AN INTRODUCTION TO GEOLOGY

Chris Johnson, Matthew D. Affolter, Paul Inkenbrandt, & Cam Mosher
Salt Lake Community College

LibreTexts™

About this Book

Written by

Chris Johnson, Matthew D. Affolter, Paul Inkenbrandt, Cam Mosher

Salt Lake Community College – 2017

Contact the authors at edits@opengeology.org with edits, suggestions, or if adopting the book.

Faculty who adopt this text for their course should contact the authors at edits@opengeology.org so that the authors can keep faculty users up to date of critical changes.

Edited and Reviewed by

Johnathan Barnes – Salt Lake Community College
Jack Bloom – Rio Tinto Kennecott Copper (retired)
Deanna Brandau – Utah Museum of Natural History
Gereld Bryant – Dixie State University
Gregg Beukelman – Utah Geological Survey
Peter Davis – Pacific Lutheran University
Renee Faatz – Snow College
Gabriel Filippelli – Indiana University-Purdue University Indianapolis
Michelle Cooper Fleck – Utah State University
Colby Ford – Dixie State University
Michael Hylland – Utah Geological Survey
Mike Kass – Salt Lake Community College
Ben Laabs – North Dakota State University
Tom Lachmar – Utah State University
Johnny MacLean – Southern Utah University
Erich Peterson – University of Utah
Tiffany Rivera – Westminster College
Leif Tapanila – Idaho State University

Special Thanks to

Jason Pickavance – Salt Lake Community College, Director of Educational Initiatives
R. Adam Dastrup – Salt Lake Community College, Geoscience Dept. Chair, Director of Opengeography.org

Open Educational Resources

This text is provided to you as an Open Educational Resource which you access online. It is designed to give you a comprehensive introduction to Geology at no or very nominal cost. It contains both written and graphic text material, intra-text links to other internal material which may aid in understanding topics and concepts, intra-text links to the appendices and glossary for tables and definitions of words, and extra-text links to videos and web material that clarifies and augments topics and concepts. Like any new or scientific subject, Geology has its own vocabulary for geological concepts. For you to converse effectively with this text and colleagues in this earth science course, you will use the language of geology, so comprehending these terms is important. Use the intra-text links to the Glossary and other related material freely to gain familiarity with this language.

Salt Lake Community College

An Introduction to Geology

Chris Johnson, Matthew D. Affolter, Paul Inkenbrandt, & Cam Mosher

The LibreTexts libraries are Powered by MindTouch® and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

This text was compiled on 04/04/2022

TABLE OF CONTENTS

This text is designed to give you a comprehensive introduction to Geology at no or very nominal cost. It contains both written and graphic text material, intra-text links to other internal material which may aid in understanding topics and concepts, intra-text links to the appendices and glossary for tables and definitions of words, and extra-text links to videos and web material that clarifies and augments topics and concepts.

Sedimentary rock and the processes that create it, which include weathering, erosion, and lithification, are an integral part of understanding Earth Science. This is because the majority of the Earth’s surface is made up of sedimentary rocks and their common predecessor, sediments. Even though sedimentary rocks can form in drastically different ways, their origin and creation have one thing in common, water.

6: METAMORPHIC ROCKS

Metamorphic rocks is one of the three rock categories in the rock cycle. Metamorphic rock material has been changed by temperature, pressure, and/or fluids. The rock cycle shows that both igneous and sedimentary rocks can become metamorphic rocks. And metamorphic rocks themselves can be re-metamorphosed. Because metamorphism is caused by plate tectonic motion, metamorphic rock provides geologists with a history book of how past tectonic processes shaped our planet

7: GEOLOGIC TIME

8: EARTH HISTORY

9: CRUSTAL DEFORMATION AND EARTHQUAKES

10: MASS WASTING

This chapter discusses the fundamental processes driving mass-wasting, types of mass wasting, examples and lessons learned from famous mass-wasting events, how mass wasting can be predicted, and how people can be protected from this potential hazard. Mass wasting is the downhill movement of rock and soil material due to gravity. The term landslide is often used as a synonym for mass wasting, but mass wasting is a much broader term referring to all movement downslope.

11: WATER

12: COASTLINES

Coastlines are the great interface between the 29% of earth's surface that is land and 71% of earth that is covered by the oceans. Therefore, it is the longest visible boundary on earth. To understand the processes that take place at this interface, we must first consider the energetic action at this boundary; namely, waves. The importance of this interface is seen in the study of ancient shorelines, and particularly for natural resources, a process called sequence stratigraphy.

13: DESERTS

Approximately 30 percent of Earth's terrestrial surface is desert. Deserts are defined as locations of low precipitation. While temperature extremes are often associated with deserts, they do not define them. The lack of moisture, including the lack of humidity and cloud cover, allow temperature extremes to occur. The sun's energy is more absorbed by the Earth's surface without cloud cover, and nighttime cooling is more drastic without cloud cover and humidity to absorb the emitted heat, so temp

14: GLACIERS

The hydrosphere, liquid water, is the single most important agent of erosion and deposition. The cryosphere, the solid state of water in the form of ice also has its own unique erosional and depositional features. Large accumulations of year-round ice on the land surface are called glaciers. Masses of ice floating on the ocean as sea ice or icebergs are not glaciers, although they may have had their origin in glaciers.

15: GLOBAL CLIMATE CHANGE

This chapter describes the systems involved in regulating Earth's temperature, its climate, geologic evidence of past climate changes, and the role humans have on today's climate change. It is critically important to be aware of the geologic context of climate change processes if we want to understand anthropogenic (human-caused) climate change.

16: ENERGY AND MINERAL RESOURCES

An important use of geologic knowledge is locating economically valuable materials for use in society. All items we use can come from only three sources: they can be farmed, hunted or fished, or they can be mined. without mining, modern civilization would not exist. Geologists are essential in the process of mining.

BACK MATTER

CHAPTER OVERVIEW

1: UNDERSTANDING SCIENCE

Learning Objectives

Contrast objective versus subjective observations, and quantitative versus qualitative observations
Identify a pseudoscience based on its lack of falsifiability
Contrast the methods used by Aristotle and Galileo to describe the natural environment
Explain the scientific method and apply it to a problem or question
Describe the foundations of modern geology, such as the principle of uniformitarianism
Contrast uniformitarianism with catastrophism
Explain why studying geology is important
Identify how Earth materials are transformed by rock cycle processes
Describe the steps involved in a reputable scientific study
Explain rhetorical arguments used by science deniers

Science is a process, with no beginning and no end. Science is never finished because a full truth can never be known. However, science and the scientific method are the best way to understand the universe we live in. Scientists draw conclusions based on objective evidence; they consolidate these conclusions into unifying models. Geologists likewise understand studying the Earth is an ongoing process, beginning with James Hutton who declared the Earth has "…no vestige of a beginning, no prospect of an end." Geologists explore the 4.5 billion-year history of Earth, its resources, and its many hazards. From a larger viewpoint, geology can teach people how to develop credible conclusions, as well as identify and stop misinformation.

1.1: What is Science?

Scientists seek to understand the fundamental principles that explain natural patterns and processes. Science is more than just a body of knowledge, science provides a means to evaluate and create new knowledge without bias [1]. Scientists use objective evidence over subjective evidence, to reach sound and logical conclusions.

An objective observation is without personal bias and the same by all individuals. Humans are biased by nature, so they cannot be completely objective; the goal is to be as unbiased as possible. A subjective observation is based on a person's feelings and beliefs and is unique to that individual.

Figure 1.1.1: This is Grand Canyon of the Yellowstone in Yellowstone National Park. An objective statement about this would be: "The picture is of a waterfall." A subjective statement would be: "The picture is beautiful." An inference would be "The waterfall is there because of erosion."

Another way scientists avoid bias is by using quantitative over qualitative measurements whenever possible. A quantitative measurement is expressed with a specific numerical value. Qualitative observations are general or relative descriptions. For example, describing a rock as red or heavy is a qualitative observation. Determining a rock's color by measuring wavelengths of reflected light or its density by measuring the proportions of minerals it contains is quantitative. Numerical values are more precise than general descriptions, and they can be analyzed using statistical calculations. This is why quantitative measurements are much more useful to scientists than qualitative observations.

Figure 1.1.2: Canyons like this, carved in the deposit left by the May 18th, 1980 eruption of Mt. St. Helens, are sometimes used by purveyors of pseudoscience as evidence for the Earth being very young. In reality, the non-lithified volcanic deposit is carved much more easily than other canyons like the Grand Canyon.

Establishing truth in science is difficult because all scientific claims are falsifiable, which means any initial hypothesis may be tested and proven false. Only after exhaustively eliminating false results, competing ideas, and possible variations does a hypothesis become regarded as a reliable scientific theory. This meticulous scrutiny reveals weaknesses or flaws in a hypothesis and is the strength that supports all scientific ideas and procedures. In fact, proving current ideas are wrong has been the driving force behind many scientific careers.

Falsifiability separates science from pseudoscience. Scientists are wary of explanations of natural phenomena that discourage or avoid falsifiability. An explanation that cannot be tested or does not meet scientific standards is not considered science, but pseudoscience. Pseudoscience is a collection of ideas that may appear scientific but does not use the scientific method. Astrology is an example of pseudoscience. It is a belief system that attributes the movement of celestial bodies to influencing human behavior. Astrologers rely on celestial observations, but their conclusions are not based on experimental evidence and their statements are not falsifiable. This is not to be confused with astronomy which is the scientific study of celestial bodies and the cosmos [2,3].

Figure 1.1.3: Geologists share information by publishing, attending conferences, and even going on field trips, such as this trip to western Utah by the Utah Geological Association in 2009.

