isotopes in question. For hydrogen and oxygen, standard mean ocean water (**SMOW**) is commonly used to produce a delta value. In some cases (e.g. where oxygen isotopes are measured in carbonate minerals) the Pee Dee Belemnite (**PDB**) is used as a reference standard (Table 10.2). The delta notation for oxygen referenced to SMOW is:

$$\delta^{18}\text{O}_{\text{SMOW}} = \left(\frac{(^{18}\text{O}/^{16}\text{O})_{\text{smpl}} - (^{18}\text{O}/^{16}\text{O})_{\text{SMOW}}}{(^{18}\text{O}/^{16}\text{O})_{\text{SMOW}}} \right) * 1000 \quad (10.2)$$

Oxygen isotope values referenced to SMOW can be converted to PDB values as follows (Friedman and O'Neil, 1977):

$$\delta^{18}\text{O}_{\text{SMOW}} = (\delta^{18}\text{O}_{\text{PDB}} * 1.03086) + 30.86 \quad (10.3)$$

A certain type of language has developed in the field of stable isotope geochemistry to describe isotope variations and trends. Often, one isotope is referred to as being enriched (or depleted) with reference to the other in the stable isotope pair. For example, rain water is enriched in ^{16}O relative to the water body from which it evaporated. When referring to delta values, stable isotope researchers frequently will indicate that a particular sample is more (or less) positive as compared to a different sample. One might say that residual water left behind by evaporation is 10 per mil more positive than the vapor that evaporated from the water body. In this case, a value such as -15‰ (the remnant surface water body) is 10‰ (per mil) more positive than the -25‰ water vapor that has evaporated from it into the atmosphere.

Table 10.2 Stable isotopes, abundances (relative %), differences in mass (relative %), delta notations and ratios of standard values of a representative (but not complete) list of the stable isotopes applied in environmental geochemistry.

Element	Isotopes	Abundance	Mass diff.	δ notation	Standard with Ratio
Hydrogen	^aH	99.985	99.8	δD	SMOW ^a : $\text{D}/\text{H} = 1.557 \times 10^{-4}$
	^bH (D)	0.015			
Boron	^{10}B	19.9	10.0	$\delta^{11}\text{B}$	SRM 951 ^b : $^{11}\text{B}/^{10}\text{B} = 4.0456$
	^{11}B	80.1			
Carbon	^{12}C	98.90	8.3	$\delta^{13}\text{C}$	PDB ^c : $^{13}\text{C}/^{12}\text{C} = 1.122 \times 10^{-2}$
	^{13}C	1.10			
Nitrogen	^{14}N	99.63	7.1	$\delta^{15}\text{N}$	Atmospheric N_2 : $^{15}\text{N}/^{14}\text{N} = 3.613 \times 10^{-3}$
	^{15}N	0.37			
Oxygen	^{16}O	99.76	12.5	$\delta^{18}\text{O}$	SMOW: $^{18}\text{O}/^{16}\text{O} = 2.0052 \times 10^{-3}$ PDB ³ : $^{18}\text{O}/^{16}\text{O} = 3.76 \times 10^{-4}$
	^{17}O	0.038			
	^{18}O	0.20			
Magnesium	^{24}Mg	78.99	8.3	$\delta^{26}\text{Mg}$	SRM 980 ^d : $^{26}\text{Mg}/^{24}\text{Mg} = 1.392 \times 10^{-1}$ DSM3 ^d :
	^{25}Mg	10.00			
	^{26}Mg	11.01			

(Continued)

Table 10.2 (Continued)

Element	Isotopes	Abundance	Mass diff.	δ notation	Standard with Ratio
Sulfur	³² S	95.02	6.2	δ ³⁴ S	CDM ^e : ³⁴ S/ ³² S = 4.43 × 10 ⁻²
	³³ S	0.750	(³⁴ S vs. ³² S)	(ratio of ³⁴ S/ ³² S)	
	³⁴ S	4.21			
	³⁵ S	0.02			
Iron	⁵⁴ Fe	5.58	3.7	δ ^{56/54} Fe	IRMM-014 ^f : ⁵⁴ Fe/ ⁵⁶ Fe = 6.37 × 10 ⁻²
	⁵⁶ Fe	91.950	(⁵⁶ Fe vs. ⁵⁴ Fe)	(ratio of ⁵⁶ Fe/ ⁵⁴ Fe)	
	⁵⁷ Fe	2.18			
	⁵⁸ Fe	0.30			
Copper	⁶³ Cu	69.15	3.2	δ ⁶⁵ Cu	ERM-AE637 ^g : ⁶⁵ Cu/ ⁶³ Cu = 4.456 × 10 ⁻¹
	⁶⁵ Cu	30.85	(⁶⁵ Cu vs. ⁶³ Cu)	(ratio of ⁶⁵ Cu/ ⁶³ Cu)	
Mercury	¹⁹⁶ Hg	0.15	1.98	δ ^{202/198} Hg	NIST-3133 ^h : ²⁰² Hg/ ¹⁹⁸ Hg = 2.9689
	¹⁹⁸ Hg	9.97	(²⁰² Hg vs. ¹⁹⁸ Hg)	(ratio of ²⁰² Hg/ ¹⁹⁸ Hg)	
	¹⁹⁹ Hg	16.87			
	²⁰⁰ Hg	23.10			
	²⁰¹ Hg	13.18			
	²⁰² Hg	29.86			
	²⁰⁴ Hg	6.87			

^a Standard mean ocean water (SMOW) is the standard for hydrogen isotopes. It is also commonly used as reference for oxygen, although oxygen in PDB is sometimes used (especially in paleoclimate studies of oxygen in carbonates).

^b The ratio of $^{11}\text{B}/^{10}\text{B}$ prepared from Searles Lake borax (SRM 951 from the National Bureau of Standards, Palmer and Slack, 1989), an evaporite deposit, is commonly used as the boron isotope standard (Hoefs, 2009).

^c PDB is a fossil from the Pee Dee Fm in South Carolina, USA used as the $^{13}\text{C}/^{12}\text{C}$ standard, and also sometimes for $^{18}\text{O}/^{16}\text{O}$.

^d SRM 980 is heterogeneous (limited to precision of $\pm 0.2 \text{ ‰}$) so DSM3 has also been adopted as a Mg isotope standard.

^e Troilite (FeS) from the Canyon Diablo Meteorite (CDM) is commonly used as a reference value for the $^{34}\text{S}/^{32}\text{S}$ ratio.

^f IRMM-014 is metallic iron provided by the EU Institute for Reference Materials and Measurements

^g Limited supply of SRM 976 led to development of new Cu isotope standards. Refer to Moeller et al. (2012) for information.

^h Details of Hg isotopic method development and standards are presented by Blum (2011).

10.3 FRACTIONATION: VIBRATIONAL FREQUENCIES, TEMPERATURE DEPENDENCE

Fractionation due to the different vibrational energies associated with light and heavy isotopes of a given element occurs when covalently

bonded compounds form – the heavier isotope of a given element will form a more stable covalent bond than will the lighter isotope. However, this type of fractionation only applies to bonds that have a predominantly covalent character. Ionic and metallic bonds do not involve sharing of electrons, so there is little or no isotopic

fractionation when ionically bonded compounds like salts crystallize from solution.

10.3.1 Stable isotopes and chemical bond strength

To explore this topic in more detail, we will briefly enter the world of quantum theory, which states that the energy of a diatomic oscillator (e.g. a gaseous molecule comprised of two atoms) can only occur at discrete energy levels, and the possible energy levels depend on the mass of the atoms that comprise the molecule. According to quantum theory, the atoms in the diatomic oscillator will vibrate at a ground frequency ν_0 (for most gases, this applies to systems at room temperature), and the energy of the molecule is:

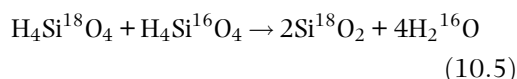
$$E = \frac{1}{2} h\nu \quad (10.4)$$

where h is Planck's constant and ν is the vibrational frequency of an atom in a molecule.

In other words, one of the important controls on the vibrational frequency (and hence the strength) of the covalent bonds in the molecule is masses of the atoms in the molecule – the vibrational energy will be lower for a bond involving the heavier isotope of an element. The energy of the molecule will be minimized when the heavier isotope occupies an atomic site, and what this means is that **bonds involving heavier isotopes will be stronger**. For systems at equilibrium, the heavier isotope (^{18}O , ^{15}N , ^{34}S , etc.) will be more likely to form a stable chemical bond within a molecule than will its lighter counterpart (^{16}O , ^{14}N , ^{32}S , etc.), and conversely, because bonds involving lighter isotopes are weaker, they have a greater probability of being broken than do their heavier counterparts. Fractionation is also controlled by the oxidation state of elements involved in bonding because the heavier isotope preferentially forms bonds with elements of higher oxidation state (Kendall and Doctor, 2003).

Accordingly, when quartz begins to crystallize from a solution that has reached saturation,

the first molecules of quartz to form will be enriched in ^{18}O relative to ^{16}O and the solution from which the quartz is being crystallized will become depleted in ^{18}O . The example below uses a simple chemical equation that depicts reaction of silicic acid to quartz. In one idealized H_4SiO_4 molecule, all oxygen atoms are the heavier ^{18}O isotope and in the other, all atoms of oxygen are ^{16}O :



In this idealized reaction, atoms of ^{18}O preferentially bond to Si^{+4} , and thus are fractionated into quartz. The ^{16}O atoms are left to form bonds with H^+ and are fractionated into water. We can say that the quartz is enriched in oxygen-18 relative to the water, and the quartz will have a higher (more positive) $\delta^{18}\text{O}$ value than will the water.

10.3.2 Temperature-dependent stable-isotope fractionation

In addition to the differences in mass among isotopes of a given element, temperature also controls stable isotope fractionation. This is because the differences in vibrational energy are most pronounced at low temperatures – the higher the temperature, the greater the vibrational energy of both the light and heavy isotope and the result is that higher temperatures produce smaller differences in vibrational energy between heavy and light isotopes of a given element. At these higher temperatures the small differences in vibrational energy produce only small fractionation effects. Based on this, we would predict that quartz, which crystallizes at low temperatures (e.g. 200°C), will exhibit greater fractionation of ^{16}O and ^{18}O than will quartz crystallized at higher temperature (e.g. 700°C) (Clayton et al., 1972).

Early theoretical studies by Urey (1947) and Bigeleisen and Mayer (1947) indicate that the temperature dependence of fractionation is

proportional to $1/T^2$, where T is in Kelvin, and where fractionation is expressed as $\delta^{18}\text{O}_{\text{phaseA}} - \delta^{18}\text{O}_{\text{phaseB}}$, using oxygen as the example. According to Savin and Lee (1988), fractionation of oxygen isotopes between most minerals and water is nearly linear when plotted as $\delta^{18}\text{O}_{\text{phaseA}} - \delta^{18}\text{O}_{\text{phaseB}}$ versus $1/T^2$. Although far less common, fractionation can be proportional to $1/T$ in some lower-temperature systems (e.g. with respect to hydrogen isotope fractionation at temperatures $< 0^\circ\text{C}$), yet regardless of the magnitude, higher temperatures produce a larger number in the denominator of $1/T^2$ (or $1/T$) and fractionation is inversely proportional to temperature.

Two illustrative examples of fractionation between minerals and water are of $\text{CaCO}_3\text{--H}_2\text{O}$ and $\text{SiO}_2\text{--H}_2\text{O}$. Figure 10.2 indicates the strong temperature dependence of oxygen isotope fractionation between quartz and water over a temperature range of $\sim 0^\circ\text{C}$ to 400°C . As is the case with many graphs in stable isotope geochemistry, the units of temperature presented on the x-axis are $10^6/T^2$, where temperature is in Kelvin. So for reference, 300 K (27°C) is equal to 11.1 on the x-axis, and 700 K (423°C) corresponds to an x-axis value of 2.0. This graph indicates that when quartz crystallizes from solution at 5 to 25°C – temperatures that are typical of many Earth surface systems (e.g. soils, surface waters and groundwaters) – the difference in $\delta^{18}\text{O}$ between quartz and water will be approximately 40 to 50‰. At 400 K (127°C), the difference in $\delta^{18}\text{O}$ between quartz and water will be approximately 15‰, and at temperatures greater than 700 K (427°C), there is essentially no fractionation between quartz and water.

Similar temperature-dependent fractionation effects occur when calcium carbonate (calcite or its polymorph aragonite) crystallize from solution. As is the case with quartz, the carbonate minerals (with carbon effectively occurring in a high oxidation state, C^{+4}) preferentially form bonds with ^{18}O , so the carbonates are enriched in ^{18}O relative to the water from which they crystallize, and as is

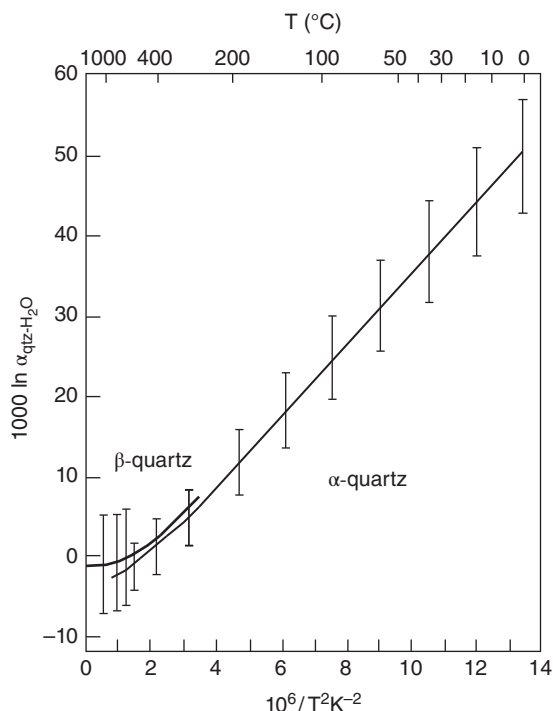


Fig. 10.2 Calculated values showing the relationship between temperature and fractionation of ^{16}O and ^{18}O in quartz (in equilibrium with water). Fractionation clearly decreases with increasing temperature. Modified from Kawabe (1978). The term $1000 \ln \alpha_{\text{qtz-H}_2\text{O}}$ is effectively equal to $\delta^{18}\text{O}_{\text{qtz}} - \delta^{18}\text{O}_{\text{H}_2\text{O}}$, where $\alpha_{\text{qtz-H}_2\text{O}} = (^{18}\text{O}/^{16}\text{O})_{\text{qtz}} / (^{18}\text{O}/^{16}\text{O})_{\text{H}_2\text{O}}$.

the case with quartz- H_2O , this effect is greatest at low temperatures.

Horibe and Oba (1972) quantified the temperature effect on isotopic fractionation between carbonates and water (Fig. 10.3). Their data are from shells of two pelecypods from Mutsu Bay, Japan, and indicate that for calcite that precipitates during formation of *Patinopecten yessoensis* (the Ezo giant scallop or Yesso scallop), the empirical equation that describes the temperature dependence of oxygen isotope fractionation (Horibe and Oba, 1972) is:

$$T(^{\circ}\text{C}) = 17.04 - 4.34 * (\delta_{\text{c}} - \delta_{\text{w}}) + 0.16 * (\delta_{\text{c}} - \delta_{\text{w}})^2 \quad (10.6)$$

In this equation, δ_c represents the $\delta^{18}\text{O}$ of CO_2 derived from the calcite by reaction with strong phosphoric acid at 25 °C, and δ_w is the $\delta^{18}\text{O}$ of CO_2 in equilibrium with water at 25 °C. Because the oxygen occurs in carbonates, the values are given relative to PDB.

The equation for aragonite (a polymorph of CaCO_3) found in shells of *Anadara broughtoni* is slightly different than the calcite equation (using PDB as the standard):

$$T(^{\circ}\text{C}) = 13.85 - 4.54 * (\delta_c - \delta_w) + 0.04 * (\delta_c - \delta_w)^2 \quad (10.7)$$

These two equations demonstrate two important principles: (1) the temperature effect on isotopic fractionation; and (2) the effect of mineral structure on isotopic fractionation, where calcite and aragonite both are CaCO_3 but possess

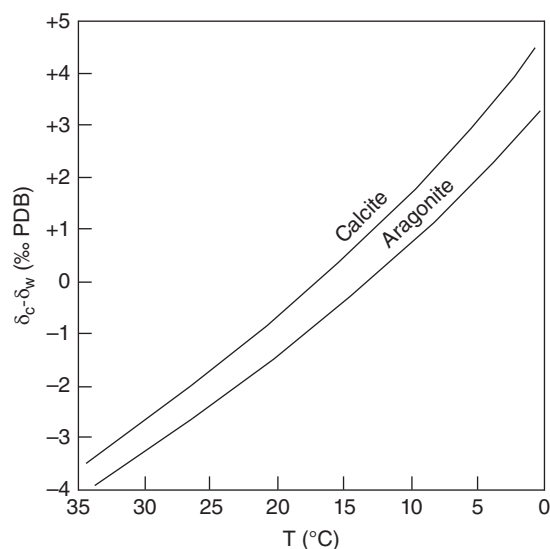


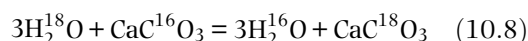
Fig. 10.3 Influence of temperature on oxygen isotope fractionation between calcium carbonates and water. Note that fractionation is greatest at lowest temperature, and that the two polymorphs of CaCO_3 exert slightly different controls on fractionation. Based on data for marine carbonates from Horibe and Oba (1972).

different crystal structures. These equations also serve to illustrate the applicability of oxygen isotopes to paleoclimate analysis.

These examples hold true for carbonate–water **systems at equilibrium**. While this state is often not achieved with respect to isotopic exchange (some examples are presented below), isotopic equilibrium *can* be achieved in closed, well-mixed systems, where the equilibrium represents the condition where the forward and backward rates of isotope exchange are equal among minerals or other phases (e.g. water, organic matter or gases). Considering that oxygen isotopes in marine water are generally well-mixed, stable isotopes geochemists consider that crystallization of carbonates in the presence of marine water is an equilibrium process.

10.3.3 Equilibrium and non-equilibrium isotope fractionation

Examining the isotopic reaction involving water and calcium carbonate as an exchange reaction:



we observe that ^{16}O and ^{18}O are exchanged between the molecules of water and calcium carbonate, and this idealized reaction represents complete exchange of 3 atoms of ^{16}O for 3 atoms of ^{18}O between the two phases. The heavier isotope is attracted to the phase where it bonds with an element of high oxidation number (C^{+4} in CaCO_3) and the light isotope is attracted to the phase where it bonds with an element of low oxidation number (H^+ in H_2O).

The equilibrium constant for this reaction can be written as

$$K_{\text{eq}} = \frac{\alpha_{\text{CaC}^{18}\text{O}_3}}{\alpha_{\text{CaC}^{16}\text{O}_3}} \times \frac{\alpha_{\text{H}_2^{16}\text{O}}^3}{\alpha_{\text{H}_2^{18}\text{O}}^3} = \alpha \quad (10.9)$$

The activities represent molar concentrations of each phase. If there is no fractionation of oxygen isotopes between water and calcium carbonate, the value of the K_{eq} will be exactly one (α is the fractionation factor, which is a term used for the K_{eq} related to stable isotope fractionation). However, for water and calcite at equilibrium at 25 °C, the K_{eq} is equal to 1.0286 (O'Neil and Epstein, 1966), indicating that the heavier isotope of oxygen is preferentially partitioned into calcite relative to water (and water is enriched in ^{16}O).

Fractionation of ^{18}O into calcite over water is predictable given the high effective oxidation state of carbon (C^{+4}) relative to hydrogen (H^{+}) and the attraction of high oxidation state ions for the heavier isotope of an element, and this is what has been observed in various studies, including those of Horibe and Oba (1972) and Friedman and O'Neil (1977). The term “effective” is used here and above because the bonds are dominantly covalent rather than ionic. The K_{eq} of 1.0286 produces a $\delta^{18}\text{O}$ value of 28.6‰ (assuming a $\delta^{18}\text{O}$ value of zero, a reasonable assumption given that the $\delta^{18}\text{O}$ value of SMOW is equal to zero). Given what we know about temperature effects on isotope fractionation (i.e. less fractionation with increasing temperature), the K_{eq} will approach 1.00 as temperature increases.

The examples above illustrate equilibrium fractionation of stable isotopes, a condition that is most commonly achieved for inorganic reactions; however, for many reactions in biological systems and in open systems, fractionation is dominated by **non-equilibrium effects**. In these cases, **kinetic factors** typically control fractionation.

Important factors that contribute to non-equilibrium, kinetic isotope effects are as follows:

- Molecules with the lighter isotope will react more rapidly than molecules with the heavier isotope – this is because the chemical bond involving the lighter isotope is less stable and hence easier to break, meaning that it requires less energy to break down molecules with the lighter isotope.

This effect is especially pronounced in biological systems where organisms expend less energy when they carry out reactions involving molecules with the lighter isotope. A good example of this is photosynthesis, where plants preferentially react with $^{12}\text{CO}_2$ molecules that are more easily decomposed than are the more-strongly bonded $^{13}\text{CO}_2$ molecules. As a consequence, photosynthetically produced organic matter is enriched in ^{12}C relative to ^{13}C . One result of this effect is that, in marine environments, phytoplankton have $\delta^{13}\text{C}$ values that are 20‰ lower than the dissolved CO_2 in ocean water (Drever, 1997).

- The lower vibrational frequency of the heavier isotope means that it tends to diffuse more slowly than its lighter counterpart and thus is less likely to come into contact with potential reactants. Furthermore, because of differences in diffusion rates, the lighter isotope can be separated from the heavier isotope (particularly in open systems) and thus equilibrium may not be established.

The Arrhenius equation, first encountered during the discussion of kinetics of chemical reactions in Chapter 1 (equation 1.76), helps to explain the kinetic isotope effect.

$$k = A * e^{(-E_a/RT)}$$

The activation energy (E_a) of ^{12}C –O bonds is lower than the E_a for ^{13}C –O bonds (i.e. the ^{13}C –C bond is stronger than the ^{12}C –O bond), and given the negative value of E_a in the exponent of the Arrhenius equation, increasing activation energy will decrease reaction rate (k). Furthermore, the lower diffusion rate of the heavier isotope will be represented in this equation as a lower value of A (the pre-exponential factor), which will also result in lower reaction rates for the heavier isotope.

A good example of non-equilibrium isotope fractionation is described by the model of **Rayleigh fractionation**. First described by Lord Rayleigh in 1896, Rayleigh fractionation is the kinetically controlled isotope fractionation that takes place during unidirectional reactions

between a source reservoir and a sink where, during the unidirectional reaction, the source reservoir decreases in magnitude – a good example of this is evaporation of water from a pond that will decrease the size of the pond (in this case the pond would be considered the “reservoir”). Because these reactions are effectively distillation reactions, this process is also referred to as **Rayleigh distillation**.

Some of the most well-known examples of Rayleigh fractionation involve the physical changes of state between liquid water and water vapor. The isotope fractionation that takes place during $\text{H}_2\text{O}_g \rightarrow \text{H}_2\text{O}_l$ or $\text{H}_2\text{O}_l \rightarrow \text{H}_2\text{O}_g$ is often kinetically controlled (rather than controlled by equilibrium considerations) because *rates* of evaporation or condensation exert a greater control on the fractionation than do equilibrium controls, and also because the product (e.g. vapor produced by evaporation) is continually removed from contact with the source reservoir; effectively, this is open-system behavior. In this example, kinetic effects are enhanced by many factors including surface temperatures (greater temperature produces greater evaporation rates), salinity of the reservoir (greater salinity can impede evaporation), wind speeds (winds produce shear at the water surface and enhance evaporation), and finally and most importantly, humidity – low humidity speeds evaporation rates and produces non-equilibrium conditions; high humidity allows exchange of vapor with the liquid phase, and as humidity approaches 100%, the vapor–liquid system approaches equilibrium. Likewise, H_2O_g in the atmosphere will typically condense to H_2O_l at rates controlled by similar non-equilibrium conditions that include contributions from factors such as humidity, temperature and presence or absence of condensation nuclei.

Considering the condensation of water from an air mass as it migrates through the atmosphere, the first molecules of water to condense from the vapor phase will be enriched in the heavy isotopes ^{18}O and D. As these droplets grow in size

and eventually fall in the form of precipitation, they are removed from contact with the water vapor reservoir. As this process continues, the vapor reservoir will become (1) progressively depleted in ^{18}O and D (and enriched in ^{16}O and ^1H), and (2) smaller in size by virtue of rainout. This then represents a classic case of Rayleigh distillation, which is quantified by the following equation (Broecker and Oversby, 1971):

$$R = R_0 * f^{(\alpha - 1)} \quad (10.10)$$

where R_0 is the initial ratio of $^{18}\text{O}/^{16}\text{O}$ (or D/H) in the reservoir (water vapor in this case), R is the ratio of $^{18}\text{O}/^{16}\text{O}$ (or D/H) in the vapor remaining in the reservoir when the fraction of vapor remaining is given by the term f (where f ranges from 0.0 to 1.0). The α term is the fractionation factor that, for a system of water vapor and liquid water, can be represented as:

$$\alpha = R_{\text{liquid}}/R_{\text{vapor}} \quad (10.11)$$

In this equation R is the ratio of $^{18}\text{O}/^{16}\text{O}$ in the liquid and vapor states, respectively. For a system of liquid water and water vapor under equilibrium conditions at 25 °C, the value of α is 1.0092. What this means is that, under these conditions, the ratio of $^{18}\text{O}/^{16}\text{O}$ in the liquid phase will be 1.0092 times greater than the $^{18}\text{O}/^{16}\text{O}$ ratio in the water vapor phase. This should intuitively make sense given that the heavier isotope will be fractionated into the liquid when H_2O undergoes condensation.

10.4 $\delta^{18}\text{O}$ AND δD

Analysis of oxygen isotopes is frequently used to examine processes within the hydrologic cycle. The following example applies the Rayleigh distillation equation to a reservoir of water vapor in the atmosphere with an initial $\delta^{18}\text{O}$ value of -10‰ , a

typical value for evaporated water vapor from the ocean. If the equation is re-arranged for the determination of $\delta^{18}\text{O}$, it is possible to determine the $^{18}\text{O}/^{16}\text{O}$ ratio of water with a $\delta^{18}\text{O}$ value of -10‰ .

$$\begin{aligned} ^{18}\text{O}/^{16}\text{O}_{\text{vapor}} &= ^{18}\text{O}/^{16}\text{O}_{\text{vapor}} \\ &+ \left[\left(^{18}\text{O}/^{16}\text{O}_{\text{SMOW}} * \delta^{18}\text{O}_{\text{vapor}} \right) \right] / 1000 \end{aligned} \quad (10.12)$$

This produces an $^{18}\text{O}/^{16}\text{O}$ ratio of 0.0019851 (for comparison, the $^{18}\text{O}/^{16}\text{O}$ ratio of SMOW is 0.002052). When this initial $^{18}\text{O}/^{16}\text{O}$ ratio is the R_0 value in the Rayleigh equation over a range of fractions of water vapor remaining (f), the result is the curve depicted in Fig. 10.4.

Also plotted on Fig. 10.4 is the oxygen isotopic composition of the liquid water that condenses from the vapor reservoir. The liquid water $\delta^{18}\text{O}$ value can be determined by the following equation (Faure, 1986):

$$\delta^{18}\text{O}_{\text{liquid}} = [\alpha * (\delta^{18}\text{O}_{\text{vapor}} + 1000)] - 1000 \quad (10.13)$$

At this point there are a few important observations to reinforce regarding these curves and water in the hydrologic cycle in general:

- 1 The initial liquid that condenses from the vapor reservoir is enriched in ^{18}O relative to the ^{18}O content of the initial vapor.
- 2 The initial liquid that condenses from vapor falls as precipitation and thus is removed from contact with the vapor, preventing the system from reaching equilibrium.
- 3 At any given time the vapor that condenses is enriched in ^{18}O relative to the ^{18}O of the remaining vapor.
- 4 As water vapor in an air mass migrates away from its source, the H_2O that condenses and falls as precipitation becomes progressively lighter.
- 5 Precipitation is always at least slightly depleted in ^{18}O relative to sea water (i.e. $\delta^{18}\text{O}$ of

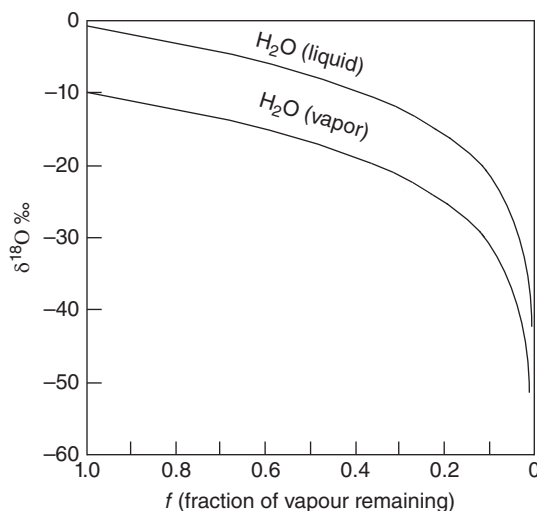


Fig. 10.4 Fractionation of oxygen isotopes during condensation at 25 °C as predicted by the Rayleigh distillation model. The initial $\delta^{18}\text{O}$ value of the vapor is -10‰ and the value of the first liquid water to form by condensation is -0.9‰ . The condensed liquid water is then removed from contact with the vapor (i.e. it falls as rain), so equilibrium cannot be established. The vapor becomes progressively lighter (more negative) with the removal of isotopically heavier liquid, and subsequent condensation of new liquid from continuously lighter vapor forms the trends illustrated above. These trends occur in nature when air masses migrate from the tropics towards higher latitudes (see Figures 7.4 through 7.8).

precipitation is always lower than $\delta^{18}\text{O}$ of sea water whose value is 0‰).

- 6 Fresh water (e.g. in lakes, rivers and groundwaters) is almost always depleted in ^{18}O and D relative to sea water, i.e. fresh water has $\delta^{18}\text{O}$ values less than zero.

The fractionation of hydrogen isotopes (H and D) is very similar to the oxygen isotopes so the first water to condense from a parcel of air will be enriched in D relative to H and the shrinking water vapor reservoir will become progressively enriched in H. A bivariate plot of $\delta^{18}\text{O}$ against δD in meteoric water results in a straight line known as the **meteoric water line** (Fig. 10.5),

“Meteoric water” refers to H_2O that falls from the atmosphere; it does not imply an outer-space origin of the water. The term meteoric is derived from the same Greek root that means “from the air” or “lofty” and gives us the term “meteorology” for weather, so it is easy to understand how the same root is applied to rain or snow as well as rocky and metallic objects that fall to Earth from outer space. Meteoric water is an important term to remember because it is used widely in geochemistry, biochemistry, geomorphology and various other fields of environmental science. The composition of meteoric water is typically quite distinct from other waters in the hydrologic cycle; for example, the relatively dilute character and slightly acidic character of meteoric water stand in stark contrast to the highly alkaline, high-pH brines derived from evaporation-dominated lakes and saline waters from sedimentary basins.

the equation of which was quantified by the pioneering isotope geochemist Harmon Craig in the early 1960s:

$$\delta\text{D} = 8 * \delta^{18}\text{O} + 10 \quad (10.14)$$

Precipitation in the tropics has $\delta^{18}\text{O}$ and δD values that are very similar to standard mean ocean water, and with increasing latitude, $\delta^{18}\text{O}$ and δD values of precipitation become progressively more negative. The only meteoric waters that do not plot on or close to the line in Fig. 10.5 are surface waters that have been appreciably affected by evaporation – they veer off the meteoric water line with lower slope. As you might imagine, evaporation of lakes in arid regions causes progressive fractionation of ^{18}O and D into the surface water, but this alone does not explain deviation from the meteoric water line. The decreased slope of the

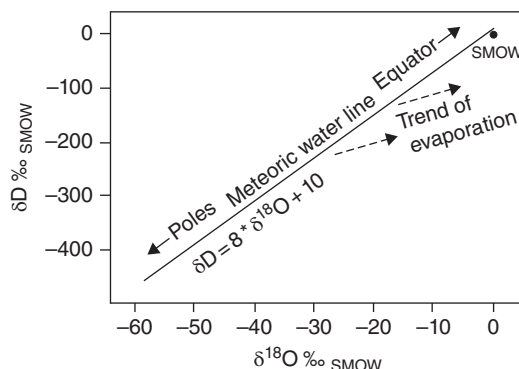
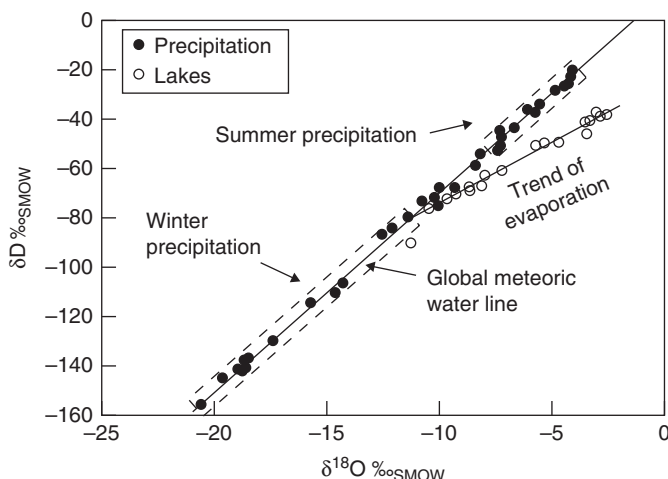


Fig. 10.5 Linear correlation of oxygen and hydrogen isotopes in meteoric water. The lighter (more negative) values towards the poles (high latitudes) is explained by Rayleigh distillation. Note the position of ocean water (SMOW) and also that evaporation has a more pronounced effect on oxygen isotope compositions than it does on hydrogen isotopes. The slope of the meteoric water line is 10; the slope of the evaporated water line is ~ 5 . (Modified from Craig 1961.)

evaporation-dominated water indicates that the surface water becomes enriched in ^{18}O at a greater rate than it does for D, a phenomenon that is due to the greater difference in mass between H_2^{18}O and H_2^{16}O (11%) as compared to D_2O and H_2O (5.6%) – H_2^{18}O will be more preferentially fractionated into the surface water reservoir than will D_2O , producing trajectories with lower slopes, like those illustrated in Fig. 10.5.

An example of local-scale variability in the isotopic composition of meteoric water is provided for northern Wisconsin, USA (Fig. 10.6). Note that while the variation in $\delta^{18}\text{O}$ and δD values are much smaller than they are for global meteoric water (Fig. 10.5), precipitation water still plots on the meteoric water line with a slope of 10, where winter precipitation contains isotopically lighter H_2O as compared to summer precipitation (the temperature control on $\delta^{18}\text{O}$ and δD is presented below). Also notable is that meteoric water derived from evaporation of surface water

Fig. 10.6 Oxygen and hydrogen isotopic values of precipitation and lake water from northern Wisconsin, USA. Note that winter precipitation is isotopically lighter than summer precipitation, and that precipitation from this region falls on the global meteoric water line. Evaporation of water from lakes causes their isotopic compositions to change in a manner similar to the trend shown in Fig. 7.4. (Data from Kendall et al. 1995.)



bodies veers off the meteoric water line with a slope of 5.2 (Kendall et al., 1995).

We can now begin to explore some of the ways that oxygen and hydrogen isotopes are applied to understanding processes within the hydrologic cycle, particularly regarding the origins and behaviors of water at various stages of the hydrologic cycle. One way to begin to visualize isotopic variation in the hydrologic cycle is to examine the fractionation of meteoric water as a function of distance from the vapor source. Figure 10.7 depicts a schematic example of orographic precipitation, where an air mass cools as it rises over hills or mountains. The uplifting of the air causes **adiabatic cooling** (which results from expansion of air as it rises into the lower air pressure associated with high levels in the atmosphere). The cooling of air vapor results in condensation which is depicted in Fig. 10.7 as the precipitation that is spatially associated with the foothills and mountains.

The first water to condense, as predicted by thermodynamic considerations and by the kinetic controls described by Rayleigh distillation, is enriched in the heavy isotopes relative to precipitation that falls further downgradient (Fig. 10.7). Accordingly, the precipitation that falls on the lee (downwind) side of the mountain range is noticeably depleted in ^{18}O and D compared to the precipitation that

falls on the upwind side of the range. Progressive Rayleigh distillation will result in progressively lighter (isotopically) water with increasing distance from the source. An example of spatial variability in $\delta^{18}\text{O}$ on a broader scale is depicted in Fig. 10.8, which illustrates changes in $\delta^{18}\text{O}$ that occur as a function of latitude (and longitude to a lesser extent) in the western hemisphere. (Also, although not presented here, the same pattern occurs with hydrogen isotope values, which likewise become progressively lighter downgradient.)

The decrease in $\delta^{18}\text{O}$ and δD values in meteoric water from the equator toward the poles has been termed the **latitude effect** (Faure, 1998), a phenomenon that is illustrated spatially in Fig. 10.8 and is primarily controlled by the following factors:

- 1 removal of ^{18}O and D from air mass by condensation of liquid water as the air mass migrates pole-ward, a process driven Rayleigh distillation and increasing fractionation factors as the air mass cools;

- 2 evaporation of water at higher latitudes occurs at lower temperature, so the fractionation factor will be greater and the vapor produced at higher latitudes will be relatively depleted in ^{18}O and D; and
- 3 transpiration of water from plants preferentially fractionates ^{18}O and D into compounds in plant tissue (the heavier isotopes will form more

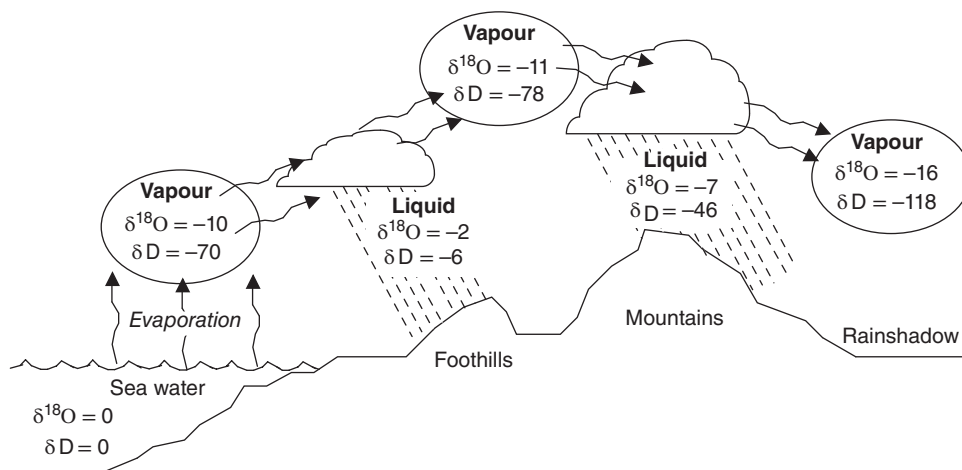


Fig. 10.7 Schematic diagram representing the fractionation of oxygen and hydrogen isotopes in water vapor and precipitation as an air mass first passes over ocean water (and gains water vapor by evaporation), then continues downwind over hills and mountains that cause adiabatic cooling, condensation and precipitation. Note that as rain falls from the air mass, the remaining water vapor becomes progressively lighter. Subsequently, water that condenses and falls as precipitation also becomes progressively lighter. The $\delta^{18}\text{O}$ and δD values are per mil relative to SMOW.

stable bonds), thus transmitting water vapor depleted in ^{18}O and D to the atmosphere.

In addition to the progressive decrease in $\delta^{18}\text{O}$ downgradient (northward and poleward) of the tropical moisture source, there is also a pronounced seasonality in the $\delta^{18}\text{O}$ pattern, which brings us to a discussion of the temperature dependence of $\delta^{18}\text{O}$ in precipitation.

The temperature control on the seasonal variation of the $\delta^{18}\text{O}$ values is related to the fractionation of oxygen (or hydrogen) isotopes when surface water evaporates to produce water vapor. Warmer summer air (July) results in precipitation that is enriched in the heavier isotopes ^{18}O and D relative to winter (January) precipitation. Part of the explanation for this is the same as the temperature effect on the isotopic composition of minerals – higher temperatures (of sea water in this case) imparts smaller differences in vibrational energy between ^{18}O and ^{16}O , or D and H, and this produces lower fractionation factors for isotopes of oxygen and hydrogen – the result is a linear correlation between temperature and $\delta^{18}\text{O}$ (Fig. 10.9).

The linear correlation originally discovered by Dansgaard (1964) yields the equation:

$$\delta^{18}\text{O}_{\text{MAP}} = 0.695 * T_{\text{MAT}} - 13.6 \quad (10.15)$$

where $\delta^{18}\text{O}_{\text{MAP}}$ is mean annual $\delta^{18}\text{O}$ of precipitation and temperature (T_{MAT}) is mean annual surface temperature in $^{\circ}\text{C}$. An increase of 10°C correlates to a $\delta^{18}\text{O}$ increase of approximately 7‰. More recent work by Rozanski et al. (1993) indicates that there are regional trends embedded within the global trend suggested by Dansgaard's data, including values for the Antarctic Peninsula and interior Canada (not shown) that deviate from Dansgaard's best-fit line. Some of the local effects that contribute to non-linearity show up in Fig. 10.8.

Similar effects are apparent with increasing elevation and increased distance from the equator (both of which correspond to lower temperature), where $\delta^{18}\text{O}$ and δD of evaporated water become lighter as temperature drops. Furthermore, because precipitation becomes progressively enriched in light oxygen as it moves toward the cold polar regions, polar

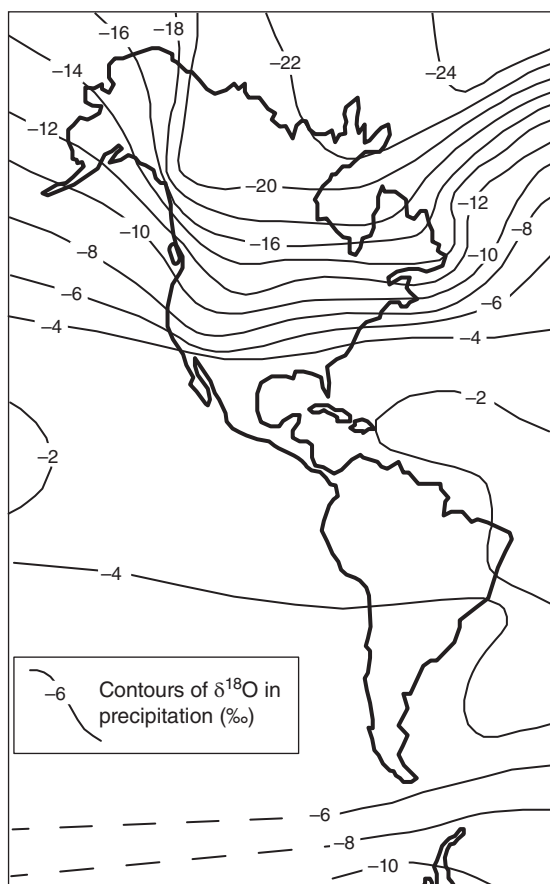


Fig. 10.8 Variation in $\delta^{18}\text{O}$ (‰) in precipitation in the western hemisphere. The general shift towards lighter (lower $\delta^{18}\text{O}$) values poleward is a result of Rayleigh distillation. (Data from Yurtsever and Gat 1981.)

ice caps are a reservoir of ^{16}O -enriched water (compared to sea water, which is relatively enriched in $\delta^{18}\text{O}$ compared to the ice caps).

10.4.1 Paleotemperature analysis using oxygen and hydrogen isotopes

The temperature dependence of $\delta^{18}\text{O}$ in meteoric water (Fig. 10.9) also varies as a function of temperature over longer time scales and is one of the most important means of analyzing

paleotemperature. Given what we know about the relationship of temperature and $\delta^{18}\text{O}$ (and δD), we would predict that colder times in the geological past should record precipitation with lower $\delta^{18}\text{O}$ (and δD) values and vice versa. One of the best records of precipitation that extends well back into the Pleistocene are the thick ice caps on Antarctica and Greenland, where meteoric water is deposited as snow that eventually transforms to ice. Glacial ice from the polar regions contains periodic layers of volcanic ash that can be radiometrically dated and geochemically fingerprinted, permitting determination of depositional age.

Two of the longest glacier paleotemperature records are from ice cores retrieved from thick ice caps in Antarctica (Petit et al., 1999; Augustin et al., 2004). Ice core data presented in Chapter 7 (Fig. 7.3a) indicate that $\delta^{18}\text{O}_{\text{atm}}$ and δD of H_2O that fell as precipitation (now preserved in layers of glacial ice) vary cyclically with global temperature changes that correspond with variations in Earth's orbit. Measurement of $\delta^{18}\text{O}$ and δD in ice cores has proven to be a very important tool in examining natural variations in temperature (globally and regionally) as well as the relationship between temperature and CO_2 prior to anthropogenic climate change.

10.4.2 Oxygen and hydrogen isotopes as tracers in the hydrologic cycle

In addition to their utility in tracing global paleotemperature, $\delta^{18}\text{O}$ and δD are also useful tracers of many other processes within the hydrological cycle. Groundwater and soil water that plot on the meteoric water line with $\delta^{18}\text{O}$ and δD values that are similar to modern precipitation are likely young waters with relatively low residence times – the assumption here is that waters that recently infiltrated the aquifer (and have not appreciably reacted with minerals or organic matter) will retain their meteoric signature. Conversely,

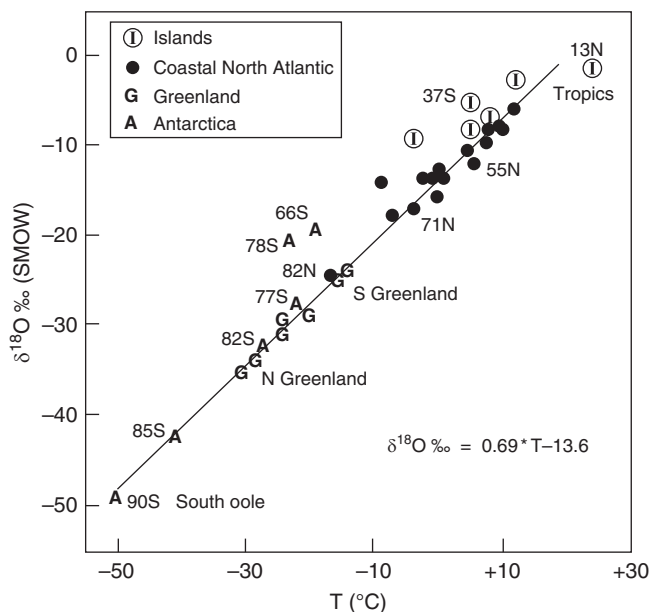


Fig. 10.9 Linear correlation between mean annual air temperature and $\delta^{18}\text{O}$ of precipitation as a function of latitude (Modified from Dansgaard, 1964).

groundwaters with $\delta^{18}\text{O}$ and δD values appreciably lighter than modern precipitation may represent older water that infiltrated the aquifer during the Pleistocene (> 12 ka) when meteoric water generally had lower $\delta^{18}\text{O}$ and δD values than modern water (Fig. 7.3a).

Groundwaters that plot to the right of the meteoric water line on Figures 10.6 or 10.9 likely indicate (1) evaporation-modified $\delta^{18}\text{O}$ and δD values (see above), or (2) oxygen exchange with isotopically heavier carbonates, clays and other minerals in the subsurface (Clayton, 1966; Savin and Lee, 1988). The hydrogen isotope values of waters are only minimally affected by reaction with minerals because minerals do not have much hydrogen. Even hydrous minerals like clays typically contain many times more moles of oxygen than hydrogen, so when minerals dissolve in water they will have a much greater impact on $\delta^{18}\text{O}$ than δD . In fact, although perhaps obvious, the dissolution of minerals that contain no hydrogen such as carbonates and sulfates (e.g. CaCO_3 , $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$) will have no impact on δD .

Hydrothermal waters, which follow a nearly flat trajectory toward higher $\delta^{18}\text{O}$ values, are

an excellent example of waters that experience appreciable isotopic exchange with ^{18}O -enriched minerals in the subsurface. They typically evolve towards greater $\delta^{18}\text{O}$ values with no noticeable change in δD from initial compositions that are nearly identical to local meteoric water (Craig, 1966). The nearly constant hydrogen isotope values and progressive increase in $\delta^{18}\text{O}$ indicate that locally infiltrating meteoric water circulates to depths where it is heated and driven to undergo oxygen isotope exchange with minerals in the subsurface (Craig, 1966). These types of data are useful in examining hydrothermal processes and origins of hydrothermal waters; the isotopic data seem to rule out the possibility that the hydrothermal waters are of deep crustal or volcanic origin and thus are useful in placing constraints on recharge and circulation in hydrothermal systems. In a more applied sense, D'Amore et al. (1987) used $\delta^{18}\text{O}$ and δD as tracers to determine that re-injection of geothermal waters at a site in Tuscany (Italy) does not negatively impact electricity generation; in fact, the re-injected waters (identified by $\delta^{18}\text{O}$ and δD), once heated up by the high geothermal gradient

and then pumped back to the surface, contain less dissolved gases than do the natural geothermal waters and are more efficient in energy generation.