Science is also a social process. Scientists share their ideas with peers at conferences, seeking guidance and feedback. Research papers and data submitted for publication are rigorously reviewed by qualified peers, scientists who are experts in the same field. The scientific review process aims to weed out misinformation, invalid research results, and wild speculation. Thus, it is slow, cautious, and conservative. Scientists tend to wait until a hypothesis is supported by an overwhelming amount of evidence from many independent researchers before accepting it as a scientific theory [46].

References

- 1. *Science literacy: concepts, contexts, and consequences*. (National Academies Press (US), 2016).
- 2. Shermer, M. What Is Pseudoscience? *Scientific American* (2011).
- 3. Popper, K. R. & Hudson, G. E. Conjectures and refutations. (1963).

Contributions and Attributions

1.2: The Scientific Method

Modern science is based on the scientific method, a procedure that follows these steps:

- Formulate a question or observe a problem
- Apply objective experimentation and observation
- Analyze collected data and Interpret results
- Devise an evidence-based theory
- Submit findings to peer review and/or publication

This has a long history in human thought but was first fully formed by Ibn al-Haytham over 1,000 years ago. At the forefront of the scientific method are conclusions based on objective evidence, not opinion or hearsay [4].

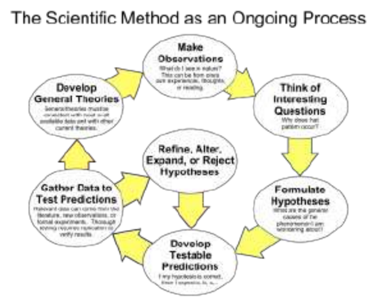


Figure 1.2.1: Diagram of the cyclical nature of the scientific method.

Step 1: Observation, Problem, or Research Question

The procedure begins with identifying a problem or research question, such as a geological phenomenon that is not well explained in the scientific community's collective knowledge. This step usually involves reviewing the scientific literature to understand previous studies that may be related to the question.

Step 2: Hypothesis

Once the problem or question is well defined, the scientist proposes a possible answer, a hypothesis, before conducting an experiment or fieldwork. This hypothesis must be specific, falsifiable, and should be based on other scientific work. Geologists often develop multiple working hypotheses because they usually cannot impose strict experimental controls or have limited opportunities to visit a field location [5; 6; 7].

Figure 1.2.2: A famous hypothesis: Leland Stanford wanted to know if a horse lifted all 4 legs off the ground during a gallop since the legs are too fast for the human eye to perceive it. These series of photographs by Eadweard Muybridge proved the horse, in fact, does have all four legs off the ground during the gallop.

Step 3: Experiment and Hypothesis Revision

The next step is developing an experiment that either supports or refutes the hypothesis. Many people mistakenly think experiments are only done in a lab; however, an experiment can consist of observing natural processes in the field. Regardless of what form an experiment takes, it always includes the systematic gathering of objective data. This data is interpreted to determine whether it contradicts or supports the hypothesis, which may be revised and tested again. When a hypothesis holds up under experimentation, it is ready to be shared with other experts in the field.

Figure 1.2.3: An experiment at the University of Queensland has been going since 1927. A petroleum product called pitch, which is highly viscous, drips out of a funnel about once per decade.

Step 4: Peer Review, Publication, and Publication

Scientists share the results of their research by publishing articles in scientific journals, such as *Science* and *Nature*. Reputable journals and publishing houses will not publish an experimental study until they have determined its methods are scientifically rigorous and the conclusions are supported by evidence. Before an article is published, it undergoes a rigorous peer review by scientific experts who scrutinize the methods, results, and discussion. Once an article is published, other scientists may attempt to replicate the results. This replication is necessary to confirm the reliability of the study's reported results. A hypothesis that seemed compelling in one study might be proven false in studies conducted by other scientists. New technology can be applied to published studies, which can aid in confirming or rejecting once-accepted ideas and/or hypotheses.

Step 5: Theory Development

In casual conversation, the word *theory* implies guesswork or speculation. In the language of science, an explanation or conclusion made in a *theory* carries much more weight because it is supported by experimental verification and widely accepted by the scientific community. After a hypothesis has been repeatedly tested for falsifiability through documented and independent studies, it eventually becomes accepted as a scientific theory.

While a hypothesis provides a tentative explanation *before* an experiment, a theory is the best explanation *after* being confirmed by multiple independent experiments. Confirmation of a theory may take years, or even longer. For example, the continental drift hypothesis first proposed by Alfred Wegener in 1912 was initially dismissed. After decades of additional evidence collection by other scientists using more advanced technology, Wegener's hypothesis was accepted and revised as the theory of plate tectonics.

The theory of evolution by natural selection is another example. Originating from the work of Charles Darwin in the mid-19th century, the theory of evolution has withstood generations of scientific testing for falsifiability. While it has been updated and revised to accommodate knowledge gained by using modern technologies, the theory of evolution continues to be supported by the latest evidence.

Figure 1.2.4: Wegener later in his life, ca. 1924-1930.

References

- 4. Sabra, A. I. & Others. *The optics of Ibn al-Haytham: Books I-III: On direct vision*. **1**, (Warburg Institute, University of London, 1989).
- 5. Railsback, B. L. T. C. Chamberlin's "Method of Multiple Working Hypotheses": An encapsulation for modern students. *Houston Geological Society Bulletin* **47**, 68–69 (2004).
- 6. Railsback, B. L. T. C. Chamberlin's 'Method of Multiple Working Hypotheses': An encapsulation for modern students. (1990). Available at: http://www.gly.uga.edu/railsback/r ailsback_chamberlin.html. (Accessed: 29th December 2016)
- 7. Chamberlin, T. C. The method of multiple working hypotheses. *Science* **15**, 92–96 (1890).

Contributions and Attributions

1.3: Early Scientific Thought

Western scientific thought began in the ancient city of Athens, Greece [8]. Athens was governed as a democracy, which encouraged individuals to think independently, at a time when most civilizations were ruled by monarchies or military conquerors. Foremost among the early philosopher/scientists to use empirical thinking was Aristotle, born in 384 BCE. Empiricism emphasizes the value of evidence gained from experimentation and observation. Aristotle studied under Plato and tutored Alexander the Great. Alexander would later conquer the Persian Empire, and in the process spread Greek culture as far east as India.

Aristotle applied an empirical method of analysis called deductive reasoning, which applies known principles of thought to establish new ideas or predict new outcomes. Deductive reasoning starts with generalized principles and logically extends them to new ideas or specific conclusions. If the initial principle is valid, then it is highly likely the conclusion is also valid. An example of deductive reasoning is if A=B, and B=C, then A=C. Another example is if all birds have feathers, and a sparrow is a bird, then a sparrow must also have feathers. The problem with deductive reasoning is if the initial principle is flawed, the conclusion will inherit that flaw. Here is an example of a flawed initial principle leading to the wrong conclusion; if all animals that fly are birds, and bats also fly, then bats must also be birds.

This type of empirical thinking contrasts with inductive reasoning, which begins from new observations and attempts to discern underlying generalized principles. A conclusion made through inductive reasoning comes from analyzing measurable evidence, rather than making a logical connection. For example, to determine whether bats are birds a scientist might list various characteristics observed in birds–the presence of feathers, a toothless beak, hollow bones, lack of forelegs, and externally laid eggs. Next, the scientist would check whether bats share the same characteristics, and if they do not, draw the conclusion that bats are not birds.

Both types of reasoning are important in science because they emphasize the two most important aspects of science: observation and inference. Scientists test existing principles to see if they accurately infer or predict their observations. They also analyze new observations to determine if the inferred underlying principles still support them [9; 10].

Figure 1.3.1: drawing of Avicenna (Ibn Sina). He is among the first to link mountains to earthquakes and erosion.

Greek culture was spread by Alexander and then absorbed by the Romans, who help further extend Greek knowledge into Europe through their vast infrastructure of roads, bridges, and aqueducts [11]. After the fall of the Roman Empire in 476 CE, scientific progress in Europe stalled [8]. Scientific thinkers of medieval time had such high regard for Aristotle's wisdom and knowledge they faithfully followed his logical approach to understanding nature for centuries. By contrast, science in the Middle East flourished and grew between 800 and 1450 CE, along with culture and the arts.

Near the end of the medieval period, empirical experimentation became more common in Europe. During the Renaissance, which lasted from the 14th through 17th centuries, artistic and scientific thought experienced a great awakening [12; 13; 14]. European scholars began to criticize the traditional Aristotelian approach and by the end of the Renaissance period, empiricism was poised to become a key component of the scientific revolution that would arise in the 17th century [15].

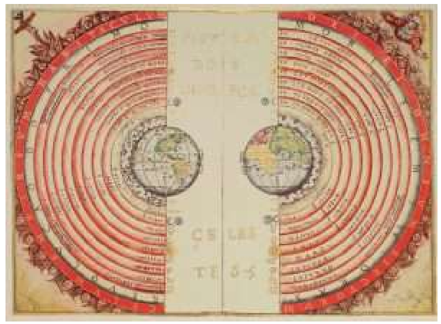


Figure 1.3.2: Geocentric drawing by Bartolomeu Velho in 1568

An early example of how Renaissance scientists began to apply a modern empirical approach is their study of the solar system. In the second century, the Greek astronomer Claudius Ptolemy observed the Sun, Moon, and stars moving across the sky. Applying Aristotelian logic to his astronomical calculations, he deductively reasoned all celestial bodies orbited around the Earth, which was located at the center of the universe. Ptolemy was a highly regarded mathematician, and his mathematical calculations were widely accepted by the scientific community. The view of the cosmos with Earth at its center is called the geocentric model. This geocentric model persisted until the Renaissance period when some revolutionary thinkers challenged the centuries-old hypothesis.

By contrast, early Renaissance scholars such as astronomer Nicolaus Copernicus (1473-1543) proposed an alternative explanation for the perceived movement of the Sun, Moon, and stars. Sometime between 1507 and 1515, he provided credible mathematical proof for a radically new model of the cosmos, one in which the Earth and other planets orbited around a centrally located Sun. After the invention of the telescope in 1608, scientists used their enhanced astronomical observations to support this heliocentric, Sun-centered, model [16; 17].

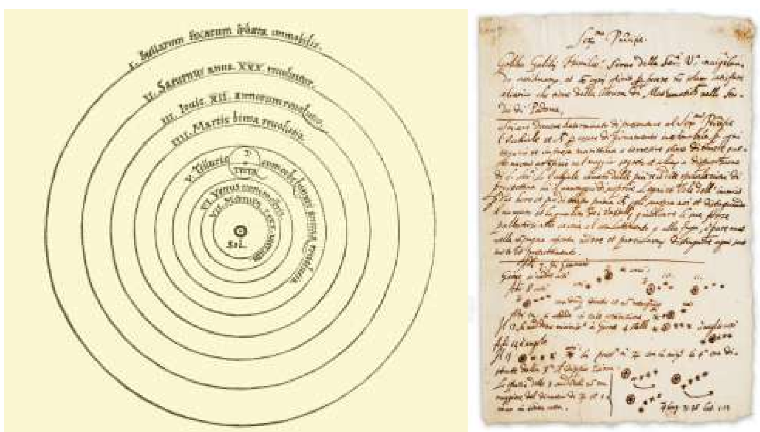


Figure 1.3.3: Copernicus' heliocentric model Galileo's first mention of moons of Jupiter.

Two scientists, Johannes Kepler and Galileo Galilei, are credited with jump-starting the scientific revolution [15]. They accomplished this by building on Copernicus' work and challenging long-established ideas about nature and science.

Johannes Kepler (1571-1630) was a German mathematician and astronomer who expanded on the heliocentric model—improving Copernicus' original calculations and describing planetary motion as elliptical paths. Galileo Galilei (1564 – 1642) was an Italian astronomer who used the newly developed telescope to observe the four largest moons of Jupiter [18]. This was the first piece of direct evidence to contradict the geocentric model since moons orbiting Jupiter could not also be orbiting Earth.

Galileo strongly supported the heliocentric model and attacked the geocentric model, arguing for a more scientific approach to determine the credibility of an idea [19]. Because of this, he found himself at odds with prevailing scientific views and the Catholic Church. In 1633 he was found guilty of heresy and placed under house arrest, where he would remain until his death in 1642 [18; 19].

Galileo is regarded as the first modern scientist because he conducted experiments that would prove or disprove falsifiable ideas and based his conclusions on mathematical analysis of quantifiable evidence—a radical departure from the deductive thinking of Greek philosophers such as Aristotle [15; 18]. His methods marked the beginning of a major shift in how scientists studied the natural world, with an increasing number of them relying on evidence and experimentation to form their hypotheses. It was during this revolutionary time that geologists such as James Hutton and Nicolas Steno also made great advances in their scientific fields of study [15].

References

- 8. Everitt, A. The Rise of Athens: The Story of the World's Greatest Civilization. (2016).
- 9. Martin, C. *Subverting Aristotle: Religion, History, and Philosophy in Early Modern Science*. (Baltimore : Johns Hopkins University Press., 2014).
- 10. Marshall, J. *A short history of Greek philosophy*. (Andrews UK Limited, 2010).
- 11. Goldsworthy, A. K. *The complete Roman army*. (Thames & Hudson, 2011).
- 12. Saliba, G. *Islamic science and the making of the European renaissance*. (MIT Press, 2007).
- 13. Hogendijk, J. P. & Sabra, A. I. *The Enterprise of Science in Islam: New Perspectives*. (MIT Press, 2003).
- 14. Beckwith, C. How western Europe developed a full scientific method. *Berfrois* (2013).
- 15. Cohen, H. F. *How modern science came into the world: Four civilizations, one 17th-century breakthrough*. (Amsterdam University Press, 2010).
- 16. Goldstein, B. R. Copernicus and the origin of his heliocentric system. *Journal for the History of Astronomy* **33**, 219–235 (2002).
- 17. Repcheck, J. *Copernicus' secret: How the scientific revolution began*. (Simon and Schuster, 2007).
- 18. Drake, S. *Galileo: Pioneer Scientist*. (University of Toronto Press, 1990).
- 19. Heilbron, J. L. *Galileo*. (Oxford University Press, 2012).

Contributions and Attributions

1.4: Foundations of Modern Geology

As part of the scientific revolution in Europe, modern geologic principles developed in the 17th and 18th centuries. One major contributor was Nicolaus Steno (1638-1686), a Danish priest who studied anatomy and geology. Steno was the first to propose the Earth's surface could change over time. He suggested sedimentary rocks, such as sandstone and shale, originally formed in horizontal layers with the oldest on the bottom and progressively younger layers on top [20].

Figure 1.4.1: Illustration by Steno showing a comparison between fossil and modern shark teeth.

In the 18th century, Scottish naturalist James Hutton (1726–1797) studied rivers and coastlines and compared the sediments they left behind to exposed sedimentary rock strata. He hypothesized that ancient rocks must have been formed by processes like those producing the features in the oceans and streams. Hutton also proposed the Earth was much older than previously thought. Modern geologic processes operate slowly. Hutton realized if these processes formed rocks, then the Earth must be very old, possibly hundreds of millions of years old [21; 22].

Hutton's idea is called the principle of uniformitarianism and states that natural processes operate the same now as in the past, i.e. the laws of nature are uniform across space and time. Geologists often state that "the present is the key to the past," meaning they can understand ancient rocks by studying modern geologic processes.

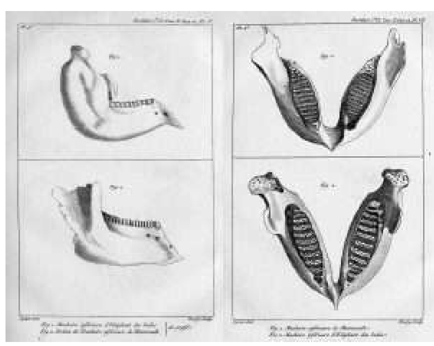


Figure 1.4.2: Cuvier's comparison of modern elephant and mammoth jawbones.

Prior to the acceptance of uniformitarianism, scientists such as German geologist Abraham Gottlob Werner (1750-1817) and French anatomist Georges Cuvier (1769-1832) thought rocks and landforms were formed by great catastrophic events. Cuvier championed this view, known as catastrophism, and stated, "The thread of operation is broken; nature has changed course, and none of the agents she employs today would have been sufficient to produce her former works." He meant processes that operate today did not operate in the past [21; 23]. Known as the father of vertebrate paleontology, Cuvier made significant contributions to the study of ancient life and taught at Paris's Museum of Natural History. Based on his study of large vertebrate fossils, he was the first to suggest species could go extinct. However, he thought new species were introduced by special creation after catastrophic floods [21; 24].

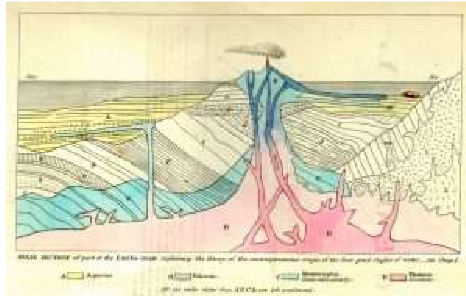


Figure 1.4.3: Inside cover of Lyell’s Elements of Geology

Hutton’s ideas about uniformitarianism and Earth’s age were not well received by the scientific community of his time. His ideas were falling into obscurity when Charles Lyell, a British lawyer and geologist (1797-1875), wrote the *Principles of Geology* in the early 1830s and later, *Elements of Geology* [21, 25]. Lyell’s books promoted Hutton’s principle of uniformitarianism, his studies of rocks and the processes that formed them, and the idea that Earth was possibly over 300 million years old. Lyell and his three-volume *Principles of Geology* had a lasting influence on the geologic community and public at large, who eventually accepted uniformitarianism and millionfold age for the Earth [21]. The principle of uniformitarianism became so widely accepted, that geologists regarded catastrophic change as heresy. This made it harder for ideas like the sudden demise of the dinosaurs by asteroid impact to gain traction.

A contemporary of Lyell, Charles Darwin (1809-1882) took *Principles of Geology* on his five-year trip on the HMS Beagle [27]. Darwin used uniformitarianism and deep geologic time to develop his initial ideas about evolution. Lyell was one of the first to publish a reference to Darwin’s idea of evolution [28].

The next big advancement, and perhaps the largest in the history of geology, is the theory of plate tectonics and continental drift. Dogmatic acceptance of uniformitarianism inhibited the progress of this idea, mainly because of the permanency placed on the continents and their positions. Ironically, the slow and steady movement of plates would fit well into a uniformitarianism model. However, much time passed and a great deal of scientific resistance had to be overcome before the idea took hold. This happened for several reasons. Firstly, the movement was so slow it was overlooked. Secondly, the best evidence was hidden under the ocean. Finally, the accepted theories were anchored by a large amount of inertia. Instead of being bias-free, scientists resisted and ridiculed the emerging idea of plate tectonics. This example of dogmatic thinking is still to this day a tarnish on the geoscience community.