10.4.3 Application of oxygen and hydrogen isotopes to paleosol climate records

Clay minerals and hydroxides that form in soils record oxygen and hydrogen isotope compositions that are dominated by local meteoric water that infiltrates soils, and for this reason, $\delta^{18}\text{O}$ and δD values of pedogenic (soil-formed) minerals contain useful information about paleoclimate. This is because the abundance of oxygen and hydrogen in soil water that is derived from precipitation is far greater than oxygen and hydrogen derived from dissolution of parent minerals (e.g. amphiboles, feldspars, micas and carbonates in sedimentary, igneous and metamorphic rocks), so the isotopic composition of the pedogenic minerals will be dominated by the composition of meteoric water, which, as we saw above, is strongly influenced by climate. An interesting application of this approach is provided by Chamberlain et al. (1999), who documented notable shifts in $\delta^{18}\text{O}$ and δD in precipitation collected from a transect across the Southern Alps mountain range on the South Island of New Zealand. Precipitation collected from the upwind side (west side) of the range has average $\delta^{18}\text{O}$ and δD values of -5.5‰ and -30‰ as compared to isotopically lighter values of -9.8‰ and -72‰ in the rain shadow on the eastern side of the range.

This type of shift in $\delta^{18}\text{O}$ and δD (indicated in Fig. 10.7) has been observed in the upwind and downwind (rain shadow) sides of many other large mountain ranges globally. Chamberlain et al. (1999) reasoned that large spatial shifts like those observed in the isotopic composition of modern orographically affected precipitation would not have existed prior to the development the

mountain range; rather, the changes would have been relatively small and more similar to the slow, progressive shifts seen in regions like the topographically flat central North America (where precipitation derived from the Gulf of Mexico undergoes Rayleigh distillation as the air mass migrates pole-ward). If this reasoning is true, the isotopic composition of precipitation should change when the mountains are uplifted to form a rain shadow, and this should be recorded in pedogenic minerals preserved in paleosols (buried ancient soils) on the downwind side of the range. Data from pedogenic kaolinites show a marked and step-wise shift from $\delta^{18}\text{O}$ values of $+18\text{‰}$ in older paleosols (5 to 10 Ma) to $+12\text{‰}$ in younger paleosols (< 5 Ma) (Chamberlain et al., 1999). Assuming the kaolinites formed in equilibrium with meteoric water, the $\delta^{18}\text{O}$ of precipitation decreased from -8‰ to -13‰ approximately 5 million years ago; this shift is attributed to the uplift of the Southern Alps and associated formation of a rain shadow approximately 5 Ma. Additional support for this interpretation is provided by the presence of smectite in paleosols younger than 5 Ma – smectite is the dominant pedogenic mineral in semi-arid rain shadow soils, so its absence prior to 5 Ma and its presence in paleosols younger than 5 Ma is also consistent with a shift to drier soil conditions at 5 Ma.

10.5 $\delta^{15}\text{N}$

Stable isotopes of nitrogen (^{14}N , ^{15}N) have been applied to various problems in low-temperature geochemistry, including analysis of natural nitrogen cycling and to the determination of anthropogenic nitrogen species. Reactive nitrogen-bearing compounds are introduced to the Earth surface environment by the following processes and sources:

- combustion of coal, hydrocarbon fuels and wood, which releases nitrogen oxides (commonly referred to as NO_x , where the variable x can be 1,

2 or 3) to the atmosphere where they contribute to formation of smog and acid rain (in the form of nitric acid, HNO_3);

- human and domesticated animal wastes, which produce organic N species (e.g. urea, $[\text{NH}_2]_2\text{CO}$) that commonly undergo transformation via microbially mediated processes to ammonium (NH_4^+), nitrite (NO_2^-) and nitrate (NO_3^-);
- addition of fertilizers (e.g. ammonium nitrate, potassium nitrate) to agricultural fields, golf courses and suburban lawns that can lead to nitrate pollution of surface water and groundwater.

Geological sources of nitrogen are generally very minor (one example of nitrogen associated with rocks and minerals is NH_4^+ substituted for K^+ in the interlayers of some micas), and the majority of nitrogen in soils, surface waters and groundwater is fixed by biological processes from atmospheric N_2 to a wide range of inorganic ions (e.g. NO_3^- , NH_4^+), organic species (dissolved and particulate organic nitrogen) and gases (N_2O , NO_x , NH_3). While some basic information on nitrogen isotope systematics, a more-detailed treatment of the nitrogen cycle is presented in **Chapter 6**.

Regardless of origin, N species at or near the Earth's surface commonly transform to the soluble inorganic species ammonium, nitrite and nitrate, the latter two of which are toxic to humans because they can cause methemoglobinemia, where insufficient O_2 reaches cells (especially in infants). Furthermore, in the oxidizing environment of the atmosphere and most surface waters and shallow groundwaters, the more-oxidized nitrate is more stable than the more-reduced species ammonium and nitrite. Nitrogen species also tend to be limiting nutrients in lakes and many ponds, so any increase of nitrogen concentrations in such water bodies can lead to eutrophication and the ecologically damaging algal blooms that result. For these reasons, understanding sources of nitrogen species in air, water and soil can be an important step in remediation, and stable isotope analysis can be an important

tool in fingerprinting N sources. For species like nitrate that contain both N and O, the combination of nitrogen and oxygen isotope compositions can be particularly useful in discriminating sources.

Given that air is very well mixed with respect to N_2 gas and that the ratio of $^{15}\text{N}/^{14}\text{N}$ is stable, the standard for reporting $\delta^{15}\text{N}$ values is the isotopic composition of N_2 in the atmosphere. Hence, the $\delta^{15}\text{N}_{\text{atm}}$ value of atmospheric N_2 is 0‰, and any nitrogen compounds derived with no appreciable fractionation from atmospheric nitrogen will also have $\delta^{15}\text{N}_{\text{atm}}$ values that are close to 0‰. This is the case for fertilizers, which are derived from atmospheric N_2 gas by the Haber process; the result is that the ammonium and nitrate in synthetic fertilizers have $\delta^{15}\text{N}_{\text{atm}}$ values that typically range from -5 ‰ to $+5$ ‰ (Fig. 10.10). Similarly, nitrogen oxides (NO_x) produced by combustion of

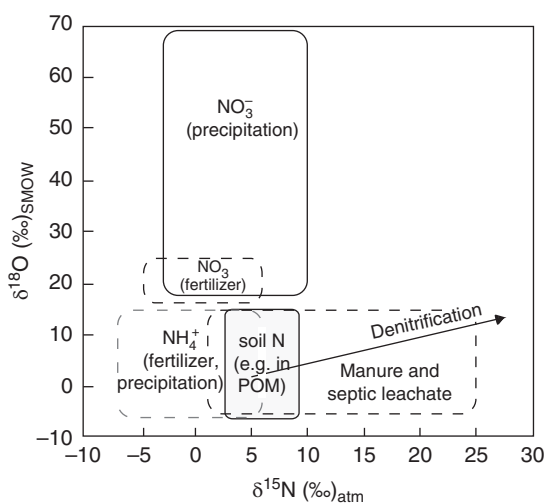


Fig. 10.10 Differentiation of sources of nitrogen based on $\delta^{15}\text{N}$ and $\delta^{18}\text{O}$. Denitrification is the reaction of nitrate or other N species to N_2 gas – this topic is covered in Chapter 6 – and the liberation of isotopically light ^{15}N (and ^{16}O) results in a trend towards isotopically heavier remaining organic solids. Modified from Kendall and McDonnell (1998) and other sources cited therein.

fuels also exhibit $\delta^{15}\text{N}_{\text{atm}}$ values that are close to zero – in the case of combustion, the source of the NO_x is chemical reaction of atmospheric N_2 and O_2 in the combustion chamber.

Plant residues also possess $\delta^{15}\text{N}_{\text{atm}}$ values that are close to zero. However, animal wastes tend to be enriched in ^{15}N relative to plant matter, atmospheric nitrogen and its direct derivatives. This enrichment is caused by two factors, the first of which is that consumers (in terms of trophic level) tend to preferentially retain ^{15}N over ^{14}N ; while this can produce urine that is slightly depleted in ^{15}N relative to the $\delta^{15}\text{N}$ of the food source, animal tissue and solid waste is typically enriched by 2 to 3‰ compared to the food source. The second factor contributing to ^{15}N -enriched animal waste is volatilization of isotopically light (^{15}N -depleted) ammonia from manure or other wastes at or near the land surface. The result of these ^{15}N -concentrating processes is that organic wastes in sewage and septic effluent and animal manure have $\delta^{15}\text{N}_{\text{atm}}$ values that typically range from +10‰ to +25‰ (Kendall and Doctor, 2003). Because solid and liquid organic fertilizers are derived from animal wastes (e.g. organic manure), they also have positive $\delta^{15}\text{N}_{\text{atm}}$ values that make them indistinguishable from their manure source.

Although there is some overlap in $\delta^{15}\text{N}$ values of synthetic fertilizers, which can be as high 8‰, and animal wastes, which in some cases can be as low as 0‰, analysis of $\delta^{18}\text{O}$ values of nitrate in conjunction with $\delta^{15}\text{N}$ values helps to produce more distinctive fields on a bivariate plot – fertilizer nitrate plots higher than animal wastes (Fig. 10.10). Given that synthetic ammonium nitrate fertilizers are derived by reacting H_2 , N_2 and O_2 from the atmosphere, the $\delta^{18}\text{O}$ of nitrate derived from synthetic fertilizers is generally +15 to +25‰, a value that is consistent with atmospheric O_2 ($\delta^{18}\text{O}_{\text{SMOW}} = +23.5\text{‰}$; Kroopnick and Craig, 1972), and the synthesis does not result in significant fractionation. The $\delta^{18}\text{O}$ value of animal wastes is typically in the range of -8 to +12‰,

reflecting incorporation of surface water with lower $\delta^{18}\text{O}$ values than those of O_2 in air.

The complex nitrogen cycle is prone to kinetic controls imparted by biologically mediated effects, yet in many cases analysis of $\delta^{15}\text{N}$ values (often in combination with $\delta^{18}\text{O}$) has proven useful in examining source of nitrogen species at or near the Earth's surface. Thorough reviews of nitrogen isotopes in soils, waters, air and biological systems are given by Kendall and McDonnell (1998) and Kendall and Doctor (2003).

10.6 $\delta^{13}\text{C}$

The ratio of carbon's two stable isotopes, ^{13}C and ^{12}C , can be used to understand sources and cycling of carbon in the Earth's atmosphere, surface environments and crust. As a general rule, carbon derived from organic matter is isotopically light (^{12}C -enriched), whereas carbon derived from inorganic sources like marine water (mainly as dissolved HCO_3^-) and carbonate, igneous and metamorphic rocks (as carbonate minerals, dissolved HCO_3^- or volatilized CO_2) is isotopically heavy. Hydrocarbons and coal, while derived from rocks in the crust of the Earth, are isotopically light because the carbon in these reservoirs was originally fixed by photosynthesis (into marine algae that have transformed to hydrocarbons or terrestrial plants that have transformed to coal).

Figure 10.11 illustrates representative $\delta^{13}\text{C}$ values of carbon in Earth-surface compounds, including land plants (C_3 vs. C_4) as well as hydrocarbons, marine organic matter and marine carbonates (which shows the ~0‰ value of limestone; in some cases, diagenesis causes deviation of $\pm 5\text{‰}$ around the initial 0‰ value).

Because the standard for carbon isotope values is a marine carbonate (the Pee Dee Belemnite), sea water and the fossils and minerals that form from dissolved inorganic carbon (DIC) in sea water have $\delta^{13}\text{C}_{\text{PDB}}$ values that are close to

0‰ – sea water is approximately -2 to $+2$ ‰ and minerals (e.g. CaCO_3 in limestone) that crystallize from sea water are generally in the range of -5 to $+5$ ‰. The $\delta^{13}\text{C}$ of atmospheric CO_2 is

approximately -8 ‰ but this value is progressively decreasing (becoming more negative) as isotopically light carbon ($\delta^{13}\text{C}_{\text{PDB}} = -20$ to -30 ‰) from fossil fuels is added to the

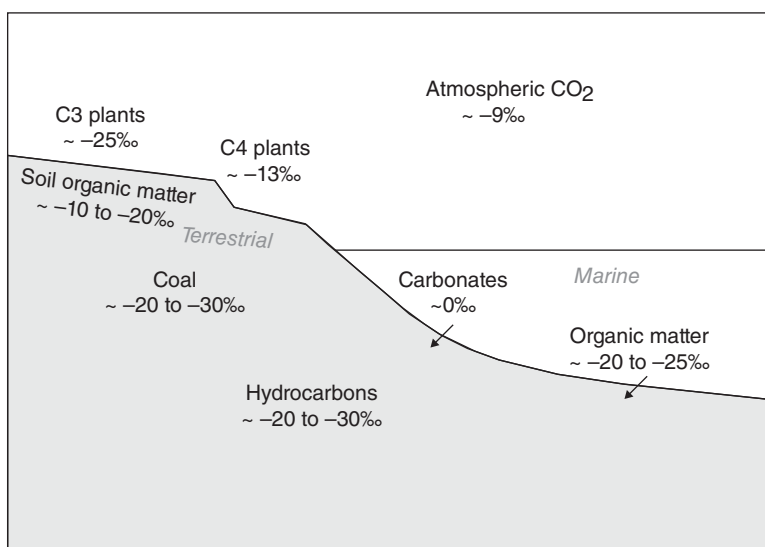


Fig. 10.11 $\delta^{13}\text{C}_{\text{PDB}}$ values for common reservoirs of carbon at and near the Earth surface. Note that all reservoirs contain $\delta^{13}\text{C}$ values ≤ 0 ‰, and that C3 plants, hydrocarbons and marine organic matter have the lowest (isotopically lightest) values; carbonates are the heaviest with values of ~ 0 ‰.

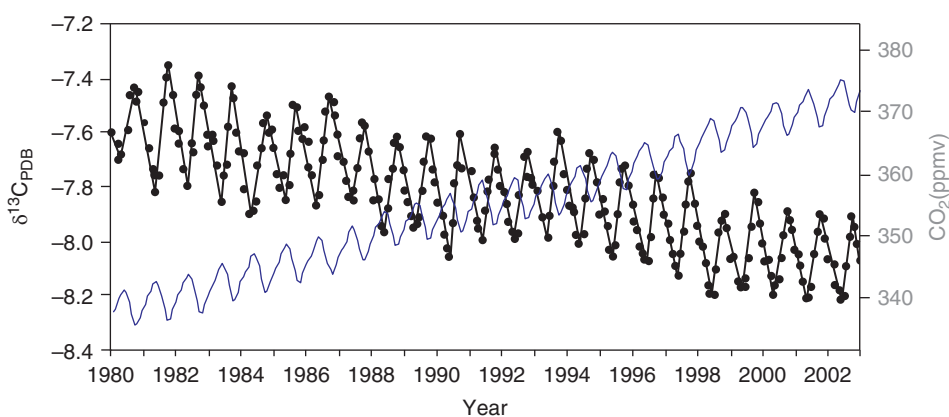


Fig. 10.12 A 23-year record of cyclically increasing atmospheric CO_2 concentration (in gray) paired with, in black, cyclically decreasing $\delta^{13}\text{C}$ of atmospheric CO_2 . Both trends result from combustion of isotopically light hydrocarbons. Data are from Oak Ridge National Lab (http://www.nicholas.duke.edu/thegreengrok/climatedisruption_solomon).

atmosphere. This trend is illustrated in Fig. 10.12, which also depicts increasing atmospheric CO_2 concentrations and annual cycles in carbon isotope ratios and CO_2 concentrations related to photosynthesis.

Atmospheric CO_2 concentrations reach annual low values in August and September, reflecting uptake of CO_2 during the growing season in the northern hemisphere – the far greater abundance of land (and hence terrestrial vegetation) in the northern hemisphere as compared to the southern hemisphere means that the annual pulse of plant growth in Asia, Europe, North America and Africa north of the equator produces a drop in atmospheric annual CO_2 values, globally. The inverse relationship on an annual basis between CO_2 and $\delta^{13}\text{C}$ also can be traced to photosynthesis, a process that selects ^{12}C over ^{13}C ; hence, $\delta^{13}\text{C}$ increases (becomes less negative) when photosynthesis peaks.

10.6.1 Carbon isotope analysis of paleoenvironment

Stable carbon isotopes are fractionated differently by the two main photosynthetic processes employed by plants. C_3 plants are the dominant forms of vegetation on the Earth's surface and tend to thrive in most climate zones (notable exceptions are arid regions and salt marshes). They are called C_3 plants because the initial organic compound synthesized by photosynthesis in these plants, 3-phosphoglyceric acid, has 3 carbon atoms. $\delta^{13}\text{C}$ values of C_3 plants are generally in the range of -23 to -34‰ with average values of -25 to -26‰ . C_3 plants, however, are not well-adapted to hot, dry conditions. Their stomata remain open all day during photosynthesis and thus they suffer from water loss and dehydration.

C_4 plants, on the other hand, have adapted to hot, dry environments by alternately opening and closing stomata during photosynthesis, which allows CO_2 to enter the stomata but limits water

loss. In C_4 photosynthesis, a 3-carbon compound (pyruvate or phospho-enol pyruvate), is carboxylated to give a 4-carbon dicarboxylic acid (Hatch and Slack 1966). This process causes a lower degree of carbon isotope fractionation, and typical C_4 plant $\delta^{13}\text{C}$ values are -9 to -16‰ with an average of -12 to -13‰ . Many terrestrial C_4 plants are tropical grasses and desert plants as well as salt marsh grasses. A third photosynthetic pathway similar to the C_4 process is known as CAM (Crassulacean Acid Metabolism) – many desert succulents are CAM plants, and their typical $\delta^{13}\text{C}$ values are -12 to -23‰ ; hence, they generally overlap with C_4 or are intermediate to C_3 and C_4 .

If organic matter from decayed plants or bone or egg shells from animals that once consumed plants is sampled from a series of paleosols or buried sediments and is found to undergo a shift, from e.g. -30‰ to -10‰ over time, it likely indicates a change from C_3 to C_4 plants, evidence for a transition from a temperate climate to a more arid climate. This approach has been applied in many regions to measure the timing and magnitudes of climate shifts. Examples include Cerling et al. (1993), Stern et al. (1994), Cerling et al. (1997), and Bechtel et al. (2008).

A pair of important points related to the application of carbon isotopic shifts to paleoclimate analysis.

1 There is a $+14\text{‰}$ shift in $\delta^{13}\text{C}$ from plant to tooth enamel of large mammals (e.g. bovids, equids, rhinocerids) which reflects the preferential fractionation of ^{13}C into mammalian tooth and bone. In other words, C_3 grasses with $\delta^{13}\text{C} = -25\text{‰}$ would produce large mammal tooth enamel with $\delta^{13}\text{C} = -11\text{‰}$. This shift has been documented in modern animals and it is applied to paleomammal-vegetation systems to determine carbon isotope compositions of paleoecosystems (Cerling et al., 1997).

2 Global ecosystems experienced a pronounced shift in $\delta^{13}\text{C}$ approximately 6 Ma. Prior to 6 Ma, the $\delta^{13}\text{C}$ of large grazing mammals from sites

across the globe cluster around -10‰ ; post-6 Ma $\delta^{13}\text{C}$ values from the same sites produce an average value of approximately $+1\text{‰}$. When corrected for fractionation, these values indicate a shift 6 million years ago from grasses with $\delta^{13}\text{C} = -24\text{‰}$ to -9‰ – these values line up very well with $\delta^{13}\text{C}$ of C_3 and C_4 plants, respectively. As you can probably surmise, Cerling et al. (1993, 1997) interpreted these data to indicate a sharp transition to a more arid climate at 6 Ma that caused a decline in the abundance in C_3 plants and a corresponding increase in C_4 plants.

10.6.2 Carbon isotopes in hydrology and chemical weathering

The difference in isotopic values of dissolved inorganic carbon (DIC) has been successfully applied to determining DIC sources in watersheds, including those affected by acid precipitation. Given the crucial role that carbonate weathering plays in acid neutralization, being able to quantify the contribution of carbonate weathering to the total DIC budget of a watershed is essential in assessing the origin and cycling of DIC and buffering potential. Carbon isotope values for DIC derived from rock weathering ($\delta^{13}\text{C} \sim 0$ to -5‰) are typically greater than the DIC derived from soil CO_2 (usually -10 to -25‰). The low values of soil CO_2 are caused by: (1) photosynthesis and microbial respiration in the soil, which produces $\delta^{13}\text{C}$ values that reflect the isotopic composition of the plants undergoing decomposition (e.g., $\delta^{13}\text{C} \sim -25\text{‰}$ in soils dominated by C_3 plants); and (2) the presence in soil of atmospheric CO_2 with $\delta^{13}\text{C}$ values of $\sim -8\text{‰}$.

The weathering of carbonate minerals, particularly calcite that occurs in limestones, calcite veins, carbonate grains in soils, and various other forms, is a vital process in buffering the effects of acidification, so being able to identify the contribution of carbonate weathering to a

watershed is an important aspect of assessing buffering potential. In watersheds with soils that contain low carbonate mineral (e.g. calcite) content, the flow paths of subsurface waters that eventually recharge lakes and rivers can be a crucial component of determining whether or not the surface water bodies will become acidic, or not. Waters that follow shallow flowpaths to surface waters through calcite-poor soils will impart $\delta^{13}\text{C}$ values that are isotopically light and indicative of soil organic matter and atmospheric CO_2 (e.g. -20 to -30‰) – in these cases the $\delta^{13}\text{C}$ values are dominated by soil respiration, reflecting limited or no contribution from isotopically heavier carbon from calcite. In contrast, waters that circulate to deeper levels and experience greater interaction with calcite in bedrock or glacial till before recharging surface waters will be characterized by relatively higher $\delta^{13}\text{C}$ values (e.g. 0 to -5‰) that reflect the contribution of carbon from isotopically heavier calcite (Kendall and Doctor, 2003). In this way, analysis of $\delta^{13}\text{C}$ values provides quantitative tracer information on soil and groundwater flow paths and can be applied to predicting response of a watershed to acidification, particularly in regions where soils (i.e. upper 1 m) are carbonate poor.

10.7 $\delta^{34}\text{S}$

There are four stable isotopes of sulfur, ^{32}S , ^{33}S , ^{34}S and ^{36}S (Fig. 10.1, Table 10.2), but the two that are most commonly measured in Earth's surface and shallow crustal environments are ^{34}S and ^{32}S . The ratio of these two isotopes is reported in standard delta notation as $\delta^{34}\text{S}_{\text{CDM}}$, where the reference for $\delta^{34}\text{S}$ values is the Canyon Diablo Meteorite (CDM) from Meteor Crater Arizona (Sharp, 2007). Values of $\delta^{34}\text{S}$ generally fall within the range of -30 to $+30\text{‰}$, although rare examples have been documented as low as -65‰ (in sulfides) and as high as 120‰ (in sulfates; Hoefs,

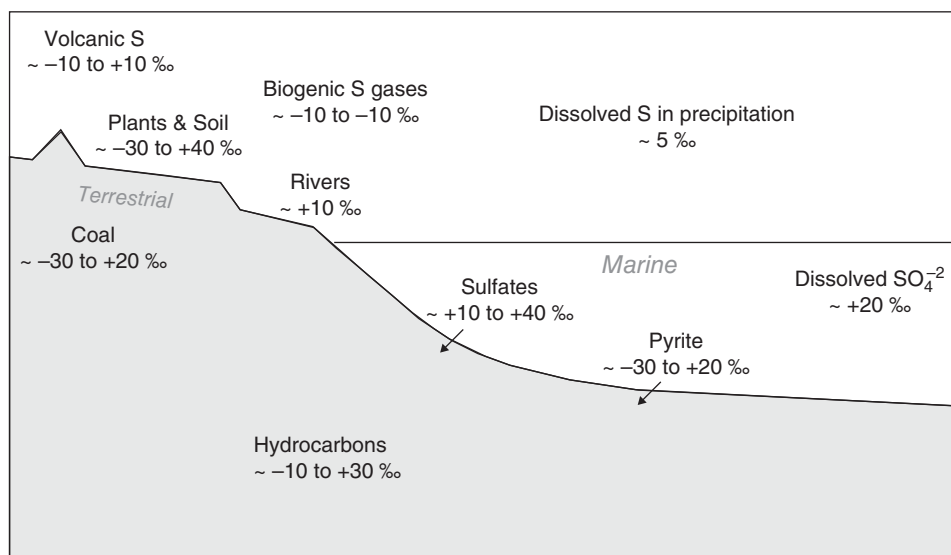


Fig. 10.13 $\delta^{34}\text{S}_{\text{CDT}}$ values for common reservoirs of sulfur at and near the Earth surface. Note that many of these reservoirs have wide ranges of $\delta^{34}\text{S}$, which in part reflects the extent of reduction–oxidation reactions in the sulfur cycle.

2009). Figure 10.13 summarizes typical sulfur isotopic compositions in various reservoirs.

In general, the lowest $\delta^{34}\text{S}$ values occur in sedimentary sulfides (e.g. pyrite), whereas values on the higher end of the scale occur in sulfate (SO_4^{2-}), including marine sulfate (+21‰; Kendall and Doctor, 2003) and evaporite sulfate minerals (+10 to +30‰; Hoefs, 2009). The partitioning makes sense intuitively: the lighter ^{32}S is partitioned into the gas phase (as H_2S) and the heavier ^{34}S remains in aqueous or solid phase in SO_4^{2-} (Fig. 10.14A). However, in very strongly reducing environments (e.g. deep anoxic marine sediments), complete reduction of sulfate can result in isotopically heavy iron sulfide with values as high as +50‰; in this case, isotopically light ^{32}S is preferentially lost to the atmosphere as H_2S and ^{34}S is preferentially partitioned into the solid phase (FeS_2) that remains in the water or sediment (Fig. 10.14B).

In addition to sulfate and sulfide minerals, sulfur also occurs in organic matter and atmospheric gases (e.g. SO_2), and hence sulfur isotope analysis

is often used to understand the sources and processes responsible for the cycling of sulfur among the atmosphere, oceans, fresh water bodies and the crust.

10.7.1 Fraction of sulfur isotopes

The fractionation of sulfur isotopes in natural systems is caused by two main processes (Drever, 1997; Hoefs, 2009):

1 Equilibrium isotopic exchange that occurs among transformations between sulfide (S^{2-}) and sulfate.

2 Kinetic isotopic exchange affected by microbially mediated redox reactions. For example, the greatest magnitude of fractionation within the sulfur isotope system occurs during dissimilatory sulfate reduction, a process that results in formation of biogenic sulfides with isotopically light (negative) values. In this process, bacteria (e.g. *desulfovibrio* or *desulfatovaculum*) take advantage

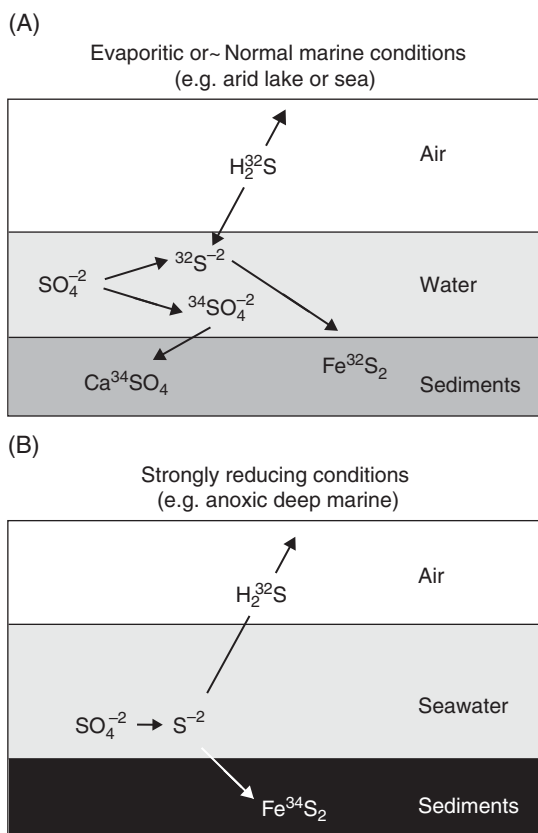
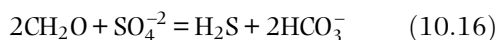


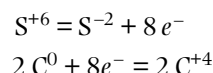
Fig. 10.14 Fractionation of sulfur isotopes in two different scenarios. In normal marine to evaporation-dominated seas or lakes (A), sulfate reduction is only partial; in this case, ^{32}S is partitioned into the sulfide anion, some of which goes on to form isotopically light H_2S , and some of which forms isotopically light iron sulfide. In terms of the solid phases in this case, ^{34}S is partitioned into the sulfate, so the sulfate is isotopically heavy and the sulfide is isotopically light. In strongly reducing conditions (B), all sulfate is reduced to sulfide and ^{34}S is partitioned in iron sulfide in sediments and ^{32}S into hydrogen sulfide gas.

of the energy associated with the transfer of 8 electrons produced by sulfate reduction:



In this example, carbon is oxidized while sulfur is reduced. Examining only the change in

oxidation states of the individual atoms, we see that:



The S^{+6} ion occurs in sulfate and the sulfide (S^{-2}) ion occurs in H_2S (and also FeS_2 and other sulfide minerals); C^0 represents reduced carbon in organic matter (here given as CH_2O) while C^{+4} represents oxidized carbon in bicarbonate.

The formation of isotopically light sulfide minerals during sulfate reduction is partly due to the liberation of H_2S gas with $\delta^{34}\text{S} \sim 25\text{‰}$ less than the sulfate that is being reduced. This leaves the residual sulfate enriched in ^{34}S , and when sulfide anions derived from H_2S participate in chemical reactions to form pyrite, the pyrite will be isotopically light.

Selected examples of sulfur isotopes applied to problems in environmental geochemistry include the following:

1 Sources and cycling of sulfur in regions affected by acid precipitation. Given that $\delta^{34}\text{S}$ of precipitation is often different from $\delta^{34}\text{S}$ of soil sulfur and bedrock sulfur, relative contributions from these sources in surface waters affected by acidification can be assessed. For example, from studies in Canada, Japan and Czech Republic, Nriagu and Coker (1978) Ohizumi et al. (1997) and Novák et al. (2005) all observed that anthropogenic sulfur (e.g. from coal combustion and ore refining) has higher $\delta^{34}\text{S}$ values than does natural soil sulfur, so analysis of $\delta^{34}\text{S}$ values of surface water can be used to assess the relative contributions of natural versus anthropogenic sulfur from the atmosphere (and hence also acidity, given that sulfuric acid is a major component of acid rain). However, higher $\delta^{34}\text{S}$ values are not always associated with anthropogenic sulfur from the atmosphere; Wadleigh and Blake (1999) observed isotopically lighter anthropogenic sulfate (relative to natural sulfate) as recorded in lichens from eastern Canada. Clearly, natural variability in both anthropogenic

and natural sources must be considered. It is also worth noting that, as is the case with hydrogen and oxygen in water and nitrogen and oxygen in nitrate, paired analysis of $\delta^{34}\text{S}$ and $\delta^{18}\text{O}$ in sulfate can enhance assessment of sources and cycling (Wadleigh et al., 1996).

2 Tracking the microbially mediated reduction of sulfate in anaerobic groundwaters. Sulfate reduction occurs in conjunction with oxidation of organic matter and this is an important control on the degradation of petroleum hydrocarbons, solvents, pesticides and similar organic compounds – the microorganisms gain energy by simultaneously reducing sulfate while oxidizing carbon in organic matter (including hydrocarbons and related organic compounds). Thus, analysis of sulfur isotopes has been applied to studies of contaminated aquifers to trace the extent of microbial reduction of sulfate. The pronounced increases in $\delta^{34}\text{S}$ of residual sulfate (i.e. the sulfate not yet reduced to sulfide) that accompany sulfate reduction have been used to quantify anaerobic microbial activity in order to assess effectiveness of hydrocarbon degradation (e.g. Schroth et al., 2001; Meckenstock et al., 2004).

3 Long-term variations in sea-water chemistry. The fractionation between dissolved sulfate in sea water and the evaporite minerals that form from evaporated sea water is nearly zero, so marine evaporites reflect the $\delta^{34}\text{S}$ of marine water and also do not affect the composition of sea water when they form (Sharp, 2007). There is no Rayleigh distillation or similar effects associated with preferential removal of one isotope from the reservoir (in this case, sea water) when marine evaporites (e.g. gypsum, $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$, or barite, BaSO_4) crystallize from solution.

10.8 NON-TRADITIONAL STABLE ISOTOPES

The focus of this chapter thus far has been the conventional or traditional stable isotopes (C, H, O, N, S), but many recent advances in stable isotope geochemistry have come in the analysis of

metal stable isotopes, including heavy metals such as Cr, Fe, Ni, Cu, Zn, Mo, Cd and Hg, as well as the lighter elements Li, Mg and Ca, and Cl. The emergence of the new avenues of stable isotope research is largely due to improved precision of instrumental analysis via development of protocols using multicollector inductively coupled plasma mass spectrometry (MC-ICP-MS) and thermal ionization mass spectrometry (TIMS). To some extent the relatively late development of some of these methods may also be attributed to the old assumption that elements that form dominantly ionic bonds (e.g. Mg, Ca, Cl) would not be affected by isotopic fractionation in the environment. This has proven to be false, but nonetheless the relatively low-magnitude fractionations of many of the non-traditional systems highlighted below require high-precision techniques and analytical tools.

The relatively small differences in mass of the heavier transition metals – as compared to more-traditional stable isotopes (e.g. ^{18}O vs. ^{16}O) – makes it difficult to precisely measure fractionation in natural systems; for example, the mass difference between ^{65}Cu and ^{63}Cu is only 3.2%, whereas for ^{18}O and ^{16}O it is 12.5% (Table 10.2). If mass-dependent fractionation (MDF) is the dominant control on the behavior of isotope pairs of two different elements, the element with greater mass (e.g. Cu) will be fractionated less. The relatively low extent of fractionation of transition metals and heavier elements was not precisely quantifiable with pre-2000 methods and instrumentation; now, with advances in MC-ICP-MS and TIMS, geochemists have the capability to precisely measure fractionations as small as 1 or 2‰ and often as low as 0.1 or 0.2‰. The principles and concepts of three metal stable isotopes (Cu, Fe, Hg) and two other non-traditional stable isotopes (Mg and $\delta^{37}\text{Cl}$) are presented in the following sections. Although not exhaustive, these examples will provide some background about how these systems operate in nature and how they can be used to answer questions regarding natural and anthropogenic processes.

10.8.1 $\delta^{65/63}\text{Cu}$

The standard notation for the two stable isotopes of copper is:

$$\delta^{65}\text{Cu}(\text{‰}) = \left(\frac{(^{65}\text{Cu}/^{63}\text{Cu}_{\text{smpl}}) - (^{65}\text{Cu}/^{63}\text{Cu}_{\text{NIST976}})}{(^{65}\text{Cu}/^{63}\text{Cu}_{\text{NIST976}})} \right) * 1000 \quad (10.17)$$

Ranges of $\delta^{65}\text{Cu}$ in natural materials presented in Borrok et al. (2008), Fujii et al. (2013) and Thompson et al. (2013) include sulfide minerals (+1 to −5‰), ocean water (0.7 to 0.9‰), oxidized carbonate minerals such as cuprite or malachite (−2 to +3‰), plants (0 to −5‰) and stream water (−1 to +3‰). When Cu is oxidized ^{65}Cu tends to be fractionated into the oxidized phase (water or mineral, hence relatively high values in surface water and minerals stable in oxidizing environments) while plants and sulfide minerals appear to fractionate ^{63}Cu preferentially over ^{65}Cu , leading to relatively low or negative values in these reservoirs.

Given that copper is influenced by redox reactions ($\text{Cu}^+ = \text{Cu}^{+2} + e^-$), stable Cu isotopes have been measured to study origins of ore deposits as well as processes controlling weathering and transport of copper in acid mine drainage areas. A study by Kimball et al. (2009) of a drainage basin in Red Mountain, Colorado, USA impacted by sulfide oxidation (Figure 4.13) and associated acid mine drainage (Chapter 9) serves to illustrate the potential application of Cu isotopes to environmental geochemistry. Data from this study demonstrate that stream water ($\sim +1.5\text{‰}$) is 1.4 to 1.6‰ enriched in ^{65}Cu relative to the parent sulfide minerals enargite and chalcopyrite, whose $\delta^{65}\text{Cu}$ values are $\sim 0.01 \pm 0.10\text{‰}$ and $0.16 \pm 0.10\text{‰}$, respectively (terrestrial Cu sulfides cluster around $0.0 \pm 1.0\text{‰}$). The oxidation of Cu^+ to Cu^{+2} has resulted in enrichment of ^{65}Cu in the oxidized Cu phases in solution,

and the solid weathering products remaining in the tailings have been enriched in ^{63}Cu . This behavior is typical of metal stable isotopes (e.g. Cr, Cu, Mo, Fe), where heavier isotopes tend to be partitioned into oxidized aqueous phases.

In order to further explore the controls on Cu isotope fractionation during sulfide weathering, Kimball et al. (2009) carried out controlled leaching experiments in the presence of, and absence of, the bacterium *Acidithiobacillus ferrooxidans* (known to accelerate weathering of many sulfides). Abiotic weathering, i.e. experiments carried out in the absence of *A. Ferrooxidans*, reproduces very similar values to those obtained from the field study; however, in the presence of *A. Ferrooxidans*, fractionation is virtually nonexistent for enargite and actually occurs in the opposite direction for chalcopyrite (waters are depleted in ^{65}Cu by $\sim 0.6\text{‰}$ relative to the parent sulfide mineral). In the absence of microbes such as *A. Ferrooxidans*, copper sulfide weathering is driven by preferential oxidation of ^{65}Cu (relative to ^{63}Cu) that releases ^{65}Cu into solution. TEM analysis shows that when copper sulfide weathering occurs in the presence of microbes such as *A. Ferrooxidans*, ^{65}Cu is preferentially incorporated into bacterial cell walls or other parts of the bacterial community (Fig. 10.15), but is not released into solution.

Given that the natural weathering of enargite and chalcopyrite led to a $\sim +1.5\text{‰}$ enrichment of ^{65}Cu in stream water (relative to the parent sulfide minerals), it appears that microbes such as *A. Ferrooxidans* do not play a significant role in copper sulfide weathering – at least not at Red Mountain – and this has implications for remediation. If bacteria do not play an important role in weathering of Cu sulfides, bactericides would not be expected to limit metal release and acid generation if the main source is Cu sulfides (this is in contrast to weathering of pyrite, where *A. Ferrooxidans* frequently enhances reaction rate, and where bactericides may be applied to limit acid generation).

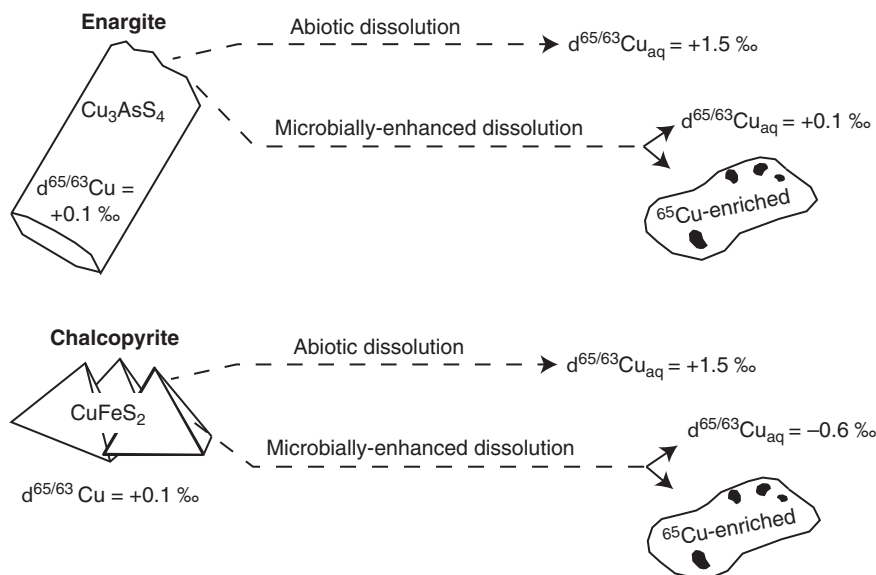
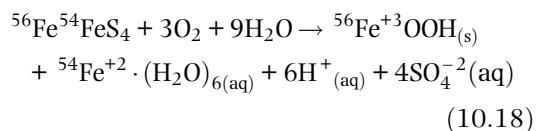


Fig. 10.15 Schematic diagram of fractionation of Cu isotopes during weathering of two Cu sulfide minerals, enargite and chalcopyrite. The bacterium *Acidithiobacillus ferrooxidans* (shown in sketch) sequesters ^{65}Cu , meaning that the extent of microbially mediated dissolution of Cu sulfides can be assessed by $\delta^{65/63}\text{Cu}$.

10.8.2 $\delta^{56/54}\text{Fe}$

As indicated in Table 10.2, there are four stable isotopes of iron, the two most abundant of which are ^{56}Fe and ^{54}Fe . Like Cu and other metal isotopes, iron isotopes are fractionated during redox reactions, resulting in enrichment of ^{56}Fe in ferric aqueous phases [e.g. $\text{Fe}(\text{III})(\text{H}_2\text{O})_6^{+3}$] and ^{54}Fe in ferrous phases [e.g. $\text{Fe}(\text{II})(\text{H}_2\text{O})_6^{+2}$]. The result of this initial fractionation and the low solubility of ferric iron is that ^{56}Fe tends to preferentially partition into iron oxides and hydroxides (Bullen, 2011), which results in ^{54}Fe enrichment in solution, all of which can be illustrated by equation 10.18.



As with Cu isotopes, abiotic reactions appear to fractionate iron isotopes differently from

biologically mediated reactions. When goethite dissolution is abiotic and proton assisted (i.e. in presence of H^+), there is no fractionation of ^{56}Fe and ^{54}Fe —both are released to solution at the same rate; however, ligand-assisted dissolution of goethite led to fractionation of ^{54}Fe into solution over ^{56}Fe (examples of ligands include the biomolecules oxalate and acetate and inorganic species like phosphate). Reductive dissolution also preferentially releases ^{54}Fe to solution (Wiederhold et al., 2006). These observations appear to have good potential for analysis of reaction mechanisms of Fe-bearing minerals.

10.8.3 $\delta^{202/198}\text{Hg}$

Mercury has seven stable isotopes (^{196}Hg , ^{198}Hg , ^{199}Hg , ^{200}Hg , ^{201}Hg , ^{202}Hg , and ^{204}Hg), the most abundant being ^{202}Hg (29.86%; refer to Table 10.2 for complete list of Hg isotopes). The

main forms of Hg at the Earth surface occur as Hg^{+2} and as the gas Hg^0 . The divalent cation Hg^{+2} often forms strong complexes with organic matter, commonly as methyl mercury, e.g. $(\text{CH}_3)_2\text{Hg}$ (Chapter 8, Section 8.11). It is a toxic element in any form although when methylated it is mobile, prone to ingestion by organisms and more toxic than other Hg forms. Hg often bioaccumulates and biomagnifies in aquatic ecosystems and mercury advisories are in place in many regions globally due to high Hg content in fish. Its average crustal abundance is very low (0.1 mg/kg) but Hg is concentrated in some sulfide-rich deposits at values from 0.1 to 2%, mainly as cinnabar (HgS), and mine and Hg-refining sites are often sources of Hg contamination. Two other major sources of Hg are emissions to the atmosphere derived from coal combustion and waste incineration – the scale of impact of these emissions are typically on scales of hundreds of km.

Blum and Bergquist (2007) discuss Hg isotope standards, which are still in development. The delta notation is $\delta^{202/198}\text{Hg}$ (Table 10.2), but values of the other Hg isotopes are expressed in different ways. One is as delta notation where the numerator of the isotope pair is ^{198}Hg .

$$\delta^{\text{xxx}}\text{Hg}\text{‰} = \left(\frac{(\delta^{\text{xxx}}\text{Hg}/\delta^{198}\text{Hg})_{\text{sample}}}{(\delta^{\text{xxx}}\text{Hg}/\delta^{198}\text{Hg})_{\text{SRM3133}}} - 1 \right) * 1000 \quad (10.19)$$

(Note that while expressed differently than other δ notations (e.g. equation 10.17), this way is mathematically identical.)

Another approach is α values (equations 10.9, 10.20), which is the ratio of an isotope pair in one phase (A) relative to that same ratio in a second phase, i.e. phase B (note: “phase” here includes minerals, bacteria, groundwater, stream water, plant leaf, etc.).

$$\alpha_{\text{A-B}} = \frac{{}^{201}\text{Hg}/{}^{198}\text{Hg}_{\text{A}}}{{}^{201}\text{Hg}/{}^{198}\text{Hg}_{\text{B}}} \quad (10.20)$$

When α is > 1.000 , the heavy isotope is fractionated into phase A; when α is < 1.000 , the heavy isotope is fractionated into phase B.

Another delta notation, this one with capital Δ , is shown below for the odd-numbered Hg isotopes:

$$\Delta^{201}\text{Hg} = \delta^{201}\text{Hg} - (\delta^{202}\text{Hg} * 0.752) \quad (10.20)$$

and

$$\Delta^{199}\text{Hg} = \delta^{199}\text{Hg} - (\delta^{202}\text{Hg} * 0.252) \quad (10.21)$$

The odd-numbered Hg isotopes ^{201}Hg and ^{199}Hg behave differently from the even-numbered Hg isotopes due to the magnetic isotope effect, a difference in orbital structure that affects the reaction/bonding rate. This is an example of **mass-independent fractionation** (MIF), a process that provides an additional angle for investigation of biotic and abiotic reactions, notably photoreduction and its effect on Hg cycling in aquatic ecosystems (Blum and Bergquist, 2007).

The Hg isotopic variability among coals from different deposits (Fig. 10.16) provides an idea of how Hg isotope analysis can be applied to assessment of Hg contamination. For example, in order to test whether Hg emitted from a coal-fired power plant (in west-central Florida, USA) is deposited close to the point source, or whether it is carried far downwind with no local effects, Sherman et al. (2012) examined Hg isotope compositions of the coal burned at the plant, of precipitation within 10 km downwind of the plant, and of precipitation unaffected by the plant (i.e. not downwind of the smokestack). They traced the distinct $\delta^{202}\text{Hg}$ signature of coal burned at the power plant ($\delta^{202}\text{Hg} = -0.72\text{‰}$) to isotopically light $\delta^{202}\text{Hg}$ ($\delta^{202}\text{Hg} = -2.6\text{‰}$) in rainfall deposited with 5–10 km of the point source. The locally deposited rainfall $\delta^{202}\text{Hg}$ is distinct from precipitation that falls upwind of the coal plant ($\delta^{202}\text{Hg} = +0.07 + 0.17\text{‰}$), demonstrating Hg from the plant is deposited close to the source. Fractionation of light ^{198}Hg during combustion

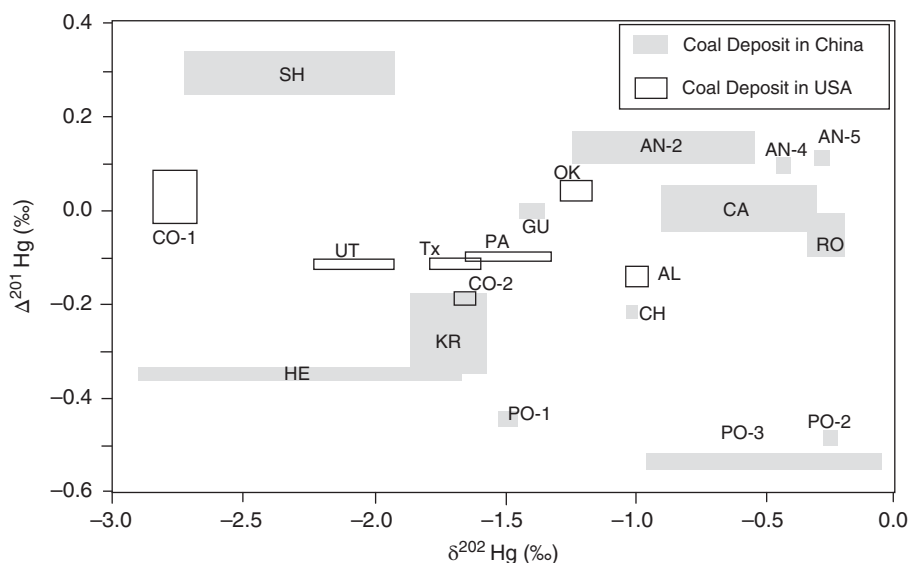


Fig. 10.16 Mercury isotope compositions of coals from China and the USA. Boxes encompass analytical uncertainty for a given deposit, and letter symbols refer to distinct coal deposits, details of which are given in Biswas et al. (2008), the source of the data shown here.

into smokestack emissions (and ^{202}Hg into ash solids) explains the lighter (-2.6‰) $\delta^{202}\text{Hg}$ in precipitation relative to the value of the coal itself (-0.72‰). The partitioning of the lighter isotope into the vapor is a typical consequence of mass-dependent fractionation.

10.8.4 $\delta^{26}\text{Mg}$ and $\delta^{44/42}\text{Ca}$

Magnesium and calcium are both major constituents of silicate and carbonate rocks as well as rivers, lakes, sea water, groundwater and (to varying extents) plants and other organisms. Isotopes of magnesium and calcium have proven useful in assessing chemical weathering, changes over geological time to the composition of sea water, and the uptake of Ca by plants. The abundances of Mg isotopes are given in Table 10.2; Ca has three stable isotopes, ^{44}Ca (0.65%), ^{42}Ca (2.01%) and ^{40}Ca (96.9%) as well as some short-lived radioactive ones. Figure 10.17 depicts known ranges of the $^{26/24}\text{Mg}$ and $^{44/42}\text{Ca}$ isotopes – note

the distinct $\delta^{26}\text{Mg}$ signal in carbonate rocks vs silicate rocks and medium to large rivers, and partitioning of light Mg and Ca isotopes into plants.

Ratios of $^{44/42}\text{Ca}$ and $^{44/40}\text{Ca}$ may be used to assess Ca cycling. Plants preferentially absorb ^{40}Ca from soil waters (Figs. 10.17 and 10.18), leaving the soil solution enriched in ^{44}Ca and ^{42}Ca .

In order to examine sources and long-term availability of exchangeable Ca in an acidic forest soil, Farkaš et al. (2011) analyzed $\delta^{44/42}\text{Ca}$ and $\delta^{44/40}\text{Ca}$ in plant tissue (red oak), Ca in exchange sites, Ca in particulate organic matter (POM) and other soluble phases, and Ca derived from biotite in the C horizon/mineral soil. Their results showed that Ca isotope values of exchangeable Ca are on a mixing line between organic soil values and biotite-derived Ca isotope values (visible in Fig. 10.18), indicating that exchangeable Ca (up to 25% of total Ca_{exch} in this study) is derived from weathering of biotite (Ca derived from biotite interlayers). The biotite Ca signature is not found in older plant matter, suggesting that this deeper Ca source may be increasing with time

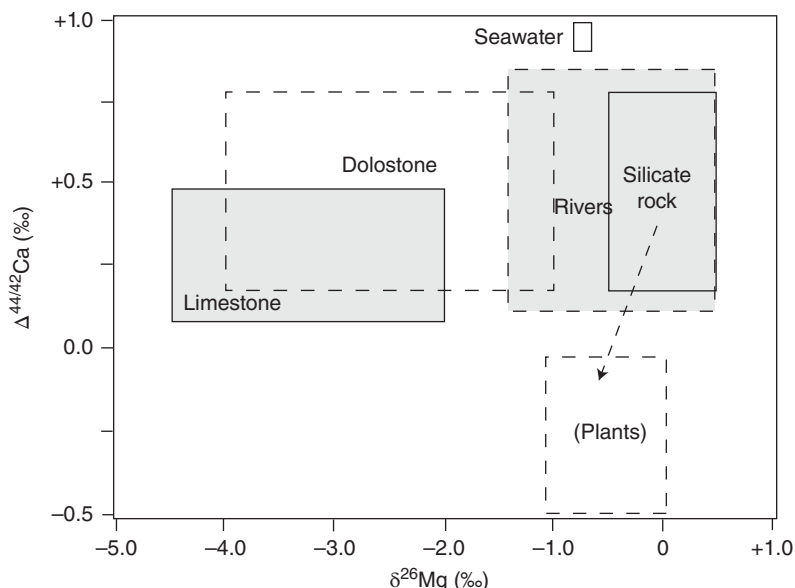


Fig. 10.17 Ranges of $\delta^{26}\text{Mg}$ and $\delta^{44/42}\text{Ca}$ values for inorganic reservoirs and plants (data from Young and Galy, 2004; Tipper, et al. 2006). Note that carbonate rocks are depleted in ^{26}Mg relative to silicate rocks. Uptake of Mg and Ca by plants fractionates the lighter isotopes into the plant, as shown by the arrow. Relative contributions of carbonate vs. clastic inputs to river dissolved or particulate loads can be assessed using Ca and Mg stable isotopes (e.g. Tipper 2008).

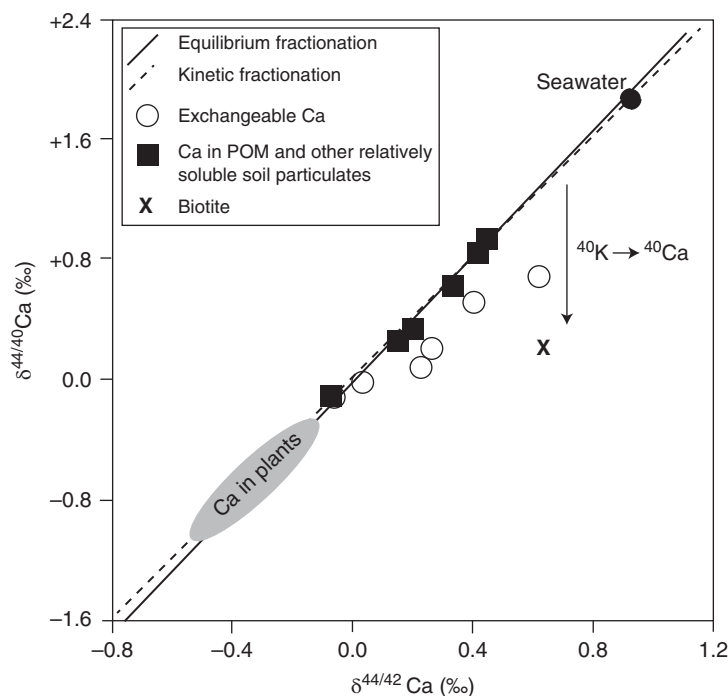


Fig. 10.18 Fractionation of Ca isotopes in a forested ecosystem. Note ^{40}Ca -enriched signal in plant tissue (red oak), and also the effect of radioactive decay on the $^{44}/^{40}\text{Ca}$ ratio. Where ^{40}K is present (e.g. in biotite), the ratio of $^{44}/^{40}\text{Ca}$ will decrease with time due to $^{40}\text{K} \rightarrow ^{40}\text{Ca}$ decay.

as Ca is depleted in the organic soil horizons (upper 15 cm).

10.8.5 $\delta^{37/35}\text{Cl}$

Chlorine has two naturally occurring stable isotopes, ^{37}Cl (24.23%) and ^{35}Cl (75.77%). Radioactive ^{36}Cl occurs in very trace amounts as a cosmogenic radionuclide and is covered in Chapter 11. The standard used in analysis is standard mean ocean chloride (SMOC) and the delta notation for $^{37}\text{Cl} : ^{35}\text{Cl}$ is $\delta^{37}\text{Cl}_{\text{SMOC}}$. One field to which stable Cl isotopes have been applied is decomposition of trichloroethylene (TCE) to degradation products (e.g. dichloroethylene, DCE, Fig. 3.25). Given the higher vibrational energy of $\text{C}-^{35}\text{Cl}$ bonds compared to $\text{C}-^{37}\text{Cl}$ bonds, ^{35}Cl should be preferentially released as TCE begins to decay to DCE, and this is what has been observed in an experimental study (Numata et al, 2002) as well as a field-based study (Sturchio et al., 1998). When TCE breaks down, initial chlorine released is enriched in ^{35}Cl , and the remaining DCE becomes enriched in ^{37}Cl (Fig. 10.19).

Zones of enhanced degradation, notably organic-rich clay-silt lenses, were characterized by increased concentration of the chloride anion (Cl^-) with an anomalously low $\delta^{37}\text{Cl}$ value, reflecting release of ^{35}Cl -enriched chlorine from TCE.

10.9 SUMMARY

The traditional or conventional stable isotopes of C, H, O, N and S have been studied for decades, providing valuable information on topics ranging from paleotemperature to nutrient cycling and contaminant sources. The development of methods and paradigms in the field of “non-traditional” stable isotope analysis, including lighter elements such as Li, B, Mg, Ca and Cl, as well as metal stable isotopes, is still in its early stages and much work remains to be done before the full significance of its application to environmental geochemistry can be appreciated. At the same time, these elements are critical to environmental geochemistry in many ways: iron controls the cycling of many other trace elements and organic compounds, Cr,

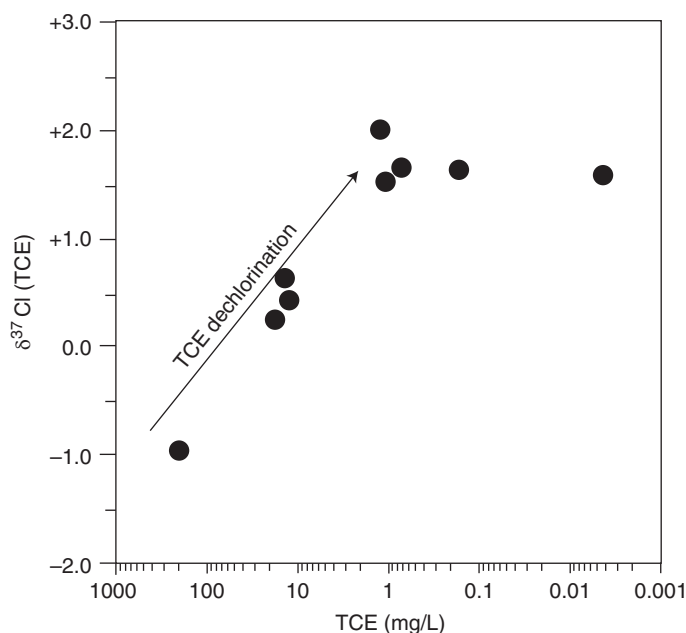
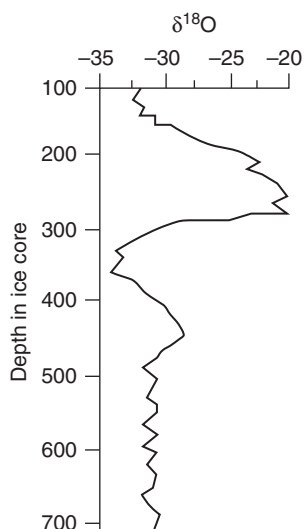


Fig. 10.19 Preferential release of ^{35}Cl during microbial degradation of TCE to DCE causes $\delta^{37}\text{Cl}$ of the Cl remaining in the metabolite DCE to increase. (Modified from Sturchio et al. 1998.)

Hg and Mo are potential contaminants of soil and groundwater, Cl is a component of many toxic organic compounds, and Cu and Mn are tracers of geochemical processes and vital plant trace nutrients, so further research into the significance of metal stable isotopes likely will prove to be an important approach to understanding the environment into the 21st century.

QUESTIONS

1. What is the $^{18}\text{O}/^{16}\text{O}$ ratio of a sample that has a $\delta^{18}\text{O}_{\text{SMOW}}$ value of +23‰?
2. For a mass of water vapor at 25 °C with an initial $\delta^{18}\text{O}$ of -12‰, determine the $\delta^{18}\text{O}$ of the precipitation that will condense, and the vapor remaining, after 50% of the initial vapor has condensed.
3. $\delta^{18}\text{O}$ values for oxygen in $\text{H}_2\text{O}(\text{s})$ sampled from an ice core are shown in the figure below (depth is in m). Which depth (or depth range) corresponds to the warmest temperatures, and which is indicative of the coolest temperatures? Explain in ~2–3 sentences the fundamental process that controls the variation in the oxygen isotope compositions of ice cores.
4. Why does isotopic fractionation decrease with increasing temperature?
5. What is the likely origin of elevated nitrate in an aquifer where the nitrate has a $\delta^{15}\text{N}$ value of +2.3‰ and a $\delta^{18}\text{O}$ value of +17.8‰? (both relative to SMOW). Explain how this source of nitrate gets its signature.
6. Provide three examples of how metal stable isotopes are applied to environmental analysis. Why has the development of this field lagged behind the more traditional analysis of δD , $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$ and $\delta^{18}\text{S}$?
7. A bedrock aquifer in pyrite-bearing, organic-rich black slate has numerous wells with elevated arsenic (30% of wells contain $\text{As} > 10 \mu\text{g/L}$). Dissolved Fe, SO_4 and As are positively correlated, and the $\delta^{34}\text{S}$ value of dissolved sulfate ranges from +31 to +38‰. The range of values of $\delta^{34}\text{S}$ in pyrite is +32 to +53‰. Suggest an explanation for the high $\delta^{34}\text{S}$ pyrite values; also, provide possible explanation(s) for the origin/cause of the elevated As. Do the geochemical correlations and isotopic data provide insight?



REFERENCES

- Augustin, L., et al., 2004, Eight glacial cycles from an Antarctic ice core. *Nature*, 429(6992), 623–628. doi:10.1038/nature02599.
- Bechtel, A., Gratzner, R., Sachsenhofer, R.F., Gusterhuber, J., Lücke, A., and Püttmann, W., 2008, Biomarker and carbon isotope variation in coal and fossil wood of Central Europe through the Cenozoic. *Palaeogeography Palaeoclimatology Palaeoecology* 262: 166–175.
- Bigeisen, J. and M. G. Mayer, 1947. Calculation of equilibrium constants for isotopic exchange reactions. *Journal of Chemical Physics*, 15: 261–267.
- Biswas, A., Blum, J.D., Bergquist, B.A., Keeler, G. J., and Xie, Z., 2008. Natural mercury isotope variation in coal deposits and organic soils. *Environmental Science and Technology*. 42: 8303–8309.

- Blum, J.D., 2011. Applications of stable mercury isotopes to biogeochemistry. In: Baskaran, M. (ed.), *Handbook of Environmental Isotope Geochemistry, Advances in Isotope Geochemistry*, Springer-Verlag, Berlin.
- Blum, J.D. and Bergquist, B.A., 2007. Reporting of variations in the natural isotopic composition of mercury. *Analytical and Bioanalytical Chemistry* 388: 353–359.
- Blum, J. D., Gazis, C. A., Jacobson, A. D., and Chamberlain, C. P., 1998. Carbonate versus silicate weathering in the Raikhot watershed within the High Himalayan Crystalline Series. *Geology* 26: 411–414.
- Borrok, D.M., Nimick, D.A., Wanty, R.B., and Ridley, W.L., 2008. Isotopic variations of dissolved copper and zinc in stream waters affected by historical mining. *Geochimica Cosmochimica Acta* 72: 329–344.
- Bullen, T.D., 2011. Stable isotopes of transition and post-transition metals as tracers in environmental studies. In *Handbook of Environmental Isotope Geochemistry, Advances in Isotope Geochemistry* (M. Baskaran, ed.), Springer-Verlag Berlin Heidelberg, pp. 177–203.
- Cerling, T.E., Harris, J.M., MacFadden, B.J., Leakey, M. G., Quade, J., Eisenmann, V., and Ehleringer, J.R., 1997. Global vegetation change through the Miocene/Pliocene boundary, *Nature* 389: 153–158.
- Cerling, T.E., Wang, Y., and Quade, J., 1993. Expansion of C₄ ecosystems as indicator of global ecological change in the Late Miocene, *Nature* 361: 344–345.
- Chamberlain, C.P., Poage, M.A., Craw, D., and Reynolds, R.C., 1999. Topographic development of the Southern Alps recorded by the isotopic composition of authigenic clay minerals, South Island, New Zealand. *Chemical Geology* 155: 279–294.
- Clayton, R. N., O'Neil, J. R. and Mayeda, T. K., 1972. Oxygen Isotope Exchange between Quartz and Water, *Journal of Geophysical Research*, 77(17): 3057–3067.
- Clayton, R.N., Friedman, I., Graf, D.L., Mayeda, T.K., Meents, and W.F., Shimp, N.F., 1966. The origin of saline formation waters. I. Isotopic composition. *Journal of Geophysical Research* 71: 3869–3882.
- Craig, H., 1961. Isotopic variations in meteoric waters. *Science*, 133: 1702–1703.
- Craig, H., 1966. Isotopic composition and origin of the Red Sea and Salton Sea geothermal brines, *Science* 154: 1544–1548.
- D'Amore, F., Fancelli, R., and Panichi, C., 1987. Stable isotope study of reinjection processes in the Larderello geothermal field. *Geochimica et Cosmochimica Acta* 51: 857–867.
- Dansgaard, W., 1964. Stable isotopes in precipitation. *Tellus* 16: 436–468.
- Drever, J.I., 1997. *The Geochemistry of Natural Waters: Surface and Groundwater Environments*, 3rd edn. Simon and Schuster, Upper Saddle River, New Jersey, USA. Europe through the Cenozoic. *Palaeogeography Palaeoclimatology Palaeoecology* 262: 166–175.
- Farkaš, J., Déjeant, A., Novák, M., and Jacobsen, S.B., 2011. Calcium isotope constraints on the uptake and sources of Ca²⁺ in a base-poor forest: A new concept of combining stable ($\delta^{44}/^{42}\text{Ca}$) and radiogenic (ϵ_{Ca}) signals. *Geochimica et Cosmochimica Acta* 75: 7031–7046.
- Faure, G., 1986. *Principles of Isotope Geology* (2nd edn). Wiley, New York.
- Faure, G., 1998. *Principles and Applications of Geochemistry* (2nd edn). Prentice-Hall, Upper Saddle River, New Jersey, USA.
- Friedman, I. and O'Neil, J.B., 1977. Compilation of stable isotope fractionation factors of geochemical interest. In: *Data of Geochemistry*, 6th edn, ed. M. Fleischer, Chapter KK. US Geol. Surv. Prof. Paper 440-KK.
- Fujii, T., Moynier, F., Abe, M., Nemoto, K., and Albarède, F., 2013. Copper isotope fractionation between aqueous compounds relevant to low temperature geochemistry and biology. *Geochimica et Cosmochimica Acta* 110: 29–44.
- Hatch, M.D. and Slack C.R., 1966. Photosynthesis in sugarcane leaves: a new carboxylation reaction and the pathway of sugar formation. *Biochemical Journal* 101: 103–111.
- Hoefs, J., 2009. *Stable Isotope Geochemistry* (6th edn). Springer, Berlin. 285 pp.
- Horibe, Y. and Oba, T., 1972. Temperature scales of aragonite-water and calcite-water systems. *Fossils* 23/24: 69–79.
- Kawabe, I., 1978. Calculation of oxygen isotope fractionation in quartz-water system with special reference to the low temperature fractionation. *Geochimica et Cosmochimica Acta* 42: 613–621.
- Kendall, C., 1988. Tracing Nitrogen Sources and Cycling in Catchments. In *Isotope Tracers in Catchment Hydrology*, C. Kendall and J. J. McDonnell (eds.). Elsevier Science, Amsterdam, The Netherlands, pp. 519–576. <http://www.rcamnl.wr.usgs.gov/isoig/isopubs/itchch16.html#16.3.3>.

- Kendall, C. and Doctor, D.H., 2003. Stable Isotope Applications in Hydrologic Studies. In *Treatise on Geochemistry*, Volume 5. ed.: J. I. Drever. Executive Editors: H. D. Holland and K. K. Turekian, pp.319–364. ISBN 0-08-043751-6. Elsevier, Amsterdam, The Netherlands, 605 pp.
- Kendall, C., Sklash, M. G., and Bullen, T. D., 1995. Isotope Tracers of Water and Solute Sources in Catchments. In *Solute Modelling in Catchment Systems*, John Wiley and Sons, New York, pp. 261–303.
- Kimball, B.E., Mathur, R., Dohnalkova, A.C., Wall, A.J., Runkel, R.L., and Brantley, S.L., 2009. Copper isotope fractionation in acid mine drainage. *Geochimica et Cosmochimica Acta* 73: 1247–1263.
- Kroopnick, P. and Craig, H., 1972. Atmospheric Oxygen: Isotopic Composition and Solubility Fractionation. *Science*, 175, 4017: pp. 54–55. DOI: 10.1126/science.175.4017.54.
- Langmuir, D., 1997, *Aqueous Environmental Geochemistry*, Prentice-Hall, Upper Saddle River, New Jersey, USA. 600 p.
- Laurent, A. et al. 2004. Eight glacial cycles from an Antarctic ice core. *Nature*, 429(6992): 623–628, doi:10.1038/nature02599.
- Meckenstock, R.U., Morasch, B., Griebler, C., and Richnow, H.H., 2004. Stable isotope fractionation analysis as a tool to monitor biodegradation in contaminated aquifers. *Journal of Contaminant Hydrology* 75 (3–4): 215–55.
- Moeller, K., Schoenberg, R., Pedersen, R.B., Weiss, D., and Dong, S.F., 2012. Calibration of the new certified reference materials ERM-AE633 and ERM-AE647 for copper and IRMM-3702 for zinc isotope amount ratio determinations, *Geostandards and Geoanalytical Research* 36: 177–199.
- Novák, M., Kirchner, J. W., Fottová, D., Prechová, E., Jäcková, I., Krám, P., and Hruska, J., 2005. Isotopic evidence for processes of sulfur retention/release in 13 forested catchments spanning a strong pollution gradient (Czech Republic, central Europe), *Global Biogeochemical Cycles*, 19: GB4012, doi:10.1029/2004GB002396.
- Nriagu J. O. and Coker R. D., 1978. Isotopic composition of sulfur in precipitation within the Great Lakes Basin. *Tellus* 30: 365–375.
- Numata, M., Nakamura, N., Koshikawa, H., J. Terashima, Y., 2002. Chlorine Isotope Fractionation during Reductive Dechlorination of Chlorinated Ethenes by Anaerobic Bacteria. *Environmental Science and Technology* 36: 4389–4394.
- O'Neil, J.R. and Epstein, S., 1966. Oxygen isotope fractionation in the system dolomite-calcite-carbon dioxide. *Science* 152: 198–201.
- Ohizumi, T., Fukuzaki, N. and Kusakabe, M., 1997, Sulfur isotopic view on the sources of sulfur in atmospheric fallout along the coast of the Sea of Japan. *Atmospheric Environment* 31 (9): 1339–1348.
- Palmer, M.R. and Slack, J.F., 1989. Boron isotopic composition of tourmaline from massive sulfide deposits and tourmalinites. *Contributions to Mineralogy and Petrology* 103: 434–451.
- Petit, J.R., Jouzel, J., Raynaud, D., Barkov, N.I., Barnola, J. M., Basile, I., Bender, M., Chappellaz, J., Davis, J., Delaygue, G., Delmotte, M., Kotlyakov, V.M., Legrand, M., Lipenkov, V., Lorius, C., Pepin, L., Ritz, C., Saltzman, E., and Stievenard, M., 1999. Climate and Atmospheric History of the Past 420,000 years from the Vostok Ice Core, Antarctica. *Nature* 399: pp. 429–436.
- Rayleigh, J.W., 1896. Theoretical considerations respecting the separation. *Rozanski, K., Araguas-Araguas, L., Giorantini, R., 1993. Isotopic patterns of gases by diffusion and similar processes. London Edinburgh Dublin Philosophical Magazine* 42: 493–498.
- Rozanski, K., Araguas-Araguas, L., and Giorantini, R., 1993. Isotopic patterns in modern global precipitation. *Geophysical Monograph* 78: 1–36.
- Savin, S. M. and Lee, M., 1988. Isotopic studies of phyllosilicates. In: *Hydrous Phyllosilicates (Exclusive of Micas)* (ed. S. W. Bailey), Reviews in Mineralogy, Vol. 19, Mineralogical Society of America, Washington, DC, pp. 189–223.
- Schroth, M.N., Kleikemper, J., Bolliger, C., Bernasconi, S. M., Zeyer, J., 2001. In situ assessment of microbial sulfate reduction in a petroleum-contaminated aquifer using push–pull tests and stable sulfur isotope analyses. *Journal of Contaminant Hydrology* 51(3–4): 179–195.
- Sharp, Z., 2007. *Principles of Stable Isotope Geochemistry*. Pearson Prentice Hall, New York, 344 pp.
- Sherman, L.S., Blum, J.D., Keeler, G.J., Demers, J.D., and Dvonch, J.T., 2012. Investigation of local mercury deposition from a coal-fired power plant using mercury isotopes. *Environmental Science and Technology* 46: 382–90. doi: 10.1021/es202793c.
- Stern, L.A., Johnson, G.D., and Chamberlain, C.P., 1994. Carbon isotope signature of environmental

- change found in fossil ratite eggshells from a South Asian Neogene sequence. *Geology* 22: 419–422.
- Sturchio, N.C., Clausen, J.L., Heraty, L.J., Huang, L., Holt, B.D., and Abrajano, T.A., 1998. Chlorine isotope investigation of natural attenuation of trichloroethene in an aerobic aquifer. *Environmental Science and Technology* 32: 3037–3042.
- Sturges W.T. and Barrie L.A., 1989. The use of stable lead 206/207 isotope ratios and elemental composition to discriminate the origin of lead in aerosols at a rural site in eastern Canada. *Atmospheric Environment* 23: 1645–1657.
- Thompson, C.M., 2013. A solvent extraction technique for the isotopic measurement of dissolved copper in seawater. *Analytica Chimica Acta* 775: 106–113.
- Tipper, E.T., Gaillardet, J., Galy, A., and Louvat, P., 2008. Ca isotope ratios in the largest rivers in the world: Implications for the global Ca budget and weathering processes. *Geochimica et Cosmochimica Acta* 72: A947–A947.
- Tipper, E.T., Galy, A., Gaillardet, J., Bickle, M.J., Elderfield, H., and Carder, E.A., 2006. The magnesium isotope budget of the modern ocean: Constraints from riverine magnesium isotope ratios. *Earth and Planetary Science Letters* 250: 241–253.
- Urey, H., 1947. The thermodynamic properties of isotopic substances. *Journal of the Chemical Society (London)*, 562–581.
- Wadleigh M. A., Schwarcz H. P., and Kramer J. R., 1996. Isotopic evidence for the origin of sulfate in coastal rain. *Tellus* 48B: 44–59.
- Wadleigh, M.A. and Blake, D.M., 1999. Tracing sources of atmospheric sulfur using epiphytic lichens. *Environmental Pollution* 106 (3): 265–271.
- Wiederhold, J.G., Kraemer, S.M., Teutsch, N., Borer, P. M., Halliday, A.N., and Kretzschmar, R., 2006. Iron isotope fractionation during proton-promoted, ligand-controlled, and reductive dissolution of goethite. *Environmental Science & Technology* 40: 3787–3793.
- World Health Organization, 2004. *Guidelines for Drinking Water Quality*, 3rd edn Geneva, Switzerland: World Health Organization.
- Young, E. and Galy, A., 2004. The isotope geochemistry and cosmochemistry of Mg. *Reviews in Mineralogy and Geochemistry* 55: 197–230.
- Yurtsever, Y., Gat, J.R., 1981. Atmospheric waters. In: *Stable isotope hydrology: Deuterium and oxygen-18 in the water cycle*. Technical Report Series 210: 103–142. IAEA, Vienna.

Radioactive and Radiogenic Isotope Geochemistry

Radioactive isotopes, whether naturally occurring (e.g. ^{14}C , ^{235}U and ^{238}U) or synthetic (e.g. ^{99}Tc , ^{239}Pu), are characterized by atoms with unstable nuclei that undergo radioactive decay to **daughter** isotopes, which, because they form by radioactive decay, are termed **radio-genic**. These daughter products also may be radioactive, or they may be stable (e.g. decay of radioactive ^{40}K to stable ^{40}Ca mentioned at the end of Chapter 10). Radioactive and radiogenic isotopes are useful in environmental analysis in much the same way that stable isotopes are – some are useful tracers of geochemical and biochemical reactions – and radioactive-radiogenic isotopes are the most common approach to quantitative dating of rocks, sediments, soils, groundwater and organic matter. Some radioactive elements, whether naturally occurring or anthropogenic, are toxic, so geochemical research has also focused on the behavior of various radionuclides in environmental systems.

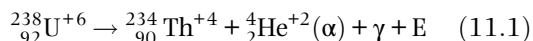
This chapter begins with some fundamentals of radioactive decay and radioactive-radiogenic elements, then moves on to examples of radionuclides applied to environmental and low-temperature geochemistry.

11.1 RADIOACTIVE DECAY

The nuclei of **radioactive isotopes** undergo spontaneous decay to **daughter atoms** (which may be stable or radioactive) via various mechanisms, all of which release energy and particles, and regardless of the mechanism, radioactive decay produces a change in both Z (# of protons) and N (# of neutrons) from parent to daughter isotope. Radioactive decay occurs by emission of three types of radiation known as alpha (α), beta (β) and gamma (γ) rays, the presence of which were discovered by Ernest Rutherford in 1899 and 1900 from groundwork laid by Henri Becquerel, Marie Curie and Pierre Curie in the late 1800s.

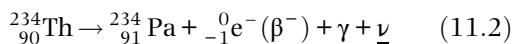
11.1.1 Decay mechanisms and products

Alpha radiation takes the form of a particle that essentially is a helium nucleus, with 2 protons and 2 neutrons and a charge of +2, and is mainly produced when atoms of atomic number ≥ 58 undergo radioactive decay. A good example of alpha decay is the transformation of ^{238}U , the most abundant isotope of uranium on the Earth, to ^{234}Th (thorium).

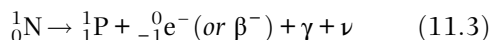


Note that alpha decay causes a reduction of 2 protons, resulting in transformation from a uranium nucleus with $Z = 92$ to thorium with $Z = 90$, and the loss of two neutrons in concert with the two protons means that atomic mass is reduced by 4 from 238 to 234 Da. In addition to the emission of an alpha particle, gamma radiation (γ) is also released, as is kinetic energy (E) associated with the fast-moving α particle plus the recoil energy imparted to the nucleus by the ejection of the α particle.

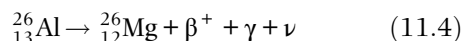
Because it is the product of a radioactive decay reaction, ^{234}Th is said to be **radiogenic**. It is also radioactive and will undergo further spontaneous transformation via **beta decay** to ^{234}Pa :



The decay reaction here results in a +1 increase in Z and no change in atomic mass, so the increase in Z must be balanced by the loss of a neutron. Viewed in this light, beta decay is the transformation (in the nucleus) of a neutron to a proton, and occurs by the spontaneous emission of an electron from the nucleus (represented below by $^0_{-1}\text{e}^{-}$ or β^{-}) plus a gamma ray (γ) and antineutrino ($\bar{\nu}$), both of which are sub-atomic particles that are also emitted during beta decay:



Radioactive decay also can occur by emission of a positively charged electron (a positron) from the nucleus, a process in which a proton is converted to a neutron plus a positron (and a gamma ray and neutrino, ν). This type of decay is known as **positron decay**, positive beta decay or beta-plus (β^{+}) decay.



where the β^{+} particle has a +1 charge and no mass and represents the transformation of a proton to a neutron.

As you might have surmised from the equations presented above, **gamma radiation** is commonly emitted during alpha or beta decay. While the emission of gamma rays does not alter the number of protons or neutrons in the nucleus, it does help to stabilize the nucleus in its spontaneous progression from a higher, less-stable state to a lower, more-stable energy state.

Because gamma rays possess a very short wavelength ($< 0.1 \text{ nm}$), they are the most energetic form of electromagnetic radiation and can penetrate skin, bone and other living tissues as well as thin sheets of metal, glass and wood. By contrast, alpha particles can be deflected by a sheet of paper and will not penetrate skin, and only tend to be problematic to living organisms when they are inhaled or ingested. The high-energy beta particles (or rays) are more penetrative than alpha particles (they can penetrate skin, for example), but far less so than gamma rays. Like alpha particles, risk to living organisms from beta radiation is greatest when inhaled or ingested.

Radioactive decay can also occur by **spontaneous fission**, where the nucleus is split into two or more new nuclei that result in two or more new atoms of approximately equal size (unlike alpha decay, which produces an α particle that is very small compared to the remaining nucleus). This process mainly applies to synthetic reactions

like those found in nuclear reactors, but a few heavy naturally occurring isotopes such as ^{238}U , ^{235}U and ^{232}Th can also undergo spontaneous decay via fission. However, the incredibly slow rate (half-lives are $\geq 10^9$ yr) relative to decay rates of alpha and beta decay mechanisms means that spontaneous fission is an insignificant natural process relative to the others. The notable exception in the geological past were the natural fission chain reactions that took place in U-rich rocks in Gabon during Proterozoic time (~ 2 Ga) (e.g. Gauthier-Lafaye et al., 1996).

11.1.2 Half-lives, decay rates and decay constants

An important consideration related to radioactive isotopes is the decay rate, which can be expressed in terms of the **half-life** ($t_{1/2}$) of the isotope. One half-life is the amount of time required to transform one-half of the mass of a radioactive isotope to its daughter product – after one half-life, 50% of the original radioactive isotope will remain, and the other 50% of it will have been transformed to its daughter product. Examples of values for half-lives of isotopes applied in geochemical analyses are presented in Table 11.1. Note that values range from very short ($t_{1/2}$ for $^{218}\text{Po} = 3.05$ min) and relatively short ($t_{1/2}$ for $^{222}\text{Rn} = 3.8$ days) to intermediate ($t_{1/2}$ for $^{14}\text{C} = 5730$ yr) and incredibly long ($t_{1/2}$ for $^{238}\text{U} = 4.5 \times 10^9$ yr).

The exponential nature of radioactive decay was presented in Chapter 1 as an example of a first-order reaction. Here, it is presented mathematically (below) and graphically (Fig. 11.1). The decay of a parent isotope (P) to a daughter isotope (D) is described as:

$$-dP/dt = \lambda_p * P \quad (11.5)$$

where P is the amount of parent isotope remaining at time t and λ_p is the decay constant associated with the parent isotope. The decay constant

represents the probability that a radioactive atom will undergo spontaneous decay, and the units of λ are reciprocal time (e.g. yr^{-1}).

If we rearrange the above equation, we can arrive at:

$$-1/P * dP = \lambda_p * dt \quad (11.6)$$

Integrating both sides gives us:

$$-\ln(P) = \lambda_p * t + C \quad (11.7)$$

where C is the constant of integration. If we consider this equation at time $= 0$, then the term $\lambda_p * t = 0$, and $C = -\ln(P_0)$, where P_0 represents the amount of parent isotope at time $t = 0$. Substituting $-\ln(P_0)$ for C in the equation above, we now have:

$$-\lambda_p * t = \ln(P) - \ln(P_0) \quad (11.8)$$

And if we take the exponential of both sides, we derive the following:

$$e^{-\lambda_p * t} = P/P_0 \quad (11.9)$$

and with one more round of rearranging, we can express the amount of parent isotope remaining at time t as:

$$P = P_0 * e^{-\lambda_p * t} \quad (11.10)$$

After one half-life, 50% of the original parent material remains, and the amount of daughter isotope that has been produced will equal the amount of parent material unless the daughter isotope is radioactive and also being depleted by radioactive decay. After two half-lives, 25% of the parent isotope remains, and after three half-lives, 12.5% remains, and so on. This type of exponential decay is described by the equation $P = P_0 * e^{-\lambda_p * t}$, where the decay constant λ is related to the half-life. After one half-life, $P = 0.5 * P_0$ (the amount of parent material is

Table 11.1 Decay mechanisms, half-lives and energies for numerous radionuclides.

Parent Isotope	Decay Mechanism	Half-Life	E (MeV)	Daughter Isotope
²³² Th	alpha	1.405 * 10 ¹⁰ a	4.081	²²⁸ Ra
²²⁸ Ra	beta	5.75 a	0.046	²²⁸ Ac
²²⁸ Ac	beta	6.25 h	2.124	²²⁸ Th
²²⁸ Th	alpha	1.9116 a	5.52	²²⁴ Ra
²²⁴ Ra	alpha	3.6319 d	5.789	²²⁰ Rn
²²⁰ Rn	alpha	55.6 s	6.404	²¹⁶ Po
²¹⁶ Po	alpha	0.145 s	6.906	²¹² Pb
²¹² Pb	beta	10.64 h	0.57	²¹² Bi
²¹² Bi	beta 64.6%	60.55 min	2.252	²¹² Po
	alpha 35.94%		6.208	²⁰⁸ Tl
²¹² Po	alpha	299 ns	8.955	²⁰⁸ Pb
²⁰⁸ Tl	beta	3.053 min	4.999	²⁰⁸ Pb
²⁰⁸ Pb	—	stable	—	—
²³⁸ U	alpha	4.468 * 10 ⁹ a	4.27	²³⁴ Th
²³⁴ Th	beta	24.10 d	0.273	²³⁴ Pa
²³⁴ Pa	alpha	6.70 h	2.197	²³⁴ U
²³⁴ U	beta	245500 a	4.859	²³⁰ Th
²³⁰ Th	alpha	75380 a	4.77	²²⁶ Ra
²²⁶ Ra	alpha	1602 a	4.871	²²² Rn
²²² Rn	alpha	3.8235 d	5.59	²¹⁸ Po
²¹⁸ Po	alpha 99.98%	3.10 min	6.115	²¹⁴ Pb
	beta 0.02%		0.265	²¹⁸ At
²¹⁸ At	alpha 99.90%	1.5 s	6.874	²¹⁴ Bi
	beta 0.10%		2.883	²¹⁸ Rn
²¹⁸ Rn	alpha	35 ms	7.263	²¹⁴ Po
²¹⁴ Pb	beta	26.8 min	1.024	²¹⁴ Bi
²¹⁴ Bi	beta 99.98%	19.9 min	3.272	²¹⁴ Po
	alpha 0.02%		5.617	²¹⁰ Tl
²¹⁴ Po	alpha	0.1643 ms	7.883	²¹⁰ Pb
²¹⁰ Tl	beta	1.30 min	5.484	²¹⁰ Pb
²¹⁰ Pb	beta	22.3 a	0.064	²¹⁰ Bi
²¹⁰ Bi	beta 99.99987%	5.013 d	1.426	²¹⁰ Po
	alpha 0.00013%		5.982	²⁰⁶ Tl
²¹⁰ Po	alpha	138.376 d	5.407	²⁰⁶ Pb
²⁰⁶ Tl	beta	4.199 min	1.533	²⁰⁶ Pb
²⁰⁶ Pb	—	stable	—	—
Plutonium and Uranium Fission Daughter Products				
¹³⁵ I		6.57 h		
¹³⁵ Xe		9.14 h		
¹³⁵ Cs		2.3 * 10 ⁶ a		
⁹⁹ Tc		2.11 * 10 ⁵ a		

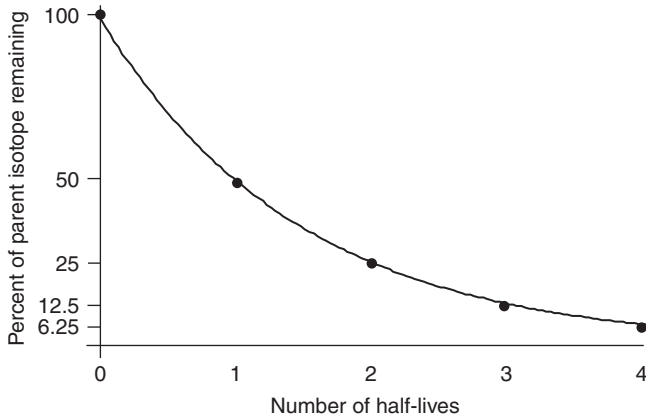


Fig. 11.1 Graphical representation of exponential decay.

equal to 50% of the original amount of parent isotope), and the relationship between half-life and decay constant is given by the following equations:

$$t_{1/2} = \ln(P_0/0.5 * P_0) * 1/\lambda = \ln 2/\lambda = 0.693/\lambda \quad (11.11)$$

So, for example, λ for the decay of ^{14}C to ^{14}N (by β^- decay) is $1.209 \times 10^{-4} \text{ yr}^{-1}$, and the half-life is 5730 yr. For ^{238}U , λ for the α decay to ^{234}Th is $1.551 \times 10^{-10} \text{ yr}^{-1}$, and the half-life is $4.47 \times 10^9 \text{ yr}$ (Faure, 1986).

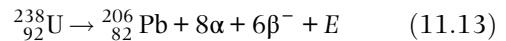
The determination of the radioactivity of a sample is typically made by measuring the **activity** of the material rather than by expressing radioactivity in terms of individual atoms, where activity is defined as:

$$A = c * \lambda * N \quad (11.12)$$

The activity A is the counting rate measured by a Geiger counter or similar detector, λ is the decay constant and N is the number of radioactive atoms remaining in the sample. The product of $\lambda * N$ is the actual decay rate, and c is a coefficient used to account for the effectiveness of the detector used in analysis (i.e. c is a multiplier applied to correct for the accuracy of the detector).

The **radioactive decay series** that begins with ^{238}U (Fig. 11.2) is an excellent example of

the varied mechanisms and half-lives that govern radioactive decay. Note that alpha decay depicts reduction of two protons and 2 neutrons, and that beta decay is shown as gain of a proton and loss of a neutron. The net reaction is:



You will observe that there are varied pathways to arrive at ^{206}Pb from ^{238}U , but regardless of the pathway, the gross quantity of radiation emitted is always eight alpha particles and 6 beta particles. The total energy emitted is 47.4 MeV per atom (Faure, 1986).

Two important radionuclides in environmental geochemistry that occur in the ^{238}U decay series are ^{226}Ra (radium), which is a groundwater contaminant in some aquifers (both from natural mineral decay and from anthropogenic sources such as U mining) and ^{222}Rn , which is the isotope of radon that is a common indoor air contaminant in areas with bedrock enriched in ^{238}U . Granites and related rocks and soils are the most common source of ^{222}Rn .

Half-lives and decay mechanisms of radionuclides in the ^{238}U and ^{232}Th series are presented in Table 11.1.

The ^{232}Th decay series (Fig. 11.3) is similar to that of ^{238}U with numerous alpha decay reactions, and the sum of the series from ^{232}Th to ^{208}Pb is:

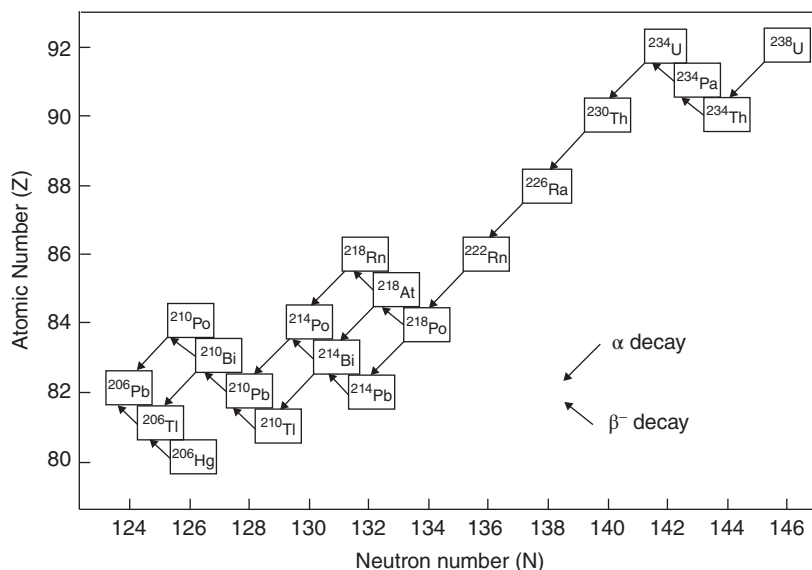


Fig. 11.2 Schematic representation of the decay series of ^{238}U .

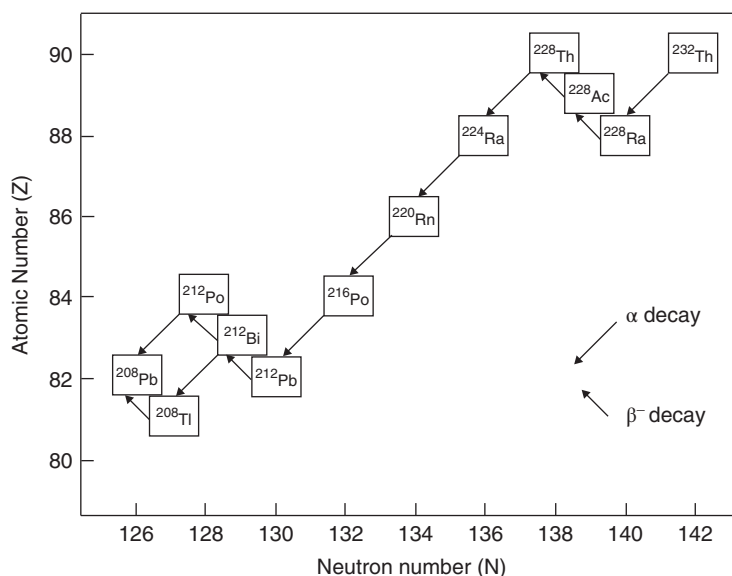
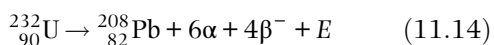


Fig. 11.3 Schematic representation of the decay series of ^{232}Th .



The total energy released is 39.8 MeV (Faure, 1986). Like ^{226}Ra , ^{228}Ra is also a groundwater contaminant (see below).

The ^{238}U , ^{235}U and ^{232}Th decay series each results in the formation of **stable isotopes of**

lead, ^{206}Pb in the case of ^{238}U decay, ^{207}Pb in the case of ^{235}U decay and ^{208}Pb in the case of ^{232}Th decay. Terminologically, these are stable, radiogenic isotopes of lead. Radioactive decay of uranium and thorium to lead over geological time helps to explain the relative abundance of lead in the crust of the Earth relative to other

heavy metals of similar atomic number (e.g. Bi, Hg, Pt). An example of Pb isotopes applied to deciphering the source of Pb in air pollution is presented in Section 11.2.

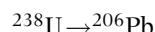
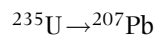
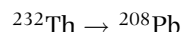
For an interesting and well-written history of the discovery of radioactive and stable isotopes as well as a very thorough presentation of detailed decay mechanisms and the application of isotopes (both stable and radioactive) to geochemistry, the reader is referred to *Principles of Isotope Geology* by Gunter Faure (Faure, 1986).

11.2 RADIONUCLIDES AS TRACERS IN ENVIRONMENTAL GEOCHEMISTRY

While radioactive isotopes often are thought of as contaminants in soil, air and water, many naturally occurring and synthesized radioisotopes are applied in studies of Earth surface systems and shallow crustal systems to measure age (e.g. dating of groundwater, rocks or sediments) or as tracers of hydrological (e.g. groundwater flow paths), mineralogical and biochemical processes. The following section presents some commonly applied radioactive isotopes (and isotope pairs) that serve as examples of the great breadth of isotopic analysis of geological–geochemical systems. This is followed by a section that examines selected radionuclides as contaminants in soils and groundwater.

11.2.1 $^{206}\text{Pb}/^{207}\text{Pb}$

Radiogenic lead isotopes, particularly the ratio $^{206}\text{Pb}/^{207}\text{Pb}$, have been used to determine the source of atmospheric lead contamination. Lead has four naturally occurring stable isotopes, three of which are produced by decay of uranium or thorium (Table 11.1; Figures 11.2 and 11.3):



The fourth, ^{204}Pb , is primordial and formed during nucleosynthesis. In the late 1960s, Murozumi et al. (1969) determined that the concentration of lead in snow and ice in Greenland had increased by a factor of 200 since pre-industrial times, and Rosman et al. (1993) sought to determine the origin of the lead contamination through an analysis of lead isotope compositions. Previous work by geochemists such as Chow et al. (1975), Shirahata et al. (1980) and Sturges and Barrie (1989) had determined that $^{206}\text{Pb}/^{207}\text{Pb}$ ratios differ by region or country depending on the source of industrial lead, some of which was an additive in gasoline (as alkyl lead) that was a major source of atmospheric lead prior to the 1980s, and some of which is added to the atmosphere by ore refining and other industrial processes. $^{206}\text{Pb}/^{207}\text{Pb}$ ratios from sites suspected by Rosman et al. (1993) to be potential sources of lead in the Greenland environment are as follows: for Canada and Europe-Asia, $^{206}\text{Pb}/^{207}\text{Pb} = 1.15 \pm 0.01$; for the USA, $^{206}\text{Pb}/^{207}\text{Pb}$ ranged from 1.15 (pre-1970) to 1.21 (post-1970). The reason for the shift in US $^{206}\text{Pb}/^{207}\text{Pb}$ ratio is due to a switch to an industrial lead source (Mississippi Valley-type lead) that is more enriched in ^{206}Pb relative to other lead sources.