Plate tectonics is most commonly attributed to Alfred Wegener, the first scientist to compile a large data set supporting the idea of continents shifting places over time. He was mostly ignored and ridiculed for his ideas, but later workers like Marie Tharp, Bruce Heezen, Harry Hess, Laurence Morley, Frederick Vine, Drummond Matthews, Kiyoo Wadati, Hugo Benioff, Robert Coats, and J. Tuzo Wilson benefited from advances in subsea technologies. They discovered, described, and analyzed new features like the mid-ocean ridge, alignment of earthquakes, and magnetic striping. Gradually these scientists introduced a paradigm shift that revolutionized geology into the science we know today.

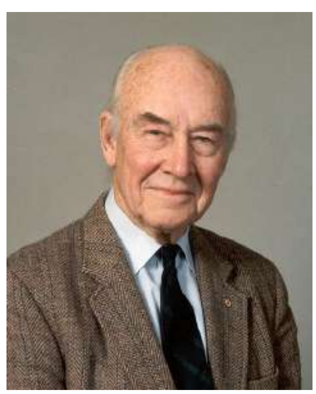


Figure 1.4.3: J. Tuzo Wilson

References

- 20. Paradowski, R. J. Nicolas Steno: Danish anatomist and geologist. *Great Lives from History: Scientists & Science* 830–832 (2012).

- 21. Repcheck, J. *The Man Who Found Time : James Hutton and the Discovery of the Earth's Antiquity*. (New York: Basic Books., 2009).
- 22. Rappaport, R. James Hutton and the History of Geology. Dennis R. Dean. *Isis* **85**, 524–525 (1994).
- 23. Van Reybrouck, D. *From Primitives to Primates: A History of Ethnographic and Primatological Analogies in the Study of Prehistory*. (Sidestone Press, 2012).
- 24. Kolbert, E. *The sixth extinction: an unnatural history*. (Henry Holt and Co., 2014).
- 25. Porter, R. Charles Lyell and the Principles of the History of Geology. *Br. J. Hist. Sci.* **9**, 91–103 (1976).
- 27. Darwin, C. *Geological Observations on South America: Being the Third Part of the Geology of the Voyage of the Beagle, Under the Command of Capt. Fitzroy, R.N. During the Years 1832 to 1836*. (Smith, Elder and Company, 1846).
- 28. Mayr, E. *Systematics and the Origin of Species, from the Viewpoint of a Zoologist*. (Harvard University Press, 1942).

Contributions and Attributions

1.5: The Study of Geology

Geologists apply the scientific method to learn about Earth's materials and processes. Geology plays an important role in society; its principles are essential to locating, extracting, and managing natural resources; evaluating environmental impacts of using or extracting these resources; as well as understanding and mitigating the effects of natural hazards.

Figure 1.5.1: A class looks at rocks in Zion National Park.

Geology often applies information from physics and chemistry to the natural world, like understanding the physical forces in a landslide or the chemical interaction between water and rocks. The term comes from the Greek word *geo*, meaning Earth, and *logos*, meaning to think or reckon with.

Why Study Geology?

Figure 1.5.2: Hoover Dam provides hydroelectric energy and stores water for southern Nevada.

Geology plays a key role in how we use natural resources—any naturally occurring material that can be extracted from the Earth for economic gain. Our developed modern society, like all societies before it, is dependent on geologic resources. Geologists are involved in extracting fossil fuels, such as coal and petroleum; metals such as copper, aluminum, and iron; and water resources in streams and underground reservoirs inside soil and rocks. They can help conserve our planet's finite supply of nonrenewable resources, like petroleum, which are fixed in quantity and depleted by consumption. Geologists can also help manage renewable resources that can be replaced or regenerated, such as solar or wind energy, and timber.

Figure 1.5.3: Coal power plant in Helper, Utah.

Resource extraction and usage impact our environment, which can negatively affect human health. For example, burning fossil fuels releases chemicals into the air that are unhealthy for humans, especially children. Mining activities can release toxic heavy metals, such as lead and mercury, into the soil and waterways. Our choices will have an effect on Earth's environment for the foreseeable future. Understanding the remaining quantity, extractability, and renewability of geologic resources will help us better sustainably manage those resources.

Figure 1.5.4: Buildings toppled from liquefaction during a 7.5 magnitude earthquake in Japan.

Geologists also study natural hazards created by geologic processes. Natural hazards are phenomena that are potentially dangerous to human life or property. No place on Earth is completely free of natural hazards, so one of the best ways people can protect themselves is by understanding the geology. Geology can teach people about the natural hazards in an area and how to prepare for them. Geologic hazards include landslides, earthquakes, tsunamis, floods, volcanic eruptions, and sea-level rise.

Figure 1.5.5: Oregon's Crater Lake was formed about 7700 years ago after the eruption of Mount Mazama.

Finally, geology is where other scientific disciplines intersect in the concept known as Earth System Science. In science, a system is a group of interactive objects and processes. Earth System Science views the entire planet as a combination of systems that interact with each other via complex relationships. This geology textbook provides an introduction to science in general and will often reference other scientific disciplines.

Earth System Science includes five basic systems (or spheres), the Geosphere (the solid body of the Earth), the Atmosphere (the gas envelope surrounding the Earth), the Hydrosphere(water in all its forms at and near the surface of the Earth), the Cryosphere (frozen water part of Earth), and the Biosphere (life on Earth in all its forms and interactions, including humankind).

Rather than viewing geology as an isolated system, earth system scientists study how geologic processes shape not only the world but all the spheres it contains. They study how these multidisciplinary spheres relate, interact, and change in response to natural cycles and human-driven forces. They use elements from physics, chemistry, biology, meteorology, environmental science, zoology, hydrology, and many other sciences.

Rock Cycle

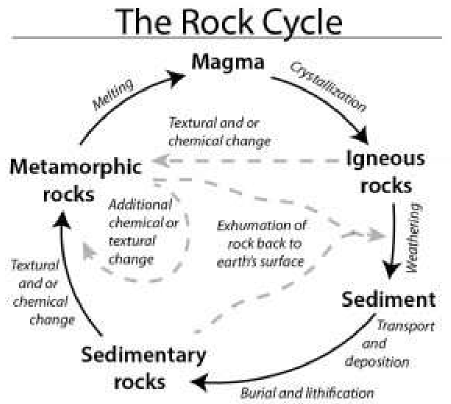


Figure 1.5.6: Rock cycle showing the five materials (such as igneous rocks and sediment) and the processes by which one changes into another (such as weathering). (Source: Peter Davis)

The most fundamental view of Earth materials is the rock cycle, which describes the major materials that comprise the Earth, the processes that form them, and how they relate to each other. It usually begins with hot molten liquid rock called magma or lava. Magma forms under the Earth's surface in the crust or mantle. Lava is a molten rock that erupts onto the Earth's surface. When magma or lava cools, it solidifies by a process called crystallization in which minerals grow within the magma or lava. The resulting rocks are igneous rocks. *Ignis* is Latin for fire.

Figure 1.5.7: Lithified raindrop impressions over wave ripples from Nova Scotia.

Igneous rocks, as well as other types of rocks, on Earth's surface was exposed to weathering and erosion, which produces sediments. Weathering is the physical and chemical breakdown of rocks into smaller fragments. Erosion is the removal of those fragments from their original location. The broken-down and transported fragments or grains are considered sediments, such as gravel, sand, silt, and clay. These sediments may be transported by streams and rivers, ocean currents, glaciers, and wind.

Sediments come to rest in a process known as a deposition. As the deposited sediments accumulate—often underwater, such as in a shallow marine environment—the older sediments get buried by the new deposits. The deposits are compacted by the weight of the overlying sediments and individual grains are cemented together by minerals in groundwater. These processes of compaction and cementation are called lithification. Lithified sediments are considered a sedimentary rock, such as sandstone and shale. Other sedimentary rocks are made by the direct chemical precipitation of minerals rather than eroded sediments and are known as chemical sedimentary rocks.

Figure 1.5.8: Migmatite, a rock which is partially molten. (Source: Peter Davis)

Pre-existing rocks may be transformed into a metamorphic rock; *meta-* means change and *-morphos* means form or shape. When rocks are subjected to extreme increases in temperature or pressure, the mineral crystals are enlarged or altered into entirely new minerals with similar chemical makeup. High temperatures and pressures occur in rocks buried deep within the Earth's crust or that come into contact with hot magma or lava. If the temperature and pressure conditions melt the rocks to create magma and lava, the rock cycle begins anew with the creation of new rocks.

Plate Tectonics and Layers of Earth

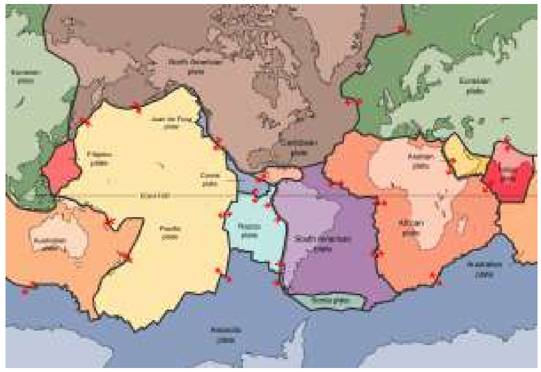


Figure 1.5.0: Map of the major plates and their motions along boundaries.

The theory of **plate tectonics** is the fundamental unifying principle of geology and the rock cycle. Plate tectonics describes how Earth's layers move relative to each other, focusing on the tectonic or lithospheric plates of the outer layer. Tectonic plates float, collide, slide past each other, and split apart on an underlying mobile layer called the **asthenosphere**. Major landforms are created at the plate boundaries, and rocks within the tectonic plates move through the rock cycle. Plate tectonics is discussed in more detail in Chapter 2.