Results show that pre-1970 $^{206}\text{Pb}/^{207}\text{Pb}$ ratio values in Greenland snow and ice are 1.16 ± 0.1 , but that the lead signature underwent a pronounced increase in the $^{206}\text{Pb}/^{207}\text{Pb}$ ratio, culminating in an average value of 1.20 in 1976. Following the peak in $^{206}\text{Pb}/^{207}\text{Pb}$ in 1976, values progressively declined throughout the late 1970s and 1980s, reaching a value of 1.15 ± 0.01 in 1989 (Rosman et al., 1993). The increase in $^{206}\text{Pb}/^{207}\text{Pb}$ values in the late 1960s and early 1970s is attributed to the change in the isotopic composition of lead added to USA gasoline; the decrease in $^{206}\text{Pb}/^{207}\text{Pb}$ from

1976 through the late 1980s is attributed to the impact of the Clean Air Act in the USA that banned leaded gasoline. The 1989 value of 1.15 likely reflects the lower but relatively constant output from Canada, Europe and Asia.

11.2.2 $^{87}\text{Sr}/^{86}\text{Sr}$

^{87}Sr is a radiogenic isotope of Sr derived from decay of ^{87}Rb . Strontium isotopes are useful as tracers of geochemical cycling in the oceans and terrestrial waters because Sr^{+2} substitutes and behaves similarly to Ca^{+2} in solution and in minerals (both are alkaline Earth metals). Sr behaves similarly to calcium because (1) it is divalent and (2) the ionic radius of Sr^{+2} (1.13 Å) is only 11% larger than the radius of Ca^{+2} (0.99 Å). As a result, Sr^{+2} can substitute for Ca^{+2} in calcium carbonate (more so in aragonite than in calcite), apatite and plagioclase feldspar. It is a trace component of rocks and sediments, typically comprising 10–1000 mg/kg (ppm), and has four naturally occurring isotopes, ^{88}Sr , ^{87}Sr , ^{86}Sr and ^{84}Sr , which comprise 82.53%, 7.04, 9.87 and 0.56%, respectively, of all Sr. All four Sr isotopes are stable, but because ^{87}Sr is radiogenic (derived from decay of ^{87}Rb), the amount of ^{87}Sr has increased slightly over geological time.

The following is an example of how $^{87}\text{Sr}/^{86}\text{Sr}$ ratios can be applied to assessment of carbonate weathering and soil water/groundwater flow paths. Let's say that strontium isotopic analysis of plagioclase feldspar and calcite in soils and bedrock of a watershed reveals a difference in $^{87}\text{Sr}/^{86}\text{Sr}$ of plagioclase relative to $^{87}\text{Sr}/^{86}\text{Sr}$ of calcite – if this is the case, and assuming that the isotopic ratio of $^{87}\text{Sr}/^{86}\text{Sr}$ of the watershed is controlled by mineral weathering, the relative contribution of calcite weathering relative to plagioclase or other silicate minerals will be revealed in the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of soil water, shallow groundwater or surface water (Clow et al., 1997; Jacobson et al., 2003). This approach was applied in a study

of chemical weathering in the Himalaya by Blum et al. (1998), where differences between strontium isotope compositions of calcite in marble layers ($^{87}\text{Sr}/^{86}\text{Sr} = 0.71$ to 0.74) compared to silicate minerals ($^{87}\text{Sr}/^{86}\text{Sr} = 0.82$ to 0.89) were used to assess what type of weathering reactions were controlling surface water compositions (Fig. 11.4).

These researchers found that even in areas where silicate minerals in crystalline metamorphic and igneous rocks comprise $\geq 99\%$ of the total mineral assemblage, the $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of rivers draining watershed soils are predominantly < 0.77 , reflecting the strong influence of weathering of carbonate minerals (with low $^{87}\text{Sr}/^{86}\text{Sr}$) on surface water strontium compositions. What this indicates is that, while calcite only comprises approximately 1% of total bedrock in this region, its breakdown to Ca^{+2} and HCO_3^- via chemical weathering is the dominant control on surface water chemistry and buffering in this region.

11.3 RADIONUCLIDES AS ENVIRONMENTAL CONTAMINANTS

The main controls on the fate and transport of ions and compounds in aqueous systems (e.g. pH, adsorption, co-precipitation, water:rock ratio, formation of polyatomic anions, ionic strength, residence time, complexation to organic matter, plant uptake, etc.) also apply to radionuclides; however, the mobility and behavior of radioactive and radiogenic species in aqueous systems is also controlled by properties and processes related to decay.

11.3.1 Controls on U, Th and their decay products

The role of alpha recoil was introduced above as a mechanism for physically ejecting uranium, thorium or their decay products (including alpha particles) out of crystal lattices and into the aqueous state, and the production of numerous elements

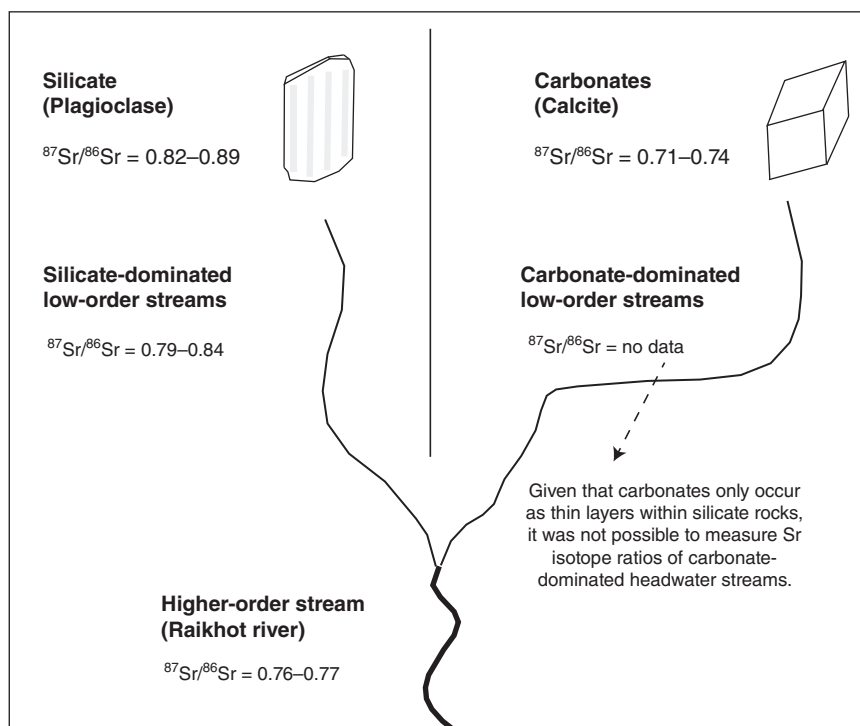


Fig. 11.4 Schematic diagram of $^{87}\text{Sr}/^{86}\text{Sr}$ as tracer of chemical weathering showing (A) isotopic signature of source rocks, (B) streams at different positions in the watershed (after Blum et al., 1998).

with varied properties in the long decay chains of ^{238}U , ^{235}U and ^{232}Th dictate that spontaneous decay will result in products that include:

- 1** insoluble components that are certain to be sorbed or co-precipitated out of solution (e.g. ^{230}Th);
- 2** relatively soluble cations likely to exist in a dissolved state (e.g. ^{226}Ra);
- 3** inert radon gas with the potential to leave the aqueous system through exsolution (e.g. ^{222}Rn);
- 4** a series of short-lived radionuclides with varied properties (e.g. ^{218}Po , ^{214}Pb , ^{214}Bi , ^{210}Pb , and others; Table 11.1), finally culminating in ~insoluble lead (the ^{206}Pb isotope in this case).

Furthermore, in order to predict the behavior of a radionuclide in solution (or suspension) we must consider the isotopic speciation of that element. Radium, for example, has three main isotopes in natural systems, ^{226}Ra (a member of the ^{238}U decay series), ^{228}Ra and ^{224}Ra (members

of the ^{232}Th decay series) and ^{223}Ra (a member of the ^{235}U decay series). The various half-lives of the radium isotopes indicate that the different isotopes will have widely varied residence times and radioactive toxicities. Atoms with short half-lives produce radioactive decay products (α , β , γ and daughter isotopes) at greater rates than isotopes with longer half-lives, so species such as ^{224}Ra ($t_{1/2} = 3.6$ d) and ^{228}Ra ($t_{1/2} = 5.75$ yr), both derived from ^{232}Th decay, will decay more rapidly and have greater radioactive toxicities than ^{238}U -derived ^{226}Ra ($t_{1/2} = 1602$ yr), but they will also persist for shorter amounts of time.

Given that uranium and thorium occur naturally in Earth surface systems, the radiogenic isotopes produced in the decay series of ^{238}U , ^{235}U and ^{232}Th all can occur as naturally derived constituents of soils, surface waters and groundwater, in some cases at concentrations that are

deemed unsafe for use as drinking water or in the production of agricultural crops. **Average concentrations of U and Th in continental crust** are 2.7 and 9.6 mg/kg (ppm), respectively (Wanty and Nordstrom, 1993). Granites and shales tend to be elevated in U and Th relative to continental crustal averages, with uranium exhibiting average concentrations of 4.4 ppm (granite) and 3.8 ppm (shale), and thorium exhibiting average concentrations of 16 ppm (granite) and 12 ppm (shale) (Wanty and Nordstrom, 1993). Mafic and ultramafic rocks are depleted in U and Th (Table 1.1), with average uranium values of 0.75 ppm (basalt) and 0.002 ppm (ultramafic rock) and average thorium values of 3.5 ppm (basalt) and 0.005 ppm (ultramafics) (Turekian and Wedepohl, 1961; Vinogradov, 1962). So clearly, the natural variability in rock, soil or aquifer lithology will exert a strong control on the abundance of uranium, thorium and their decay products in natural systems.

Beyond natural variability in the amounts of uranium and thorium in bedrock, soil or sediment, the numerous controls on aqueous composition covered in Chapters 4 and 5 apply to any radionuclide in solution with the exception of radon gas, where the potential to de-gas from water or rock is an additional variable. Compared to uranium (Fig. 4.14), which in oxidizing environments is soluble as U^{+6} (which forms the UO_2^{+2} uranyl ion) in various hydroxyl-based and carbonate-based polyatomic anions (e.g. $[UO_2]_2CO_3[OH]_3^{-1}$, $UO_2[CO_3]_3^{-4}$), thorium is almost completely insoluble. Regardless of redox conditions it only exists as Th^{+4} , an ion that forms compounds which are insoluble in water. The same is true for U^{+4} (uranous ion), the oxidation state of uranium in reducing conditions where U^{+4} is similarly insoluble to Th^{+4} .

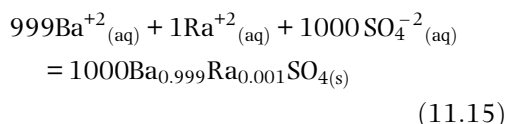
The relatively high solubility of uranium in natural waters produces concentrations of 2–4 $\mu\text{g/L}$ (ppb) in ocean water and, locally, > 100 ppb in streams and groundwaters affected either by (1) uranium-rich soil, sediment or bedrock or (2) evaporative concentration in arid regions.

In some areas, relatively low concentrations of uranium in aquifer sediments or bedrock (e.g. < 10 mg/kg of uranium in bulk sediment or rock) can still produce elevated uranium in groundwater due to the effect of carbonate-enhanced uranium solubility (Langmuir, 1997). In areas with uranium-poor bedrock or soils, uranium might occur in trace amounts (e.g. ≤ 1 ppb) in streams and groundwaters, especially where the lack of carbonate or bicarbonate inhibits formation of soluble anions with uranyl. For reference, the USEPA maximum contaminant level for uranium in public water supplies is 30 ppb, and the third edition of the World Health Organization (WHO) *Guidelines for Drinking-water Quality* (2004) indicated a provisional guideline value of 15 ppb.

The very slow rate of decay of ^{238}U means that its toxicity is derived not from alpha decay (to ^{234}Th), but rather from the inherent toxicity of U as a trace metal and the associated carcinogenic effects and potential to cause kidney disease. Spontaneous decay of uranium and thorium and their daughter products can result in elevated concentrations of radium, a relatively soluble component of many groundwater systems, as well as radon gas and alpha particles. These all are known to occur at elevated concentrations in many soils and groundwaters and as such are regulated by the WHO and USEPA.

Radium occurs as the Ra^{+2} cation and the two dominant isotopes in natural systems are ^{226}Ra and ^{228}Ra . The ^{224}Ra isotope is not abundant because its half-life is so short (3.6 day) and the ^{223}Ra isotope is rare because the parent isotope that ultimately produces ^{223}Ra , ^{235}U , is far less abundant than ^{238}U and ^{232}Th , the decay chains of which result in formation of ^{226}Ra and ^{228}Ra , respectively (Figures 11.2 and 11.3). Soluble Ra^{+2} behaves much like other alkali and alkaline-earth metals in solution, particularly Ba^{+2} . Its relatively high solubility indicates that in some groundwater systems, Ra^{+2} will exhibit conservative behavior; however, in aquifers with dissolved barium (Ba^{+2}) and sulfate (SO_4^{2-}) at

concentrations that approach saturation for barite (BaSO_4), chemical precipitation of barite will cause co-precipitation of Ra^{+2} into the barite structure, thus limiting its mobility in the aquifer, as follows schematically:



Also, aquifers with clay minerals or organic matter that possess high cation exchange capacities also can limit Ra^{+2} mobility by sorption.

A recent study of naturally derived radium in groundwater is presented by Vengosh et al. (2009), who document high concentrations of naturally occurring radium in groundwater produced from Paleozoic Nubian sandstones that form the framework of the Disi aquifer in south-west Jordan. Concentrations of dissolved radium are approximately an order of magnitude greater than WHO standards for drinking water and data suggest that release of radium into groundwater is primarily caused by alpha recoil, a process where the kinetic energy of an alpha particle produced by spontaneous decay may propel it into solution (Fig. 11.5).

In the shallow unconfined zone, average ^{226}Ra and ^{228}Ra activities are 0.53 and 0.91 Bq/L, respectively, whereas in the deep, confined aquifer, average ^{226}Ra and ^{228}Ra activities are 0.68 and 2.0 Bq/L, respectively. In the unconfined zone, $^{228}\text{Ra}/^{226}\text{Ra}$ activity ratios have an average value of 1.6 and are similar to the $^{228}\text{Ra}/^{226}\text{Ra}$ ratio measured in the host aquifer rocks (~ 1.6); however, in the deeper confined zone of this groundwater system, the $^{228}\text{Ra}/^{226}\text{Ra}$ ratio of groundwater (2.9) is higher than shallow groundwater and is also higher than host rock (1.6). Given that ^{228}Ra is derived from ^{232}Th , whereas ^{226}Ra is derived from ^{238}U , the high $^{228}\text{Ra}/^{226}\text{Ra}$ ratio might indicate a higher Th/U ratio at depth, but a more likely interpretation is that loss of ^{238}U daughter products results in

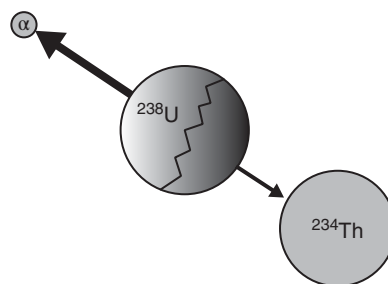
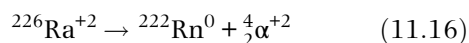


Fig. 11.5 Schematic representation of alpha recoil upon decay of an atom of ^{238}U . Note that decay products (α particle and ^{234}Th) are propelled in opposite directions, a process that can result in introduction of such decay products (either α or ^{234}Th in this case) from the edge of a crystal into solution.

excess ^{232}Th over ^{230}Th (the direct parents of ^{228}Ra and ^{226}Ra), thus producing a higher $^{228}\text{Ra}/^{226}\text{Ra}$ ratio in groundwater. The absence of clay minerals in this coarse-grained, sandy aquifer system leads to very little sorption of Ra^{+2} to mineral surfaces, contributing to high radium concentrations in the groundwater.

$^{226}\text{Ra}^{+2}$ decays to the inert gas $^{222}\text{Rn}^0$, by emission of an alpha particle:



When this decay reaction occurs, the behavior of the radionuclide shifts dramatically from a soluble divalent cation (Ra^{+2}) to an inert gas that has the potential to exsolve from water into air; however, given the relatively short half-life of ^{222}Ra (3.8 days), this is only likely to happen in soils or in the shallow vadose zone. A given ^{222}Rn atom typically diffuses a few cm in water before decaying to ^{218}Po , so radon gas should not diffuse into the atmosphere from the saturated zone. In reality, the two main pathways by which radon can contaminate indoor air are (1) seepage from soil into buildings through cracks in foundations or old dirt cellars, or (2) by pumping of radon-enriched groundwater from depth into a building, where ^{222}Rn then may enter air, especially when the water is used in

Radionuclide Drinking Water Guidelines

WHO guidelines for radium in drinking water recommend maximum activities of 0.1 Bq/L for ^{228}Ra and 1 Bq/L for ^{226}Ra . The order of magnitude difference reflects the shorter half-life (and thus greater decay rate) of ^{228}Ra as compared to ^{226}Ra . The US EPA maximum contaminant level (MCL) of 5 pCi/L (0.185 Bq/L) is for combined ^{228}Ra and ^{226}Ra and does not differentiate among the two isotopes. Alpha decay of ^{226}Ra produces ^{222}Rn , the isotope of radon that often occurs as an indoor air contaminant where buildings are constructed on soil or bedrock with elevated uranium. The WHO guideline for ^{222}Rn is 100 Bq/m³ ([http://](http://www.euro.who.int/document/aicq/8_3radon.pdf)

www.euro.who.int/document/aicq/8_3radon.pdf), while the USEPA recommended limit for ^{222}Rn in indoor air is 4 pCi/L (where the radiation for WHO and EPA guidelines is measured per unit volume of air). Converting pCi to Bq and liters to m³, the USEPA value corresponds to 148 Bq/m³. Remediation for radon contamination is relatively simple and typically involves improved ventilation systems that replace Rn-enriched air with fresh air. The WHO guideline for gross alpha radiation in water is 0.5 Bq/L, while the EPA value is 15 pCi/L (0.56 Bq/L). Guidelines for gross beta activity are 1 Bq/L (WHO) and 4 millirems per year (USEPA).

showers (the misting action of showers increases the probability that Rn will exsolve into air). Following alpha decay of ^{222}Rn to ^{218}Po , relatively rapid decay of ^{218}Po and daughter products produces ^{206}Pb via the branched series presented in Fig. 11.2 and Table 11.1.

Alpha particles emitted by spontaneous decay of ^{238}U , ^{232}Th , ^{226}Ra , ^{222}Rn and other radionuclides in groundwater are monitored (1) because they can be used to indicate the presence of radioactive elements in water supplies, and (2) because they are a form of ionizing radiation that is harmful to humans. Because of the high charge to small radius of alpha particles, they tend to strongly adsorb to solids.

11.3.2 Refined uranium ores and associated nuclear wastes

Uranium is mined from ore deposits that typically contain 0.1 to 1% uranium and are most commonly hosted in granitic rocks or diagenetically

altered sandstones. In U-rich sandstone deposits, elevated concentrations of uranium occur at the boundary of oxidized and reduced sediment (Fig. 11.6). These deposits (known as roll-front deposits) form when oxidized paleogroundwaters containing dissolved uranium carbonate complexes flow into reduced sediments (often caused by presence of organic matter), where soluble U^{+6} is reduced to insoluble U^{+4} . Reduction to U^{+4} results in formation of uranium oxide minerals such as uraninite and pitchblende (both approximately UO_2) and autunite ($\text{Ca}[\text{UO}_2]_2[\text{PO}_4]_2 \cdot x\text{H}_2\text{O}$).

Although only approximately 0.72% of naturally occurring uranium is the ^{235}U isotope (> 99% of natural uranium is ^{238}U), it is necessary to enrich uranium ore so that ^{235}U concentrations reach 2–4% for nuclear power and to > 80% ^{235}U for weapons (although “dirty” nuclear bombs can be made with as little as 20% ^{235}U). Centrifuges are used to do this.

When ^{235}U undergoes fission in a reactor it forms two approximately equal-sized daughter

Notes on Units of Radioactivity

According to Lawrence Berkeley Laboratory of the US Department of Energy, 1 curie (Ci) is equal to 3.7×10^{10} radioactive decays per second, and was originally defined based on the radioactivity produced by 1 g of ^{226}Ra . The unit applied to measuring radioactivity of water supplies in the United States is the picocurie per liter (pCi/L), where a pCi is equal to 3.7×10^{-2} radioactive decays per second (or the amount of radioactivity produced by 10^{-12} g (one picogram) of ^{226}Ra . The currently accepted SI unit of radioactivity is the becquerel, where one Bq

is equal to 2.703×10^{-11} Ci, meaning that $1 \text{ Bq} = 27.03 \text{ pCi}$. A becquerel is also equal to 1 count per second. A millirem, the unit used by the USEPA to quantify beta radiation in drinking water, measures the radiation dose to tissue, attempting to factor in the different biological effects of different types of ionizing radiation. For this reason, it is difficult to apply a unit conversion for Bq (or pCi/L), which are direct measurements of radioactivity in water, to millirems. (<http://www.lbl.gov/abc/wallchart/chapters/15/2.html>)

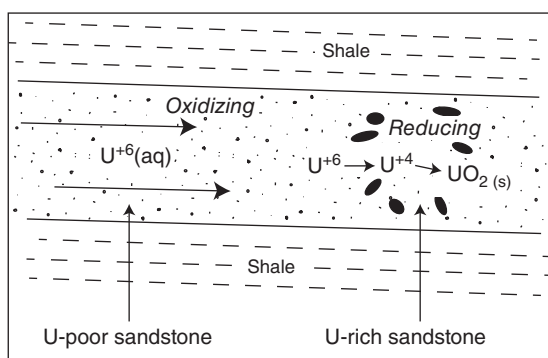
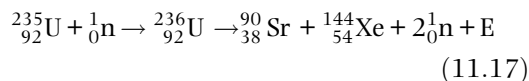


Fig. 11.6 Schematic sketch of how transition from oxidizing to reducing conditions in sandstone can lead to formation of a uranium roll-front deposit. Arrows indicate direction of paleo-groundwater flow; black circles and ellipses represent organic matter in reducing zone. Similar processes cause U and many trace metals to precipitate in wetlands.

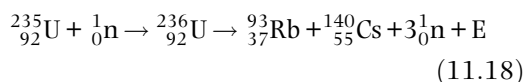
products with atomic masses in the ranges 90–100 and 130–145, examples of which include ^{90}Sr , ^{99}Mo , ^{99}Tc , ^{131}I , ^{135}Cs , ^{137}Cs and ^{141}Ba (Table 11.1). Fission of ^{235}U also produces lighter isotopes such as ^{14}C and ^3H (tritium) plus

neutrons and very large amounts of energy. Tritium atoms are so small that they are difficult to retain and they often combine with oxygen to form tritiated water that can escape into local air and water bodies surrounding nuclear power plants. Tritium has a half-life of 12.3 yr and decays to ^3He by low-energy β decay – the low-energy β radiation lacks the energy needed to penetrate skin (as does high-energy β radiation) but tritium can be problematic when ingested or inhaled.

The fission of ^{235}U is initiated by neutron bombardment that first transforms ^{235}U into the highly unstable ^{236}U isotope, which undergoes spontaneous decay to many combinations of daughter products. The example presented below shows the formation of the daughter products ^{90}Sr and ^{144}Xe plus two neutrons and energy:



Another example of a ^{235}U fission reaction, depicted to show the varied possibilities of fission products, is:



The liberation of neutrons by ${}^{235}\text{U}$ fission leads to a chain reaction, a process in which fission-produced neutrons bombard other nearby ${}^{235}\text{U}$ atoms, which in turn transform to ${}^{236}\text{U}$ atoms that then undergo fission to produce daughter elements and more neutrons (and much energy), and so on. In nuclear power plants, the fission reactions are controlled; in nuclear bomb explosions, the fission reactions are not controlled and proceed at a very high rate, releasing a tremendous amount of heat and radioactive daughter products nearly instantaneously.

The energy produced by nuclear fission is liberated in the form of heat that is used to turn liquid water to steam that in turn runs turbines that produce electricity. In this way, the main difference between nuclear power and more conventional carbon-based energy sources is the origin of the heat – in the case of carbon-based fuel, the heat is produced by combustion.

After one to three years, ${}^{235}\text{U}$ -enriched fuel becomes unsuitable for energy production. This is because the accumulation of daughter products in the fuel rods adsorbs and deflects neutrons, thus decreasing the amount of energy that can be produced from a given rod. At this point, the rods are removed from the nuclear reactor and either stored or disposed of in a waste repository. However, the spent fuel is highly radioactive and produces high amounts of heat energy, so any disposal options must prevent leakage of radioactive isotopes and their decay products until the harmful isotopes have been depleted by radioactive decay, and must also be able to account for heat produced by decay reactions within the spent fuel. In some cases, spent fuel can be reprocessed and used again as fuel in a nuclear reactor – this is the case in many countries, but not the United States.

With half-lives in the range of 1 to 30 yr, the relatively short-lived daughter products of fission

reactions (e.g. ${}^{90}\text{Sr}$, ${}^{99}\text{Tc}$, ${}^{106}\text{Ru}$, ${}^{131}\text{I}$, ${}^{134}\text{Cs}$ and ${}^{137}\text{Cs}$) cause the greatest amount of damage to humans and ecosystems when released into the environment. These short-lived isotopes are the main elements of concern in **high-level radioactive waste**. The two main reasons for their high toxicity and potential to cause damage are:

- 1 Their decay rates are orders of magnitude greater than longer-lived isotopes such as ${}^{99}\text{Tc}$, ${}^{233}\text{U}$, ${}^{237}\text{Np}$, ${}^{238}\text{U}$ and ${}^{239}\text{Pu}$, which possess half-lives in the range of 210 000 yr (${}^{99}\text{Tc}$) and 4.5 billion years (${}^{238}\text{U}$).

- 2 They persist for far longer than some of the very short-lived daughter products such as ${}^{89}\text{Sr}$, ${}^{133}\text{Xe}$ and ${}^{103}\text{Ru}$ (while highly radioactive, their half-lives are ≤ 50 days). These highly unstable, very short-lived daughter isotopes decay so rapidly that they are very quickly (e.g. within one year) transformed into less-hazardous isotopes, and really only pose a problem in bomb explosions or during uncontrolled fission and melting of radioactive fuel at a power plant (often referred to as a meltdown).

In the wake of the 2011 earthquake/tsunami-caused disaster at Fukushima Daiichi in Japan, ${}^{131}\text{I}$ and ${}^{137}\text{Cs}$ were the main short-lived radionuclides released to the air and sea. They are the most highly volatile of the fission products, and because Fukushima was not affected by a major fire (as was the case at Chernobyl), most other less-volatile radionuclides were contained on-site. Thus, while the Fukushima disaster released much less radioactivity compared to Chernobyl, emissions of ${}^{131}\text{I}$ and ${}^{137}\text{Cs}$ are still of serious concern because they are easily taken up by humans and other organisms (e.g. ${}^{131}\text{I}$ tends to concentrate in the thyroid where it undergoes decay and can cause cancer).

The fate and transport of the varied isotopes in high-level radioactive wastes are controlled by decay rates and mechanisms as well as by typical controls on aqueous species (pH, inherent solubility of the species, sorption, alkalinity, dissolved oxygen, etc.). Given their abundance in spent

fuel and half-lives of 28 and 30 yr, respectively, ^{90}Sr and ^{137}Cs are the main source of radioactivity in wastes that are less than 100 years old. Wastes older than 100 years will be depleted in ^{90}Sr and ^{137}Cs (by decay to less-radioactive daughter isotopes) relative to some of the more long-lived isotopes, particularly ^{237}Np , ^{239}Pu , ^{240}Pu , ^{241}Am and ^{243}Am (Langmuir, 1997). These isotopes of Np, Pu and Am will then dominate radioactive decay reactions in the spent fuel for approximately 100 000 yr, after which isotopes of lead and radium (e.g. ^{210}Pb , ^{226}Ra) become the main sources of radioactivity. By this time emitted radiation is close to background levels.

11.3.3 Geological disposal of high-level radioactive wastes

According to the International Atomic Energy Agency (IAEA, 2006), the 441 nuclear power reactors on Earth create 10 500 metric tons of spent fuel per year, a quantity that would fill a football field to a depth of 1.5 m – while this a relatively small volume ($\sim 10\,000\text{ m}^3$), environmental and security risks dictate that appropriate disposal is crucial. Currently, most wastes are temporarily being stored on-site at nuclear plants in pools of water, yet there is a growing consensus that deep underground storage is the best long-term solution. Finland, France and Sweden are in the process of creating geological repositories for the long-term disposal of spent nuclear fuel and high level waste. Sweden appears to be the most advanced in this respect – in 2009 a site 150 km north of Stockholm near the small city of Östhammar was selected for a high-level waste repository which, pending review, will be hosted 500 m below the land surface in a homogeneous granite pluton. In addition to granite, other rock types being considered in Europe for deep repositories are thick clay deposits (Belgium France, Germany, Spain and Switzerland) and salt deposits (Germany).

In 2008, the U.S. Department of Energy (DOE) applied to the U.S. Nuclear Regulatory Commission (NRC) for authorization to construct a deep geologic repository in Miocene-age volcanic tuffs at Yucca Mountain, Nevada to host high-level radioactive waste from sites across the United States (NRC, 2008, available at <http://www.nrc.gov/waste/hlw-disposal.html>). While this issue is sure to be debated politically for many years to come, the site is favorable for many reasons, including a very deep water table (500–750 m below land surface; Montazer and Wilson, 1984) which should help to avoid water-based corrosion of storage containers, and the presence of smectite clays and zeolites with high cation exchange capacities (Bish et al., 2003) that will retard migration of cations such as Cs^+ , $\text{Pu}^{+4,+5,+6}$ and Sr^{+2} . However, critics point out that fractures in the tuff may permit water to penetrate downward toward the repository, potentially leading to interaction with the waste, a point that is countered by the expectation that any infiltrating water is likely to be driven away from the wastes by vaporization caused by the heat of the wastes themselves (Langmuir, 1997).

Corrosion of waste-bearing containers within the first 100 yr of storage of spent fuel could potentially release soluble $^{90}\text{Sr}^{+2}$ and $^{137}\text{Cs}^+$. However, it is far more likely that corrosion would occur after ^{90}Sr and ^{137}Cs have decayed to insignificant amounts, the point at which the actinides (mainly Np, Pu and Am) plus ^{99}Tc and ^{129}I will be the dominant radioactive species in the spent fuel. Many of these isotopes exist as soluble anions that are less likely to sorb to smectites and zeolites in the volcanic tuffs than would be cationic species (e.g. Cs^+ , Sr^{+2}), so for this reason ^{99}Tc , ^{129}I , ^{233}U , ^{234}U and ^{233}Np are assumed to be soluble and the most likely to contaminate groundwater or reach the surface before the approximately 100 000 yr required to decay to background radioactivity levels. Examples of the soluble species of these isotopes include TcO_4^- , I^- , the various polyatomic anions formed by uranium with hydroxide

and carbonate anions, and $\text{Am}(\text{CO}_3)_2$. Given the alkaline character of the groundwater in the arid Yucca Mountain region, these radionuclides are likely to form OH^- and CO_3^{2-} compounds.

While many actinides occur in only one oxidation state in natural systems (e.g., Am^{+3} , Cm^{+3} , Th^{+4}), Np and Pu can exist in several oxidation states (you will also recall that uranium can occur as U^{+4} and U^{+6}). In oxidizing groundwaters, neptunium mainly occurs as Np^{+5} in the soluble and mobile species NpO_2^+ below a pH of 10 (Langmuir, 1997). The low ratio of ionic charge to radius of NpO_2^+ indicates that it will not be strongly adsorbed to mineral surfaces. In very strongly oxidizing water, Np can occur as Np^{+6} in the less soluble NpO_2^{+2} , but this species is far less common than NpO_2^+ . Plutonium occurs in three common oxidation states in soils and groundwaters, Pu^{+4} , Pu^{+5} and Pu^{+6} – there is a trivalent ion (Pu^{+3}), but it only occurs in extremely reducing or highly acidic environments. PuO_2^+ is the dominant soluble plutonium ion in most natural waters. The pentavalent cation XO_2^+ (where X is Np or Pu) comprises approximately 95% of soluble neptunium and plutonium species in solution non-alkaline waters where carbonate concentrations are $\leq 10^{-4}$ M (< 10 mg/L). However, they differ in their behaviors because the solubility of plutonium is commonly limited to 10^{-8} to 10^{-6} M (0.2 to 20 $\mu\text{g/L}$ or ppb) by reduction of Pu^{+5} to Pu^{+4} (a reaction that can occur in relatively oxidizing, slightly alkaline groundwater), and this facilitates formation of the insoluble species $\text{Pu}(\text{OH})_4$ (Fig. 11.7); conversely, neptunium cannot be reduced to a tetravalent ion and thus does not form an insoluble hydroxide, so its concentration in solution (as NpO_2^+) can be as high as 10^{-4} M, or ~ 20 mg/L (Choppin, 2007).

Accordingly, plutonium is relatively insoluble compared to most actinides (with the exception of the notoriously insoluble Th^{+4}). Nonetheless, formation of colloidal plutonium $\text{Pu}(\text{OH})_4$ can enable transport of plutonium tens to thousands of meters in groundwater (McCarthy and Zachara, 1989), illustrating the complex

relationships between aqueous phases, solid phases, and the potential to transport extremely fine-grained solids (i.e. colloids with diameters of 10^{-8} to 10^{-6} m) in suspension. In this case, although insoluble, the particles of $\text{Pu}(\text{OH})_4$ are examples of solids that are so fine grained that geochemists must consider their potential for transport as a non-aqueous colloidal phase, particularly in neutral to alkaline conditions that favor formation of $\text{Pu}(\text{OH})_4$.

11.4 GEOCHRONOLOGY

Shortly following the discovery of radioactivity by Henri Becquerel in 1896 and rapid advances in understanding of radioactivity by Pierre and Marie Curie, Ernest Rutherford and others at the turn of the 19th–20th centuries, the decay of ^{238}U to ^{206}Pb was applied to geological dating. The first published paper, by British geologist Arthur Holmes, provided the first geochemically quantitative determination of the age (370 ma) of a rock, in this case a granite in Norway (Holmes, 1911). Since then the U-Pb method has been refined and many others have been developed, including other decay pairs (e.g. K–Ar, Rb–Sr), radiocarbon dating and dating of surficial Earth features by cosmogenic radionuclides. The following section presents some of the main principles, methods and some selected applications of radionuclides applied to dating of geological and biological materials.

11.4.1 ^{14}C , cosmogenic radionuclides and earth-surface dating techniques

Cosmogenic radionuclides are produced in the atmosphere by nuclear reactions initiated by bombardment of stable atoms by cosmic rays, and also at the surface of the Earth in rocks, soils and sediments that are exposed to cosmic rays. These two modes of formation mean that cosmogenic isotopes can be used to date (1) deposition

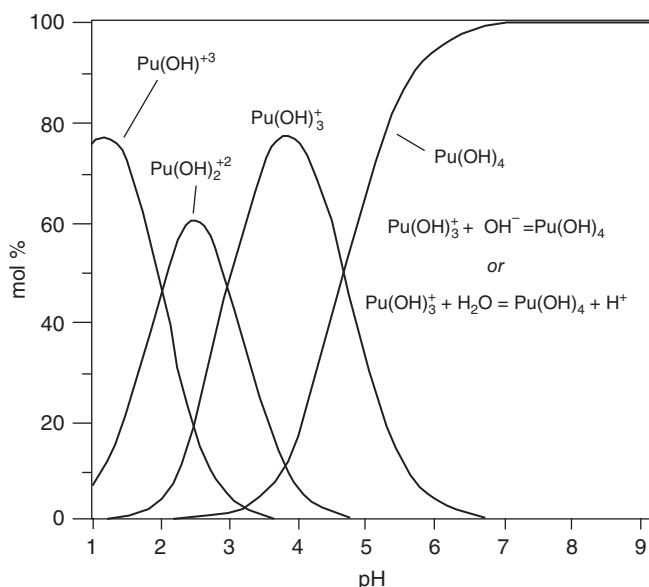


Fig. 11.7 Speciation of Pu^{+4} as a function of pH in mildly oxidized waters, a redox state that is typical of shallow groundwater. Note that Pu solubility is pH dependent, where decreasing pH results in transformation of the insoluble Pu(OH)_4 to relatively soluble Pu(OH)^{+3} and other cationic plutonium hydroxides. (Modified after Choppin 2007.)

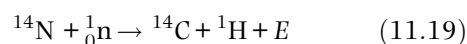
of sediments in lakes, ponds or other surficial environments, and (2) exposure ages of boulders, bedrock surfaces, sediments and other materials that can determine when rocks or sediments were first exposed to cosmic ray bombardment. When cosmogenic radionuclides formed in the atmosphere fall to the surface in precipitation, they may be incorporated into aquifers, snow (and eventually glacial ice) or sediments, which enables dating of materials as varied as groundwater and deep-ocean water to layers in glacial ice and sediments. Cosmic ray bombardment of rock or sediment exposed at the Earth's surface forms cosmogenic isotopes that can date the amount of time elapsed since exposure of the rock or sediment at the surface, allowing dating of surface processes such as landslides, river erosion, fault activity, glacial ice retreat, and cooling/crystallization of lava flows.

^{14}C (Radiocarbon)

The radiocarbon dating technique was developed by Willard Libby and University of Chicago colleagues in the late 1940s, and the accuracy of the method was first demonstrated by dating wood

of a known age from an ancient Egyptian barge (Arnold and Libby, 1949). Radiocarbon dating was the first cosmogenic method to be developed and it is now widely applied in geology and allied fields like archeology and biology. For his pioneering work, Libby was awarded the Nobel Prize in chemistry in 1960. Since its inception, radiocarbon dating has become one of the most well-known and most commonly applied method for dating carbon-bearing materials with ages $< 60\,000$ yr.

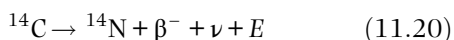
^{14}C is a cosmogenic radionuclide because it forms as the result of the interaction of cosmic rays and their products with nitrogen in the atmosphere – cosmic rays (mainly originating from the sun) collide with atmospheric gases and one of the consequences is that neutrons or protons can be ejected from a nucleus of a molecule in the atmosphere. The reaction of atmospheric nitrogen (nitrogen atoms in N_2 gas) to radiocarbon occurs according to this reaction:



The ^{14}C generated in the atmosphere very rapidly reacts with O_2 to produce $^{14}\text{CO}_2$, and atmospheric

circulation causes rapid dispersion and mixing of $^{14}\text{CO}_2$, so there is little or no variability in concentration of ^{14}C as a function of latitude. The $^{14}\text{CO}_2$ can be incorporated into plants by photosynthesis or it can combine with H_2O to form $\text{H}_2^{14}\text{CO}_3$ and is delivered to the Earth's surface in precipitation (just like "normal" CO_2 with ^{12}C), from where it can be incorporated into living organisms, minerals or water bodies. The ratio of $^{12}\text{C}/^{14}\text{C}$ in CO_2 or H_2CO_3 is governed mainly by the atmospheric ratio of these two isotopes. If incorporated as CO_2 into a plant, ^{14}C is no longer incorporated following death of the plant, so the amount of ^{14}C progressively decreases as it undergoes radioactive decay, so in this way, the amount of ^{14}C in organic matter (especially wood, seeds, pine cones and other readily identifiable materials) can be used to determine time elapsed since death of the organism. The same can be said for the dating of groundwater, where H_2CO_3 or decaying organic matter that infiltrates the water table will no longer incorporate ^{14}C , so the progressive decay of ^{14}C makes it possible to determine groundwater age, an important parameter in assessing recharge rates and sustainable use of groundwater resources.

^{14}C decays to ^{14}N according to the following spontaneous beta decay reaction:



where, as above, $\bar{\nu}$ is an antineutrino and E represents energy released during the decay reaction.

The decrease in the amount of ^{14}C with time is given as:

$$[^{14}\text{C}]_t = [^{14}\text{C}]_0 * e^{-\lambda_p * t} \quad (11.21)$$

This is a more specific form of the exponential decay equation presented in the previous section, and $\lambda = 1.209 \times 10^{-4} \text{yr}^{-1}$ corresponds to $t_{1/2} = 5730 \text{yr}$. This relatively short half-life is the reason that materials older than $\sim 50\,000 \text{yr}$

cannot be dated by the radiocarbon method – after 10 half-lives (57 300 yr), only about 0.1% of the original ^{14}C remains in the material to be dated. However, the positive attribute of the short $t_{1/2}$ is that it enables relatively precise dating of geologically young materials.

Another important point to realize about radiocarbon dating is that, unlike most radioactive dating methods, radiocarbon dating does not involve measuring the ratio of parent isotope (^{14}C) to daughter isotope (^{14}N). The abundance of ^{14}N in the atmosphere and at the Earth surface and shallow sub-surface means that it is effectively impossible to distinguish between ^{14}N derived from ^{14}C and the large reservoir of ^{14}N in the environment (i.e. in atmosphere as N_2 , in soils as component of organic matter, in water as NO_3^- , etc.). So, for this reason, age measurements are made by counting the decay of individual C atoms by, e.g., gas proportional counting (de Vries and Barendsen, 1953) or liquid scintillation counting (Noakes et al., 1967), or, more recently, by the advent of accelerator mass spectrometry (AMS). The older methods are still appropriate for dating samples with either high C content or young ages, but the advantage of AMS is that the ratio of ^{14}C : ^{12}C atoms can be directly measured, and the result is that this method can facilitate dating of materials that contain only a few mg of C (e.g. lake sediments with small amounts of organic matter or shells with small amounts of incorporated organic carbon) or have ages $> 30\,000 \text{yr}$.

Radiocarbon dating is based on a few important premises and recognition of potential sources of error in age analysis. They are:

- Any variability in the rate of neutron production by cosmic radiation will alter the rate of ^{14}C production, and these changes must be accounted for or incorporated into analytical uncertainty. Examples of factors that tend to increase the rate of neutron flux, and hence ^{14}C production, include increases in sun spot activity and solar wind, and decreases in the magnetic

moment of Earth's magnetic field. The opposites tends to decrease neutron flux. Greater rates of ^{14}C production from some time in the past compared to the modern rate would yield relatively high ^{14}C concentrations compared to the sample's actual age, thus likely producing anomalously young ages, and vice versa.

- Radiocarbon dating assumes an atmosphere that is spatially well mixed (homogeneous) with respect to CO_2 concentration. This is a good assumption.
- Changes in atmospheric CO_2 that have occurred during the late Pleistocene and Holocene are quantifiable and can be applied to calibration of ^{14}C ages. A cooler Earth means cooler oceans and greater oceanic uptake of CO_2 ; this decreases atmospheric CO_2 , which in turn could increase the concentration of ^{14}C in atmospheric CO_2 . Variations in atmospheric ^{14}C concentration, including increases of approximately 2% during stages of the Little Ice Age 500 and 700 yr ago (1500 and 1700 AD), are known as the **de Vries effect**, after deVries (1958).
- Lastly, once the material to be dated has been removed from the carbon source (e.g. an organism dies and no longer incorporates CO_2), it should be a closed system with respect to ^{14}C – there should be no exchange of ^{14}C with its surroundings, for any gains or losses of ^{14}C by a pine needle or a fragment of charcoal in a buried soil will bias the age determination.

Specimens that contain inorganic carbon that has formed by chemical or biological precipitation can also be dated by the ^{14}C method. In general, the carbon in these materials is contained in the minerals calcite or apatite, polymorphs of CaCO_3 . The inorganic carbon incorporated in these instances consists of carbonic acid (CO_2 dissolved in water), bicarbonate (HCO_3^-) and carbonate (CO_3^{2-}) in seawater, soil water, groundwater or fresh surface water (e.g. lakes or ponds), and occurs in shells of organisms (e.g. mollusks) as calcium carbonate or as chemically precipitated carbonate minerals in soil water, groundwater

or sea water. Although inorganic, the carbonate in these materials is datable if it was in equilibrium with ^{14}C from the atmosphere. This assumption is often where analytical errors arise. For example, contributions of inorganic carbon from weathered limestone could contain bicarbonate with no ^{14}C , and this would dilute the initial concentration of ^{14}C in the shell when it formed, producing a raw radiocarbon date that is anomalously old. This is known as the **reservoir effect**, and the inorganic carbon source with no ^{14}C is often known as “dead carbon”.

In addition to carbon derived from weathering of ancient rocks, the organic carbon in petroleum also contains no ^{14}C (the petroleum is far too old to contain ^{14}C dating back to the initial burial of organic matter), so the vast quantities of CO_2 from fossil-fuel combustion have introduced large quantities of dead carbon to the atmosphere since the 1800s, lowering atmospheric ^{14}C by about 2%. This phenomenon is known as the “Suess effect” (Suess, 1955). Age bias can also be introduced by slow exchange of CO_2 between atmosphere and the waters from which shells or carbonate minerals form, particularly those from deep marine or deep groundwater regimes, and by isotope fractionation, a process whereby an organism or mineral may exhibit preference for one isotope of an element over another. For example, plants tend to preferentially incorporate ^{12}C over ^{13}C and ^{14}C , and this type of fractionation must be accounted for in age determinations.

Raw radiocarbon ages are commonly reported in years before present (BP), where the present is defined by convention as 1950. So, if a pine cone retrieved from lake sediments in the year 2010 contained a ^{14}C content that yielded a date of 4070 yr old, its age would be adjusted by 60 yr (2010–1950) to 4010 B.P. (to make the date relative to 1950). However, it is also important to recognize that variability in atmospheric ^{14}C through the Holocene and Pleistocene caused by the mechanisms described above plus potential sources of dead carbon makes it necessary to

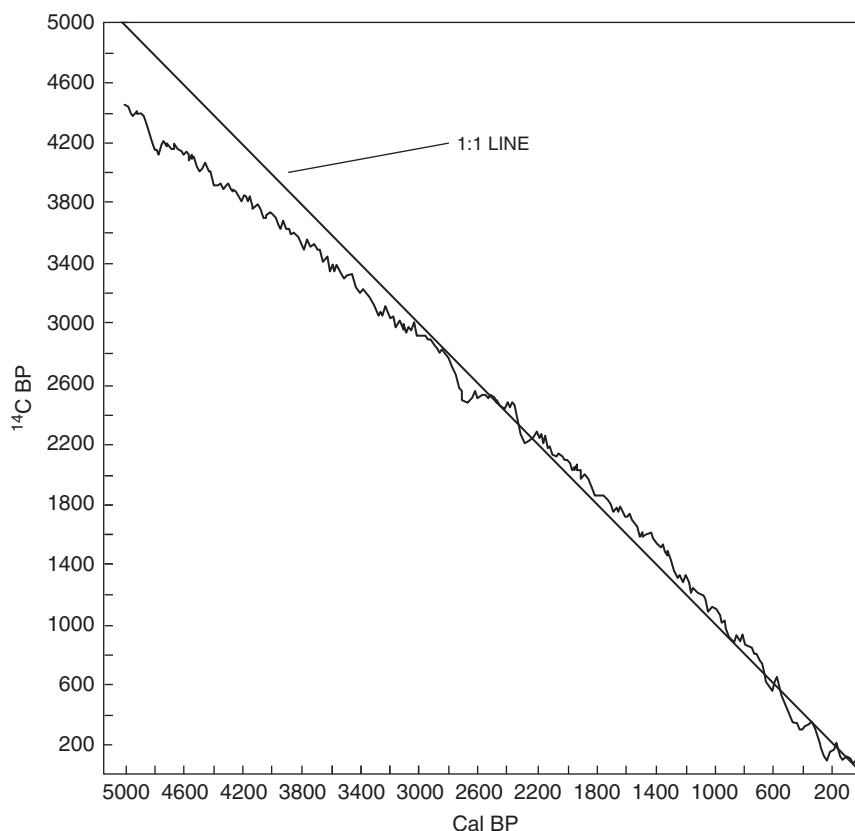


Fig. 11.8 Calibrated radiocarbon ages (*x*-axis) plotted against raw radiocarbon ages. (ages here were calibrated against tree ring and marine coral records). The 1:1 line is presented for reference to illustrate deviations, and the axes are relative to 1950, by convention (see text for explanation). Compiled from data in Stuiver et al., 1998.

calibrate raw radiocarbon ages, so when interpreting radiocarbon dates, be aware that raw dates differ from (and typically are older than) calibrated dates, in some cases by up to 10% or more.

The curve in Fig. 11.8 illustrates the relationship between raw ages and radiocarbon years (both adjusted to pre-1950 scale) corrected for factors such as atmospheric variability and the de Vries effect. Calibration of radiocarbon ages includes analysis of historical materials of known age (e.g. 10 000-year old trees, wood from Egyptian tombs) as well as comparison to coral records (coral produces annual growth rings) and coral ages determined by $^{230}\text{Th}/^{234}\text{U}$ and ^{14}C . For detailed

information on calibration of radiocarbon dates the reader is referred to Stuiver et al. (1998), Hughen et al. (2004) and Reimer et al. (2004).

Cosmogenic ^{10}Be , ^{26}Al and ^{36}Cl

If you do a search for information on cosmogenic radionuclides, particularly in geological publications and institutions, you may find no mention of ^{14}C . What you will likely find is abundant reference to ^{10}Be , ^{26}Al , ^{36}Cl , isotopes that are produced in a manner similar to ^{14}C in the atmosphere, but that in many ways are fundamentally different from ^{14}C .

1 Their half-lives are 2 to 3 orders of magnitude greater than ^{14}C (^{10}Be is 1.5×10^6 yr; ^{26}Al is 7.16×10^5 yr; ^{36}Cl is 3.1×10^5 yr), and thus they are particularly well suited for dating materials that are 1 to 2 orders of magnitude older than those that can be dated by ^{14}C . This provides a mechanism for dating Earth surface and shallow crustal materials with ages on the order of 10^5 to 10^6 yr.

2 Unlike ^{14}C , they are not partitioned into a gas phase and thus they are not well mixed in the atmosphere. This necessitates certain assumptions and corrections for production and deposition rates.

3 They occur in inorganic materials such as quartz and feldspar. One of the limitations of classical ^{14}C dating is the need for datable organic matter – many a geomorphologist, ecologist or geochemist has been known to state phrases like “let’s hope these sediments contain pine cones or sticks so we can get a radiocarbon date” – and this is where ^{10}Be , ^{26}Al and ^{36}Cl can prove especially useful. They can be used to date inorganic materials that lack carbon, and given the paucity or absence of organic matter or datable inorganic carbon in many sediments, waters and rocks, this is an important advantage of these cosmogenic isotopes over ^{14}C .

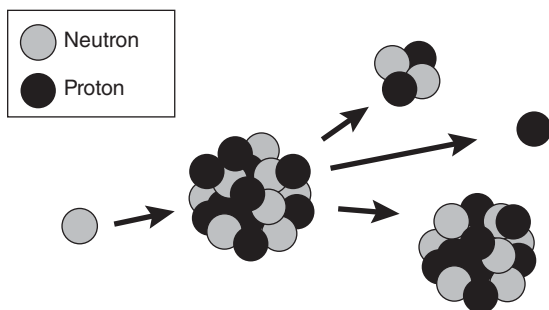
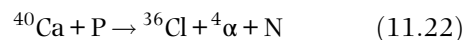


Fig. 11.9 Schematic sketch of spallation reaction where a nucleus (produced by cosmic radiation) crashes into a nucleus of an atom (e.g., ^{40}Ca), yielding a smaller atom (e.g. the cosmogenic radionuclide ^{36}Cl), an alpha particle (2P, 2N) and a proton.

In the atmosphere, ^{10}Be is produced by spallation of O and N when these atoms are bombarded by cosmic rays (where spallation is the ejection of protons or neutrons from the nucleus; Fig. 11.9).

Similarly, ^{26}Al and ^{36}Cl are produced by spallation of atmospheric Ar. A representative spallation reaction is the formation of ^{36}Cl by collision of a proton with a ^{40}Ca atom (in particular, with the nucleus of the atom).



In addition to the ^{36}Cl atom, products also include an alpha particle and a neutron.

Cosmogenic radionuclides formed in the atmosphere are rapidly removed from air with precipitation (e.g. within one week of formation) because unlike ^{14}C , which forms CO_2 and becomes well mixed in the atmosphere, ^{10}Be and ^{26}Al adsorb onto aerosols that act as condensation nuclei for rain droplets, and ^{36}Cl can dissolve in rainwater. One consequence of their rapid deposition out of the atmosphere is that cosmogenic nuclides are not deposited equally over the Earth’s surface – their low residence times in the atmosphere prevent thorough mixing (Granger and Reibe, 2007).

Cosmogenic radionuclides are produced in unequal amounts spatially and temporally. As is the case with ^{14}C production, greater influx of solar radiation over low latitudes and at high elevations means that greater amounts of these isotopes are generated closer to the equator and in mountains and elevated plateaus, and because they fall out of the atmosphere rapidly, they are unevenly delivered to the Earth’s surface. Unlike ^{14}C , a constant deposition rate across the surface of the Earth for a given period in time cannot be assumed. Sites at high latitude and low elevation will receive lower concentrations of cosmogenic radionuclides than will high-elevation sites closer to the equator. This same issue applies to production of cosmogenic radionuclides in rocks exposed at the Earth’s

surface. Like ^{14}C , production rates of ^{10}Be , ^{26}Al and ^{36}Cl are also prone to temporal variations due to changes in cosmic ray flux and variations in Earth's geomagnetic field.

Dating Groundwater with ^{14}C and ^{36}Cl

An example of how cosmogenic isotope analysis can be applied to dating groundwater is provided by Cresswell et al. (1999), who examined a sandstone aquifer system in the Amadeus Basin in central Australia using a combination of ^{14}C and ^{36}Cl . Relatively high $^{36}\text{Cl}:\text{Cl}_{\text{total}}$ ratios occur in shallow groundwater ($\sim 175 \times 10^{-15}$) but appreciably lower $^{36}\text{Cl}:\text{Cl}_{\text{total}}$ ratios (as low as 8×10^{-15}) are present in groundwater from hundreds of meters of depth. Qualitatively, the decrease in ^{36}Cl with depth reflects radioactive decay of ^{36}Cl to ^{36}Ar (or less frequently, to ^{36}S) that occurs as water travels to deeper portions of the aquifer system, and quantitatively, the lower ratios at depth indicate that the deep waters infiltrated the aquifer as long ago as 400 000 yr, following dipping sedimentary bedding planes along a 40–50 km flow path. This ^{36}Cl -determined groundwater age is consistent with what is known about the aquifer hydraulic conductivity (0.1 m per year) and groundwater flow paths.

The absolute age of groundwater can be determined by the equation

$$^{36}\text{Cl} = ^{36}\text{Cl}_0 * e^{-\lambda_p * t} \quad (11.23)$$

where ^{36}Cl is the amount of ^{36}Cl remaining at time t , $^{36}\text{Cl}_0$ is the amount of ^{36}Cl at time $t = 0$ (i.e. when the water infiltrated the aquifer) and λ_p is the decay constant for Cl ($2.25 \times 10^{-6} \text{ yr}^{-1}$).

The determination of $^{36}\text{Cl}_0$ requires certain **assumptions** regarding rate of production and deposition of $^{36}\text{Cl}_0$ from the atmosphere (latitude and elevation dependence) as well as assumptions about the transport of ^{36}Cl through soil to the aquifer, and potential additional sources of Cl that

might contaminate the atmospheric deposition signal. One critical assumption in groundwater dating is that the Cl is not retained in soil (e.g. plant uptake or anion exchange) or the vadose zone for significant periods of time. Researchers must also address potential contamination from ^{36}Cl derived from rocks exposed at or near the surface, and through which infiltrating water comes into contact; in other words, cosmic-ray bombardment of rocks at the surface may be an additional source of ^{36}Cl infiltrating the aquifer and thus must be accounted for. An additional factor is the abundance of Cl in coastal air derived from sea spray, which dilutes ^{36}Cl and produces low $^{36}\text{Cl}:\text{Cl}_{\text{total}}$ ratios (Keywood et al., 1997), indicating that understanding (or assumptions) of past weather systems and transport of coastal air masses must be factored into the age analysis. For reasons like these, many studies rely on two or more isotope systems in their analyses as a cross-check of the accuracy of results.

An additional consideration that makes ^{36}Cl useful in dating groundwater is that chlorine exhibits **conservative** behavior in water, a term that describes its high solubility and tendency to remain dissolved – conservative components like Cl are unlikely to be removed from the aqueous phase by mineral crystallization or ion exchange. Thus, ^{36}Cl that infiltrates the saturated zone of an aquifer is likely to remain dissolved, and analysis of the ratio of ^{36}Cl to total dissolved chloride will facilitate groundwater age analysis. Conversely, Al and Be exhibit **nonconservative** behavior, as they are effectively insoluble at typical soil and groundwater pH values and thus will adsorb onto surfaces of minerals or organic matter or – especially in the case of Al – become incorporated into low-temperature minerals such as the Al-hydroxide gibbsite or clays such as kaolinite or smectite.

In addition to eons of production by natural atmospheric reactions, ^{36}Cl was also produced in relatively large quantities during the 1950s by atmospheric nuclear bomb explosions that irradiated seawater, and this pulse is a useful

means of identifying groundwater younger than 50 yr old. This is because groundwater greater than 50 yr old will lack the signature of this pulse. A similar spike in ^3H (**tritium**) was produced in the atmosphere in the 1950s and 1960s in association with hydrogen bomb testing, with a peak at 1963, making tritium a similarly useful method for determining groundwater age.

The dating of buried sediments and of layers in glacial ice cores is similar to dating of groundwater given that progressive deposition of new material removes sediments from active cosmic ray bombardment once buried to a depth of ~ 1 m. Thus, we only need to consider decay of the isotope, but do not need to concern ourselves with continued production, as is the case with dating of materials exposed at the Earth's surface. However, the nonconservative behavior of ^{10}Be and ^{26}Al is not a problem in sediments or glacial ice. In fact, elemental immobility is advantageous in systems where meltwater or porewater could redistribute the mobile ^{36}Cl but not the relatively immobile ^{10}Be and ^{26}Al .

Determining exposure ages is made possible because in addition to reactions in the atmosphere, ^{10}Be , ^{26}Al and ^{36}Cl also form by direct bombardment of soils and rocks at the Earth's surface (to a depth of ~ 1 m). Spallation of Si and O in quartz produces ^{10}Be , spallation of Si produces ^{26}Al , and spallation of K and Ca in feldspars produces ^{36}Cl . Quartz is especially valuable because it is abundant, chemically stable (and thus not prone to chemical weathering), and it lacks ^{27}Al , which comprises 10–18% of feldspar and thus may serve as a contaminant in feldspar with respect to detection of ^{26}Al .

Minerals at depths greater than 1 m do receive the cosmic radiation needed to form these cosmogenic isotopes, and any cosmogenic nuclides at depths > 1 m will undergo progressive decay. However, when exposed at the surface rocks and sediments begin to accumulate these radionuclides, and thus they are frequently applied to determination of exposure ages, e.g. to answer

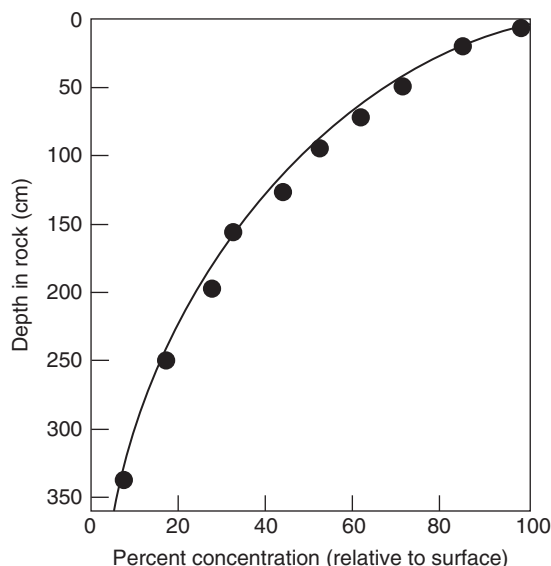


Fig. 11.10 Representative example of a cosmogenic radionuclide depth profile. Units of concentration on the x-axis are typically on the order of 10^2 to 10^5 atoms (of ^{10}Be , ^{26}Al , ^{36}Cl) per gram of mineral (e.g. quartz).

questions of timing of glacier retreat, mass wasting (landslides, slumps, etc.) and exposure of bedrock and sediment surfaces by wind or river erosion. A typical depth–concentration curve for cosmogenic radionuclides in rocks exposed at the surface is shown in Fig. 11.10.

Once exposed to cosmic rays, quartz grains begin to accumulate ^{10}Be and ^{26}Al and initial production rates are high. However, production rates decrease approaching saturation, the point where the rate of production of the radionuclide is equal to its rate of decay. Given that annual production rates of ^{10}Be and ^{26}Al in quartz are of the order of tens to hundreds of *atoms* per gram and that saturation of ^{10}Be and ^{26}Al in quartz is typically reached at concentrations of 10^7 to 10^8 atoms per gram (at mid to high latitudes at sea level), rocks exposed at the surface will become saturated after approximately 10^6 – 10^7 yr (Nishiizumi et al., 1986). Saturation is typically reached after a rock

has been irradiated for approximately four half-lives (Faure, 1998), so in the case of ^{10}Be , saturation is reached after approximately 6 million years, whereas saturation with respect to ^{26}Al is reached after approximately 3 million years, and the saturation time for ^{36}Cl is about 1 million years. Dunai and Stuart (2009) present guidelines for reporting of exposure age and erosion rate determinations in cosmogenic nuclide dating.

For specimens below saturation, exposure age can be calculated by balancing production rate with decay rate:

$$P = R/\lambda * (1 - e^{-\lambda_P * t}) \quad (11.24)$$

In this equation, P is the quantity of cosmogenic isotope produced (in atoms per gram), R = rate of production of cosmogenic isotope (in atoms per gram of sample per year), λ is the decay constant (in yr^{-1}) and t is the exposure time.

In addition to rate of production and decay of a cosmogenic isotope, analysis of surface exposure ages must also consider potential effects of physical or chemical erosion. Any removal of material from the surface of a boulder by wind abrasion, rolling or chemical weathering will remove some portion of the cosmogenic atoms that have accumulated – the resistance to chemical erosion of quartz makes it a desirable mineral for exposure dating, but physical erosion and slow chemical erosion may still act upon even this resistant mineral. In cases where erosion has affected target minerals, cosmogenic isotope-derived exposure ages will be anomalously young. Because erosion removes the outer layer, with the greatest concentration of cosmogenic atoms, cosmogenic exposure ages will underestimate the actual exposure age. Lal (1991) developed an equation for correcting exposure ages based on estimates of erosion rate:

$$P = R/(\lambda + \rho\epsilon/\Lambda) * (1 - e^{(-\lambda_P + \rho\epsilon/\Lambda) * t}) \quad (11.25)$$

Error may also be introduced if boulders or other surface-exposed materials had been previously

exposed and still contain cosmogenic atoms from prior exposure. These latent atoms will result in anomalously old exposure ages.

An example of ^{10}Be dating applied to analysis of river incision into a bedrock channel is provided by Reusser et al. (2007). They sought to determine the timing and rate of incision of the Susquehanna River into its bedrock channel in the foothills of the central Appalachian Mountains of eastern North America. Fluvial down-cutting has produced a series of strath (eroded bedrock) terraces with a total relief of 10 m that have not been buried by sediment or soil and apparently have been only minimally affected by post-fluvial weathering. In other words, the main erosive force appears to have been high-energy fluvial action that formed the terraces, and since then chemical and physical weathering of the terraces is negligible; the terraces are presently exposed at the Earth's surface and are assumed to have been exposed to cosmic rays (i.e. not buried) since the time of their original formation. ^{10}Be ages from 80 samples indicate exposure ages ranging from 36 000 yr on the uppermost terraces to 14 000 yr on the lowest terraces (those closest to the modern stream level). These data are interpreted to indicate rapid erosion into the bedrock channel during the last glacial cycle – 10 m of incision over 22 ka – a process that the authors attribute to greater frequency and magnitude of floods during the late Pleistocene compared to the hydrologically calmer interglacial periods before and after the last glacial cycle.

The determination of glacial mass balance changes in response to rapid shifts in the climate of western Tibet has been examined using a combination of ^{10}Be surface exposure dating and surface mapping (Seong et al., 2009). Results from dating of boulders on moraines and sediment in depth profiles indicate that glaciers advanced and retreated at least 12 times over the past 17 ka with a periodicity of 1.5 ± 1 ka that seems to correlate with northern hemisphere climate oscillations and, to a lesser extent, the south Asian monsoon. The rapid response of climate in the Tibetan Plateau to

global shifts may have implications for predicting future climate in the region.

While ^{10}Be , ^{26}Al and ^{36}Cl (and ^{14}C) are the most common cosmogenic isotopes applied to dating Earth surface processes, there are others, including:

- ^3He – this isotope has certain advantages over the common cosmogenic isotopes ^{10}Be , ^{14}C , ^{26}Al and ^{36}Cl . First, it is stable, i.e. does not undergo radioactive decay, meaning that it does not reach saturation in an exposed rock; secondly, it has the highest production rate of the cosmogenic radionuclides ($\geq 10^2$ atoms/g/yr), making it useful for dating surfaces as young as 600 yr old (much younger than is datable using ^{10}Be , ^{26}Al or ^{36}Cl methods; Kurz et al., 1990); and thirdly, the production of ^3He from spallation of magnesium, calcium and other elements in calcite, olivine and pyroxene makes ^3He dating useful for carbonate and mafic rocks where other cosmogenic approaches are precluded by the absence of quartz (Cerling, 1990; Licciardi et al., 1999; Amidon and Farley, 2011).

- ^{21}Ne – bombardment of igneous olivine and plagioclase feldspar produces ^{21}Ne (from Mg, e.g.), the rate of which has proven useful in dating exposure ages of young volcanic rocks (Poreda and Cerling, 1992).

- ^{32}Si – bombardment of Ar in the atmosphere produces ^{32}Si that then rapidly falls to the Earth's surface in precipitation. With a half-life of 140 ± 15 yr, it has the potential to become an important isotope because its half-life is intermediate to short-lived ^3H (12.3 yr) and ^{210}Pb (22.3 yr) on the one hand, and the longer-lived ^{14}C (5730 yr) on the other hand, making it well suited for dating materials with age ranges of ~ 50 –2000 yr, including glacial layers, groundwaters, ocean waters and late Holocene sediments. However, application of this approach is complicated by a half-life that is only known to $\pm 10\%$ and low ^{32}Si concentrations in sediments, water and ice that make it difficult to measure (Fifield and Morgenstern, 2009).

- ^{129}I – spallation of ^{129}Xe in the atmosphere produces ^{129}I which, like ^{36}Cl , exhibits conservative

behavior in aqueous systems. The long ^{129}I half-life (15.7 million years) has enabled it to date groundwaters with ages ranging from ~ 1 to 100 million years (Fabryka-Martin et al., 1985). ^{129}I is also produced by ^{238}U fission, so it exhibits a late 1950s – early 1960s nuclear bomb spike (much like ^3H and ^{36}Cl), and it was used in the mid-1980s to detect fall-out from the Chernobyl nuclear power plant accident (Paul et al., 1987).

11.4.2 Common radioactive decay methods of dating sediments and minerals

^{210}Pb

The decay of radioactive ^{210}Pb , with a half-life of 22 yr, has been applied to dating very young sediments or glacial ice generally with ages ≤ 150 yr. ^{210}Pb forms in the atmosphere from decay of the noble gas ^{222}Rn via short-lived intermediate products that include ^{218}Po , ^{218}At , ^{218}Rn and ^{214}Po , all with half-lives of the order of milliseconds to minutes (Table 11.1). ^{210}Pb is a member of the ^{238}U radioactive decay series shown in Fig. 11.2, and it is not a cosmogenic isotope because it forms via spontaneous decay in the ^{238}U series rather than from interactions with cosmic rays or their products. For this reason it is referred to as a **fall-out radionuclide**, a group that also includes ^7Be and ^{137}Cs . The rapid transition from gaseous ^{222}Rn to ^{210}Pb , combined with the tendency of ^{210}Pb to strongly sorb to aerosols, means that ^{210}Pb is rapidly rained out of the atmosphere (within 10 days of formation, much like cosmogenic isotopes). Because it is adsorbed to particle surfaces, it is assumed to rapidly settle out of low-energy water (lakes, swamps, marine bodies) and accumulate in sediments soon after it rains out of the atmosphere.

The activity of atmospherically derived ^{210}Pb (unsupported ^{210}Pb) in sediments or layers of glacial ice will decrease with time at a rate in accordance with its half-life, and typically the

age of a given layer of sediment or ice is determined using the following formula:

$$t(\text{yr}) = 1/\lambda * \ln(A^{210}\text{Pb}_0/A^{210}\text{Pb}_h) \quad (11.26)$$

where λ is the decay constant 0.0311 yr^{-1} , $A^{210}\text{Pb}_0$ is the unsupported lead activity in disintegrations per minute at time zero (the present) and $A^{210}\text{Pb}_h$ is the activity in disintegrations per minute at depth h . Substituting the decay constant into this equation produces a simple equation for calculating age:

$$t(\text{yr}) = 1/0.0311 * \ln(A^{210}\text{Pb}_0/A^{210}\text{Pb}_h) \quad (11.27)$$

Ideally, a plot of ^{210}Pb activity will exhibit a logarithmic decrease with depth. However, error can be introduced from ^{210}Pb that is derived from decay reactions in the sediment, commonly originating with decay of ^{226}Ra to, eventually, ^{210}Pb . This lead is known as “Ra-supported ^{210}Pb ” and because it is being added to the sediment after deposition, it will produce higher ^{210}Pb concentrations in the sediment than would be yielded by atmospherically derived ^{210}Pb alone. The presence of ^{210}Pb supported by ^{226}Ra decay in sediments or glacial ice is often evident in deeper layers where a log-normal plot of ^{210}Pb activity vs. depth strays from linearity (Fig. 11.11).

Here, the supported ^{210}Pb is evident in the deeper, older sediments where atmospherically derived (unsupported) ^{210}Pb has already decayed to its daughter products, leaving ^{210}Pb derived from decay of ^{226}Ra in the sediment or ice. In this example, the supported ^{210}Pb level is 8–10 decays per minute per gram (dpm/g). In addition to correcting for supported ^{210}Pb , ^{210}Pb dates can be confirmed or adjusted by ^{137}Cs dating, which takes advantage of a notable spike in ^{137}Cs deposition in the late 1950s and early 1960s related to nuclear bomb testing.

An example of how ^{210}Pb dating is used in environmental analysis (the impact of acid

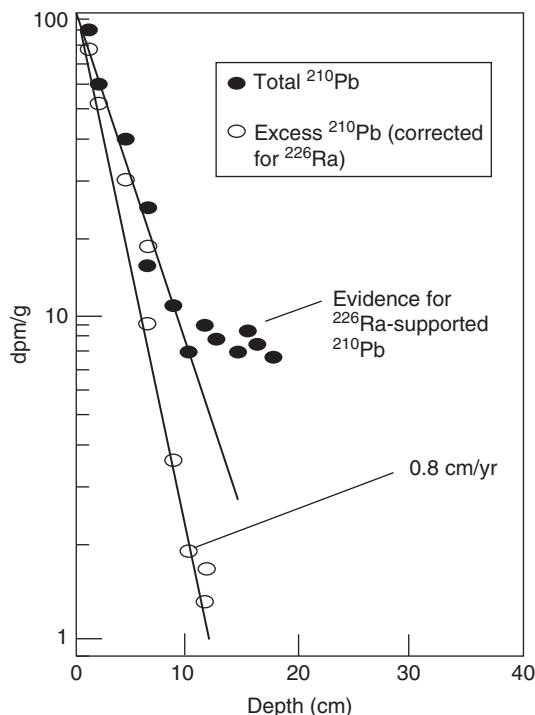


Fig. 11.11 Example of the type of curve showing deviation from linearity in ^{210}Pb activity with depth produced by $^{226}\text{Ra} \rightarrow ^{210}\text{Pb}$ decay. This non-atmospherically derived (“supported”) ^{210}Pb must be accounted for in age determinations. In this hypothetical example, the corrected ^{210}Pb activities give a sedimentation rate of 0.8 cm/year.

deposition on a watershed) is presented in Chapter 9 (Section 9.4) – in that case, ^{210}Pb was used to determine ages of layers in a sediment core back to the year 1850, and the dates obtained helped to determine timing of lake acidification and changes to the calcium–aluminum chemistry of the watershed.

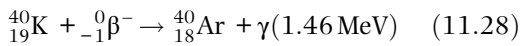
K–Ar

The decay of radioactive ^{40}K to ^{40}Ar is a well-established method for dating K-rich minerals such as illite, muscovite, biotite and K-feldspar in rocks ranging from shales and sandstones to

a wide range of igneous and metamorphic rocks. The half-life of ^{40}K is 1.25×10^{10} yr, and the method has been applied to dating sites as young as Pleistocene (e.g. basalts at Olduvai Gorge) to as old as Precambrian.

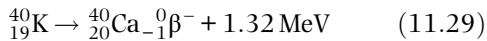
When a K-bearing mineral (e.g. a mica) crystallizes from solution it does not incorporate Ar into the structure. Ar is a noble gas that does not form ions, and thus is not present in K-minerals when they form, so ideally all ^{40}Ar present in a K-bearing mineral is derived via ^{40}K decay.

The decay of ^{40}K to ^{40}Ar occurs by electron capture and can be visualized as:



This process accounts for 11% of the decay of radioactive ^{40}K .

The remaining 89% of ^{40}K atoms undergo transformation to stable ^{40}Ca by beta decay, which is sometimes referred to as negatron emission (Faure, 1986):



However, because ^{40}Ca is the most abundant isotope of calcium, the amount produced by ^{40}K decay relative to the non-radiogenic ^{40}Ca in the mineral makes ^{40}K -derived ^{40}Ca difficult to determine. Thus, K–Ar dating is far more common than the K–Ca method.

The decay constants for the two spontaneous decay reactions (Steiger and Jager, 1977) are:

$$\begin{array}{l} ^{40}\text{K} \rightarrow ^{40}\text{Ar} \quad \lambda(^{40}\text{Ar}) = 0.581 \times 10^{-10} \text{ yr}^{-1} \\ ^{40}\text{K} \rightarrow ^{40}\text{Ca} \quad \lambda(^{40}\text{Ca}) = 4.962 \times 10^{-10} \text{ yr}^{-1} \\ \hline \text{Total} \quad \lambda_{\text{tot}} = 5.543 \times 10^{-10} \text{ yr}^{-1} \end{array}$$

The total decay constant for the branched decay of ^{40}K to ^{40}Ar and ^{40}Ca yields a half-life of:

$$t_{1/2} = \ln(2) / (5.543 \times 10^{-10} \text{ yr}^{-1}) = 0.693 / (5.543 \times 10^{-10} \text{ yr}^{-1}) = 1.250 \times 10^{10} \text{ yr}$$

The K–Ar date of the mineral provides a valid crystallization age if the following criteria are met (or addressed) – they mainly require that the mineral remains a closed system after the mineral crystallizes.

- The mineral must contain no inherited ^{40}Ar introduced into the crystal structure after crystallization.

- The mineral must not lose ^{40}Ar following mineral crystallization. This criterion is harder to achieve because uncharged ^{40}Ar will not be electrostatically attracted to crystallographic sites that are suited to bonding with K^{+1} and may leak out above the **closure temperature** (T_c) or **blocking temperature** (for K–Ar, the temperature below which a mineral will retain Ar). Values of T_c range from as low as 130 °C for microcline (Harrison and McDougall, 1982) to 200–250 °C for low crystallinity illites, 300–350 °C for biotites, 350–400 °C for muscovites (Dahl, 1996), and 500–600 °C for hornblende (Faure, 1986). Closure temperatures increase with increasing crystal size, reflecting longer diffusion pathways and slower Ar loss. While Ar loss during heating by geological fluids can produce K–Ar ages younger than mineral crystallization ages, the tendency of minerals to lose Ar above their T_c values can also be an advantage to geochemists as it can reveal timing of geological events that produce heat flow such as orogenies (Elliott and Aronson, 1987), hydrothermal events (Bish and Aronson, 1993) and fluid flow along faults (Ylagin et al., 2002).

- Argon does not diffuse out of the mineral lattice at temperatures below the T_c (Fig. 11.12).

- Mineral crystallization should occur rapidly, because slow crystallization tends to foster argon loss. However, cooling rates of igneous plutons or metamorphic rocks can be measured if K–Ar ages are obtained for a suite of minerals with different closure temperatures – as the region cools through the closure temperatures of various minerals, they should record an age progression, with high T_c minerals (e.g. hornblende) recording the oldest

dates and low T_c minerals (e.g. fine-grained illite) recording youngest dates (Fig. 11.12).

- No new potassium can be incorporated into the mineral once it has crystallized.

One of the means of dealing with partial loss of Ar is to make use of the **Ar–Ar method** of dating. In this method, the total amount of stable ^{39}K in the mineral of interest is measured and then the ^{39}K is partially converted to unstable ^{39}Ar ($t_{1/2} = 269$ yr) by neutron irradiation in a nuclear reactor. Although unstable, ^{39}Ar is sufficiently long-lived for lab analyses that take place on the order of days or weeks that its decay is negligible, and ultimately the age is determined by the ratio of ^{40}Ar produced by decay of ^{40}K (like in the K–Ar method) to ^{39}Ar produced by irradiation of ^{39}K in the nuclear reactor. The specimen is heated stepwise and the $^{40}\text{Ar}/^{39}\text{Ar}$ ratio of the unknown specimen is compared to a standard with known $^{40}\text{Ar}/^{39}\text{Ar}$ ratio, K concentration and age that has been simultaneously analyzed with the unknown specimen. In this way, the Ar–Ar age is a relative age compared to a standard, rather than a result of a direct measurement of parent and daughter isotope, but the stepwise nature of the Ar release yields information about the distribution of K and Ar in the mineral structure that can inform researchers about Ar loss.

Rb–Sr

Rubidium is an alkali metal (Rb^{+1}) and its behavior in geochemical systems is much like the common alkali metals sodium and potassium, so it substitutes into minerals in place of K^{+} . This substitution is facilitated by the identical charge and the similarity in ionic radii (12-fold coordination) of Rb^{+} (1.72 Å) and K^{+} (1.64 Å) (Table 2.1). The method is very similar to K–Ar but the decay product ^{87}Sr is divalent and Sr^{+2} is held more strongly in the mineral lattice than Ar^0 . For this reason, closure temperatures for Sr are higher than for Ar, and Rb–Sr ages tend to be closer to crystallization

ages than K–Ar ages. Also, Sr^{+2} tends not to diffuse out of crystal matrices as rapidly as Ar, and therefore Rb–Sr dates are less prone to resetting by subsequent thermal events; for this reason, K–Ar used in conjunction with Rb–Sr has been applied to dating paleohydrology/fluid flow in sedimentary basins Clauer et al. (2003).

U–Th–Pb

Uranium in nature consists primarily of ^{238}U (99.27%) with the remainder occurring as ^{235}U (0.72%) and trace amounts of ^{234}U (0.0057%). ^{238}U and ^{235}U have long half-lives of 4.47×10^9 yr and 7.04×10^8 yr, but ^{234}U is less stable and decays to ^{230}Th by alpha decay with a half-life of 2.47×10^5 yr (Fig. 6.3; Table 6.1).

The main naturally occurring isotope of thorium is ^{232}Th with a very long half-life of 1.41×10^{10} yr. The other thorium isotopes (^{227}Th , ^{228}Th , ^{230}Th , ^{231}Th , ^{234}Th) are all radiogenic and are produced in the decay chains of ^{232}Th , ^{235}U and ^{238}U . Half-lives of the naturally occurring radiogenic thorium isotopes are 18.7 days (^{227}Th), 1.92 yr (^{228}Th), 7.54×10^4 yr (^{230}Th), 25.5 h (^{231}Th) and 24.1 days (^{234}Th).

While uranium occurs as U^{+4} and U^{+6} ions, thorium only occurs as Th^{+4} . The tetravalent ions have similar radii in 12-fold coordination ($\text{U}^{+4} = 1.17$ Å, $\text{Th}^{+4} = 1.21$ Å) and co-substitute in mineral structures; therefore, U and Th coexist in many minerals, notably as trace elements in zircon, sphene, monazite and apatite.

Radioactive decay of U and Th eventually results in stable isotopes of Pb, and these decay series (Figures 11.2 and 11.3) are the basis of U–Th–Pb dating methods. The only Pb isotope that is not radiogenic is ^{204}Pb , so it is used for reference. The decay rates of the parent ^{238}U , ^{235}U and ^{232}Th are many orders of magnitude greater than their intermediate daughter products, meaning that the abundance of the parent isotope remains effectively unchanged for many half-lives of the daughter isotopes.

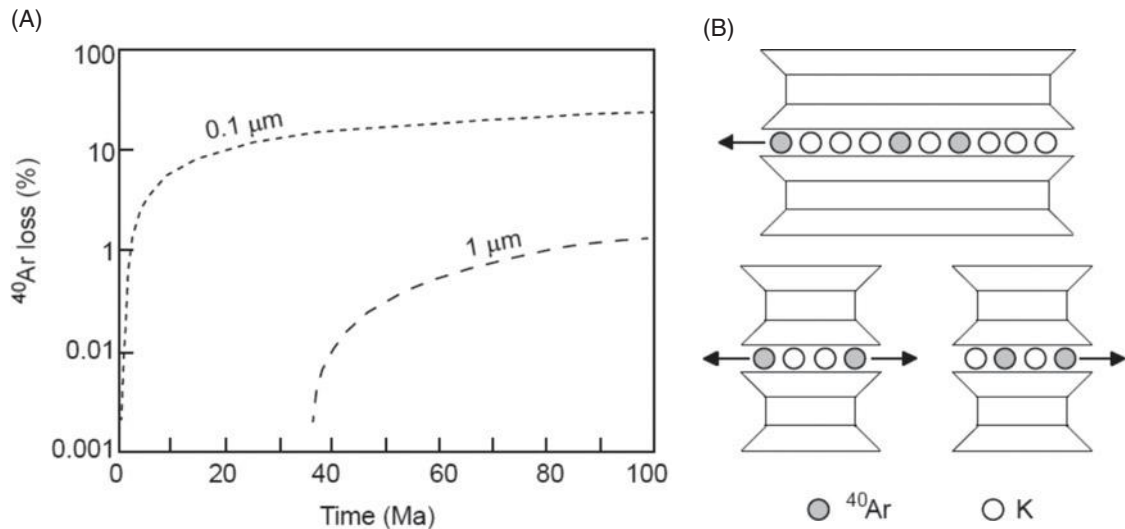


Fig. 11.12 Diffusion of radiogenic ^{40}Ar out of illite crystals. (A) Grain size exerts a strong control on Ar loss, where finer grained crystals (e.g. $0.1\ \mu\text{m}$ effective radius) are more prone to Ar loss than are coarser-grained crystals (e.g. $1\ \mu\text{m}$ effective radius). This is because fine-grained crystals have shorter diffusion domains than do coarser crystals (Modified after Clauer et al., 2003). (B) The smaller the crystal, the more likely it is to have a greater proportion of Ar atoms at crystal edges where diffusion out of the crystal lattice is more likely to occur.

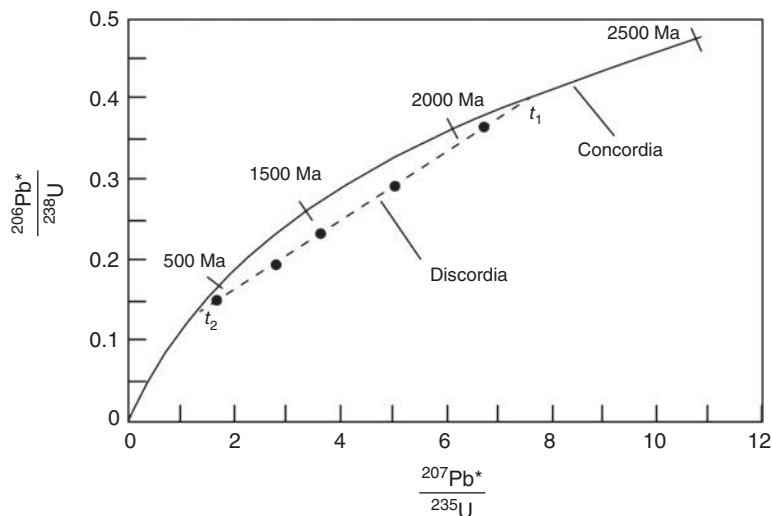


Fig. 11.13 Representative concordia–discordia diagram. The older date (t_1) is typically interpreted to be the date when the system became closed (e.g. timing of crystallization of zircon) and the younger date (t_2) corresponds to the age of an event (commonly metamorphism) that caused Pb loss. In this case, one possible interpretation would be that the zircon first crystallized 2200 Ma (e.g. during cooling of granitic magma) and then was affected by metamorphism at ~ 400 Ma.

Dating by U–Th–Pb methods relies on two important **assumptions**:

1 The existence of secular equilibrium (which occurs when the ratio of parent isotope to daughter isotope reaches a constant value).

2 The decay chain occurs within a closed system – this is because loss of daughter products would (1) alter the secular equilibrium and (2) produce erroneously low measurements of the daughter product needed to quantify age (e.g. ^{206}Pb), which would result in an anomalously young age. However, minerals used for U and Th series dating tend to be refractory (stable) silicates such as zircon (ZrSiO_4) and the phosphate monazite ($[\text{CeLaNdTh}][\text{PO}_4]$). Other than in the case of zircon, where the crystal structure allows substitution of U^{+4} and Th^{+4} for Zr^{+4} and inhibits incorporation of Pb^{+2} , an initial ratio of inherited radiogenic lead (i.e. present at the time of crystallization) must also be assumed. For this reason, zircon is the most well-suited mineral for U–Th–Pb dating.

The three equations used in dating igneous or metamorphic minerals using U or Th decay are:

$$^{206}\text{Pb}/^{204}\text{Pb} = \left(^{206}\text{Pb}/^{204}\text{Pb}\right)_i + \left(^{238}\text{U}/^{204}\text{Pb}\right) * e^{-\lambda_a * t} \quad (11.30)$$

$$^{207}\text{Pb}/^{204}\text{Pb} = \left(^{207}\text{Pb}/^{204}\text{Pb}\right)_i + \left(^{235}\text{U}/^{204}\text{Pb}\right) * e^{-\lambda_b * t} \quad (11.31)$$

$$^{208}\text{Pb}/^{204}\text{Pb} = \left(^{208}\text{Pb}/^{204}\text{Pb}\right)_i + \left(^{232}\text{Th}/^{204}\text{Pb}\right) * e^{-\lambda_c * t} \quad (11.32)$$

λ_a , λ_b and λ_c correspond to decay constants for ^{238}U , ^{235}U and ^{232}Th , respectively, and the first term to the right of the equal sign represents the initial ratio of inherited radiogenic lead (^{206}Pb , ^{207}Pb or ^{208}Pb , depending on parent isotope) to stable ^{204}Pb present in the mineral when it crystallized.

Ideally, the decay chains of ^{238}U , ^{235}U and ^{232}Th to ^{206}Pb , ^{207}Pb and ^{208}Pb , respectively, should result in identical mineral crystallization ages, but this is usually not the case and the ages are **discordant**. The main cause of discordance is loss of daughter isotopes in the long decay chains from U or Th parent isotope to the formation of stable isotopes of lead at the end of the decay chains. Some of the daughter products are reasonably mobile, including the various isotopes of radon, which can be lost from the system as a gas, and radium, which is moderately soluble in many fluids – the loss of these daughter products from the mineral to be dated will result in anomalously low lead concentrations and dates that are younger than the crystallization age. Radioactive decay also causes damage to crystal lattices and this too can lead to loss of daughter products. Alpha decay, for example, produces a high-energy particle with an atomic mass of 4 Daltons, and the migration of an α particle will disrupt the crystal lattice and lead to greater solubility that enhances loss of certain daughter products.

The benefit of the combined U–Th system is that it provides internal checks on age determinations. As depicted in Fig. 11.13, there are two separate isotope pairs that yield individual ages, and loss of lead is indicated by points that fall below the concordia curve. If discordant ages from zircons or similar U-bearing minerals in a rock plot as a straight line, the upper intercept with the concordia curve corresponds to the crystallization age of the minerals (e.g. initial igneous crystallization age), whereas the intersection of the discordia line with the concordia curve may represent the amount of time elapsed since lead was lost from the sample.

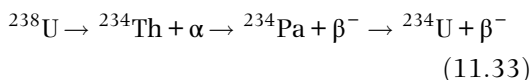
11.4.3 $^{234}\text{U}/^{238}\text{U}$ and ^{234}U disequilibrium

While the above-described U–Th–Pb methods are applicable for determining crystallization ages of igneous, sedimentary and metamorphic minerals

millions to billions of years old, U–Th decay can also be applied to low-temperature systems such as soils and groundwater where the ages in question are thousands to hundreds of thousands of years. These methods are based on the disequilibrium that can develop between parent and daughter isotopes in the long decay chains of ^{238}U and ^{232}Th , where the varied geochemical behaviors of the many decay products of ^{238}U and ^{232}Th mean that some of the more soluble species can become separated from their parent isotopes, particularly in aqueous systems.

One of the more common low-temperature geochronometers and tracers of geochemical processes in the U and Th decay series is the $^{234}\text{U}/^{238}\text{U}$ method, an approach that is based on the disequilibrium that develops between ^{238}U and its mobile daughter product ^{234}U .

^{234}U is produced by decay of ^{238}U in a series that includes intermediate isotopes of thorium and protactinium whose half-lives are 24.1 days and 6.7 hours, respectively (Table 6.1). Therefore, the decay of ^{238}U to ^{234}U via ^{234}Th and ^{234}Pa occurs with a combined half-life of a mere 24.38 days.



Unlike the conditions of secular equilibrium that tend to occur in minerals older than a few million years old (where the ratio of $^{234}\text{U}/^{238}\text{U}$ becomes constant), disequilibrium between ^{234}U and ^{238}U tends to develop when uranium-bearing minerals undergo spontaneous decay in soil or groundwater. An important term applied to describe systems with respect to secular equilibrium or disequilibrium is the **activity ratio** of parent to daughter that in the case of a ^{234}U – ^{238}U system in secular equilibrium is defined as:

$$\text{AR} = \lambda(^{234}\text{U}) * N(^{234}\text{U}) / \lambda(^{238}\text{U}) * N(^{238}\text{U}) = 1 \quad (11.34)$$

Given the decay constants for ^{234}U and ^{238}U of $2.81 \times 10^{-6} \text{ yr}^{-1}$ and $1.55 \times 10^{-10} \text{ yr}^{-1}$, respectively,

the ratio of atoms of ^{234}U and ^{238}U is approximately 5.5×10^{-5} under conditions of secular equilibrium, i.e. when $\text{AR} = 1$.

Upon decay of ^{238}U to ^{234}U , secular equilibrium tends to be lost due to preferential release of ^{234}U into solution, which increases the $^{234}\text{U}/^{238}\text{U}$ activity ratio of the newly liberated uranium (in solution) to values that are > 1 , and in some cases as high as $\text{AR} = 10$. On the other hand, preferential loss of ^{234}U from minerals can result in minerals with AR values < 1 , which is one of the potential sources of error in the U–Th–Pb methods described above.

^{234}U tends to be partitioned into the aqueous phase at a greater rate than ^{238}U for the following reasons:

- 1 Damage to the crystal lattice caused by alpha decay that facilitates leaching of newly formed ^{234}U out of crystals into solution.
- 2 Radiation produced by decay of ^{238}U to ^{234}U via ^{234}Th and ^{234}Pa can oxidize ^{234}U to the soluble U^{+6} ion, enhancing its migration into solution.
- 3 The kinetic energy of alpha particles emitted upon spontaneous decay of ^{238}U can cause newly formed ^{234}Th daughter product to be blasted out of the crystal and into solution – this particularly applies to cases where the parent ^{238}U atom is located along the margins of a crystal. Rapid decay of ^{234}Th to ^{234}U through ^{234}Pa (Table 11.1) then can result in dissolved ^{234}U in solution and ^{238}U that remains locked in the crystal lattice. A schematic of alpha recoil is provided in Fig. 11.5.

The concentration and isotopic composition of uranium that is leached into groundwater can be used as a tracer of groundwater recharge zones and flow patterns. For example, in a study designed to examine groundwater flow patterns in and around the proposed Yucca Mountain high-level radioactive waste repository in Nevada, USA, Paces et al. (2002) noted relatively low uranium concentrations (0.8 to 2.5 ppb) and high $^{234}\text{U}/^{238}\text{U}$ ratios (7 to 8) in groundwater from the aquifer beneath Yucca Mountain as compared to groundwater from surrounding aquifers

that contain relatively high and variable uranium concentrations (0.6 to 10 ppb) and lower $^{234}\text{U}/^{238}\text{U}$ ratios (1.5 to 6). These differences in $[\text{U}]$ and $^{234}\text{U}/^{238}\text{U}$ were interpreted to indicate that the groundwater beneath Yucca Mountain is isolated from surrounding aquifers. Furthermore, $^{234}\text{U}/^{238}\text{U}$ ratios were very similar in saturated zone and unsaturated zone groundwater beneath Yucca Mountain, implying that recharge of the aquifer beneath Yucca Mountain is controlled by localized down-gradient flow. If confirmed by independent hydrogeological analyses, the apparent isolation of groundwater beneath Yucca Mountain from other regional aquifers would comprise an important factor when considering site suitability for waste storage or disposal.

Flow of groundwater into the oceans either directly or via streams delivers uranium to the oceans with a homogenized $^{234}\text{U}/^{238}\text{U}$ ratio of 1.15 (Miyake et al., 1966). When calcium carbonate (aragonite or calcite) precipitates from ocean water it incorporates uranium (typically in the range of 1 ppm), and provided a $^{234}\text{U}/^{238}\text{U}$ ratio (typically 1.15) can be assumed, the decay of ^{234}U to ^{230}Th by alpha decay can be used to take crystallization of marine carbonates. With a half-life of 2.48×10^5 yr, ^{234}U decay is well-suited for dating coral reefs and carbonates with ages of tens of thousands to hundreds of thousands of years. The important assumptions here are (1) that the system has remained closed to uranium, and (2) that the initial ratio of $^{234}\text{U}/^{238}\text{U}$ is known.

The rapid decay of ^{234}U to ^{230}Th as compared to the slow decay of ^{238}U to ^{234}U (via ^{234}Th and ^{234}Pa isotopes) means that with time the activity ratio of $^{234}\text{U}/^{238}\text{U}$ will exponentially decrease from 1.15 to 1.0 – when the $^{234}\text{U}/^{238}\text{U}$ AR reaches a value of 1.0, the isotope pair will have returned to a state of secular equilibrium (Fig. 11.14). For an initial $^{234}\text{U}/^{238}\text{U}$ AR of 1.15, reaching AR = 1.0 requires approximately 1 million years, thus bracketing the applicability of this method to specimens with maximum ages of hundreds of thousands of years.

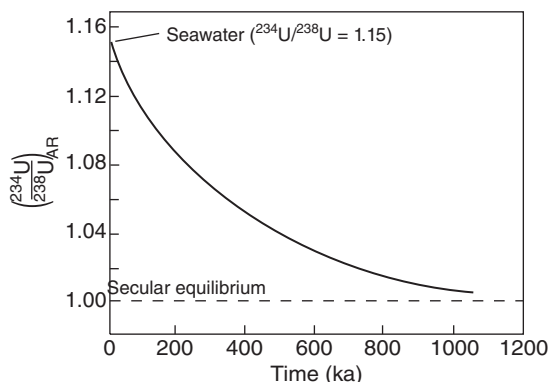


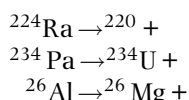
Fig. 11.14 Decrease in ratio of $^{234}\text{U}/^{238}\text{U}$ with time. The decrease is related to relatively rapid decay of ^{234}U compared to ^{238}U , and the decreasing ratio with time facilitates dating of marine carbonates (e.g. corals).

Unlike uranium, which typically comprises 1 ppm of marine carbonates, thorium is excluded from carbonates when they initially crystallize, meaning that any ^{230}Th in a carbonate mineral was almost certainly produced by decay of ^{234}U . Assuming an initial $^{234}\text{U}/^{238}\text{U}$ activity ratio of 1.15, the decay of both inherited (excess) ^{234}U and ^{238}U to ^{230}Th can be used to determine the age of carbonate crystallization. These methods and their assumptions and corrections were developed by Broecker (1963) and are presented in detail by Faure (1986).

The radioactive isotopes presented here include representative cases of radionuclide contaminants as well as some of the most commonly applied methods for dating sediments, waters and rocks. Some of the dating methods not included in this chapter include the Sm–Nd, Lu–Hf and Re–Os methods, fission track age analysis, and various methods pertaining to specific isotope pairs (e.g. ^{230}Th – ^{226}Ra , ^{230}Th – ^{232}Th) contained within the lengthy ^{238}U , ^{235}U and ^{232}Th decay chains. Information on these methods is presented in more detailed accounts by Faure (1986) and Dickin (2005).