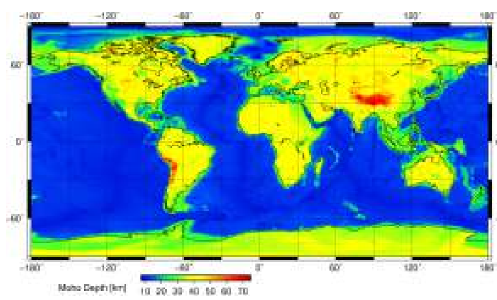


Figure 1.5.10: The global map of the thickness of the crust.

Earth's three main geological layers can be categorized by chemical composition or the chemical makeup: crust, mantle, and core. The crust is the outermost layer and composed of mostly silicon, oxygen, aluminum, iron, and magnesium [29]. There are two types: continental crust and oceanic crust. **Continental crust** is about 50 km (30 mi) thick, composed of low-density igneous and sedimentary rocks, **Oceanic crust** is approximately 10 km (6 mi) thick and made of high-density igneous basalt-type rocks. Oceanic crust makes up most of the ocean floor, covering about 70% of the planet [30]. Tectonic plates are made of crust and a portion of the upper mantle, forming a rigid physical layer called the lithosphere.

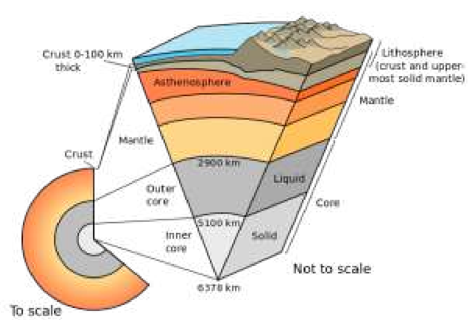


Figure 1.5.11: The layers of the Earth. Physical layers include lithosphere and asthenosphere; chemical layers are crust, mantle, and core.

The **mantle**, the largest chemical layer by volume, lies below the crust and extends down to about 2,900 km (1,800 mi) below the Earth's surface [31]. The mostly solid mantle is made of peridotite, a high-density composed of silica, iron, and magnesium [32]. The upper part of the mantel is very hot and flexible, which allows the overlying tectonic plates to float and move about on it. Under the mantle is the Earth's core, which is 3,500 km (2,200 mi) thick and made of iron and nickel. The core consists of two parts: a liquid **outer core** and solid **inner core** [33; 34; 35]. Rotations within the solid and liquid metallic core generate Earth's magnetic field (see figure above) [36; 37].

Geologic Time and Deep Time

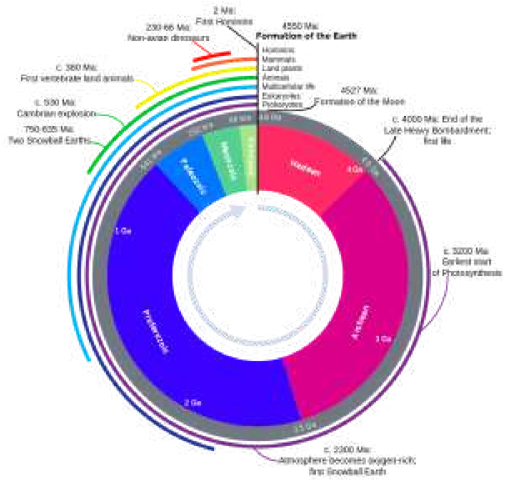


Figure 1.5.12: Geologic time on Earth, represented circularly, to show the individual time divisions and important events. Ga=billion years ago, Ma=million years ago.

> *"The result, therefore, of our present enquiry is, that we find no vestige of a beginning; no prospect of an end." (James Hutton, 1788) [22]*

One of the early pioneers of geology, James Hutton, wrote this about the age of the Earth after many years of geological study. Although he wasn't exactly correct—there is a beginning and will be an end to planet Earth—Hutton was expressing the difficulty humans have in perceiving the vastness of geological time. Hutton did not assign an age to the Earth, although he was the first to suggest the planet was very old.

Today we know Earth is approximately 4.54 ± 0.05 billion years old. This age was first calculated by Caltech professor Clair Patterson in 1956, who measured the half-lives of lead isotopes to radiometrically date a meteorite recovered in Arizona [38]. Studying geologic time, also known as deep time, can help us overcome a perspective of Earth that is limited to our short lifetimes. Compared to the geologic scale, the human lifespan is very short, and we struggle to comprehend the depth of geologic time and slowness of geologic processes. For example, the study of earthquakes only goes back about 100 years; however, there is geologic evidence of large earthquakes occurring thousands of years ago. And scientific evidence indicates earthquakes will continue for many centuries into the future.

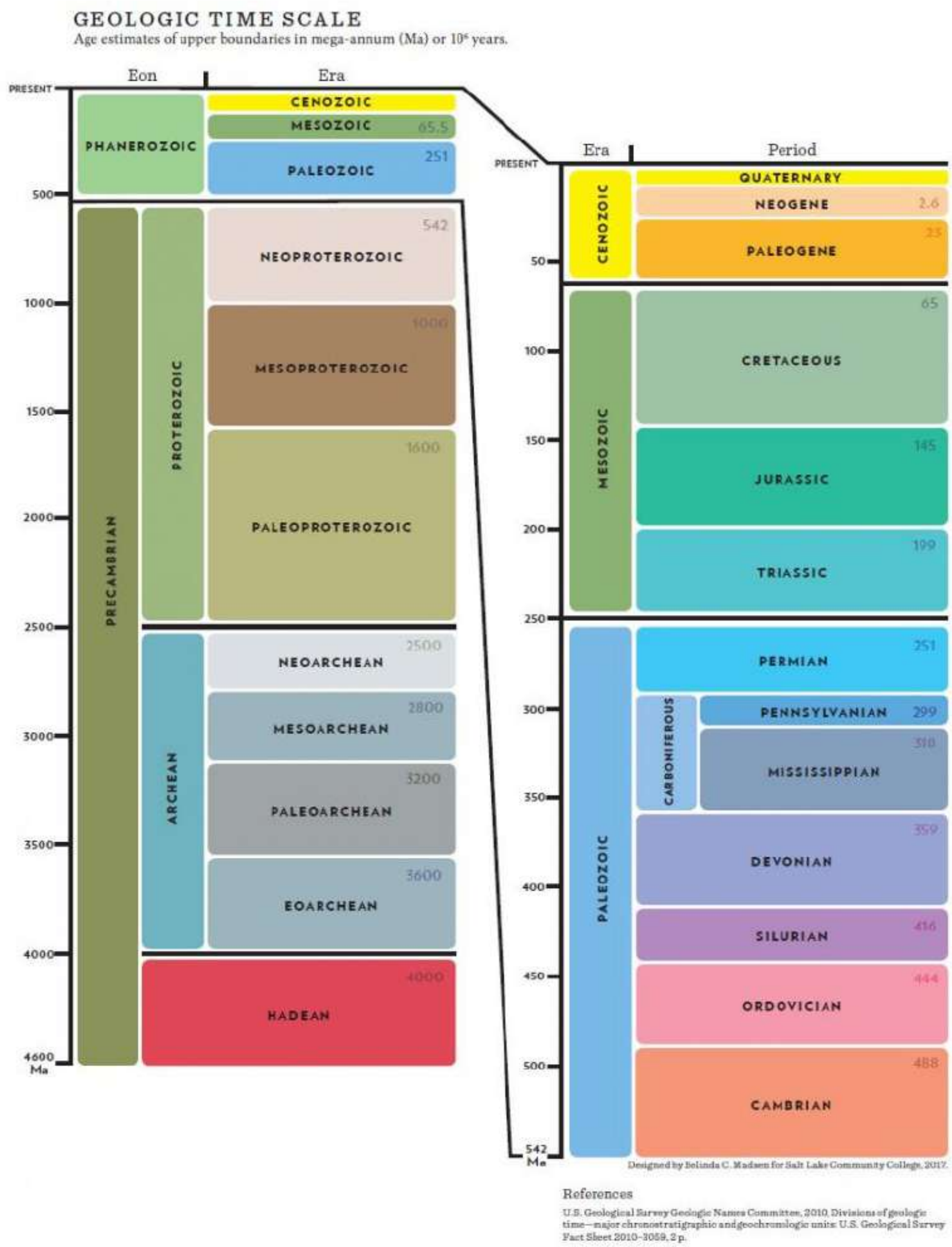


Figure 1.5.12: Geologic time scale showing time period names and ages. (Source: Belinda Madsen)

Eons are the largest divisions of time, and from oldest to youngest are named Hadean, Archean, Proterozoic, and Phanerozoic. The three oldest eons are sometimes collectively referred to as Precambrian time.

Life first appeared more than 3,800 million years ago (Ma). From 3,500 Ma to 542 Ma, or 88% of geologic time, the predominant life forms were single-celled organisms such as bacteria. More complex organisms appeared only more recently, during the current Phanerozoic Eon, which includes the last 542 million years or 12% of geologic time.

The name Phanerozoic comes from *phaneros*, which means visible, and *zoic*, meaning life. This eon marks the proliferation of multicellular animals with hard body parts, such as shells, which are preserved in the geological record as fossils. Land-dwelling animals have existed for 360 million years, or 8% of geologic time. The demise of the dinosaurs and subsequent rise of mammals occurred around 65 Ma, or 1.5% of geologic time. Our human ancestors belonging to the genus *Homo* have existed since approximately 2.2 Ma—0.05% of geological time or just $\frac{1}{2{,}000}^{th}$ the total age of Earth.

The Phanerozoic Eon is divided into three eras: Paleozoic, Mesozoic, and Cenozoic. Paleozoic means *ancient life*, and organisms of this era included invertebrate animals, fish, amphibians, and reptiles. The Mesozoic (*middle life*) is popularly known as the Age of

Reptiles and is characterized by the abundance of dinosaurs, many of which evolved into birds. The mass extinction of the dinosaurs and other apex predator reptiles marked the end of the Mesozoic and beginning of the Cenozoic. Cenozoic means *new life* and is also called the Age of Mammals, during which mammals evolved to become the predominant land-dwelling animals. Fossils of early humans, or hominids, appear in the rock record only during the last few million years of the Cenozoic. The geologic time scale, geologic time, and geologic history are discussed in more detail in chapters 7 and 8.

The Geologist's Tools

Figure 1.5.13: Iconic Archaeopteryx lithographica fossil from Germany.

In its simplest form, a geologist's tool may be a rock hammer used for sampling a fresh surface of a rock. A basic toolset for fieldwork might also include:

- Magnifying lens for looking at mineralogical details
- Compass for measuring the orientation of geologic features
- Map for documenting the local distribution of rocks and minerals
- Magnet for identifying magnetic minerals like magnetite
- Dilute solution of hydrochloric acid to identify carbonate-containing minerals like calcite or limestone.

In the laboratory, geologists use optical microscopes to closely examine rocks and soil for mineral composition and grain size. Laser and mass spectrometers precisely measure the chemical composition and geological age of minerals. Seismographs record and locate earthquake activity, or when used in conjunction with ground-penetrating radar, locate objects buried beneath the surface of the earth. Scientists apply computer simulations to turn their collected data into testable, theoretical models. Hydrogeologists drill wells to sample and analyze underground water quality and availability. Geochemists use scanning electron microscopes to analyze minerals at the atomic level, via x-rays. Other geologists use gas chromatography to analyze liquids and gases trapped in glacial ice or rocks.

Technology provides new tools for scientific observation, which leads to new evidence that helps scientists revise and even refute old ideas. Because the ultimate technology will never be discovered, the ultimate observation will never be made. And this is the beauty of science—it is ever-advancing and always discovering something new.

References

- 22. Rappaport, R. James Hutton and the History of Geology. Dennis R. Dean. *Isis* **85**, 524–525 (1994).
- 29. Hans Wedepohl, K. The composition of the continental crust. *Geochim. Cosmochim. Acta* **59**, 1217–1232 (1995).
- 30. Mooney, W. D., Laske, G. & Masters, T. G. CRUST 5.1: A global crustal model. *J. Geophys. Res. [Solid Earth]* **103**, 727–747 (1998).
- 31. Birch, F. Elasticity and constitution of the Earth's interior. *J. Geophys. Res.* **57**, 227–286 (1952).
- 32. Wyllie, P. J. Ultramafic rocks and the upper mantle. in *Fiftieth anniversary symposia: Mineralogy and petrology of the Upper Mantle; Sulfides; Mineralogy and geochemistry of non-marine evaporites* (ed. Morgan, B. A.) 3–32 (Mineralogical Society of America, 1970).
- 33. Alfe, D., Gillan, M. J. & Price, G. D. Composition and temperature of the Earth's core constrained by combining ab initio calculations and seismic data. *Earth Planet. Sci. Lett.* **195**, 91–98 (2002).
- 34. Lehmann, I. P', Publ. *Bur. Centr. Seism. Internat. Serie A* **14**, 87–115 (1936).

- 35. Engdahl, E. R., Flynn, E. A. & Masse, R. P. Differential PkiKP travel times and the radius of the core. *Geophysical J Royal Astro Soc* **40**, 457–463 (1974).
- 36. de Wijs, G. A. *et al.* The viscosity of liquid iron at the physical conditions of the Earth's core. *Nature* **392**, 805–807 (1998).
- 37. Jakosky, B. M. *et al.* MAVEN observations of the response of Mars to an interplanetary coronal mass ejection. *Science* **350**, aad0210 (2015).
- 38. Patterson, C. Age of meteorites and the earth. *Geochim. Cosmochim. Acta* **10**, 230–237 (1956).

Contributions and Attributions

1.6: Science Denial and Evaluating Sources

Figure 1.6.1: Anti-evolution league at the infamous Tennessee v. Scopes trial.

Introductory science courses usually deal with accepted scientific theory and do not include opposing ideas, even though these alternate ideas may be credible. This makes it easier for students to understand complex material. Advanced students will encounter more controversies as they continue to study their discipline.

Some groups of people argue that some established scientific theories are wrong, not based on their scientific merit but rather on the ideology of the group. This section focuses on how to identify evidence-based information and differentiate it from pseudoscience.

Science Denial

Figure 1.6.2: 2017 March for Science in Salt Lake City. This and other similar marches were in response to funding cuts and anti-science rhetoric.

Science denial happens when people argue that established scientific theories are wrong, not based on scientific merit but rather on subjective ideology—such as for social, political, or economic reasons. Organizations and people use science denial as a rhetorical argument against issues or ideas they oppose. Three examples of science denial versus science are:

1. Teaching evolution in public schools
2. Linking tobacco smoke to cancer
3. Linking human activity to climate change.

Among these, denial of climate change is strongly connected with geology. A climate denier specifically denies or doubts the objective conclusions of geologists and climate scientists.

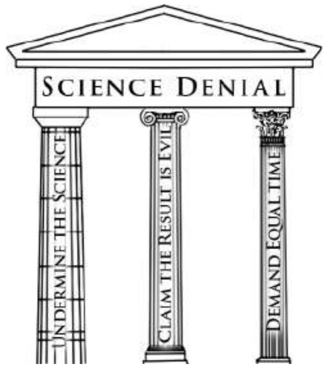


Figure 1.6.3: Three false rhetorical arguments of science denial (Source: National Center for Science Education)

Science denial generally uses three false arguments. The first argument tries to undermine the credibility of the scientific conclusion by claiming the research methods are flawed or the theory is not universally accepted—the science is unsettled. The notion that scientific ideas are not absolute creates doubt for non-scientists; however, a lack of universal truths should not be confused with scientific uncertainty. Because science is based on falsifiability, scientists avoid claiming universal truths and use language that conveys uncertainty. This allows scientific ideas to change and evolve as more evidence is uncovered.

The second argument claims the researchers are not objective and motivated by ideology or economic agenda. This is an *ad hominem* argument in which a person's character is attacked instead of the merit of their argument. They claim results have been manipulated so researchers can justify asking for more funding. They claim that because the researchers are funded by a federal grant, they are using their results to lobby for expanded government regulation.

The third argument is to demand a balanced view, equal time in media coverage and educational curricula, to engender the false illusion of two equally valid arguments. Science deniers frequently demand equal coverage of their proposals, even when there is little scientific evidence supporting their ideology. For example, science deniers might demand religious explanations to be taught as an alternative to the well-established theory of evolution [39; 40]. Or that all possible causes of climate change be discussed as equally probable, regardless of the body of evidence. Conclusions derived using the scientific method should not be confused with those based on ideologies.

Furthermore, conclusions about nature derived from ideologies have no place in science research and education. For example, it would be inappropriate to teach the flat earth model in a modern geology course because this idea has been disproved by the scientific method. Unfortunately, widespread scientific illiteracy allows these arguments to be used to suppress scientific knowledge and spread misinformation.

The formation of new conclusions based on the scientific method is the only way to change scientific conclusions. We wouldn't teach Flat Earth geology along with plate tectonics because Flat Earthers don't follow the scientific method. The fact that scientists avoid universal truths and change their ideas as more evidence is uncovered shouldn't be seen as meaning that the science is unsettled. Because of widespread scientific illiteracy, these arguments are used by those who wish to suppress science and misinform the general public.

Figure 1.6.4: The lag time between cancer after smoking, plus the ethics of running human trials, delayed the government in taking action against tobacco.

In a classic case of science denial, beginning in the 1960s and for the next three decades, the tobacco industry and their scientists used rhetorical arguments to deny a connection between tobacco usage and cancer. Once it became clear scientific studies overwhelmingly found that using tobacco dramatically increased a person's likelihood of getting cancer, their next strategy was to create a sense of doubt about the science. The tobacco industry suggested the results were not yet fully understood and more study was needed. They used this doubt to lobby for delaying legislative action that would warn consumers of the potential health hazards [39, 41]. This same tactic is currently being employed by those who deny the significance of human involvement in climate change.

Evaluating Sources of Information

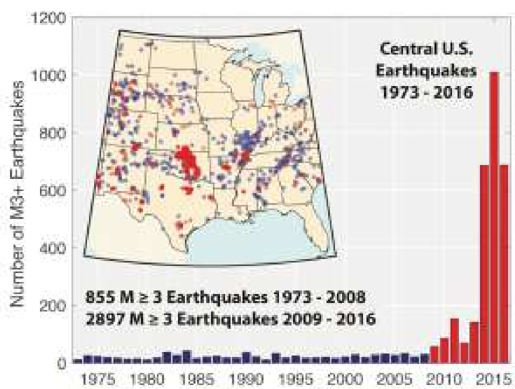


Figure 1.6.5: This graph shows the earthquake data. To call this data induced, due to fracking/drilling, would be an interpretation. Since the USGS (a reputable institution) has interpreted these earthquakes are caused by humans, it is a more reliable interpretation.

In the age of the internet, information is plentiful. Geologists, scientists, or anyone exploring scientific inquiry must discern valid sources of information from pseudoscience and misinformation. This evaluation is especially important in scientific research because scientific knowledge is respected for its reliability [42]. Textbooks such as this one can aid this complex and crucial task. At its roots, quality information comes from the scientific method [43], beginning with the empirical thinking of Aristotle. The application of the scientific method helps produce unbiased results. A valid inference or interpretation is based on objective evidence or data. Credible data and inferences are clearly labeled, separated, and differentiated. Anyone looking over the data can understand how the author's conclusion was derived or come to an alternative conclusion. Scientific procedures are clearly defined so the investigation can be replicated to confirm the original results or expanded further to produce new results. These measures make a scientific inquiry valid and its use as a source reputable. Of course, substandard work occasionally slips through and retractions are published from time to time. An infamous article linking the MMR vaccine to autism appeared in the highly reputable journal *Lancet* in 1998. Journalists discovered the author had multiple conflicts of interest and fabricated data, and the article was retracted in 2010.