QUESTIONS

1. Complete the following radioactive decay reactions:



2. List, describe and indicate the differences among the three main forms of radiation released by radioactive decay.
3. Would you predict uranium to be more mobile in shallow, oxidizing groundwater or deeper, reducing groundwater? How does U redox chemistry lead to formation of roll-front U ore deposits?
4. Explain why the insolubility of thorium makes ${}^{234}\text{U}/{}^{238}\text{U}$ dating possible.
5. Compare the expected mobility of two classes of high-level radioactive wastes that could leak out of storage areas, cationic forms (e.g. ${}^{90}\text{Sr}^{+2}$ and ${}^{137}\text{Cs}^{+}$) and anionic forms (e.g. ${}^{99}\text{TcO}_4^{-}$ and ${}^{131}\text{I}^{-}$). If their solubilities are approximately equal, consider two other factors that will influence their mobilities.
6. Describe the main principles behind K–Ar dating, including the K–Ar decay reaction and important assumptions. What are two criticisms of these assumptions?
7. What is a spallation reaction? What instigates spallation? Provide an example.
8. What is the origin of ${}^{14}\text{C}$ and how is it incorporated into living material? What is the daughter product of ${}^{14}\text{C}$, and how are radiocarbon ages determined if the decay product escapes as a gas? What is the radiocarbon age in years B.P. of a shell determined to be 2230 yr old in 2012?
9. Describe how ${}^{36}\text{Cl}$ can be applied to dating groundwater. In particular, indicate how it forms, how it decays and why Cl is a better choice of a cosmogenic radionuclide for dating groundwater than ${}^{26}\text{Al}$ or ${}^{10}\text{Be}$.

10. What are three advantages to ${}^{10}\text{Be}$, ${}^{26}\text{Al}$ and ${}^{36}\text{Cl}$ dating methods compared to ${}^{14}\text{C}$ dating?
11. What correction needs to be made when determining sediment depositional ages by the ${}^{210}\text{Pb}$ method? In your reply include an equation and brief description.

REFERENCES

- Amidon, W.H. and Farley, K.A., 2011. Cosmogenic ${}^3\text{He}$ dating of apatite, zircon and pyroxene from Bonneville flood erosional surfaces: Quaternary Geochronology, 6: 10–21.
- Arnold, J. R. and Libby, W. F., 1949. Age Determinations by Radiocarbon Content: Checks with Samples of Known Age. *Science* 110 (2869): 678–680.
- Bish, D.L. and Aronson, J.L., 1993. Paleogeothermal and paleohydrologic conditions in silicic tuff from Yucca Mountain, Nevada. *Clays and Clay Minerals* 41: 148–161.
- Bish, D.L., Vaniman, D.T., Chipera, S.J., and Carfey, J. W., 2003. The distribution and importance of zeolites at Yucca Mountain, Nevada, USA, *American Mineralogist* 88: 1889–1902.
- Blum, J. D., Gazis, C. A., Jacobson, A. D., and Chamberlain, C. P., 1998. Carbonate versus silicate weathering in the Raikhot watershed within the High Himalayan Crystalline Series. *Geology* 26: 411–414.
- Broecker, W.S., 1963. A preliminary evaluation of uranium series in equilibrium as a tool for absolute age measurements on marine carbonates. *Journal of Geophysical Research* 68: 2817–2834.
- Cerling, T. E. 1990. Dating geomorphic surfaces using cosmogenic ${}^3\text{He}$. *Quaternary Research* 33: 148–156.
- Cerling, T. E., Poreda, R. J., and Rathburn, S. L., 1994. Cosmogenic ${}^3\text{He}$ and ${}^{21}\text{Ne}$ age of the Big Lost River Flood, Snake River Plains, Idaho. *Geology* 22: 227–230.
- Choppin, G.R., 2007. Actinide speciation in the environment. *Journal of Radioanalytical and Nuclear Chemistry* 273: 695–703. DOI: 10.1007/s10967-007-0933-3.
- Chow, T.J., Snyder, C.B., and Earl, J.L., 1975. Isotope ratios of lead as pollutant source indicators.

- Proceedings of the International Atomic Energy Agency (Vienna) 191: 95–108.
- Clauer, N., Zwingmann, H., and Gorokhov, I., 2003. Postdepositional evolution of platform claystones based on a simulation of thermally induced diffusion of radiogenic ^{40}Ar from diagenetic illite. *Journal of Sedimentary Research* 73: 58–63.
- Clow, D.W., Mast, M.A., Bullen, T.D., and Turk, J.T., 1997. Strontium 87/strontium 86 as a tracer of mineral weathering reactions and calcium sources in an Alpine/Subalpine Watershed, Loch Vale, Colorado. *Water Resources Research* 33: 1335–1351.
- Cresswell, R.G., Jacobson, G., Wischusen, J., and Fifield, L.K., 1999. Ancient groundwaters in the Amadeus Basin, Central Australia: evidence from the radioisotope ^{36}Cl . *Journal of Hydrology* 223: 212–220.
- Dahl, P.S., 1996. The crystal-chemical basis for Ar retention in micas: inferences from interlayer partitioning and implications for geochronology. *Contributions to Mineralogy and Petrology* 123: 22–39. DOI: 10.1007/s004100050141.
- deVries, H. and Barendsen, G.W., 1953. Radiocarbon dating by a proportional counter filled with carbon dioxide. *Physica* 19: 987–1003.
- deVries, H., Barendsen, G.W., and Waterbolk, H. T., 1958. Groningen radiocarbon dates II: *Science* 127: 129–137.
- Dickin, A.P., 2005. *Radiogenic Isotope Geology* (2nd edn). Cambridge University Press, Cambridge.
- Dunai, T.J. and Stuart, F.M., 2009. Reporting of cosmogenic nuclide data for exposure age and erosion rate determinations. *Quaternary Geochronology* 4: 437–440.
- Elliott, W.C. and Aronson, J.L., 1987. Alleghanian episode of K-bentonite illitization in the southern Appalachian Basin: *Geology* 15: 735–739.
- Fabryka-Martin, J., Bentley, H., Elmore, D., and Airey, P. L., 1985. “Natural iodine-129 as an environmental tracer”. *Geochimica et Cosmochimica Acta*, 49: 337.
- Faure, G., 1986, *Principles of Isotope Geology* (2nd edn). Wiley, New York.
- Faure, G., 1998. *Principles and Applications of Geochemistry* (2nd edn). Prentice-Hall, Upper Saddle River, New Jersey, USA.
- Fifield, K. and Morgenstern, U., 2009. Silicon-32 as a tool for dating the recent past. *Quaternary Geochronology* 4: 400–405.
- Gauthier-Lafaye, F., Holliger, P., and Blanc, P.-L., 1996. Natural fission reactors in the Franceville basin, Gabon: A review of the conditions and results of a “critical event” in a geologic system. *Geochimica et Cosmochimica Acta* 60: 4831–4852.
- Granger, D.E. and Riebe, C.S. 2007. Cosmogenic Nuclides in Weathering and Erosion. In *Treatise on Geochemistry, Volume 5: Surface and Ground Water, Weathering, and Soils*. J. I. Drever (ed.) Elsevier, London.
- Harrison, T.M. and McDougall, I., 1982. The thermal significance of potassium feldspar K–Ar ages inferred from ^{40}Ar – ^{39}Ar age spectrum results. *Geochimica et Cosmochimica Acta* 46: 1811–1820.
- Holmes, A., 1911. The association of lead with uranium in rock-minerals, and its application to the measurement of geological time. *Proceedings of the Royal Society of London A* 85: 248–256.
- Hughen, K., Lehman, S.J., Southon, J.R., Overpeck, J.T., Marchal, O., Herring, C., and Turnbull, J., 2004. ^{14}C activity and global carbon cycle changes over the past 50,000 years. *Science* 303: 202–207.
- Jacobson A.D., Blum J.D., Chamberlain C.P., Craw D., and Koons P.O., 2003. Climatic and tectonic controls on chemical weathering in the New Zealand Southern Alps. *Geochimica et Cosmochimica Acta* 67: 29–46.
- Keywood, M.D., Chivas, A., Fifield, L.K., Cresswell, and R.G. Ayers, G., 1997. The accession of salt to the Western half of the Australian continent. *Australian J. Soil Science* 35: 177–1189.
- Kurz, M.D., Colodner, D., Trull, T.W., Moore, R.B., and O'Brien, K., 1990. Cosmic-ray exposure dating with in situ produced cosmogenic ^3He : results from young Hawaiian lava flows. *Earth and Planetary Science Letters* 97: 177–189.
- Lal, D., 1991. Cosmic ray labeling of erosion surfaces: In situ production rates and erosion models. *Earth and Planetary Science Letters* 104: 424–439.
- Langmuir, D., 1997. *Aqueous Environmental Geochemistry*. Prentice Hall, Upper Saddle River, New Jersey, USA, 600 p.
- Licciardi, J.M., Kurz, M.D., Clark, P.U., and Brook, E.J., 1999. Calibration of cosmogenic ^3He production rates from Holocene lava flows in Oregon, USA, and effects of the Earth's magnetic field. *Earth and Planetary Science Letters* 172: 261–271.

- Miyake, Y., Sugimura, Y., and Uchida, T., 1966. Ratio U^{234}/U^{238} and the uranium concentration in seawater in the western North Pacific. *Journal of Geophysical Research* 71: 3083–3087.
- Montazer, P. and Wilson, W.E., 1984. Conceptual Hydrologic Model of Flow in the Unsaturated Zone, Yucca Mountain, Nevada. USGS Water-Resources Investigations Report 84–4345: 55 p.
- Murozumi, M.T., Chow, J., and Patterson, C.C., 1969. Chemical concentration of pollutant lead aerosols, terrestrial dusts and sea salts in Greenland and Antarctic snow strata. *Geochimica et Cosmochimica Acta* 33: 1247–1294.
- Nishiizumi, K., Lai, D., Klein, J., Middleton, R., and Arnold, J.R., 1986. Production of ^{10}Be and ^{26}Al by cosmic rays in terrestrial quartz in situ and implications for erosion rates. *Nature* 319: 134–136.
- Noakes, J.F., Kim, S.M., and Akers, L.K., 1967. Recent improvements in benzene chemistry for radiocarbon dating. *Geochimica et Cosmochimica Acta* 31: 1094–1096.
- Paces, J.B., Ludwig, K., Peterman, Z., and Neymark, L., 2002. $^{234}U/^{238}U$ evidence for local recharge and patterns of ground-water flow in the vicinity of Yucca Mountain, Nevada, USA. *Applied Geochemistry*, 17 (6): 751–779.
- Paul, M., Fink, D., Hollos, G., Kaufman, A., Kutschera, W., and Magaritz, M., 1987. Measurement of ^{129}I in the environment after the Chernobyl reactor accident. *Nuclear Instruments and Methods – Physics Research*, B29: 341–345.
- Poreda, R. J. and T. E. Cerling, 1992. Cosmogenic neon in recent lavas from the western United States, *Geophysical Research Letters*, 19(18): 1863–1866.
- Quade, J., Cerling, T.E., and Bowman, J.R., 1989. Development of Asian monsoon revealed by marked ecological shift during the latest Miocene in northern Pakistan, *Nature* 342: 163–166.
- Reimer, P.J., Baillie, M.G.L., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P.G., Buck, C.E., Burr, G.S., Cutler, K. B., Damon, P.E., Edwards, R.L., Fairbanks, R.G., Friedrich, M., Guilderson, T.P., Herring, C., Hughen, K. A., Kromer, B., McCormac, F.G., Manning, S.W., Ramsey, C.B., Reimer, P.J., Reimer, R.W., Remmele, S., Southon, J.R., Stuiver, M., Talamo, S., Taylor, F. W., van der Plicht, J., and Weyhenmeyer, C.E., 2004. IntCal04 Terrestrial radiocarbon age calibration, 0–26 cal kyr BP. *Radiocarbon* 46: 1029–1058.
- Reusser, L., Bierman, P., Pavich, M., Larsen, J., and Finkel, R., 2007. An episode of rapid bedrock channel incision during the last glacial cycle, measured with ^{10}Be . *American Journal of Science* 306: 69–102.
- Rosman, K.J.R., Chisholm, W., Boutron, C.F., Candelone, J.P., and Görlach, U., 1993. Isotopic evidence for the source of lead in Greenland snows since the late 1960s. *Science* 362: 333–335.
- Seong, Y.B., Owen, L.A., Yi, C., and Finkel, R.C., 2009. Quaternary glaciation of Muztag Ata and Kongur Shan: Evidence for glacier response to rapid climate changes throughout the Late Glacial and Holocene in westernmost Tibet. *Geological Society of America Bulletin* 121: 348–365. doi: 10.1130/B26339.1.
- Shirahata H., Elias R.W., and Patterson C.C., 1980. Chronological variations in concentrations and isotopic compositions of anthropogenic atmospheric lead in sediments of a remote subalpine pond. *Geochimica et Cosmochimica Acta* 49: 149–162.
- Steiger, R.H. and Jäger, E., 1977. Subcommission on geochronology: Convention on the use of decay constants in geo- and cosmochronology. *Earth and Planetary Science Letters* 36: 359–362.
- Sturges W.T. and Barrie L.A., 1989. The use of stable lead 206/207 isotope ratios and elemental composition to discriminate the origin of lead in aerosols at a rural site in eastern Canada. *Atmospheric Environment* 23: 1645–1657.
- Stuiver, M., Reimer, P.J., Bard, E., Beck, J.W., Burr, G.S., Hughen, K.A., Kromer, B., McCormac, G., van der Plicht, J., and Spurk, M., 1998. IntCal98 radiocarbon age calibration, 24,000–0 cal BP. *Radiocarbon* 40: 1041–83.
- Suess, H.E., 1955. Radiocarbon concentration in modern wood: *Science*, 122: 415–417.
- Turekian, K.K. and Wedepohl, K.H., 1961. Distribution of the Elements in some major units of the Earth's crust. *Geological Society of America Bulletin* 72: 175–192.
- Urey, H., 1947. The thermodynamic properties of isotopic substances, *Journal of the Chemical Society (London)*: 562–581.
- Vengosh, A., Hirschfeld, D., Vinson, D.S., Dwyer, G.S., Raanan, H., Rimawi, O., Al-Zoubi, A., Akkawi, E., Marie, A., Haquin, G., Zaarur, S., and Ganor, J.,

2009. High naturally occurring radioactivity in fossil groundwater in the Middle East. *Environmental Science and Technology* 43: 1769–1775. DOI: 10.1021/es802969r.
- Vinogradov, A.P., 1962. Average contents of chemical elements in the principal types of igneous rocks of the Earth's crust. *Geochemistry* 7: 641–664.
- Wanty, R. C. and Nordstrom, D. K., 1993. Natural radionuclides. In Alley, W. M., ed., *Regional Ground-Water Studies*, Van Nostrand Co., New York, Chapter 17: 423–441.
- World Health Organization. 2004. *Guidelines for Drinking Water Quality*, 3rd edn, Geneva, Switzerland: World Health Organization.
- Ylagan, R.F., Kim, C.S., Pevear, D.R., and Vrolijk, P.J., 2002. Illite polytype quantification for accurate K–Ar age determination. *American Mineralogist* 87: 1536–1545.

Appendix I

Case Study on the Relationship of Volatile Organic Compounds (VOCs), Microbial Activity, Redox Reactions, Remediation and Arsenic Mobility in Groundwater

This case study presents an interesting example of interactions and reactions among inorganic species (e.g. iron hydroxides, arsenic), organic compounds (e.g. organochlorine solvents) and biotic factors (e.g. microbially mediated mineral dissolution). In order to provide the reader with the opportunity to analyze and interpret, results are presented prior to – and separately from – interpretations. This case study was selected because it integrates concepts presented throughout the text, including kinetics and thermodynamics, the roles of minerals and microbes in elemental cycling, naturally derived inorganic and anthropogenic organic contaminants in aqueous systems, oxidation–reduction and sorption–desorption reactions, stable isotopes (ongoing investigation) and numerous analytical approaches (XRD, XAS, ICP-MS, ion and liquid chromatography). The following information is mainly from US EPA reports (e.g. Ford et al., 2008) and a pair of peer-reviewed journal articles

(He et al., 2010; Ford et al., 2011), with additional updates from recent conference abstracts (Hildum et al., 2012; Ahmed et al., 2013). It is worth examining the original sources for the data and interpretations they contain and also because they illustrate ways in which scientific results are published (e.g. focused 12-page journal article vs. comprehensive 193-page EPA report vs. 300-word conference abstract).

I.1 SITE INFORMATION, CONTAMINANT DELINEATION

The site is located at Fort Devens, a former military base in central Massachusetts (USA) (Fig. I.1). This appendix focuses on two areas: (1) Shepley's Hill Landfill (Fig. I.2A), a disposal site for household waste and construction debris from 1917 to 1992; and (2) an airfield 2 km north of the landfill where VOCs (chlorinated solvents

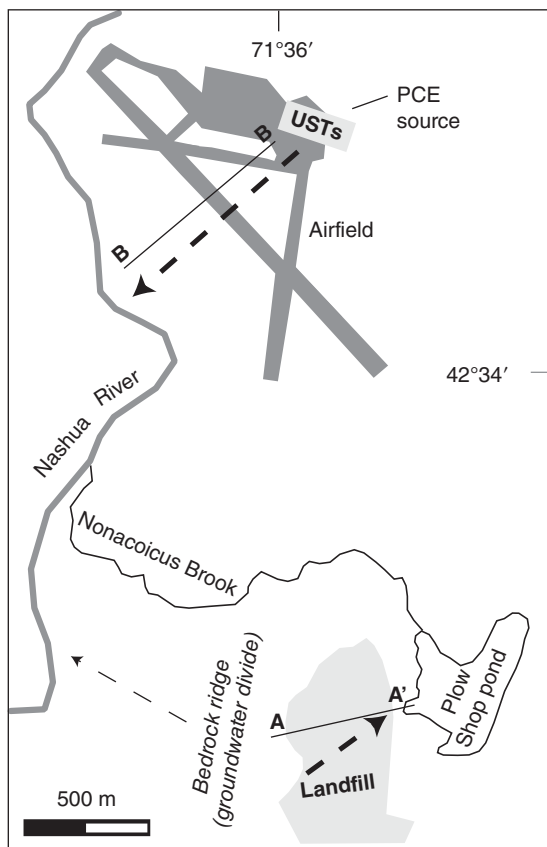


Fig. I.1 Location map of Shepley's Hill landfill (bottom center) and the airfield at Fort Devens US EPA Superfund site. Note cross Sections A–A' and B–B', which are shown in Fig. I.2.

and fuels) were stored in underground storage tanks (USTs) (Fig. I.2B). The airfield is situated on a Pleistocene terrace.

In 1991 the EPA began investigations to delineate contamination:

- Monitoring wells drilled through the landfill into the sandy aquifer below, and also in groundwater downgradient from the landfill, revealed elevated concentrations of numerous inorganic trace elements, including As (as high as 15 000 µg/L; Hildum et al., 2012), Cd, Fe, Mn and Pb. For reference, the US EPA maximum

contaminant level (MCL) for As in drinking water is 10 µg/L (10 ppb). VOCs were also detected at relatively low levels in groundwater below the landfill in the 1990s: dichlorobenzenes comprised 11 mg/L in monitoring well SHL-20, and 1,2-dichloroethane was reported at 9.9 ppb in monitoring well SHM-93-10C. The landfill occupies a low point in the pre-landfill topography, and some of the landfill is below the water table (Fig. I.2A), suggesting reducing conditions in the deep part of the landfill. Interestingly, sediment cores into the aquifer below the landfill do not indicate elevated As in the solids of the aquifer, only in the groundwater itself (i.e. As is not elevated in the solid phase but is in the aqueous phase). Elevated As, Cd, Mn, Fe, Mn and Pb also occur in sediments and deep water of Plow Shop Pond, located down-gradient of the landfill;

- Monitoring wells installed in the shallow sandy aquifer downgradient of the potential VOC source at the airfield documented the presence of PCE as well as PCE-derived chlorinated metabolites (TCE, dichloroethene, vinyl chloride, dichloropropane, methylene chloride and dichloroethane). Close to the leaking UST, PCE concentration in groundwater was as high as 30 000 µg/L (the EPA MCL is 5 µg/L). Concentrations of PCE in soil were as high as 5.5 micrograms per kilogram (µg/kg or ppb) at 3 m below the surface and 3.2 µg/kg in groundwater sediments 15 m below surface. Given that PCE is synthetic, all PCE detected is anthropogenic (in contrast to As, which could be either naturally occurring or introduced by human activities).

I.2 REMEDIATION EFFORTS

Remediation at Shepley's Hill Landfill includes improving the cap on the landfill (to limit infiltration) and a pump-and-treat system downgradient from the landfill to remove As from groundwater. In order to address VOC contamination at the

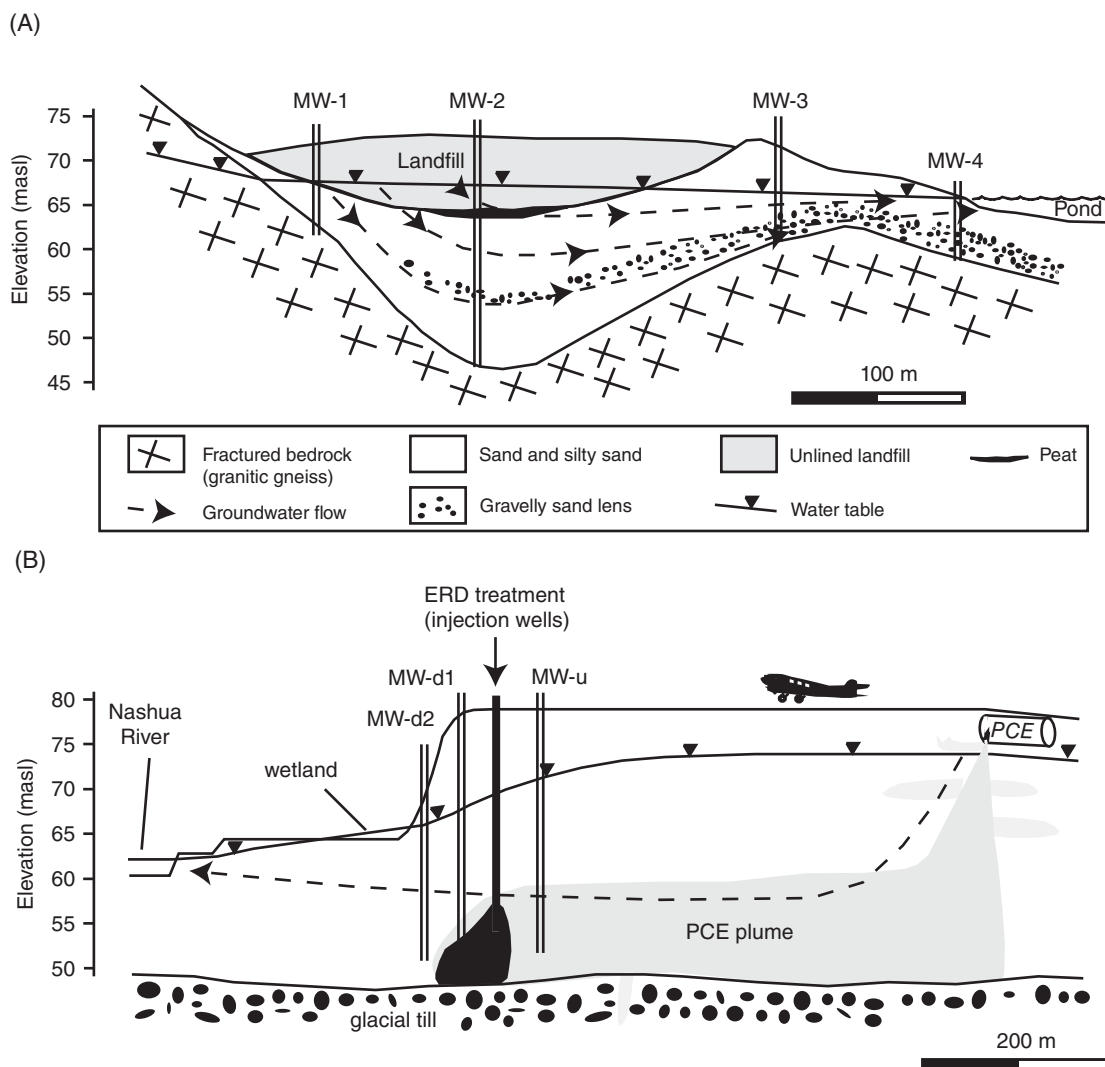


Fig. I.2 Cross sections through Shepley's Hill Landfill (A) and the airfield (B) at Fort Devens. "MW" indicates locations of some of the dozens of monitoring wells. (Ford et al. 2008. US EPA.)

airfield, a soil vapor extraction (SVE) system was installed adjacent to the former UST area in 1993 and later, after documenting the PCE plume migrating downgradient (Fig. I.2B), an enhanced reductive dechlorination (ERD) remediation system was installed to address VOC

contamination (refer to Fig. 3.25 and adjacent section for background on reductive breakdown of VOCs).

After many years of ERD treatment for VOCs at the airfield, monitoring showed the following (He et al., 2010):

1 upgradient of ERD injection (and within the PCE plume), concentrations of redox sensitive elements dissolved in groundwater are: As <1 µg/L, Fe < 0.1 mg/L, Mn = 0.1 mg/L, and SO₄ = 16.4 mg/L (MW-u, Fig. I.2B);

2 downgradient of ERD injection (MW-d1), As, Fe and Mn were significantly elevated (As = 680 µg/L, Fe = 420 mg/L and Mn = 5.0 mg/L), whereas SO₄ had decreased to 5.0 mg/L.

3 far downgradient of ERD injection (MW-d2), conditions are similar to the upgradient monitoring well: As, Fe and Mn are undetectable and SO₄ = 17.7 mg/L.

I.3 SOURCES OF PCE AND AS

The source of the **VOCs** (PCE and its metabolites) is clear – leaking USTs and disposal of used PCE into a dry well allowed these compounds to seep down into soil and then into groundwater, where the relatively high solubility of PCE (Table 3.4) and potential to behave as a DNAPL (Fig. 3.24) facilitated its dispersion in the aquifer.

The source of elevated **As and other inorganic elements** below and downgradient of the landfill, and in the vicinity of the ERD treatment, is open to interpretation and is still under investigation (Ahmed et al., 2013). If As, Cd, Fe, Mn and Pb were among the materials disposed of in the landfill, then leachate from the landfill itself could be an important source; however, unconsolidated sand and gravel aquifers with 1–20 mg/kg As are known to produce groundwater with concentrations of dissolved As like those seen below and downgradient of the Shepley's Hill landfill (Zheng et al., 2005), so the possibility that the As and other inorganics may be emanating from the natural sediments or bedrock of the aquifer system must also be considered.

One way to assess potential source of As and other inorganics is to examine whether or not the aquifer below the landfill contains a landfill

signature (Strudwick, 1999; Christensen et al., 2001). In landfill leachate:

- As tends to occur within the range of 10–1000 µg/L.
- Cl[−] is often indicative of leachate – it tends to occur in the range of 150–4500 mg/L.
- Organic compounds tend to be elevated; for example, aromatic hydrocarbons (e.g. benzene) usually comprise 10–10 000 µg/L and organochlorines (e.g. PCE, vinyl chloride) tend to range from 3–4000 µg/L.

In groundwater below the landfill:

- As concentrations as high as 15 000 µg /L (i.e. at least 15× that of leachate) suggest that landfill leachate is not the sole source of As.
- Cl[−] concentrations < 40 mg/L suggest that landfill leachate is a minor component of the groundwater below the landfill.
- The presence of VOCs confirms infiltration of leachate through the peat layer and into groundwater.

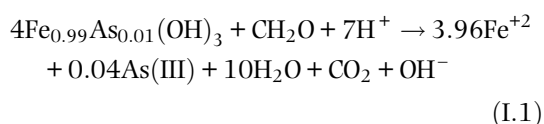
The evidence from monitoring wells installed in the aquifer below the landfill suggest that leachate from the unlined landfill has infiltrated the aquifer, but also that the leachate is not responsible for transporting the high levels of As and other inorganics from landfill into aquifer; rather, it appears that processes within the aquifer may be causing As and other inorganics to be mobilized. To examine the issue quantitatively, Ahmed et al. (2013) are using stable isotope analysis (δD and δ¹⁸O) to determine the signature of leachate water (relative to water derived from bedrock or other infiltration pathways) as well as potential isotopic fractionation (e.g. methanogenesis).

I.4 MOBILIZATION OF ARSENIC

High amounts of dissolved organic carbon (DOC; from decomposing food, paper, septic/sewage waste, etc.) in landfill leachate support diverse microbial populations and foster reducing/anaerobic conditions in aquifers below landfills

(Christensen et al., 2001). In reducing environments like this, iron hydroxides are thermodynamically unstable and undergo microbially mediated reductive dissolution. The iron hydroxide is used by the microbes as an electron acceptor and if As, Pb or other elements are sorbed to the iron hydroxide, they will be released to solution when the iron hydroxide dissolves. Kinetic factors that contribute to dissolution rate are likely abundance and species of microbes, abundance of DOC, and other aqueous factors such as pH and detachment rate/solubility of resulting Fe species.

Sequential chemical extraction and X-ray absorption spectroscopy indicate that As is sorbed to Fe-hydroxides in Fort Devens aquifer sediments (He et al., 2010), so a likely mechanism for release of As into solution (in the presence of microbes) is:



The downward seepage of organic matter (here shown as CH_2O) from the landfill, perhaps augmented by organic matter derived from peat sediments below the landfill, is a probable cause of the high concentrations of As observed. The effect of reduction on iron hydroxide is shown in the Eh–pH diagram in Fig. I.3, and given that Cd and Pb commonly sorb to Fe-hydroxides, this too could explain their high values. High Mn in the aquifer is probably due to reductive dissolution of Mn oxides by a process similar to that shown in Equation I.1.

Reductive dissolution of iron hydroxides as an important As source is also supported by the positive correlation of Fe and As in aquifer sediments (Fig. I.4).

The fate of As mobilized below the landfill is also related to redox processes. Given that groundwater below the landfill flows into Plow Shop Pond (Fig. I.1, right side of Fig. I.2a) or into Nonacoicus Brook, both of which are relatively oxidizing

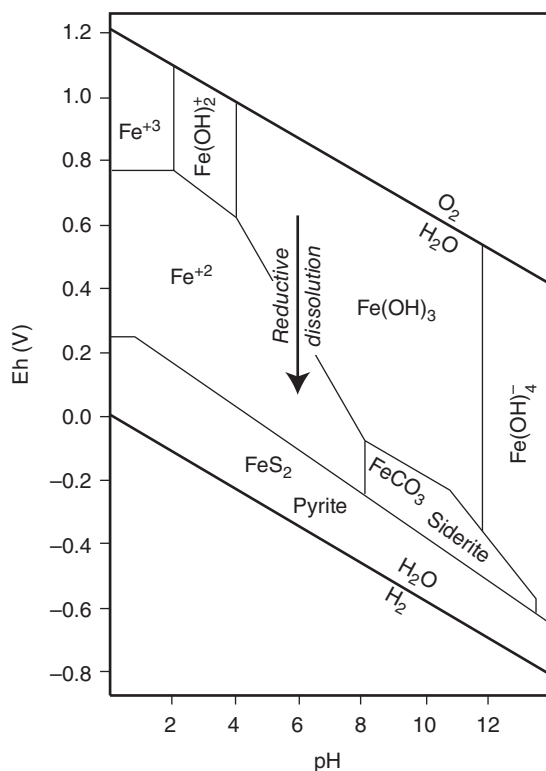


Fig. I.3 Eh–pH diagram illustrating destabilization of Fe-hydroxides associated with shift from oxidizing to reducing conditions. Refer to Fig. 4.15 for details.

environments compared to the groundwater, iron hydroxide precipitates at in sediments at the interface of groundwater and pond water (Ford et al., 2011). Reaction I.1 proceeds in reverse, incorporating As, Pb and other trace elements into solid iron hydroxides at the redox boundary (of reducing groundwater and oxidizing surface water).

The effect of ERD on concentrations of SO_4 , Mn, TOC, As and Fe in groundwater is shown in Fig. I.5, which depicts a transect through the ERD injection area, from upgradient of the injection to within the ERD zone to downgradient. The ingredient used to stimulate ERD was molasses, which was injected into wells installed into the PCE plume (Fig. I.2B) – molasses is the source of organic carbon that fosters reducing conditions and stimulates microbial activity.

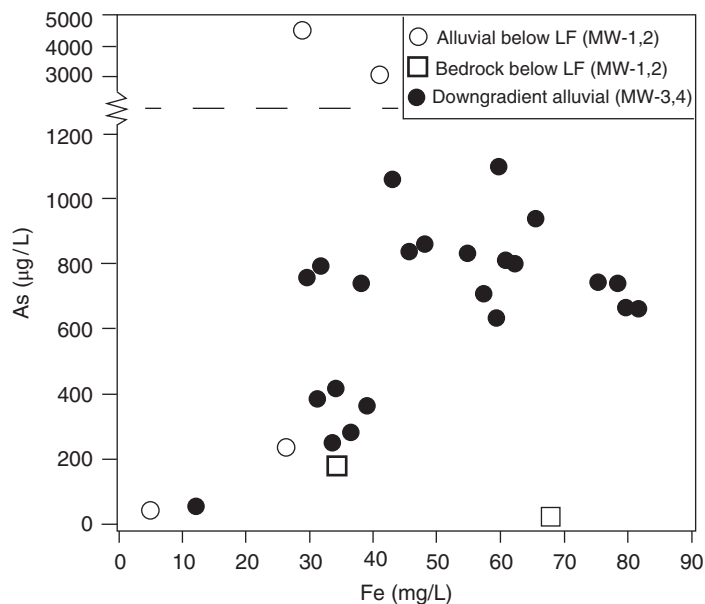


Fig. I.4 Positive correlation of Fe and As is consistent with dissolution of iron hydroxides as source of As in alluvial (sand and gravel) aquifer. “MW” = monitoring wells, locations of which are shown in Fig. I.2. “LF” = landfill.

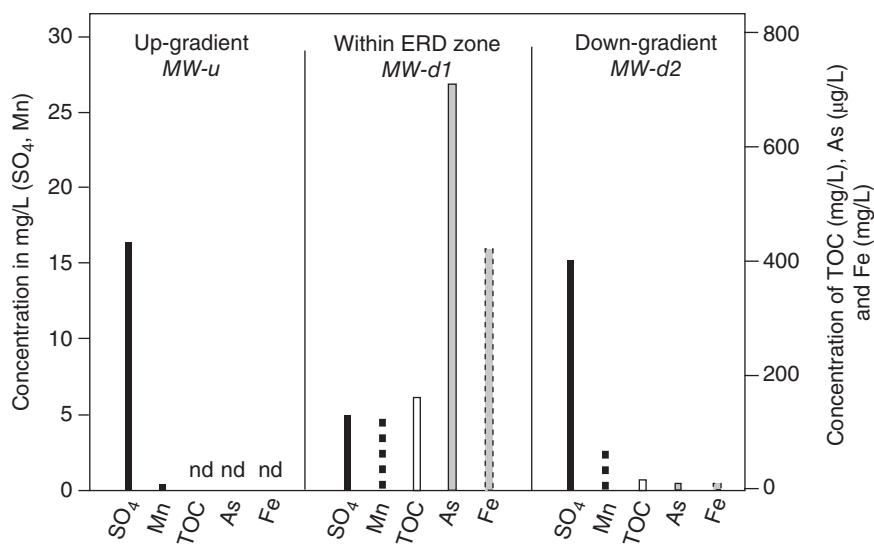


Fig. I.5 Concentrations in solution of redox-sensitive species SO₄, Mn, TOC, As and Fe, from upgradient of ERD injection (MW-u), through the main ERD injection area (MW-d1), and downgradient of ERD injection (MW-d2). Decreasing SO₄ in conjunction with increasing Mn, TOC, As and Fe are indication of strongly reducing conditions that cause SO₄ → S⁻² and reduction of Fe-hydroxides and Mn-oxides. Far downgradient (MW-d2), Fe-hydroxides are stable and As is sorbed out of solution. (From data in He et al. 2010.)

The role of microbes in enhancing reduction of Fe-hydroxides is supported by experiments of He et al. (2010) that reproduced bacteria-facilitated reduction of Fe-hydroxide in the presence of molasses. The enhanced mobility of As following ERD is not surprising – the EPA recognizes this as a potential consequence of ERD – but when faced with a complicated problem like a deep plume of PCE migrating in the direction of a river, sometimes there are no fool-proof solutions, as is often the case with complex environmental problems.

REFERENCES

- Ahmed, S., Hon, R., Hildum, B., and Xie, Y., 2013. Stable isotopes (δD and $\delta^{18}O$) fractionation in groundwater in the vicinity of an arsenic contaminant landfill plume in central Massachusetts. *Geological Society of America* 45 (1): p. 67.
- Christensen, T.H., Kjeldsen, P., Bjerg, P.L., Jensen, D.L., Christensen, J.B., Baun, A., Albrechtsen, H.J., and Heron, G., 2001. Biogeochemistry of landfill leachate plumes. *Applied Geochemistry* 16, no. 7–8: 659–718.
- Ford, R.G., Acree, S.D., Lien, B.K., Scheckel, K.G., Luxton, T.P., Ross, R.R., Williams, A.G., Clark, P., 2011. Delineating landfill leachate discharge to an arsenic contaminated waterway. *Chemosphere* 85(9): 1525–37.
- Ford, R.G., Scheckel, K.G., Acree, S.D., Ross, R., Lein, B., Luxton, T., Clark, P., 2008. Final Report *Arsenic Fate, Transport and Stability Study Groundwater, Surface Water, Soil and Sediment Investigation, Fort Devens Superfund Site, Devens, Massachusetts*.
- He, Y.T., Fitzmaurice, A.G., Bilgin, A., Choi, S., O'Day, P.A., Horst, J., Harrington, J., Reisinger, H.J., Burris, D.R., Hering, J.G., 2010. Geochemical processes controlling arsenic mobility in groundwater: A case study of arsenic mobilization and natural attenuation. *Applied Geochemistry* 25: 69–80.
- Hildum, B., Hon, R., Ahmed, S., and Simeone, R., 2012. Arsenic speciation and groundwater chemistry at Shepley's Hill landfill, Devens, Massachusetts. *Geological Society of America* 44 (2): p. 46.
- Strudwick, D.G., 1999. Leachate chemistry of two modern municipal waste landfills in Melbourne, Victoria. Masters Research thesis, School of Earth Sciences, The University of Melbourne.
- Zheng, Y., van Geen, A., Stute, M., Dhar, R., Mo, Z., Cheng, Z., Horneman, A., Gavrieli, H., Simpson, J., Versteeg, R., 2005. Geochemical and hydrogeological contrasts between shallow and deeper aquifers in two villages of Arai-hazar, Bangladesh: Implications for deeper aquifers as drinking water sources. *Geochimica et Cosmochimica Acta* 69: 5203–5218.

Appendix II

Instrumental Analysis

Environmental geochemistry is characterized by dozens of methods of instrumental analysis applied to analysis of concentrations of ions or molecules in solution, of mineral types or compositions, of speciation of trace elements in minerals, of isotopic compositions of minerals, waters or gases, and more. The following section provides some fundamental concepts and applications of instrumental analysis in geochemistry. It is not all-inclusive but does provide insight into some common methods.

II.1 ANALYSIS OF MINERALS AND CRYSTAL CHEMISTRY

Minerals in soils and sediments are commonly too fine grained to be identified in hand-sized specimens or by standard petrographic microscopes. Instead, their identification and analysis requires high-powered electron microscopy for imaging of crystal forms and diffraction and vibrational spectroscopy for determination of crystallographic structures. Regardless of crystal size, elemental analysis is achieved using approaches that typically measure

elemental absorbance, emission, fluorescence or mass. Some methods require destruction of the sample for analysis, particularly those that employ strong acids to dissolve minerals into solution; while these methods require dissolution into an aqueous solution for chemical analysis of solids, their benefit is that they are also capable of determining compositions of natural waters (e.g. in soils, streams, aquifers, seas). Other methods, like electron microscopy, X-ray fluorescence and X-ray diffraction, are capable of analyzing mineral powders and hence are not sample destructive but they are not capable of measuring dissolved ions in water. The following sections present some basic information on analysis of minerals and elements by instrumental methods.

II.1.1 Electron microscopy (SEM, TEM and many other acronyms)

Samples analyzed by electron microscopy commonly consist of single mineral grains, polished sections of rock, or dispersed grains of sediment or soil that are coated with carbon, copper or gold.

There are two main approaches to electron microscopy, both of which involve directing a high-energy electron beam at the sample. The electron beam is scattered by the sample in *scanning electron microscopy* (SEM) and it is transmitted through the sample in *transmission electron microscopy* (TEM).

One of the main advantages of SEM and TEM is the imaging of morphological properties such as crystal shape, pore size and spatial relationships of crystals and other solids (including microbes and amorphous material). Electron microscopy, particularly high-resolution TEM (HRTEM) also facilitates microanalysis of single mineral grains, even when they are exceedingly fine grained, an attribute lacking in other methods. Secondary X-rays produce an elemental spectrum that yields compositional information that enables determination of major element compositions of minerals – this type of microanalysis is known as energy-dispersive X-ray analysis (variably abbreviated as EDX, EDS, and EDAX). An SEM image of gibbsite $[\text{Al}(\text{OH})_3]$ from a tropical soil and its elemental spectrum are presented in Fig. II.1. Note the form of the 1–5 micrometer gibbsite crystals as well as

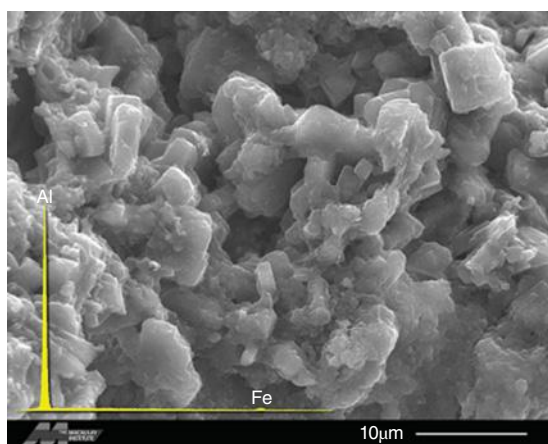


Fig. II.1 Scanning electron microscope (SEM) image and energy-dispersive X-ray (EDX) spectrum (in yellow) of gibbsite. (Evylyene Delbos, Macaulay Institute (now Hutton Institute).)

the spectrum indicating trace amounts of Fe in the gibbsite.

Resolution of images produced by TEM is much greater than SEM. In Fig. II.2, individual hexagonal crystals of 0.1 μm kaolinite are visible.

Electron microprobe analysis (EMP or EMPA) is similar to EDX analysis in an SEM or TEM. A solid sample (powder or rock chip) is bombarded with an electron beam, and the emitted X-rays occur at wavelengths characteristic of given elements.

II.1.2 X-ray diffraction

XRD provides a method for identification of crystalline phases, not by direct imaging as in electron microscopy, but rather by diffraction peaks produced by constructive interference of coherently scattered X-rays. It was the method used by Linus Pauling in the 1920s and 1930 to determine the structures of micas and other phyllosilicates, and by Watson and Crick in the 1950s to decipher the structure of DNA. In modern environmental geochemistry its main application is the determination of mineral assemblages of soils, sediments

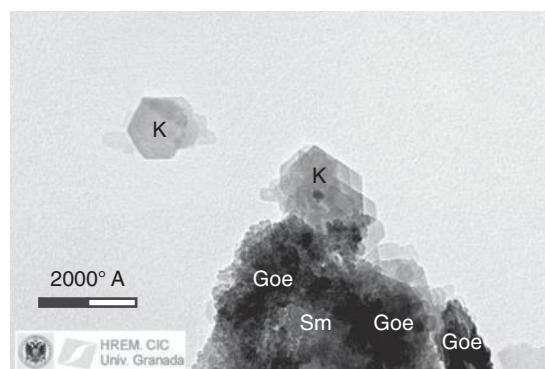


Fig. II.2 TEM image of tropical soil minerals, including kaolinite (note hexagonal crystal form) and smectite (lower center). Compare the scales of these SEM and TEM images – 2000 Å = 0.2 μm, so the field of view in Fig. II.1 (SEM) is ~50 times that of the TEM image. (María del Mar, Centro de Instrumentación Científica (Granada, Spain).)

and powdered rocks, hence it is sometimes referred to as X-ray powder diffraction or variations thereof, as opposed to single-crystal XRD methods used in crystallography. Sample preparation is generally minimal except where mineral quantification is desired.

Bragg's Law

The X-rays used in XRD are typically generated from a copper tube (sometimes cobalt or other metals), collimated and aimed at the sample. Bragg's Law, formulated in 1912 by the father-son team of W. H. and W. L. Bragg, describes how diffraction of an X-ray beam can be applied to analysis of crystal structure. The geometry of XRD as it pertains to Bragg's Law is presented below. Note that an incoming X-ray beam is depicted very schematically as only three in-phase rays that are diffracted by planes of atoms in a crystalline lattice (Fig. II.3):

Bragg's Law permits determination the d -spacing of a unit cell. Most minerals have a relatively unique set of d -spacings that, when combined with a relatively unique set of XRD peak intensities, provides data that enable mineral identification. When analyzing an unknown mineral, d is the variable (Fig. II.3). We know θ

(theta), the angle between the incident beam and the crystallographic planes in question, because the diffractometer has a read-out that displays the angle of incidence. In terms of this geometric approach, the diffracted beam is effectively reflected, so the angle θ^* is equal to θ , and the angles EAB and BAF are also equal to θ . We also know the wavelength λ , which for CuK α radiation is 1.5418 Å (Cu tubes are the most common X-ray source in powder XRD).

In order for the incoming, in-phase X-ray beam to be diffracted coherently (as opposed to destructively, where out-of-phase X-rays cancel and there is no peak), the reflected rays must be in-phase – in this example, they are depicted as in-phase when they cross line O–O', so a diffraction peak will be produced. In order for this to occur, rays 2–2* and 3–3* (and others) must travel one, two, three (or some other whole integer) wavelength more than is travelled by ray 1–1* – this is the case in Fig. II.3, where ray 2–2* travels two complete wavelengths between points E and F, and ray 3–3* travels four complete wavelengths between points D and G.

All that remains to derive Bragg's Law is to determine the relationship between λ , θ and d . Working with triangles AEB and AFB (keeping

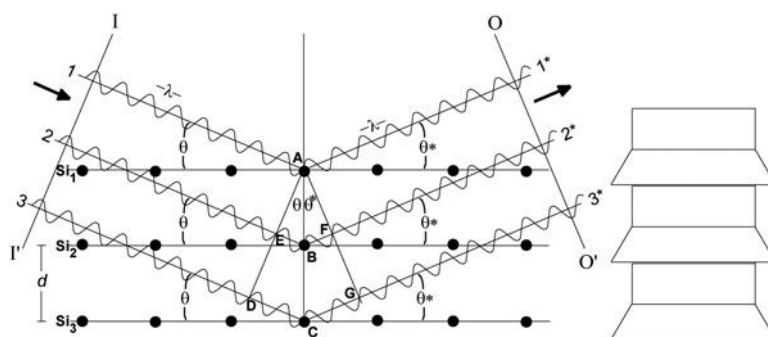


Fig. II.3 Schematic representation of diffraction from three rows of atoms at equivalent positions in three unit cells (where the equivalent positions are conceptually represented here as planes of Si atoms in kaolinite unit cells). The numerals 1, 2 and 3 represent rays of the incident beam, and 1*, 2* and 3* are rays from the diffracted beam. θ is the angle of incidence, λ is wavelength and d is the d spacing along the c -axis. I and I' represent a plane perpendicular to the incident beam and O and O' represent a plane perpendicular to the outgoing (diffracted or reflected) beam.

in mind that $AB = d$ spacing), $\sin\theta = EB/AB$ and that $\sin\theta^* = BF/AB$. Rearranged slightly, $AB\sin\theta = EB$ and $AB\sin\theta^* = BF$. Given that $AB = d$ and that $\sin\theta = \sin\theta^*$, then $2d\sin\theta = EB + BF$. Finally, for coherent diffraction, $EB + BF$ must be equal to some integer multiple of λ , so $EB + BF = n\lambda$, yielding Bragg's Law:

$$n\lambda = 2d\sin\theta \quad (\text{II.1})$$

With modern diffractometers and software there is no need to sit down with a calculator to determine d , but understanding Bragg's Law means makes XRD analysis more than a total "black-box" approach.

Mineral Identification by XRD

Figure II.4 contains an XRD pattern of the minerals kaolinite and illite that includes information on d spacings (in Å) and peaks (in Miller indices). At most angles from 2 to 40° 2θ , X-rays are

diffracted by the specimen out-of-phase, producing destructive interference and hence no peaks – only at very specific angles are peaks produced. In this case, platy kaolinite and illite crystals are oriented perpendicular to the c -axis, so the only peaks are from d spacings in the c crystallographic direction. The convention in powder diffraction uses the angle of diffraction as 2θ rather than θ . Also note the 7.16 Å d -spacing of kaolinite (along the c -axis) and the 10 Å spacing of mica, each indicated by the 001 peak – these are the same spacings as indicated schematically in Fig. II.3.