Figure 1.6.6: Logo for The Geological Society of America, one of the leading geoscience organizations. They also publish GSA Bulletin, a reputable geology journal.

In addition to methodology, data, and results, the authors of a study should be investigated. When looking into any research, the author(s) should be investigated [44]. An author's credibility is based on multiple factors, such as having a degree in a relevant topic or being funded from an unbiased source.

The same rigor should be applied to evaluating the publisher, ensuring the results reported come from an unbiased process [45]. The publisher should be easy to discover. Good publishers will show the latest papers in the journal and make their contact information and identification clear. Reputable journals show their peer review style. Some journals are predatory, where they use unexplained and unnecessary fees to submit and access journals. Reputable journals have recognizable editorial boards. Often, a reliable journal will associate with a trade, association, or recognized open-source initiative.

One of the hallmarks of scientific research is peer review. Research should be transparent to peer review. This allows the scientific community to reproduce experimental results, correct and retract errors, and validate theories. This allows the reproduction of experimental results, corrections of errors, and proper justification of the research to experts.

Citation is not only imperative to avoid plagiarism, but also allows readers to investigate an author's line of thought and conclusions. When reading scientific works, it is important to confirm the citations are from reputable scientific research. Most often, scientific citations are used to reference paraphrasing rather than quotes. The number of times a work is cited is said to measure the influence an investigation has within the scientific community, although this technique is inherently biased [46].

References

- 39. National Center for Science Education. Climate change denial. (2016). Available at: http://ncse.com/climate/denial. (Accessed: 6th April 2016)
- 40. Dreifus, C. Naomi Oreskes Imagines the Future History of Climate Change. *The New York Times* (2014).
- 41. Oreskes, N., Conway, E. & Cain, S. *Merchants of doubt: how a handful of scientists obscured the truth on issues from tobacco smoke to global warming*. (Bloomsbury Press, 2010).
- 42. Bocking, S. *Nature's experts: science, politics, and the environment*. (Rutgers University Press, 2004).
- 43. Alkin, M. C. *Evaluation Roots: Tracing theorists' views and influences*. (SAGE, 2004).
- 44. Krimsky, S. Do financial conflicts of interest bias research? An inquiry into the "funding effect" hypothesis. *Sci. Technol. Human Values* **38**, 566–587 (2013).
- 45. Spier, R. The history of the peer-review process. *Trends Biotechnol.* **20**, 357–358 (2002).
- 46. Moustafa, K. Aberration of the Citation. *Account. Res.* **23**, 230–244 (2016).

Contributions and Attributions

CHAPTER OVERVIEW

2: PLATE TECTONICS

Learning Objectives

At the end of this chapter, students should be able to:

Describe how the ideas behind plate tectonics started with Alfred Wegener's hypothesis of continental drift
Describe the physical and chemical layers of the Earth and how they affect plate movement
Explain how movement at the three types of plate boundaries causes earthquakes, volcanoes, and mountain building
Identify convergent boundaries, including subduction and collisions, as places where plates come together
Identify divergent boundaries, including rifts and mid-ocean ridges, as places where plates separate
Explain transform boundaries as places where adjacent plates shear past each other
Describe the Wilson Cycle, beginning with continental rifting, ocean basin creation, plate subduction, and ending with ocean basin closure
Explain how the tracks of hotspots, places that have continually rising magma, is used to calculate plate motion

Revolution is a word usually reserved for significant political or social changes. Several of these idea revolutions forced scientists to re-examine their entire field, triggering a paradigm shift that shook up their conventionally held knowledge. Charles Darwin's book on evolution, *On the Origin of Species*, published in 1859; Gregor Mendel's discovery of the genetic principles of inheritance in 1866; and James Watson, Francis Crick, and Rosalind Franklin's model for the structure of DNA in 1953 did that for biology. Albert Einstein's relativity and quantum mechanics concepts in the early twentieth century did the same for Newtonian physics.

The concept of plate tectonics was just as revolutionary for geology. The theory of plate tectonics attributes the movement of massive sections of the Earth's outer layers with creating earthquakes, mountains, and volcanoes. Many earth processes make more sense when viewed through the lens of plate tectonics. Because it is so important in understanding how the world works, plate tectonics is the first topic of discussion in this textbook.

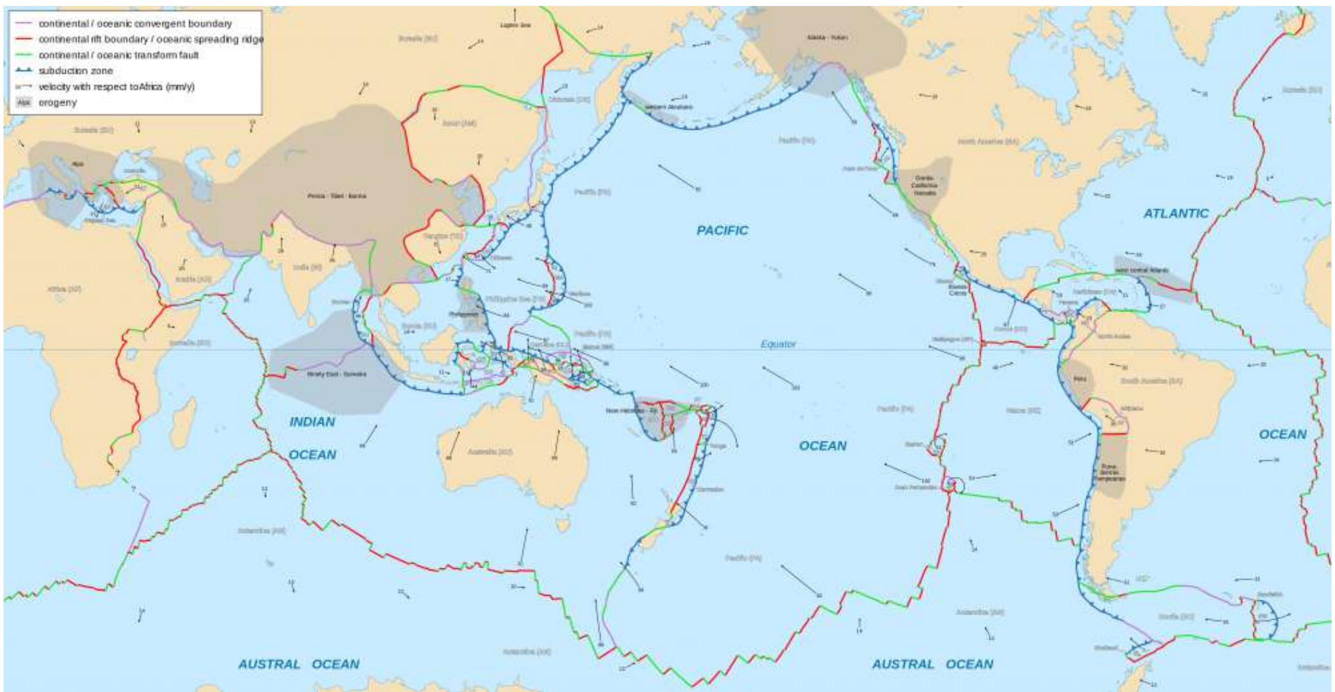


Figure \PageIndex1: Detailed map of all known plates, their boundaries, and movements.

2.1: ALFRED WEGENER'S CONTINENTAL DRIFT HYPOTHESIS

Alfred Wegener (1880-1930) was a German scientist who specialized in meteorology and climatology. His knack for questioning accepted ideas started in 1910 when he disagreed with the explanation that the Bering Land Bridge was formed by isostasy and that similar land bridges once connected the continents. After reviewing the scientific literature, he published a hypothesis stating the continents were originally connected and then drifted apart.

2.2: LAYERS OF THE EARTH

In order to understand the details of plate tectonics, it is essential to first understand the layers of the earth. Firsthand information about what is below the surface is very limited; most of what we know is pieced together from hypothetical models, and analyzing seismic wave data and meteorite materials. In general, the Earth can be divided into layers based on chemical composition and physical characteristics.

2.3: CONVERGENT BOUNDARIES

Convergent boundaries, also called destructive boundaries, are places where two or more plates move toward each other. Convergent boundary movement is divided into two types, subduction and collision, depending on the density of the involved plates. Continental lithosphere is of lower density and thus more buoyant than the underlying asthenosphere. Oceanic lithosphere is denser than continental lithosphere, and, when old and cold, may even be denser than asthenosphere.

2.4: DIVERGENT BOUNDARIES

At divergent boundaries, sometimes called constructive boundaries, lithospheric plates move away from each other. There are two types of divergent boundaries, categorized by where they occur: continental rift zones and mid-ocean ridges. Continental rift zones occur in weak spots in the continental lithospheric plate. A mid-ocean ridge usually originates in a continental plate as a rift zone that expands to the point of splitting the plate apart, with seawater filling in the gap.

2.5: TRANSFORM BOUDARIES

A transform boundary, sometimes called a strike-slip or conservative boundary, is where the lithospheric plates slide past each other in the horizontal plane. This movement is described based on the perspective of an observer standing on one of the plates, looking across the boundary at the opposing plate. Dextral, also known as right-lateral, movement describes the opposing plate moving to the right. Sinistral movement describes the opposing plate moving to the left.