Given the geometry of diffraction and conditions that produce coherent diffraction, peaks also occur at angles that are whole integer multiples of the angle where the 001 peak occurs, so for kaolinite we observe the 001 peak at 12.3° 2θ (7.16 Å), the 002 peak at 24.8° 2θ (3.58 Å), and the 003 peak at 37.7° 2θ (2.38 Å), and would expect to see higher-order peaks at (1.79, 1.43 Å, etc.). Similarly, illite produces peaks at 10 Å, 5.0 Å,

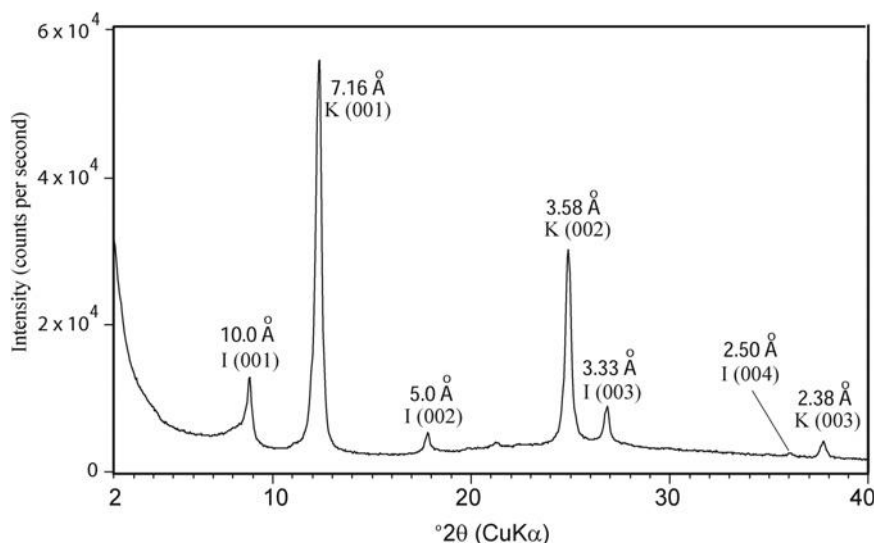


Fig. II.4 X-ray diffraction pattern (spectrum) of the < 2 micrometer (clay size) fraction of a hydrothermally altered schist. The presence of illite is indicated by the series of 10 Å, 5 Å, 3.33 Å and 2.50 Å peaks. Kaolinite is indicated by the 7.16, 3.58 and 2.38 Å peaks. Note that, as dictated by Bragg's Law, d spacing decreases as the angle 2θ increases (to the right). The data in this figure were generated from an oriented clay preparation, so other than a very weak dioctahedral 02,11 peak at ~21.3 degrees 2θ , the only peaks visible are those of the 001 series for kaolinite and illite.

3.33 Å, 2.5 Å, 2.0 Å and so on. The intensity of the peaks is controlled by mineral composition as well as the structure factor and other more detailed aspects of diffraction not covered here.

When smectites are heated to $\geq 100^\circ\text{C}$, inter-layer water is driven out, producing a dehydrated smectite with $d(001)$ of ~ 10 Å. The ability of water (and hydrated cations) to enter and leave the smectite interlayer is the source of much of this mineral group's high cation exchange capacity. It is also a characteristic property that aids in the identification of this mineral group. While smectites, chlorites and vermiculites all can have $d(001)$ values of ~ 14 Å, only smectite will fully collapse to 9–10 Å when heated to 250°C , and in the presence of the polar organic compounds ethylene glycol, smectites expands to 17 Å.

For further information on X-ray diffraction and mineral analysis the reader is referred to an excellent and thorough treatment of the topic presented in Moore and Reynolds (1997).

II.1.3 FTIR

Infrared radiation (IR) occurs in the electromagnetic spectrum between the long-wavelength end of the visible spectrum ($0.75\ \mu\text{m}$) and the short-wavelength part of the microwave spectrum ($\sim 1000\ \mu\text{m}$). Units of IR wavelength are typically presented either as units of length, e.g. nm or μm , or as wave number (cm^{-1}), which is the inverse of wavelength (i.e. $1/\lambda$). So, a wavelength of $500\ \mu\text{m}$ ($0.5\ \text{cm}$) corresponds to a wave number of $2\ \text{cm}^{-1}$, and a wavelength of $1\ \mu\text{m}$ ($10^{-3}\ \text{cm}$) equals $1000\ \text{cm}^{-1}$. The common unit in mineralogical analyses is the wave number, so that will be used from here on, with wavelengths in parentheses.

The infrared spectrum is often sub-divided as follows:

- The near-infrared, from $\sim 13\ 000$ – $4000\ \text{cm}^{-1}$ (0.075 – $0.25\ \mu\text{m}$);
- The mid-infrared, from ~ 4000 – $400\ \text{cm}^{-1}$ (0.25 – $2.5\ \mu\text{m}$);

- The far-infrared, from ~ 400 – $10\ \text{cm}^{-1}$ (2.5 – $100\ \mu\text{m}$);

Mid-infrared absorption can be used to study fundamental vibrations and associated rotational–vibrational structures and is the region mainly used to study minerals. Figure II.5 summarizes the different ways that components of minerals, soils and other materials (e.g. H_2O) stretch, vibrate and rotate in the presence of IR. Water molecule molecules absorb at many points in the IR spectrum (Fig. 7.13); the examples shown in Fig. II.5 include absorption at $1595\ \text{cm}^{-1}$ ($0.627\ \mu\text{m}$) due to flexing of the O–H bond, $3652\ \text{cm}^{-1}$ ($0.274\ \mu\text{m}$) due to symmetrical stretching of the O–H bond, and $3756\ \text{cm}^{-1}$ ($0.266\ \mu\text{m}$) due to asymmetrical stretching of the O–H bond (Fig. II.5). In Fig. II.6, note that water adsorbed to mineral surfaces produces a broad peak between ~ 3800 and $3100\ \text{cm}^{-1}$.

Absorption of IR at relatively precise wavelengths provides data on the types of components in minerals (Fig. II.6). The region from 3800 to $3400\ \text{cm}^{-1}$ is known as the OH stretching region,

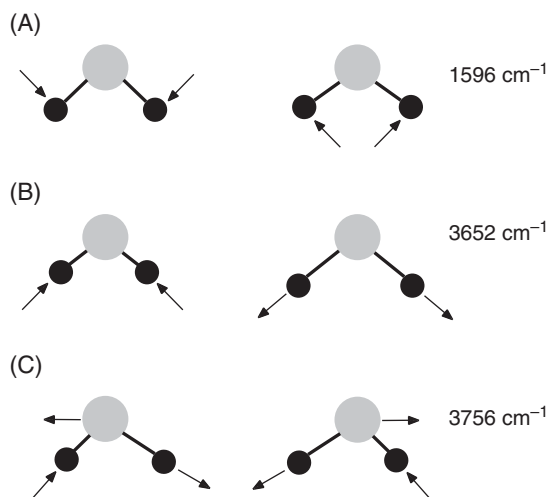


Fig. II.5 Schematic sketches of the ways in which water absorbs IR radiation: (A) flexing, (B) symmetrical stretching, and (C) asymmetrical stretching/rotation. Molecular distortions are exaggerated to make them more apparent.

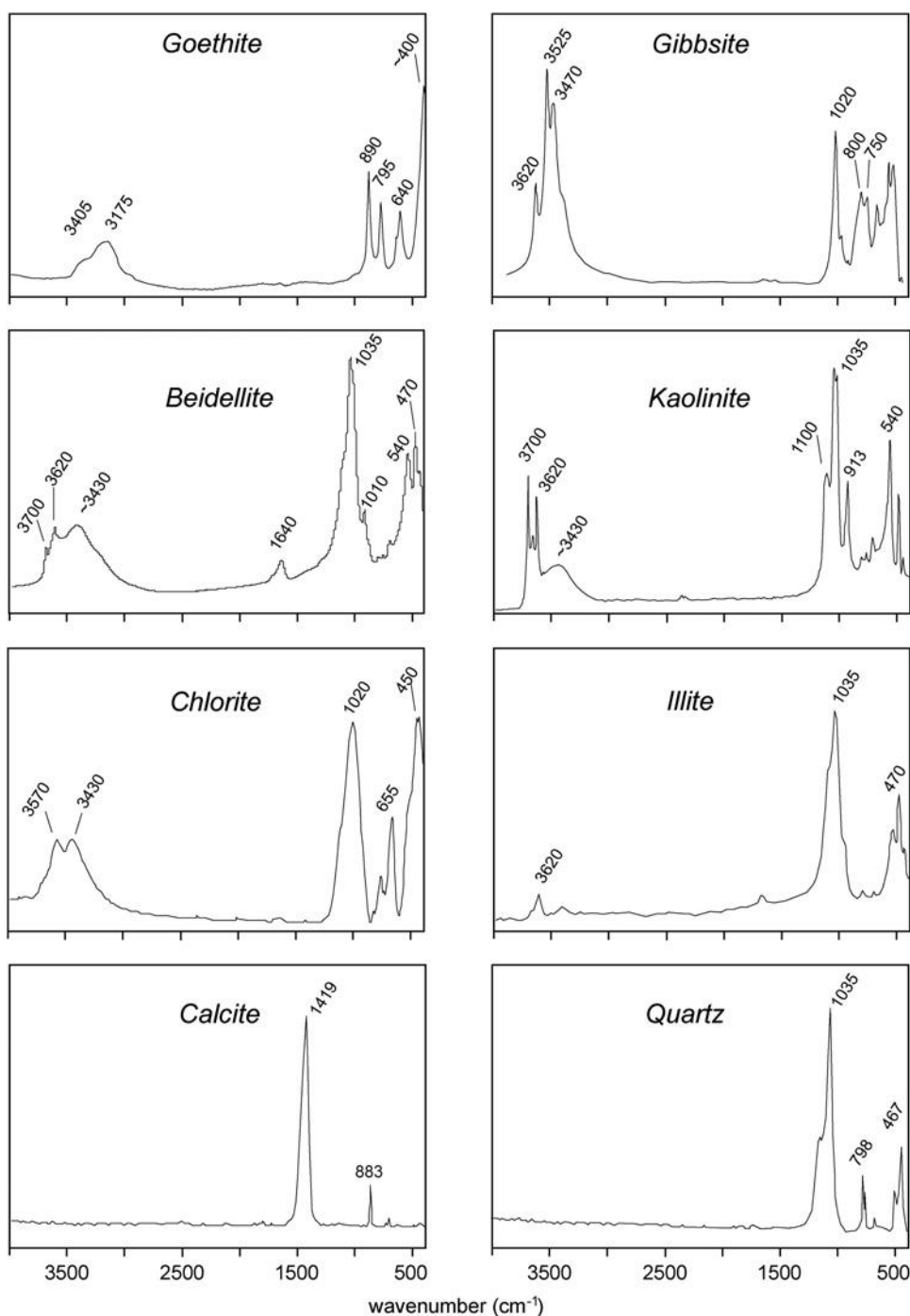


Fig. II.6 FTIR spectra for selected minerals. Peaks in the OH stretching region ($3800\text{--}3400\text{ cm}^{-1}$) are from cation–OH bonds. All silicate minerals contain an Si–O stretching band, which in the examples provided here occurs at $1035\text{ to }1020\text{ cm}^{-1}$ – note that this peak is absent in goethite and calcite. A peak in this range does occur in gibbsite because Al–OH bending vibrations absorb IR in this range. In quartz and other silicates, the 798 cm^{-1} peak is produced by Si–Si stretching and the $\sim 470\text{ cm}^{-1}$ peak is from O–Si–O bending. 3430 and 1640 cm^{-1} peaks in beidellite and, to a lesser extent, in kaolinite and illite, are H–O–H stretching bands produced by adsorbed water. The 1419 and 883 peaks in calcite are from vibrations in the CO_3^{2-} molecule.

a part of the spectrum that provides information on bonds between hydroxyls and cations (and that also includes the broad H_2O band mentioned in the previous paragraph). The hydroxides goethite and gibbsite contain peaks in this region, as do all of the clay minerals shown in Fig. II.6. One distinctive peak at 3620 cm^{-1} is produced by inner hydroxyl groups located between the tetrahedral and octahedral sheets. Calcite and quartz lack hydroxyls, so spectra for these minerals do not contain OH stretching peaks. The region from 1300 to 400 cm^{-1} is known as the OH bending region. The composition of the octahedral sheet in clay minerals strongly influences the position of OH bending bands; for example, in dioctahedral minerals such as kaolinite and illite, absorption is centered around the 950 to 800 cm^{-1} region, whereas OH absorption in trioctahedral minerals (e.g. chlorite) occurs in the 700 to 600 cm^{-1} range. Al–OH–Al bending bands of kaolinite occur at $\sim 915\text{ cm}^{-1}$, while Mg–OH–Mg bending bands of trioctahedral clays like chlorite and serpentine occur at $\sim 650\text{--}600\text{ cm}^{-1}$. Mineral composition can also be assessed; for example, substitutions in the octahedral sheet of clay minerals or hydroxides (e.g. Fe or Mg for Al) can be determined based on positions and intensities of IR peaks.

Many modern IR spectrometers are Fourier transform IR (FTIR) spectrometers that provide greater resolution than older IR spectrometers. For further reading on FTIR in mineral analysis, the reader is referred to the classic book by pioneering researcher V. Colin Farmer (1974) as well as the more recent, undergraduate-oriented chapter by Schroeder (2002).

II.1.4 Elements in solution by ICP-AES, ICP-MS, AAS

The compositions of rocks, soils, sediments and minerals can be determined by spectrometry provided that the solid material can be dissolved into solution (typically some form of acid). Three

common techniques are inductively coupled plasma – atomic emission spectrometry (ICP-AES), which is also known as ICP-optical emission spectrometry (ICP-OES) and inductively coupled argon plasma spectrometry (ICAPS); ICP-mass spectrometry (ICP-MS); and atomic absorption (AA) and graphite furnace AA (GFAA). These methods are also useful for analysis of aqueous solutions such as surface waters, groundwaters, soil water and waste water.

ICP spectrometers contain a very hot argon plasma ($\sim 7000^\circ\text{C}$) into which a solution is injected at a constant rate by a peristaltic pump. At the extremely high temperature of the argon plasma, all compounds are atomized and the resulting electronically excited atoms emit electromagnetic radiation at wavelengths characteristic of individual elements. For example, chromium is commonly detected by radiation emitted with a wavelength of 267.7 nm and nickel by radiation at 231.6 nm . The intensity of the emitted radiation is proportional to the concentration of the element in the sample, and the instrument is calibrated by running standards of known concentration and creating a calibration curve like the one shown for calcium in Fig. II.7. Given that many elements produce spectra with overlapping peaks, it is necessary to correct for interferences, a task made possible by the fact that the spectrum for each element is unique. One of the great advantages of ICP-AES is the ability to analyze dozens of elements simultaneously, a feat not possible with ICP-MS and AA. One of the disadvantages is that certain elements, notably arsenic, lead and uranium, are difficult to detect at low concentrations because they emit relatively weak signals in the ICP compared to other metals and metalloids. For this reason, AA or ICP-MS are commonly used to detect As, Pb and U.

ICP-MS uses an argon plasma in the same manner as ICP-AES, but rather than a bank of detectors calibrated to measure emitted radiation from excited atoms, as is the case with ICP-AES, atoms in an ICP-MS are accelerated towards detectors calibrated to receive atoms of discrete atomic

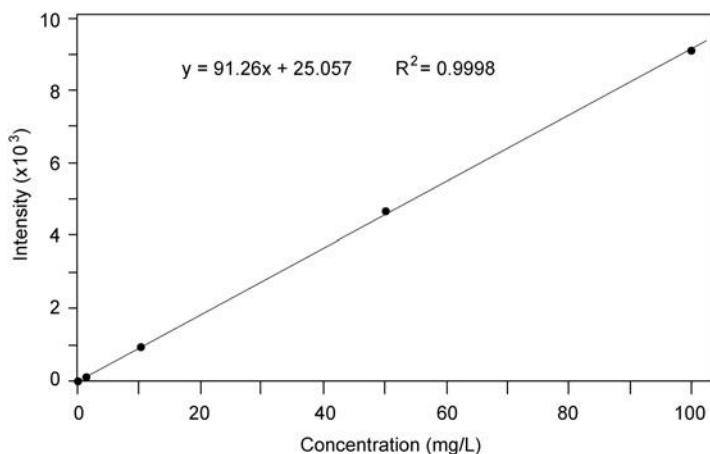


Fig. II.7 Calibration curve for calcium analysis (peak at 317.9 nm) by ICP-AES. Note that points become slightly non-linear at higher concentrations, i.e. while the linear fit produces an excellent R^2 value, the intensity at 50 mg/L is slightly above the line, the intensity at 100 mg/L is slightly below the line, and the y -intercept is not zero.

mass. Most ICP-MS units contain a quadrupole that separates elements based on their mass-to-charge ratio, and each detector is bombarded by ions of known mass whose signal is proportional to concentration. Given that ICP-MS quantifies based on mass, it can be used in isotopic analysis, a distinct advantage over ICP-AES. Detection limits are also typically lower in ICP-MS than they are in ICP-AES. AA spectroscopy is, in a way, the inverse of ICP-AES because elemental abundance is determined by absorption, rather than emission, of radiation, and for this reason, certain elements with weak emission signals (e.g. As, Pb, U) can be detected at lower concentrations with AA than with ICP-AES. Like ICP spectrometers, AA commonly is applied to measurement of concentration of ions in solution, where solutions are aspirated into a chamber. Lamps that emit radiation of known wavelengths are used for individual elements known to absorb radiation at that given wavelength, and the greater the absorption, the greater the amount of element in question.

II.1.5 XRF

X-ray fluorescence spectrometry exposes powdered solids to short-wavelength X-radiation that causes ionization of atoms in the sample. When

ionization occurs, one or more electrons are ejected from the atom, often from inner orbitals. When this occurs, the electron configuration of the atom is unstable, and electrons in higher orbitals move into the lower orbital to stabilize the atom (Figure II.8). The released energy occurs in the form of a photon whose energy is equal to the difference in energy of the two orbitals involved. In this way, individual atoms in the sample emit radiation characteristic of the atoms present, and the intensity of radiation at a given energy level is proportional to the abundance of that particular element. Fluorescence describes the two-step process in which an atom absorbs energy (in this case, from X-rays), and then re-emits radiation of a different energy that is typically of a lower level than the initially absorbed energy.

II.1.6 X-ray absorption spectroscopy (XAS) techniques (EXAFS, XANES)

X-ray absorption spectroscopy (XAS) uses a synchrotron to produce an intense and adjustable source of X-rays, and the main application is determination of speciation of elements, particularly oxidation state and bond lengths, bond angles and coordination numbers. This can be

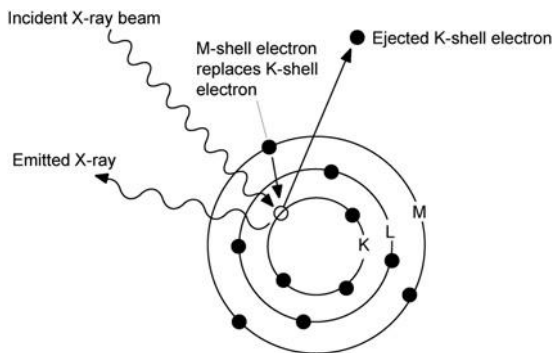


Fig. II.8 Schematic representation of fluorescence caused by ejection of a K-shell electron by an incident X-ray beam, followed by an M-shell electron falling into the K-shell to replace the ejected electron. The energy of the X-ray emitted by the falling electron is characteristic of a given element and is used in determining chemical composition by XRF.

very useful when trying to determine the speciation of trace elements.

X-ray absorption fine structure (XAFS) refers to the ways in which X-rays are absorbed by an atom at energies just below, at and immediately above the binding energies of the atom. The two main regions of the XAS spectrum from which useful information can be obtained are known as X-ray absorption near-edge spectroscopy (XANES) and extended X-ray absorption fine-structure spectroscopy (EXAFS). Figure II.9A illustrates the difference between the two using a XAS spectrum of iron oxide (FeO). The region between XANES and EXAFS is the near EXAFS region (NEXAFS). When the incident X-ray energy is low (left side of Fig. II.9A), little to no absorption of X-rays occurs. As the energy is increased to a point, a step-wise increase in absorption occurs; this is known as the absorption edge, and it corresponds to increased absorption of X-rays by an excited atom whose core electron structure (core electrons are those that are not valence electrons) has been altered by the incident X-rays. Beer's Law describes absorption of X-rays as follows:

$$I = I_0 e^{-\mu t} \quad (\text{II.2})$$

I_0 is the intensity (energy) of the incident X-ray beam, t is the sample thickness, μ is the absorption coefficient and I is the intensity of the X-ray beam once it has passed through the sample. Important controls on μ are atomic number (Z) and I_0 , which are related approximately as:

$$\mu \sim Z^4 / I_0^3 \quad (\text{II.3})$$

At low energies, I is roughly equal to I_0 , but as I_0 increases, μ decreases, causing I (intensity of transmitted X-rays) to decrease. This is the absorption edge. Given that Z is raised to the fourth power, small differences in atomic mass produce values of μ that are very different for different elements, thus enhancing resolution.

The location of the edge (i.e. x-axis position) and its intensity are related to the binding energies of the central atom and its neighbors. Every atom has core-level electrons with well-defined binding energies, and by adjusting the energy of the X-ray beam, analysts can select the element of interest. The XANES part of the spectrum occurs near the absorption edge, and provides information on the oxidation state of the absorbing atom. Figure II.9b depicts location of the absorption edge in the XANES region for various arsenic compounds. For a given element, higher absorption edge energies generally indicate higher oxidation states because the greater (more positive) charge makes it more difficult to photoionize core electrons. This is visible in Fig. II.9b, where As(V) is indicated by a higher absorption edge than As(III). On the higher-energy (right-hand) side of the absorption edge, absorbance decreases in an oscillatory manner with increasing energy – this is the EXAFS region. The oscillations are controlled by the coordination number (i.e. number of neighboring atoms), the type of neighboring elements, and bond distance.

One way in which XANES and EXAFS are applied to problems in environmental geochemistry is illustrated in Fig. II.9B, which shows how the oxidation state and coordination chemistry of arsenic can be determined.

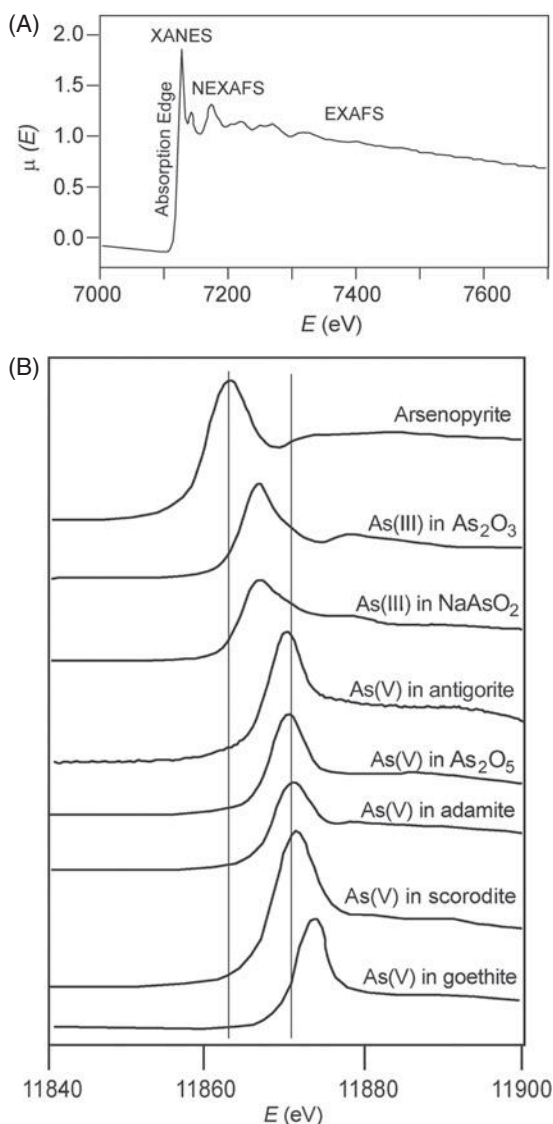


Fig. II.9 XAFS spectra: (A) FeO spectrum that shows the absorption edge, XANES, NEXAFS and EXAFS regions, as well as the oscillations in the EXAFS region useful in determining coordination number, neighboring elements, and bond distance. The y -axis is the absorption coefficient – note marked increase across absorption edge; (B) XANES region showing spectra of arsenic-bearing compounds. Image A is adapted from Newville (2004) and image B was created from data presented in Foster et al. (1997) and Hattori et al. (2005). Scorodite is $\text{FeAsO}_4 \cdot 2\text{H}_2\text{O}$ and adamite is $\text{Zn}_2(\text{AsO}_4)(\text{OH})$, both minerals that contain As(V), i.e. As^{+5} . Arsenopyrite contains $\sim \text{As}^{-1}$.

II.1.7 Isotopic analysis: mass spectrometry

The main analytical method for isotope ratios and concentrations is mass spectrometry (Fig. II.10). The left-hand side of the diagram shows the source of the material to be analyzed – this source typically consists of either ionized gas (e.g. CO_2 in the case of carbon isotope analysis) produced by strong electromagnetic field or a hot filament, or in the case of multicollector inductively coupled plasma mass spectrometry (MC-ICP-MS), as ionized atoms produced in a 10 000 °C plasma.

The gaseous ions are accelerated by a vacuum toward the magnet depicted in the central part of the diagram, and the magnet bends or pulls the components of the ion beam to an extent that is determined by their respective masses. Lighter isotopes are more greatly affected by the magnet and thus are bent to a greater extent. Shown in the lower right-hand side of the diagram are the collectors that measure the quantity of atoms or molecules in the beams (corresponding to the different isotopes) as a current that is amplified, allowing determination with high precision. The positions of the collectors are very precisely calibrated to the isotopes in question, and the two main types of detectors are Faraday cups (or Faraday cages), which measure ion current incident upon a flat surface, and ion counters that measure individual ions impacting the surface of the counter.

Very low concentrations of isotopes such as ^{14}C , ^{10}Be and ^{26}Al are detected by tandem accelerator mass spectrometers. As is the case with traditional mass spectrometry, the ion source produces a beam of ions, but for users the really important application is that tandem accelerator mass spectrometers can work with very small samples (e.g. a few milligrams of a fragment of a pine needle or other solids). They make use of multiple magnets to hone in on the target isotope; the tandem accelerator part of the mass spectrometer contains two accelerating gaps with high voltage in the center, and individual ions are counted by a gas ionization detector.

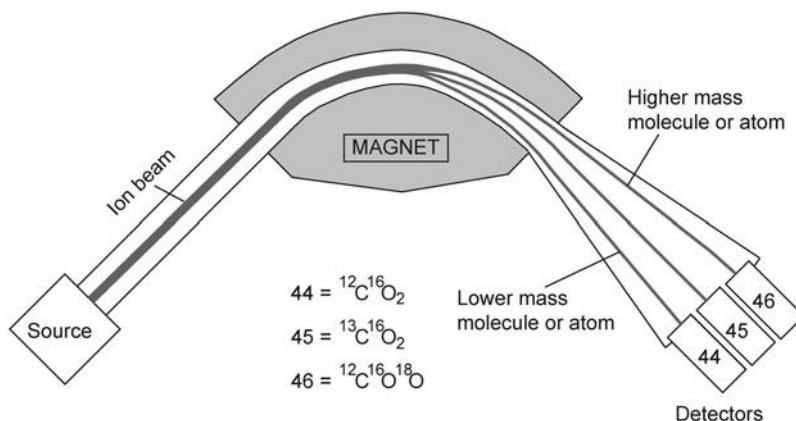


Fig. II.10 Schematic diagram of a mass spectrometer illustrating the effect of the magnet on the trajectories of CO_2 molecules (in ion beams) of different masses. Note that the rare oxygen isotope ^{17}O is not included in the example of isotopes in CO_2 .

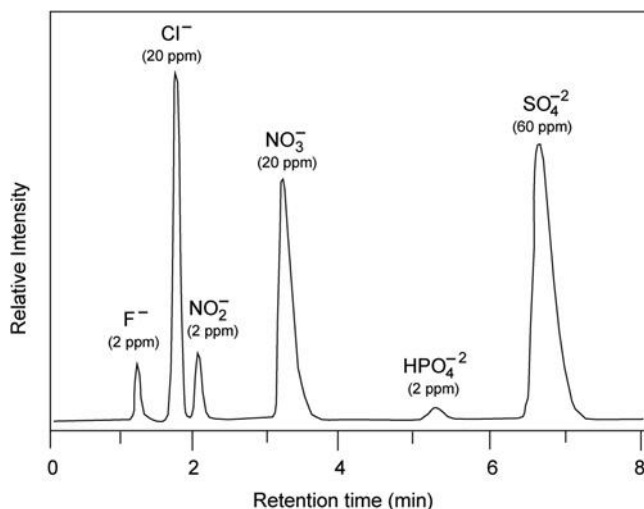


Fig. II.11 Ion chromatogram showing controls on retention time and relative intensity of peaks for different ions (note concentrations), using a Dionex AS4a column (P/N 37041). (Adapted from Pfaff 1993, US EPA Method 300.0.)

Two recent advances in stable isotope mass spectrometry applicable to stable isotope analysis are thermal ionization mass spectrometry (TIMS) and multicollector inductively coupled plasma mass spectrometry (MC-ICP-MS), the development of which provide greater precision analysis for traditional isotope systems (e.g. hydrogen, carbon, nitrogen, oxygen, sulfur) while also fostering new advances in systems (e.g. metal stable isotopes) previously limited by detection limits or precision of analysis.

II.1.8 Chromatography

Anions and volatile organic compounds are commonly analyzed by chromatography (Figure II.11). Ion chromatography (IC) is commonly used in analysis of nitrate, chloride, sulfate, phosphate and other anions (e.g. using USEPA method 300.0), whereas gas chromatography (GC) is used for VOCs. In both cases samples are injected into columns packed with beads (e.g. resin or silica), and the rate at which compounds pass through

the column depends on charge and molecular mass. With certain columns and beads, small, low-charge ions (e.g. F^- , Cl^-) pass through the column more rapidly than larger molecules with the same charge (e.g. NO_3^-), and those that move most slowly through the column are even higher-charged molecules such as HPO_4^{2-} and SO_4^{2-} . Controls on the rate at which organic compounds flow through a column include volatility (more volatile = more rapid flow through column), polarity and presence/absence of functional groups (OH groups are attracted to resin beads and thus retard flow) and molecular mass (larger molecules tend to migrate more slowly). Common detectors for GC are electron-capture detector (GC-ECD) and mass spectrometer (GC-MS), flame ionization detector (GC-FID) and photoionization detector (GC-PID).

REFERENCES

- Farmer, V.C., 1974. *The Infrared Spectra of Minerals*. Monograph 4, Mineralogical Society, London, 539 pp.
- Foster, A.L., Brown Jr., G.E., Parks, G.A., Tingle, T.N., Voigt, D.E., and Brantley S.L., 1997. XAFS determination of As(V) associated with Fe(III) oxyhydroxides in weathered mine tailings and contaminated soil from California, U.S.A. *Supplément au Journal de Physique IV de France* 7: 815–816.
- Hattori, K., Takahashi, Y., Guillot, S., and Johanson, B., 2005. Occurrence of arsenic (V) in forearc mantle serpentinites based on X-ray absorption spectroscopy study. *Geochimica et Cosmochimica Acta* 69: 5585–5596.
- Moore, D. and R.C. Reynolds, Jr., 1997. *X-Ray Diffraction and the Identification and Analysis of Clay Minerals*, 2nd edn, Oxford University Press, New York.
- Pfaff, J.D., 1993. US EPA Method 300.0, Methods for the determination of inorganic substances in environmental samples, EPA-600/R-93-100, NTIS PB94-121811. Environmental Monitoring Systems Laboratory, Office Of Research and Development, U.S. Environmental Protection Agency Cincinnati, Ohio, USA.
- Schroeder, P.A. (2002) Infrared Spectroscopy in clay science: In *CMS Workshop Lectures*, Vol. 11, Teaching Clay Science, A. Rule and S. Guggenheim, eds., The Clay Mineral Society, Aurora, CO, 181–206.

Appendix III

Table of Thermodynamic Data of Selected Species at 1 atm and 25 °C

Species	State	Name	H (kJ/mol)	S (J/mol K)	G (kJ/mol)	Source/Comment
Al ³⁺	aq	aluminum ion	−531.4	−321.7	−485.3	Dean (1979)
Al ₂ O ₃	s	γ-corundum	−1656.9	59.8	−1562.7	Dean (1979)
Al ₂ O ₃	s	α-corundum	−1675.3	50.9	−1582.0	Dean (1979)
Al(OH) ₃	s	gibbsite	−1281.4	70.1	−1143.7	Dean (1979)
AlOOH	s	boehmite	−987.4	48.4	−912.7	Dean (1979)
AlOOH	s	diaspore	−1000.0	35.3	−920.5	Dean (1979)
Al ₂ Si ₂ O ₅ (OH) ₄	s	halloysite	−4079.8	203.3	−3759.3	Dean (1979)
Al ₂ Si ₂ O ₅ (OH) ₄	s	kaolinite	−4098.6	202.9	−3778.2	Dean (1979)
Al ₂ (SO ₄) ₃	s	aluminum sulfate	−3435.1	239.3	−3506.6	Dean (1979)
H ₂ AsO ₄ ^{−1}	aq	dihydrogen arsenate ion	−909.6	117.2	−753.3	Dean (1979)
HAsO ₄ ^{−2}	aq	hydrogen arsenate ion	−906.3	−1.7	−714.7	Dean (1979)
As ₂ O ₃	s	arsenolite	−657.0	107.4	−576.0	Dean (1979)
AsS	s	realgar	−71.3	63.5	−70.3	Dean (1979)
As ₂ S ₃	s	orpiment	−169.0	163.6	−168.4	Dean (1979)
FeAsS	s	arsenopyrite	−41.8	121.3	−50.2	Dean (1979)
Ba ⁺²	aq	barium ion	−537.6	9.6	−560.7	Dean (1979)
BaCO ₃	s	witherite	−1216.3	112.1	−1137.6	Dean (1979)
BaSO ₄	s	barite	−1473.2	132.2	−1362.3	Dean (1979)
C	s	diamond	1.9	2.4	2.9	Dean (1979)
C	s	graphite	0	5.7	0	Dean (1979)

(Continued)

Species	State	Name	H (kJ/mol)	S (J/mol K)	G (kJ/mol)	Source/Comment
C ₆ H ₅ CH ₃	l	toluene	12.0	221.0	113.8	Dean (1979)
C ₆ H ₆	l	benzene	49.0	173.3	124.3	Dean (1979)
CH ₃ (CH ₂) ₆ CH ₃	l	octane	-250.0	357.7	7.4	Dean (1979)
CH ₃ (CH ₂) ₆ CH ₃	g	octane	-208.4	466.7	16.4	Dean (1979)
CO	g	carbon monoxide	-110.5	197.9	-137.3	Dean (1979)
CO ₂	g	carbon dioxide	-393.5	213.7	-394.4	Dean (1979)
CO ₂	aq	(undissociated)	-413.8	117.6	-386.0	Dean (1979)
H ₂ CO ₃	aq	carbonic acid	-699.1	189.3	-623.1	Drever (1997)
HCO ₃ ⁻¹	aq	bicarbonate	-692.0	91.2	-586.8	Dean (1979)
CO ₃ ⁻²	aq	carbonate ion	-677.1	-56.9	-527.9	Dean (1979)
Ca ⁺²	aq	calcium ion	-542.8	-53.1	-553.5	Dean (1979)
Ca ₅ (PO ₄) ₃ (OH)	s	hydroxyapatite	-6738.5	390.4	-6337.1	Wolery and Jove-Colon (2007)
Ca ₅ (PO ₄) ₃ F	s	fluorapatite	-6872.2	387.9	-6491.5	Dean (1979)
Ca(OH) ₂	s	portlandite	-986.1	83.4	-898.5	Wolery and Jove-Colon (2007)
CaCO ₃	s	aragonite	-1207.4	88.0	-1127.8	Robie et al. (1979)
CaCO ₃	s	calcite	-1207.4	91.7	-1128.8	Robie et al. (1979)
CaMg(CO ₃) ₂	s	dolomite	-2326.3	155.2	-2163.5	Dean (1979)
CaSiO ₃	s	wollastonite	-1630.0	82.0	-1544.8	Wolery and Jove-Colon (2007)
CaSO ₄	s	anhydrite	-1434.1	106.7	-1321.8	Wolery and Jove-Colon (2007)
CaSO ₄ · 2H ₂ O	s	gypsum	-2022.6	194.1	-1797.2	Wolery and Jove-Colon (2007)
CaMgSi ₂ O ₆	s	diopside	-3206.2	142.9	-3032.1	Dean (1979)
CaAl ₂ Si ₂ O ₈	s	anorthite			-4012.8	Dean (1979)
Ca _{0.165} Al _{2.33} Si _{3.67} O ₁₀ (OH) ₂	s	Ca-beidellite		244.1	-5357.3	Wolery and Jove-Colon (2007)
Ca _{0.165} Mg _{0.33} Al _{1.67} Si ₄ O ₁₀ (OH) ₂	s	Ca-montmorillonite		250.4	-5323.3	Wolery and Jove-Colon (2007)
Ca _{0.165} Fe ₂ Al _{0.33} Si _{3.67} O ₁₀ (OH) ₂	s	Ca-nontronite		282.6	-4513.2	Wolery and Jove-Colon (2007)
Ca _{0.165} Mg ₃ Al _{0.33} Si _{3.67} O ₁₀ (OH) ₂	s	Ca-saponite		260.8	-5591.6	Wolery and Jove-Colon (2007)
Cl ⁻¹	aq	chloride ion	-167.2	56.5	-131.3	Dean (1979)
F ⁻¹	aq	fluoride ion	-255.6	145.5	-262.3	Dean (1979)
Cu ⁺¹	aq	copper(I) ion	71.7	40.6	50.0	Dean (1979)
Cu ⁺²	aq	copper(II) ion	64.8	-99.6	65.5	Dean (1979)
CuCO ₃ · Cu(OH) ₂	s	malachite	-1051.4	186.2	-893.7	Dean (1979)

(Continued)

(Continued)

Species	State	Name	H (kJ/mol)	S (J/mol K)	G (kJ/mol)	Source/Comment
$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	s	chalcantite	-2279.7	300.4	-1880.1	Dean (1979)
Cu_2S	s	chalcocite	-80.1	120.8	-86.9	Dean (1979)
Fe^{+2}	aq	ferrous ion	-89.1	-137.7	-78.9	Dean (1979)
Fe^{+3}	aq	ferric ion	-48.5	-315.9	-4.6	Dean (1979)
$\text{Fe}(\text{OH})^{+2}$	aq		-291.2		-233.2	Drever (1997)
$\text{Fe}(\text{OH})_2^{+1}$	aq		-548.9		-450.5	Drever (1997)
$\text{Fe}(\text{OH})_3$	s	iron hydroxide	-832.6	104.6	-696.5	Wolery and Jove-Colon (2007)
FeOOH	s	goethite	-558.1	60.4	-488.6	Wolery and Jove-Colon (2007)
Fe_2O_3	s	hematite	-824.2	87.4	-742.2	Dean (1979)
Fe_3O_4	s	magnetite	-1118.4	146.4	-1015.5	Dean (1979)
FeCr_2O_4	s	chromite	-1458.6	146.9	-1355.9	Wolery and Jove-Colon (2007)
Fe_7S_8	s	pyrrhotite	-736.4	485.8	-748.5	Dean (1979)
FeS	s	pyrrhotite	-100.0	60.3	-100.4	Dean (1979)
FeS	s	mackinawite			-93.0	Drever (1997)
FeS_2	s	pyrite	-178.2	52.9	-166.9	Dean (1979)
Fe_2SiO_4	s	fayalite	-1479.9	145.2	-1379.0	Dean (1979)
$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	s	melanterite	-3014.6	409.2	-2510.3	Dean (1979)
$\text{Fe}_{4.8}(\text{SO}_4)_6(\text{OH})_2 \cdot 20\text{H}_2\text{O}$	s	ferricopiapite	-12045.1	1449.2	-10089.8	Majzlan et al. (2006)
H^{+1}	aq	hydrogen ion	0	0	0	Dean (1979)
H_2O	g	water vapor	-241.8	188.7	-228.6	Dean (1979)
H_2O	l	liquid water	-285.8	69.9	-237.2	Dean (1979)
K^{+}	aq	potassium ion			-282.6	Dean (1979)
KAlSi_3O_8	s	microcline			-3732.7	Dean (1979)
$\text{K}_{0.6}\text{Mg}_{0.25}\text{Al}_{2.3}\text{Si}_{3.5}\text{O}_{10}(\text{OH})_2$	s	illite		266.4	-5455.7	Wolery and Jove-Colon (2007)
$\text{K}_{0.33}\text{Al}_{2.33}\text{Si}_{3.67}\text{O}_{10}(\text{OH})_2$	s	K-beidellite			-5360.3	Wolery and Jove-Colon (2007)
$\text{KFe}(\text{SO}_4)_2(\text{OH})_6$	s	jarosite	257 ??		-3299.5	Arslan and Arslan (2006)
MgCO_3	s	magnesite	-1113.3	65.1	-1029.5	Robie et al. (1979)
MnO_2	s	pyrolusite	-520.0	53.1	-465.1	Dean (1979)
MnCO_3	s	rhodochrosite	-889.3	100.0	-816.0	Dean (1979)
Na^{+}	aq	sodium ion			-261.7	Dean (1979)
$\text{NaAlSi}_3\text{O}_8$	s	albite (low albite)	-3935.1	207.4	-3711.7	Robie et al. (1979)
$\text{Na}_{0.33}\text{Al}_{2.33}\text{Si}_{3.67}\text{O}_{10}(\text{OH})_2$	s	Na-beidellite		249.6	-5354.2	Wolery and Jove-Colon (2007)
$\text{Na}_{0.96}\text{Al}_{0.96}\text{Si}_{2.04}\text{O}_6 \cdot \text{H}_2\text{O}$	s	Analcime	-3296.9	226.7	-3077.2	Wolery and Jove-Colon (2007)
NaHCO_3	s	nahcolite	-936.2	101.2	-852.9	Dean (1979)
$\text{H}_2\text{PO}_4^{-1}$	aq	dihydrogen phosphate ion	-1296.3	90.4	-1130.3	Dean (1979)

(Continued)

Species	State	Name	H (kJ/mol)	S (J/mol K)	G (kJ/mol)	Source/Comment
HPO_4^{2-}	aq	hydrogen phosphate ion	-1292.0	-33.5	-1089.1	Dean (1979)
PO_4^{3-}	aq	phosphate ion	-1284.4	-221.0	-1025.5	Dean (1979)
PbCO_3	s	cerussite	-699.2	131.0	-625.3	Robie et al. (1979)
PbS	s	galena	-977.1	91.4	-960.8	Robie et al. (1979)
H_4SiO_4	aq	dissolved silica	-1460	180	-1316	Drever (1997)
SiO_2	s	quartz	-910.7	41.5	-856.3	Robie et al. (1979)
SiO_2	s	cristobalite	-908.4	43.4	-854.5	Robie et al. (1979)
ZnCO_3	s	smithsonite	-812.8	82.4	-731.5	USGS
ZnS	s	sphalerite	-206.9	58.7	-202.5	USGS

Sources:

- Arslan, C. and Arslan, F., 2003. Thermochemical Review of Jarosite and Goethite Stability Regions at 25 and 95 °C. *Turkish Journal of Engineering and Environmental Science* 27: 45–52.
- Dean, J.A., 1979. *Lange's Handbook of Chemistry*, 12th edn McGraw-Hill, New York.
- Drever, J.I., 1997. *The Geochemistry of Natural Waters: Surface and Groundwater Environments*, 3rd edn Simon and Schuster, Upper Saddle River, New Jersey, USA.
- Majzlan, J., Navrotsky, A., McCleskey, R. B., Alpers, C. N., 2006. Thermodynamic properties and crystal structure refinement of ferricopiapite, coquimbite, rhomboclase, and $\text{Fe}_2(\text{SO}_4)_3(\text{H}_2\text{O})_5$. *European Journal of Mineralogy* 18: 175–186.
- Robie, R.A., Hemingway, B.S., Fisher, J.R., 1979. Thermodynamic properties of minerals and related substances at 298.15 K and 1 bar 105 pascals pressure and at higher temperatures. United States Geological Survey Bulletin 1452. Accessed on-line 5-December 2013: <http://pubs.usgs.gov/bul/1452/report.pdf>.
- Wolery, T. and Jove-Colon, C., 2007. Qualification of Thermodynamic Data for Geochemical Modeling of Mineral-Water Interactions in Dilute Systems. Report Prepared for US Department of Energy, ANL-WIS-GS-000003 REV 01. Accessed online 5-December 2013: <http://pbadupws.nrc.gov/docs/ML0907/ML090770163.pdf>.

Index

- Absorption, 66, 98, 99, 141, 219–21, 245
- Acid
 - definition, 19–21
 - dissociation, 20, 88, 150, 151, 167, 189, 236, 249
- Acid mine drainage, 70, 123, 130, 200, 263, 325
- Acid rain (acid deposition), 2, 95, 138, 181, 249–52
- Activation energy, 31, 36, 308
- Activity, activity coefficient, 125, 126
- Adsorption
 - anions, 66, 141
 - cations, 64–70, 137–41, 182, 284
 - control on metals, 140, 287, 292
 - definition, 65
 - edge, 140–142
 - isotherms, 142–3
 - organic
 - compounds, 99–103, 295
- Advection, 75–7
- Aerobic conditions, 23, 79, 105, 156, 162, 170, 182
- Age dating
 - ^{14}C , 350–354
 - cosmogenic
 - radionuclides, 350–359
 - K–Ar, 360–362
 - ^{210}Pb , 359–60
- Rb–Sr, 362
- ^{234}U , 364–6
- U–Th–Pb, 362–4
- Albite
 - chemical weathering of, 265, 277
 - composition and structure, 58
 - mineral stability
 - diagram, 278–80
- Alcohols, 87
- Alfisol, 268, 273, 277, 294, Plate 13
- Aliphatic hydrocarbons, 103
- Aldehyde, 97, 240, 244
- Alkaline Earth metals, 12, 13, 74, 119–21, 129, 374
- Alkalinity, 143–6, 152–5, 286, 287
- Alkanes, 84–6, 94–7, 103, 163, 243
- Alkenes, 83, 86
- Alkynes, 84
- Alpha decay, 336, 339, 344, 346, 362, 364–6
- Aluminum
 - ^{26}Al , 300, 354–6
 - hydrolysis, 123, 124
 - hydroxide solubility, 123, 124
 - speciation, 116, 124
 - toxicity, 250, 283
- Amines, 87, 92, 97
- Ammonia, 20, 21, 123, 181, 183, 184, 186–7, 208, 251, 318
- Ammonium nitrate, 180, 184, 317–18
- Amorphous phases, 31, 78, 130, 136–7, 379
- Amorphous silica solubility, 136–7
- Amphiboles, 39, 50–53, 57–8, 241, 260, 316
- Amphoteric, 124
- Andisol, 273–4
- Anaerobic conditions, 23, 79, 83, 105, 159, 162, 170, 182–3, 198–9, 274, 322–4, 374
- Anion adsorption, 142
- Anion exchange, 66, 69
- Anorthite, 259–61
- Anoxic conditions *see* Anaerobic conditions
- Aqueous solutions, 19–20, 34, 40, 74–8, 115–17, 137–8, 151–3, 384
- Aquifer, 11, 61, 76, 103–4, 112, 140, 155–7, 292, 344–5, 351, 356, 366, 372
- Aragonite, 70–71, 155, 158, 161, 172, 306–7, 342, 366
- Aridisol, 274, 277, 279, 283, 287, 294
- Aromatic hydrocarbons, 85–9, 94–7, 103, 230, 241–3, 374
- Arrhenius equation, 36, 269, 308

- Arsenic, 12, 57, 117, 120–124, 129–35, 196, 199, 254, 293, 371–6
 Asbestiform minerals
 chrysotile, 56, 57
 crocidolite, 51, 241
 tremolite, 51, 57, 261
 Atmosphere
 composition, 208, 213
 photochemical reactions, 208, 231, 237, 242–8
 structure, 207–8
 variation over geologic time, 208–13
 Atomic absorption analysis, 384
 Atomic mass, defined, 6
 Atomic number, defined, 6
 Atomic orbitals, 6–8, 17, 83
 Atomic weight, defined, 6
 Aufbau principle, 7
 Avogadro's number, 9, 15

 Bacteria, 79, 104, 162, 171, 181–7, 195, 199–200, 208, 264, 322, 325, 377
 Basalt, 5, 49, 55, 266–71, 344, 361
 Base, defined, 19–21
 Base cations, 13, 64, 129, 140, 266–76, 285–9
 Base saturation, 273, 285–7, 291
 Bauxite, 267–8, 275
 Beidellite
 composition and structure, 60, 70
 mineral stability diagram, 278–80
 Benzene, 88–90, 97, 100, 104, 230–231, 234, 254, 294–5
 Berthierine, 55, 63, 175
¹⁰Be, 300, 354
 Beta decay, 4, 336–9, 352, 361
 Bicarbonate, 21, 70–71, 116, 124, 149–62, 167, 259, 286, 323, 344
 Big Bang, 2–3
 Big Moose Lake, 286

 Biological uptake, 192, 292, 295
 Biomagnification, 98–9, 253
 Biomineralization, 78–9
 Bioremediation, 105
 Biotite, 55, 62, 64, 129, 259, 260, 276, 329, 360
 Bitumen, 163
 Bjerrum plot, 143, 152
 Black shale, 292
 Bohr model of atom, 6
 Boltzmann's Constant, 207
 Bonding (chemical)
 covalent, 16–19, 48, 73, 99, 113–15, 169, 196, 237, 259, 263, 301, 305
 hydrogen, 18, 56, 114–15, 169
 ionic, 14–18, 45, 74, 115, 259–60, 262, 324
 metallic, 18, 304
 polar covalent, 16–18, 48, 115
 van der Waals, 18, 54, 66, 69, 99, 159
 Bowen's reaction series, 259–60
 Brackish water, 127
 Brine, 125, 127, 133
 Buffering, 70, 88, 123, 150–153, 263–4, 283, 286–7, 321, 342

 Cadmium, 124, 130, 230, 254
 Calcite
 acid neutralization, 286–7
 biogenic vs. inorganic, 79
 chemical weathering, 260, 261
 covalent and ionic bonds, 17
 dissolution
 thermodynamics, 156–7
 effect of CO₂ on solubility, 157–8
 infrared spectrum, 383
 marine sediments, 162
 pedogenic, 274–6
 source of Ca in weathering, 260, 289, 342–3
 stable isotopes, 306–8, 321
 structure and composition, 71
 ²³⁴U/²³⁸U age dating, 366
 Calcium
 depletion by acid rain, 69, 283–5
 isotopes, 328–9
 Calorie, 27
¹³C and carbon isotope
 fractionation, 308, 318–21
¹⁴C dating, 351–4
 Carbon dioxide
 absorption of infrared radiation, 223–6
 atmospheric
 concentration, 208–14, 227
 in groundwater, 156–7
 solubility in water, 168–9
 Carbonate rock composition, 5
 Carbonic acid, 20–21, 123, 149–52, 155–6, 167, 259, 262, 353
 Carboxylic acid, 87–8, 93
 Catalyst, 86, 183, 199, 246, 253
 Cation exchange
 and acid deposition, 250, 286, 287
 capacity, 66, 69, 139, 275, 287, 294, 345, 349, 382
¹³⁷Cs, 347–9, 359, 360
 Charge balance, 46, 55, 69, 77
 Chelates and chelation, 124, 125, 192, 270
 Chemical bonds *see* Bonds, chemical
 Chemical potential energy, 26
 Chemical weathering
 and acid deposition, 284–7
 dissolution, 262
 hydration, 263
 hydrolysis, 262
 oxidation, 263
 rate, 260–261
 Chlorinated hydrocarbons, 90–92, 97, 100, 103–6, 330, 371–7
 Chlorofluorocarbons (CFCs)
 absorption of infrared radiation, 225–6
 composition, 85–6
 effect on stratospheric ozone, 246–8
³⁶Cl, 354–6
 Chlorite (silicate clay), 55–6
 Chromatography, 102, 371, 389

- Chromium, 120, 130
- Clay minerals
- cation exchange capacity, 66–70
 - crystal structure, 58–64
 - interstratified clays (mixed-layer clays), 64–5
 - octahedral sheet, defined, 52
 - tetrahedral sheet, defined, 52
- Climate
- greenhouse gas control, 223–6
 - and soil composition, 268–9
 - and surface water composition, 109, 111–13
- Closed system, 26, 96, 118, 177, 353, 361, 364
- Closure temperature, 361–2
- Coal
- acid rain, 196, 249–4
 - air contaminant, 99, 234–5
 - carbon cycle, 160–164
 - carbon isotopic signature, 318–19
 - geological origin, 95–6
 - mercury source, 206, 217
 - nitrogen content, 186
 - PAHs, 88, 103
 - source of C to atmosphere, 173, 202
 - sulfur content, 200
- Colloid/colloidal, 66, 115–17, 295
- Common ion effect
- Complexes, complex formation, 116, 125, 287, 291, 293, 327, 346
- Component (in thermodynamics), 24–6
- Concentration
- definition, 8
 - units, 8–10
- Concordia diagram, 364
- Conservative species, 76, 344, 356
- Contaminant transport, 59, 64, 78, 99–104, 217, 371–7
- Continental crust (composition), 4–5
- Convection, 217
- Coordination number, 43
- Copper
- in aqueous solution, 130, 134
 - Cu in sulfides, 325–6
 - stable Cu isotopes, 325–6
- Co-precipitation, 128–31
- Cosmogenic radionuclides, 350–358
- Coupled substitution (paired substitution), 47
- Covalent bonds *see* Bonding
- Daltons (atomic mass unit, Da), defined, 6, 12
- Darcy's Law, 76
- Dating *see* Age dating
- DDT, 90, 93, 97–100, 217, 255, 293
- Debye–Hückel equation, 125
- Degrees of freedom (thermodynamic), 25
- Delta notation (for stable isotope pairs), 302, 303, 327
- Denitrification, 181–3, 185, 317
- Deuterium (^2H), 3, 301
- Diagenesis, 39, 53, 81, 95
- Diffuse double-layer model, 68
- Diffusion, 33–4, 75–8, 308, 361–3
- Dimethyl sulfide (DMS), 92, 197–9
- Di octahedral clays, defined, 54–5, 60–64
- Dioxin, 91–3, 252
- Disequilibrium, 24, 258, 365
- Dissociation constants, 196, 236
- Dissolved organic carbon (DOC) *see* Organic matter
- DNAPLs, 96, 103–4, 374
- Dobson units (for ozone), 245
- Dolomite, 25, 70, 158, 260–261, 289
- Double layer *see* Diffuse double-layer model
- Dry deposition, 249
- Earth
- abundance of elements, 5, 48
 - black body radiation, 220–222
 - composition of crust, 5
- EDTA, 124
- Eh, 122
- Eh-pH diagrams
- Al, 135
 - As, 135
 - Cu, 134
 - Fe, 136
 - Pb, 134
 - U, 135
- Electromagnetic spectrum, 219–21
- Electronegativity, 17, 47, 134, 231–2
- Endothermic reaction, 28, 32, 118, 232, 237
- Energy
- of chemical bonds, 26, 45, 305
 - of fission reactions, 23, 336–8, 346–8
 - of fusion reactions, 23
 - heat, 26–7, 348
 - levels in atoms, 7, 305
 - mechanical, 26
- Enthalpy, 27–32, 117, 232
- Entropy, 27–32, 76
- Equilibrium, defined, 24
- Equilibrium constant, 31, 32, 102, 126, 128, 133, 142, 150–158, 277–80, 307
- Equilibrium control in weathering, 258, 272, 277–80
- Ethanol, 87
- Eutrophication, 122, 180, 182, 186–9, 194–6, 238, 317
- Evaporation, 72–4, 108–12, 123, 127, 216, 218, 271, 289, 301–3
- Evaporative concentration, 74, 111, 344
- Evapotranspiration, 109, 169
- Exchange reaction, 65, 69, 101, 108, 137, 140, 284, 307
- Exchangeable cations, 69, 284, 290
- Exothermic reaction, 3, 28, 32, 117–18
- Extensive property, 30

- Feldspar
 chemical weathering, 260–262, 265, 277–80
 mineral stability
 diagram, 277–82
 structure and composition, 234
- Ferrihydrite, 72, 79
- Fertilizers, 74, 180–191, 194, 198, 201, 273–4, 288, 294, 317–18
- Fick's laws, 77
- Field weathering rates, 260–261
- First law of thermodynamics, 26–7
- First-order reaction, 34–5, 337
- Fission, 23, 336–8, 346–8, 359
- Fixed layer (ion exchange), 68
- Flocculation, 116
- Fluorescence, 385–6
- Flux
 carbon cycle, 159–61, 164–8, 171–3
 diffusion of particles, 77–8
 electron, 22, 132
 elemental cycles, 177–80
 heat, 25–6
 hydrologic cycle, 109, 118
 iron from tropical forest, 95
 neutron, 4
 nitrogen, 184–93
 phosphorus, 194
 solar flux, 221
 sulfur, 198–201
- Forest damage, 242, 251, 285–7
- Formaldehyde, 97, 240, 244
- Forsterite, 49, 171, 175, 258–61
- Fossil fuels *see* Coal; Gasoline; Petroleum
- Fourier-transform infrared (FTIR) spectroscopy, 382–4
- Fractionation (of stable isotopes), 199, 300–313, 320–329
- Free radicals (in atmospheric chemistry), 232–4, 237
- Freundlich isotherm, 142–3
- Fugacity, 31
- Fulvic acid, 94, 125, 155, 164, 191–2, 269, 283
- Functional groups, 86–94
- Gamma decay, 3, 335–6
- Gamma rays, 336
- Gasoline, 1, 84–5, 89, 93, 96, 164, 242, 341–2
- Gibbs free energy, 30–32, 183, 278–80
- Gibbsite
 acidic soils, 20, 32
 anion exchange capacity
 composition and structure, 53
 mineral stability
 diagram, 277–80, 282
- Glauconite, 59, 63
- Global meteoric water
 line, 310–312
- Global warming potential (of greenhouse gases), 225
- Goldschmidt's Rules, 45, 48
- Gouy layer, 68
- Granite
 bedrock aquifer
 composition, 110
 whole-rock composition, 5
- Greenhouse gas, 181, 206, 210–213, 219–28, 230, 247
- Greenhouse effect, 171, 219, 222–8
- Groundwater
 in carbonate terrains, 110, 146, 155–8
 chemical evolution, 103–6, 122–3, 144–6, 371–7
 contamination, 103–6, 122, 371–7
 granite aquifer, 110
 redox conditions, 120, 122–3, 133–6
- Gypsum, 73, 78, 111, 196, 197, 219, 240, 261, 262, 269, 272–6, 290, 324
- Half-life, 337–9
- Half-cell reactions (in redox chemistry), 120–122
- Halite, 14, 40–43, 46, 47, 74, 111, 115
- Halloysite, 30, 55, 59, 63–4, 69, 271, 275
- Heat, defined, 26–30
- Heat capacity, 29–30
- Heavy metals
 acid deposition, 287
 adsorption, 140–142
 behavior, 129–32, 134–5, 253
 ionic radii, 44
 stable isotopes, 325–7
 sulfides, 72
- Hematite
 anion exchange capacity, 68
 composition and structure, 71–2
 enthalpy of formation, 28
 soil mineral, 263, 266–7, 272, 275–6
- Henry's law, 118, 166–7
- Henry's gas constant, 150
- Herbicides, 92, 93, 101, 294
- Heterogeneous
 reaction, 33, 237–8
- Homogeneous reaction, 237
- Hornblende, 51, 129, 259, 261, 361
- Hubbard Brook Experimental Forest, 285
- Humic acid, 94, 116, 162, 164, 184, 291
- Human health, 39, 51, 98, 189, 200, 241
- Humin, 94, 160, 162
- Hydraulic conductivity, 76, 356
- Hydrogen bonding *see* Bonding (chemical)
- Hydrologic cycle, 109
- Hydrogen electrode, 121, 132–3
- Hydrogen isotopes
 (8D), 301–3, 311–16
- Hydrolysis, 123, 184, 253, 261–5
- Hydrophilic, 115
- Hydrophobic, 97, 98, 115
- Hydroxyl radical (in atmospheric chemistry), 233, 236–40, 243, 244, 249, 255

- Illite
 composition and structure, 55, 62–3
 K-Ar, Ar-Ar dating, 360–362
 mineral stability diagram, 282
 Incongruent dissolution, 265, 271
 Inconsistency in thermodynamic data, 281
 Infrared radiation, 215, 219–24, 382
 Infrared absorption by greenhouse gases *see* Greenhouse effect
 Inner-sphere complex, 68, 131, 192
 Intensive property, 30
 Interlayer (in clays), 43, 47, 53–6, 59–64, 69, 139, 186, 291, 328
¹²⁹I, 349, 359
 Ion activity product, 126, 128, 156
 Ionic bonds *see* Bonding (chemical)
 Ionic radii, 44
 Ionic strength, 125–6
 Iron
 Eh-pH diagram, 136
 microbial effect on redox, 79, 375
 oxyhydroxide, 69, 78, 132
 stable Fe isotopes, 326
 Irrigation, 112, 255, 274, 289–90
 Isoelectric point, 67–8
 Isotherm, 102, 142–3
 Isotope, defined, 6, 12
 Isotopic systems *see* individual elements

 Joule, 15, 27

 Kaolinite
 composition and structure, 55, 56, 63, 64
 mineral stability diagram, 277–82
 pedogenic mineral, 261, 262, 265, 267, 271–3, 275–7, 316
 K-feldspar, 259–62
 Kinetic energy, 26, 117, 130, 215, 218, 336, 345, 365
 Kinetic isotope effect, 308
 Kinetics, 33–7
 Lakes, 74, 109, 187, 192, 195
 Langmuir isotherm, 102–3, 142–3
 Lanthanides, 268
 Law of mass action, 20, 102, 138–9
 Layer charge (in clays), 55, 59–70
 Lead (Pb), 1, 71, 127–8, 134, 197, 242, 255, 340–342
 Ligand, 124, 326
 Limestone, 70, 76, 78, 112, 144, 157, 164, 250, 318–21, 329, 353
 LNAPLs, 103–4
 Magnesium isotopes, 328–30
 Magnesite, 46, 158, 175
 Magnetite, 30–31, 72, 129
 Mass spectrometry, 384, 387
 Manganese, 120, 288, 292, 299
 Mass balance, 12, 165, 215, 301
 Mass-dependent fractionation, 300
 Mercury
 coal emissions, 206
 dimethylmercury, 99
 isotopes, 304, 326–8
 toxic effects, 253
 Metalloids, 130
 Metamorphism, 28, 53, 55, 172, 363
 Methemoglobinemia, 182, 189, 317
 Meteoric water line, 310–312
 Methane
 absorption of infrared radiation, 223, 225
 atmospheric concentration, 211–14, 227
 atmospheric residence time, 231
 biogenic, 162
 clathrates (hydrates), 169–71
 composition and structure, 84, 159
 Mica (composition and structure), 53
 Microbial activity, 23, 177, 255, 263, 292, 324, 375
 Mixed-layer clays, 64, 275
 Molar units, 9
 Mollisol, 274, 277, 279, 287, 294
 Molybdenum, 288
 Monazite, 191, 362, 364
 Montmorillonite, 60–62, 282
 Multicollector inductively coupled plasma mass spectrometry (MC-ICP-MS), 324, 387
 Muscovite, 55, 62–3, 129, 259–61, 282, 360–361

 NAPLs, 103
 Neutron, 3, 4, 6, 299–301, 335, 340, 347, 351–3, 355, 362
 Nickel, 4, 9, 60, 64, 130, 254, 288, 384
 Nitrate, 19, 21, 67, 68, 74, 110, 116, 122, 162, 180–189, 195, 238, 250, 284, 301, 317
 Nitric acid, 19–21, 88, 181, 238, 249, 251, 283, 317
Nitrobacter, 183
 Nitrogen isotopes, 316–18
 Nitrogen oxides (NO, NO₂), 180, 181, 184, 186, 188, 190, 237, 317
Nitrosomonas, 183
 Nitrous oxide
 absorption of infrared radiation, 181, 223, 226
 atmospheric concentration, 188–90, 213, 227
 ozone destruction, 246
 structure, 226
 Nontronite, 60
 Normality, 154
 Nuclides, 299, 300
 Nutrients
 acid rain effect on, 138, 284
 eutrophication, 187, 191–5

- iron as limiting marine nutrient, 175
 plant nutrients, 39, 59, 62, 65, 95, 184, 269, 288–90
- Ocean
 acidification, 167–8
 carbonate precipitation, 161
 composition, 110
 residence time of water in oceans, 118
 as sink for CO₂, 166–70
 Octahedral coordination, 42, 48, 52, 67, 131
 Octahedral sheet (in clay minerals), 52–63, 129, 262, 263, 293, 384
 Octanol-water partition coefficient (K_{OW}), 97–100
 Oddo-Harkins rule, 3,4
 Oligotrophic lake, 187
 Olivine, 30, 33, 49, 172, 175, 259, 260, 271, 359
 Open system, 26, 118, 178, 308
 Orbitals, electron *see* Atomic orbitals
 Order/disorder, 27, 30, 40, 59, 65, 76, 78, 110, 199
 Order of reaction, 34–6
 Organic acids and anions, 20, 87–8, 94, 155, 164, 192, 268–70
 Organic matter
 adsorption onto, 68–9, 99–103, 116
 dissolved organic carbon (DOC), 94, 155, 161–2, 291, 374
 humus, 68–9, 94–5, 122, 159, 192, 200, 268, 274, 284, 286, 292–5
 particulate organic carbon (POC), 155, 162
 Organic pollutants, 96–106, 242–5, 330, 371–7
 Organophosphate, 92
 Oxidation reaction, 21–4, 31, 73, 79, 104, 120–122, 131, 159, 183–4, 199–200, 231, 234–7, 263, 292, 295
 Oxide minerals, 22, 48, 67, 69, 71, 129, 263, 266, 276
 Oxide conversions, 9
 Oxisol, 273–4, 277, 279
 Oxygen
 concentration in atmosphere, 208–9, 213
 dissolved oxygen (in water), 118, 122, 130, 263, 283
 stable isotopes, 302–3, 309–16
 Ozone
 stratospheric (ozone layer), 189, 207, 208, 220, 224, 245–8
 tropospheric (pollutant), 223–6, 243–4
 Parabolic rate law, 35, 118, 272
 Partial pressure, 150, 155–7, 166, 174, 207, 214
 Particulate, 115–19, 124, 129, 178, 193–6, 201, 214, 218, 286, 329
 Particle matter (in air), 240–241, 252, 254
 Particulate organic carbon (POC) *see* Organic matter
Pe, 132
pe-pH diagrams *see* Eh-pH diagrams
 Pelecypods, 306
 Pentachlorophenol, 89–90, 97, 101
 Permeability, 61, 75–6, 96, 112, 262, 270, 290–293
 Pesticides, 91–3, 96, 122, 252–3, 292
 Petroleum, 28, 84–6, 89, 95–6, 122, 163–4, 200, 234, 324, 353
 Phenols, 89
 Phlogopite, 261
 pH
 of acid mine drainage, 123
 of acid rain, 249
 and cation/anion exchange capacity, 67–8
 definition, 20
 Eh-pH diagrams *see* Eh-pH diagrams
 of fresh waters, 110
 of unpolluted precipitation, 150–151
 of seawater, 110, 167–8
 Phase (in thermodynamics), 25
 Phase diagram, 25
 Phosphate, 47, 78, 92, 101, 116, 124, 191–4, 288, 293, 389
 Phosphorus
 adsorption to ferrihydrite, 192
 cycle, 190–195
 and eutrophication *see* Eutrophication
 as plant nutrient, 288–9
 Photochemical
 reactions, 208, 242–5
 Photosynthesis, 96, 159–62, 166, 171–4, 208, 288–9
 Photosynthesis and stable isotopes, 308, 318, 320–321
 Phyllosilicates, 52–70
 Physical erosion, 200, 270, 358
 Piper diagrams, 143–4
 Plagioclase weathering, 259–60, 266, 271–2, 277–80, 343
 Planck's constant, 305
 Point of zero charge (PZC), 67–8
 Polar (covalent) bonds *see* Bonding (chemical)
 Polyatomic ions, 116
 Polychlorinated biphenyls (PCBs), 90, 93
 Polymerization, 49, 52, 259
 Polymorphs, 49, 70, 72, 83, 129, 155, 307
 Polynuclear (or polycyclic) aromatic hydrocarbons (PAHs), 88, 103, 230, 234, 241–2, 254
 Porosity, 11, 75–6, 96
 Potassium-argon (K-Ar) dating *see* Age dating
 Potential energy, 26, 104

- ppm, ppb, units of concentration
see Concentration
- Precipitation
 composition, 110
 Rayleigh fractionation associated with, 308–14
- Proteins, 94, 162, 164, 183, 254, 288, 300
- Pseudomonas*, 183
- Pump and treat, 106, 372
- Pyrite
 acid mine drainage, 264
 Archean pebbles, 171, 208
 arsenic, 131
 coal, 235
 composition, 72, 197
 formation in marine sediment, 199–200
 microbially-mediated decomposition, 75
 oxidation, 73
 stability, 122, 136, 375
 sulfur cycle, 198
 sulfur isotopes, 323
 thermodynamics, 23–4
 weathering in soil, 263
- Pyrophyllite, 54
- Pyroxene, 50, 259–60, 359
- Quartz
 composition and structure, 58
 (in)solubility, 13, 18, 115, 136–7, 259–61, 266
 oxygen isotope fractionation, 305–6
- Radioactive decay *see* Alpha decay, Beta decay; Gamma decay
- Radiogenic isotopes, 335, 336, 340–344, 362–4
- Radium, 129, 339, 343–6
- Radius ratio, 43, 45–6
- Radon, 339, 343–6
- Rare Earth elements, 268
- Rates of chemical weathering, 259–61
- Rayleigh fractionation
see Precipitation
- Reaction order *see* Order of reaction
- Reduction reactions, 21–2, 72, 106, 120–123, 131
- Residence time
see Ocean, residence time
- Reversible reaction, 27–8
- Rhizobium, 183, 185
- Rb-Sr age dating, 362
- Saline soils, 289–91
- Sandstone composition, 5
- Saponite, 60
- Saturation
 barite and radium, 344
 biomineralization, 78
 carbonates, 127, 155
 CO₂ in seawater, 168, 174
 cosmogenic dating, 357–9
 ion activity product, 126
 organic compounds, 84
 soil mineral formation, 269
 sulfate and halide precipitation, 73–4, 111
- Seawater *see* Ocean
- Second law of thermodynamics, 27
- Selenium, 123, 130–132
- Serpentine
 minerals, 55–7, 64
 CO₂ sequestration, 175
- Shale
 composition, 5
 groundwater Mn source, 92
 mineralogy, 61–4
 petroleum source, 95, 163
- SI units, 27–8
- Silica
 amorphous solid, 78, 136–7
 dissolved form, 117, 136–7, 262, 278, 281
- Silicate dissolution
 kinetics, 259–61
- Smectite
 composition and structure, 55, 59–62, 64–70, 74, 75
 high-level waste disposal, 349
- mineral stability diagram, 277–82
- organic compound sorption, 294–5
- paleosols, 316
- sodic soils, 290–291
- weathering and soil formation, 266, 275–7
- Smog, 93, 181, 208, 230, 231, 237–8, 242–5
- SMOW, 303–4
- Sodic soils, 289–91
- Sodium
 cation exchange reaction, 69
 exchangeable sodium percentage, 290
 Lewis electron dot, 7
- Soil
 horizons, Plate 13
 orders (classification), 273–74
 soil forming factors, 268–73
- Solar constant, 221
- Solubility *see* specific mineral or compound
- Solute, 127
- Solutions (aqueous), 19, 20, 34, 40, 74, 75, 115, 117, 137, 151, 153, 384
- Solvent (organic), 84–7, 103, 294–5, 371–7
- Solvent (property), 97, 113, 115
- Soot, 224, 237, 240–242, 254
- Sorbate (adsorbate), 66, 99–102, 142–3
- Sorbent, 66, 99, 100–102, 142, 143
- Sorption, 33–4, 58, 64–70, 75–7, 99–105, 131, 137–43, 182, 193, 284, 290–295, 345
- Sorption isotherms *see* Isotherm
- Spallation reactions, 355–9
- Speciation
 aluminum, 123–4
 arsenic, 131, 141
 carbonate, 151–2
 Eh-pH diagrams (Al, As, Cu, Fe, Pb, U), 132–7

- nitrogen, 181
 polyatomic ions, 116–17
 phosphorus, 191
 sulfur, 197
 Species (chemical), 13, 14, 20,
 24–7 *see also* specific element
 or compound
 Spectrum of solar energy,
 219–20
 Spodosol, 273–4, 277, 283,
 287, 294
 Stability diagrams (silicates and
 hydroxides) *see* specific
 element
 Stable isotopes *see* specific element
 Standard electrode potential, 121
 Standard hydrogen electrode, 121
 Standard state, 27–8
 Stern layer, 68
 Stiff diagrams, 145–6
 Stream composition
 acidic, 130, 283–5, 264, 325
 arid, 110–113
 dissolved organic matter, 94
 dissolved oxygen, 122
 nitrate increase, 189
 phosphorus transport, 193–5
 temperate, 112–13
 total dissolved solids, 125
 tropical, 110–113
 Stratosphere, 189, 201, 207–8,
 242, 245–8
 Stratospheric ozone *see* Ozone
⁸⁷Sr, 342
 Sulfate
 acid rain
 anion, 250–252, 284–5
 adsorption of, 141
 anion in aqueous solutions, 21,
 68, 122, 196, 200
 minerals, 47, 72–4, 78, 127–9,
 260–262, 269
 plant nutrient, 289
 scatterer in atmosphere, 92,
 201, 219, 220, 224,
 235–6, 241
 seawater, 198
 sulfur isotopic
 values, 315, 322–4
 Sulfate minerals *see* Sulfate
 Sulfide anion, 72–3, 122, 129–31,
 196, 264, 323
 Sulfide minerals, 72–4, 122,
 130–131, 196–9 *see also* Pyrite
 Sulfur
 in atmosphere, 200–201,
 234–7, 249–52
 cycle, 196–201
 stable isotopes, 322–4
 Sulfuric acid, 19, 21, 23, 73, 95,
 123, 196–7, 236, 252,
 264, 323
 Surface
 charge, 66–7, 99, 101, 141
 complexation, 66
 Swelling (shrink-swell)
 clays, 61, 291
 System, defined, 25

 Talc, 54, 59
 Tectosilicates, 49, 58
 Temperature
 and Arrhenius equation *see*
 Arrhenius
 of atmosphere, 207, 211, 221–8
 as degree of freedom, 25
 effect on reaction rate, 269
 of metamorphism, 29, 63
 Tetrahedral coordination, 43–5, 73
 Thermal ionization mass
 spectrometry, 324, 388
 Thermodynamics, definition, 24–5
Thiobacillus ferrooxidans, 75, 200,
 264, 325–6
Thiobacillus thiooxidans, 200
 Thiophene, 235–6
 Thorium, 50, 336, 340–344,
 362, 365
 Time *see* Rate or kinetics
 Total dissolved solids (TDS), 109,
 111–13, 126–7, 218
 Toxicity, 86, 92, 104–5,
 241, 254, 344, 348
 Trace elements
 atmospheric, 252–5
 concentrations in Earth, 5
 controls on
 behavior, 292–3, 372
 in iron hydroxide, 78, 375
 ionic and crystal radii, 44, 48
 units of concentration, 8
 Transition metals, 44, 59, 72, 324
 Trigonal coordination, 42–5
 Trioctahedral clays, 54–5, 60–63
 Tritium, 300, 347, 357
 Troposphere, 189, 200–201,
 207–8, 227, 238, 242, 246–7

 Ultisol, 273, 274, 277, 279
 Ultramafic rocks, 5, 61, 344
 Ultraviolet radiation, 207,
 219–20, 245
 Uncertainty, 10–11, 209, 212,
 278, 281, 328, 352
²³⁴U age dating *see* Age dating
 Uranium
 Autunite, 191
 Eh-pH diagram, 135
 ionic forms, 116, 120, 130
 particulate vs. dissolved, 116
 radioactive decay, 336–44
 radioactive waste, 346–50
 uraninite, 72
 zircon, 50
 U-Th-Pb age dating *see* Age dating
 Urea, 181, 184, 187, 251, 317
 Urease, 184

 Vadose zone, 110, 112, 114,
 345, 356
 Valence, 7, 12–18, 22, 41, 83,
 113, 120, 181, 196, 232, 386
 Van der Waals force *see* Bonding
 Van't Hoff equation, 32–3
 Vapor extraction
 (remediation), 105, 373
 Vegetation and weathering, 33,
 75, 269–70
 Vermiculite
 adsorption of organic
 compounds, 101, 294

- Vermiculite (*cont'd*)
 cation exchange, 66, 69, 294
 composition and
 structure, 55, 62
 formation in soil, 261, 273,
 276, 277
- Vertisol, 273–4, 279
- Vinyl chloride, 105–6, 372, 274
- Vostok core, 210–212
- Water
 absorption of infrared
 radiation, 221–6
- H and O isotopes *see* Hydrogen;
 Oxygen
 structure, 113–14
- Water table, 76, 103–4, 110,
 112, 269–71, 284, 291,
 349, 372
- Weathering and acid
 deposition, 282–7
- Weathering reactions
 see individual minerals
- Wetlands, 72, 75, 111, 120–123,
 130, 163, 187, 189, 249,
 274–5, 347
- Work (in
 thermodynamics), defined,
 26–7
- X-ray diffraction, 43, 292,
 379–82
- X-ray fluorescence, 378, 385
- Zeolites, 58, 66, 349
- Zero point of charge
 see Point of zero charge
- Zinc, 71, 124, 130, 288
- Zircon, 49–50, 260, 272, 362–4



Plate 1 Oxidized mine tailings showing colors typical of hematite (right) and goethite and more-disordered Fe-hydroxides (left). These oxides and hydroxides formed from oxidation of pyrrhotite and chalcopyrite.

(A)



(B)

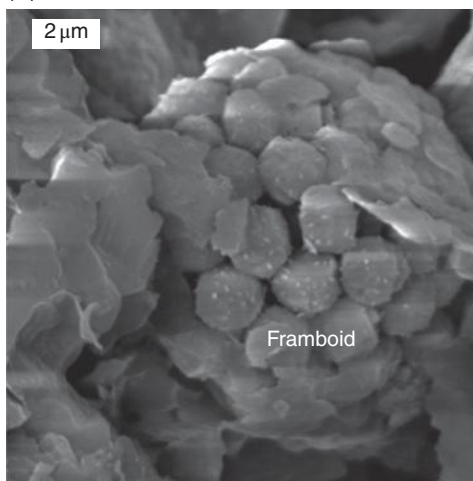


Plate 2 Macroscopic cubic pyrite crystals in Cambrian schist (A) and microscopic framboidal pyrite (B) occurring in the saturated zone of a landfill. (Photo by J. Kim, Vermont Geological Survey (Plate 2A); Photo by G. Breit, US Geological Survey (Plate 2B).)



Plate 3 Photograph of chalcopyrite (CuFeS_2)-rich rock that is weathering into iron hydroxide (orange coating) and chalcocite ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$). Both minerals are secondary – the iron hydroxide forms when iron liberated from chalcopyrite is oxidized; the chalcocite (tan crystals) forms on the rock surface when water containing dissolved Cu and sulfate (derived from chalcopyrite oxidation) is drawn upward through fractures and then evaporates, producing a solution saturated with respect to Cu and SO_4 .

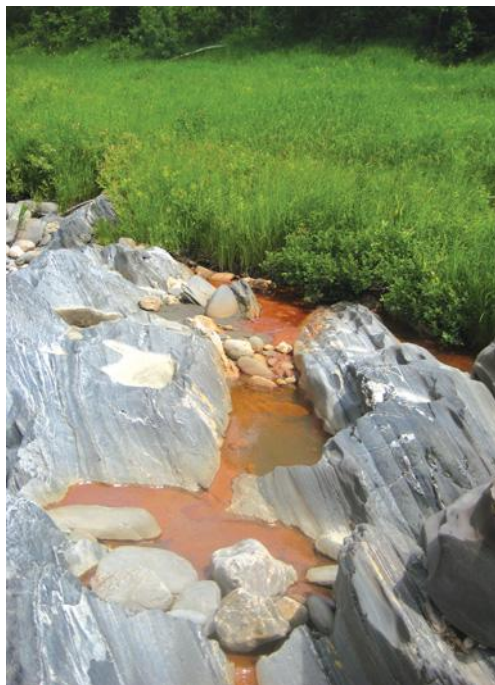


Plate 5 Orange iron hydroxide formed when dissolved Fe(II) in reduced groundwater seeps out at the surface and is oxidized to Fe(III) .



Plate 4 Humic and/or fulvic acids imparting a tea-like light-brown color to river water tumbling over a waterfall. (Photo by E.C. Ryan.)

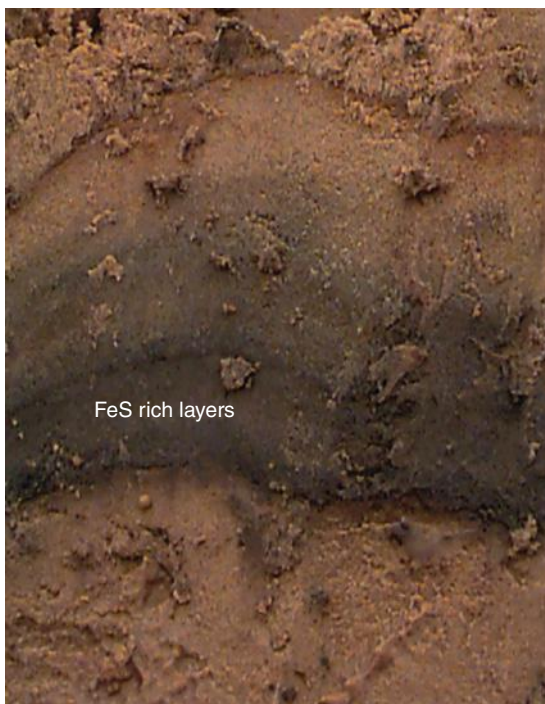


Plate 6 Black layers of poorly-crystalline iron sulfide interbedded with tan sand. The colors of these layers are indicative of their respective redox states – black reduced sediment surrounded by relatively oxidized tan sediment. Photo by Michele Tuttle, US Geological Survey.



Plate 7 Effect of mineral dissolution rate displayed in calcite and dolomite layers in Ordovician carbonate rock. The more-resistant dolomite layers have positive relief relative to calcite layers, which decompose more rapidly and hence are recessed.



Plate 8 Algal bloom in a eutrophic lake.
(Photo by Nara Souza, Florida Fish and Wildlife Commission.)



Plate 9 Effect of naturally-derived acid rain on the flanks of Volcan Poas in Costa Rica. Prevailing winds transport nitric and sulfuric acid emission westward (to the left), creating a barren landscape.

Calidad del Aire		
Cont	mic/m3	Calidad
CO	1303	Buena
NO2	85	Buena
O3	19	Buena
Par	66	Mala
SO2	29	Buena

Logo: JUNTA DE ANDALUCÍA
Consejería de Medio Ambiente

Plate 10 Public display of air quality in Granada, Spain showing acceptable levels of CO, NO, O and SO but unhealthy levels of particulates.



Plate 11 Air inversion in southern Spain. Cold air has sunk into the valley, causing condensation (fog) and trapping of contaminants. The warm air above the inversion layer prevents air circulation. Air temperature where the photograph was taken was about 5 °C warmer than air below. (Photo by F. Javier Huertas, Instituto Andaluz de Ciencias de la Tierra.)



Plate 12 Acid mine drainage caused by oxidation of sulfides, Rio Tinto region of SW Spain. Precipitation of iron hydroxide in the streambed is visually obvious – the cause is increasing Eh and/or pH (Figure 4.15).

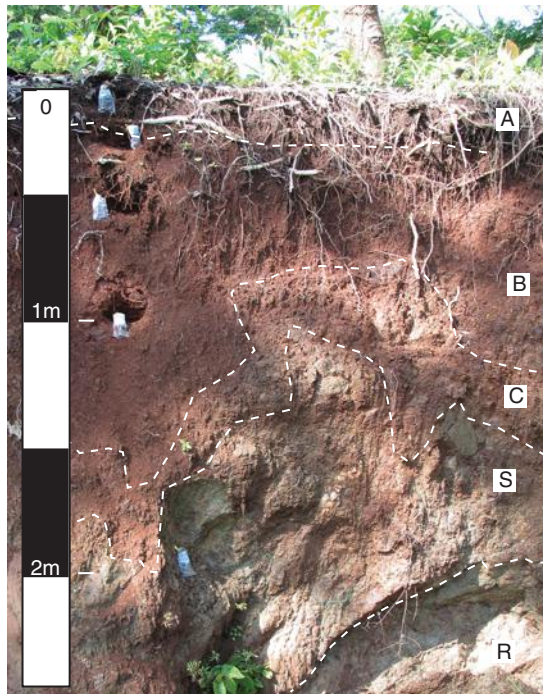


Plate 13 Soil profile formed via weathering of basalt in northwestern Costa Rica. The soil is a tropical Alfisol and horizons are indicated by A-B-C-S-R on the right, where S = saprolite and R = rock. Note the wavy nature of horizon boundaries and the progressive increase up-profile in red colors indicating formation of iron hydroxides (e.g. goethite) via liberation of Fe from silicate minerals.

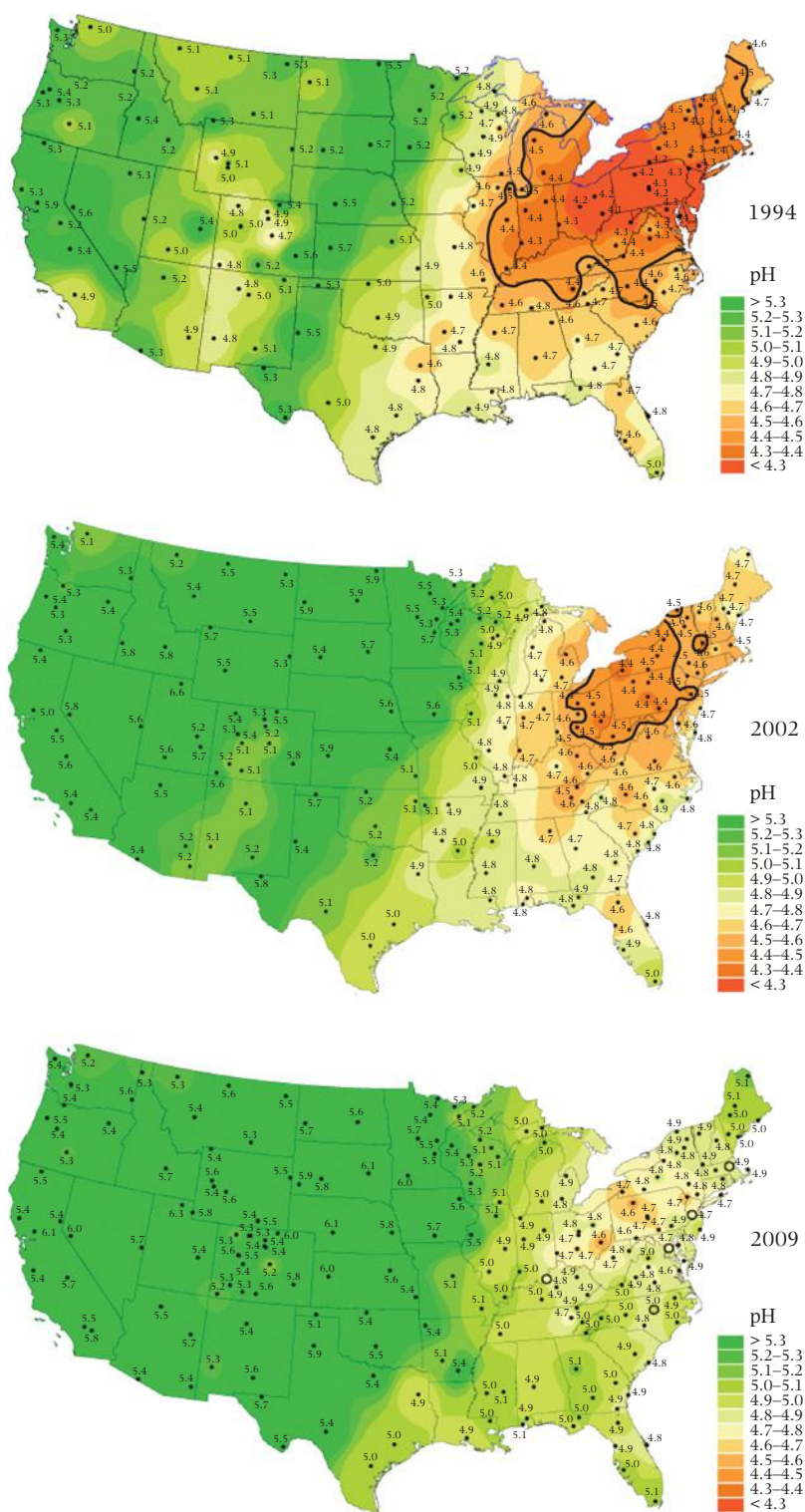


Plate 14 Average annual pH in precipitation in the conterminous United States for the years 1994, 2002 and 2009. The black contour line is drawn to highlight areas with mean annual pH < 4.5. (US EPA.)

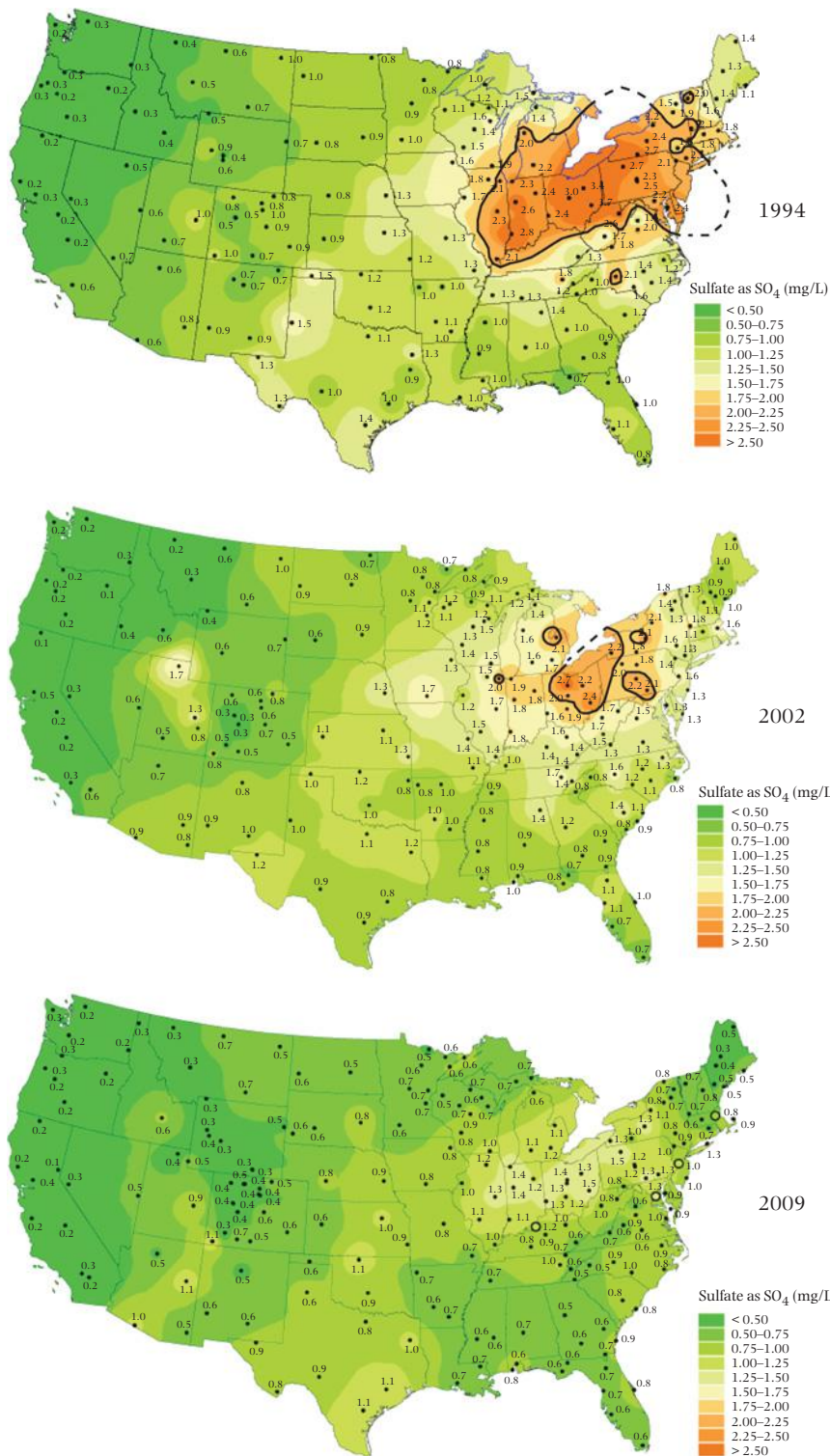


Plate 15 Average annual nitrate concentrations in precipitation in the conterminous United States for the years 1994, 2002 and 2009. The black contour line is drawn to highlight areas with mean annual nitrate concentrations > 1.5 mg/L. Note that these values are given for concentration of NO_3 , not $\text{NO}_3\text{-N}$, which is commonly the case with surface and ground waters. (US EPA.)

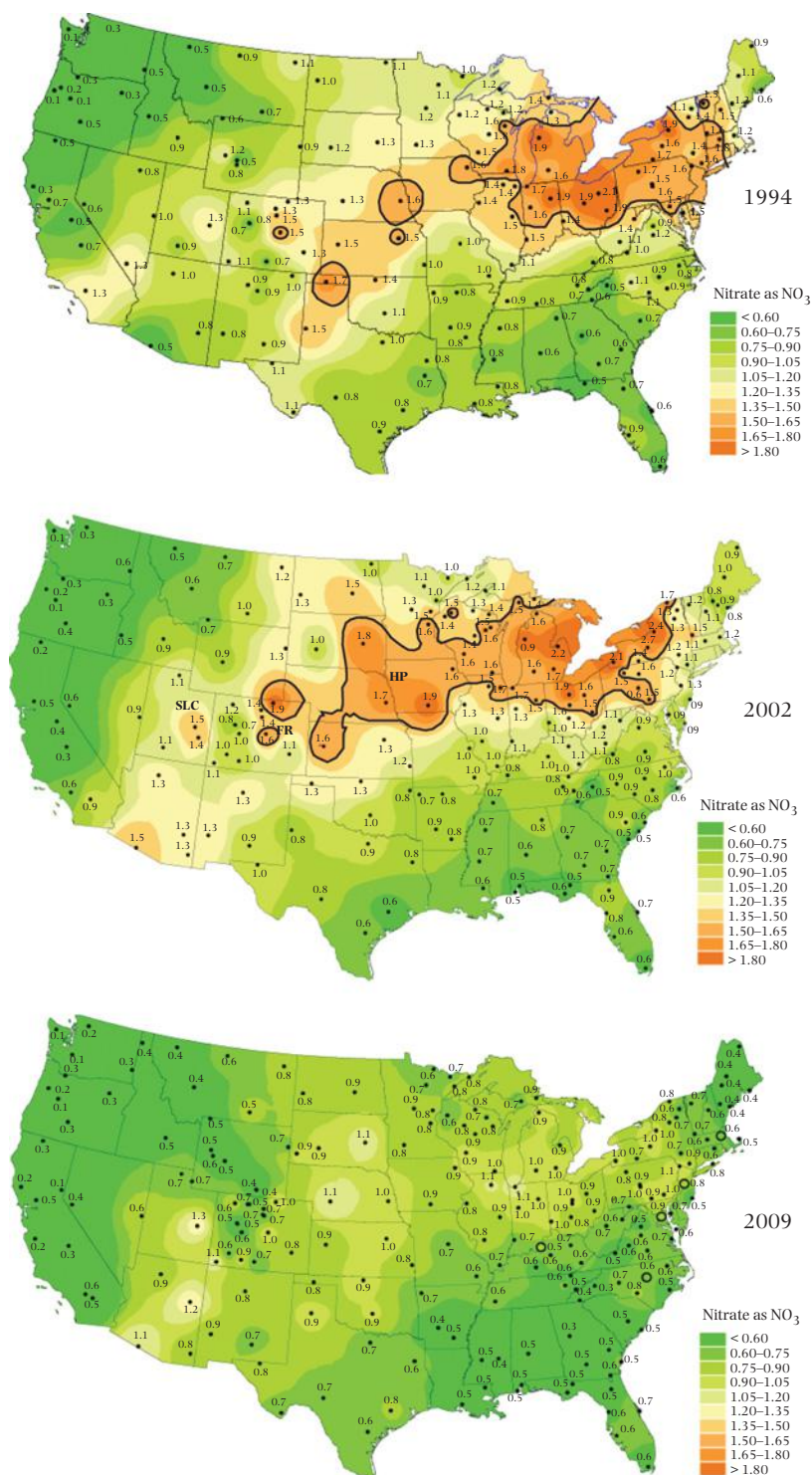


Plate 16 Average annual sulfate concentrations in precipitation in the conterminous United States for the years 1994, 2002 and 2009. The black contour line is drawn to highlight areas with mean annual sulfate concentrations > 2.0 mg/L. Note that these values are given for concentration of SO_4 , not $\text{SO}_4\text{-S}$, which is sometimes the case. (US EPA.)