2.6: THE WILSON CYCLE

The Wilson Cycle is named for J. Tuzo Wilson who first described it in 1966, and it outlines the ongoing origin and breakup of supercontinents, such as Pangea and Rodinia. Scientists have determined this cycle has been operating for at least three billion years and possibly earlier.

2.7: HOTSPOTS

The Wilson Cycle provides a broad overview of the tectonic plate movement. To analyze plate movement more precisely, scientists study hotspots. First postulated by J. Tuzo Wilson in 1963, a hotspot is an area in the lithospheric plate where molten magma breaks through and creates a volcanic center, islands in the ocean and mountains on land.

2.1: Alfred Wegener’s Continental Drift Hypothesis

Alfred Wegener (1880-1930) was a German scientist who specialized in meteorology and climatology. His knack for questioning accepted ideas started in 1910 when he disagreed with the explanation that the Bering Land Bridge was formed by isostasy and that similar land bridges once connected the continents [1]. After reviewing the scientific literature, he published a hypothesis stating the continents were originally connected and then drifted apart. While he did not have the precise mechanism worked out, his hypothesis was backed up by a long list of evidence.

Figure 2.1.1: Wegener later in his life, ca. 1924-1930.

Early Evidence for Continental Drift Hypothesis

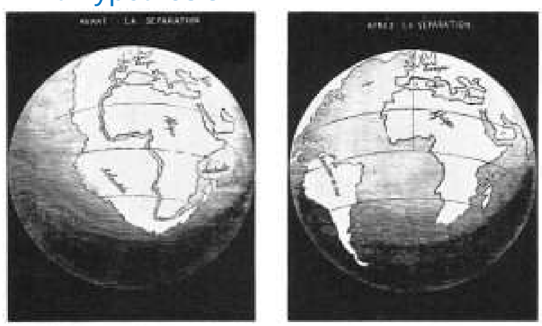


Figure 2.1.2: Snider-Pellegrini’s map showing the continental fit and separation, 1858.

Wegener’s first piece of evidence was that the coastlines of some continents fit together like pieces of a jigsaw puzzle. People noticed the similarities in the coastlines of South America and Africa on the first world maps, and some suggested the continents had been ripped apart [3]. Antonio Snider-Pellegrini did preliminary work on continental separation and matching fossils in 1858.

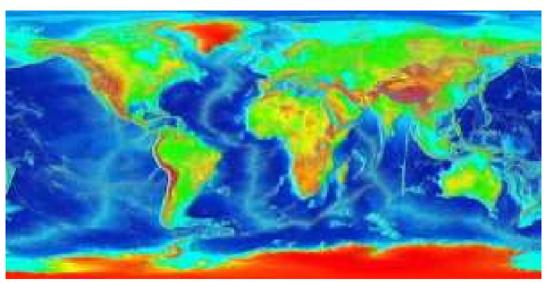


Figure 2.1.3: Map of world elevations. Note the light blue, which are continental shelves flooded by shallow ocean water. These show the true shapes of the continents.

What Wegener did differently was synthesizing a large amount of data in one place. He used the true edges of the continents, based on the shapes of the continental shelves [4]. This resulted in a better fit than previous efforts that traced the existing coastlines [5].

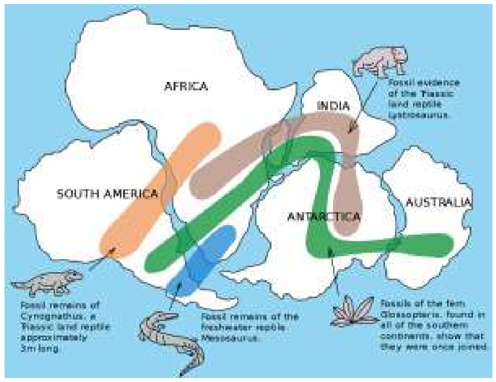


Figure 2.1.4: Image showing fossils that connect the continents of Gondwana (the southern continents of Pangea).

Wegener also compiled evidence by comparing similar rocks, mountains, fossils, and glacial formations across oceans. For example, the fossils of the primitive aquatic reptile *Mesosaurus* were found on the separate coastlines of Africa and South America. Fossils of another reptile, *Lystrosaurus*, were found in Africa, India, and Antarctica. He pointed out these were land-dwelling creatures could not have swum across an entire ocean.

Opponents of continental drift insisted trans-oceanic land bridges allowed animals and plants to move between continents [6]. The land bridges eventually eroded away, leaving the continents permanently separated. The problem with this hypothesis is the improbability of a land bridge being tall and long enough to stretch across a broad, deep ocean.

More support for continental drift came from the puzzling evidence that glaciers once existed in normally very warm areas in southern Africa, India, Australia, and Arabia. These climate anomalies could not be explained by land bridges. Wegener found similar evidence when he discovered tropical plant fossils in the frozen region of the Arctic Circle. As Wegener collected more data, he realized the explanation that best fit all the climate, rock, and fossil observations involved moving continents.

Proposed Mechanism for Continental Drift

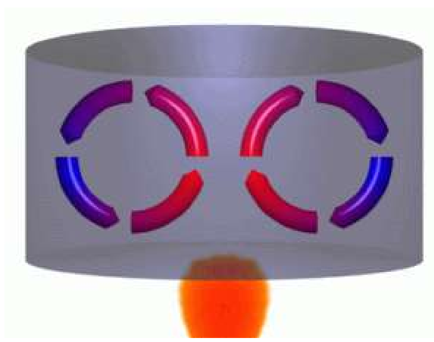


Figure 2.1.5: [Click to Animate] Animation of the basic idea of convection: an uneven heat source in a fluid causes rising material next to the heat and sinking material far from the heat.

Wegener's work was considered a fringe science theory for his entire life. One of the biggest flaws in his hypothesis was the inability to provide a mechanism for how the continents moved. Obviously, the continents did not appear to move, and changing the conservative minds of the scientific community would require exceptional evidence that supported a credible mechanism. Other pro-continental drift followers used expansion, contraction, or even the moon's origin to explain how the continents moved. Wegener used centrifugal forces and precession, but this model was proven wrong [7]. He also speculated about seafloor spreading, with hints of convection, but could not substantiate these proposals [8]. As it turns out, current scientific knowledge reveals convection is the major force in driving plate movements.

Development of Plate Tectonic Theory

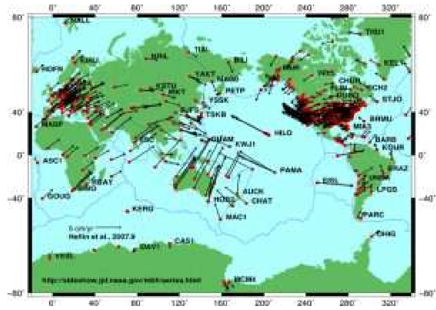


Figure 2.1.6: GPS measurements of plate motions.

Wegener died in 1930 on an expedition in Greenland. Poorly respected in his lifetime, Wegener and his ideas about moving continents seemed destined to be lost in history as fringe science. However, in the 1950s, evidence started to trickle in that made continental drift a more viable idea. By the 1960s, scientists had amassed enough evidence to support the missing mechanism—namely, seafloor spreading—for Wegener's hypothesis of continental drift to be accepted as the theory of plate tectonics. Ongoing GPS and earthquake data analyses continue to support this theory. The next section provides the pieces of evidence that helped transform one man's wild notion into a scientific theory.

Mapping of the Ocean Floors

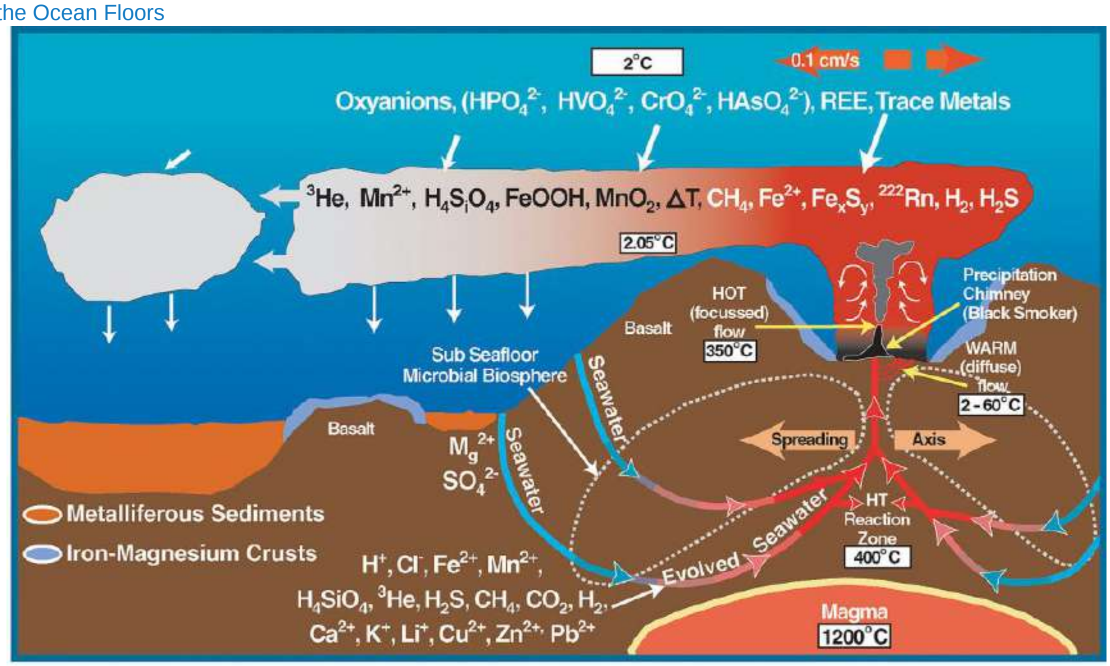
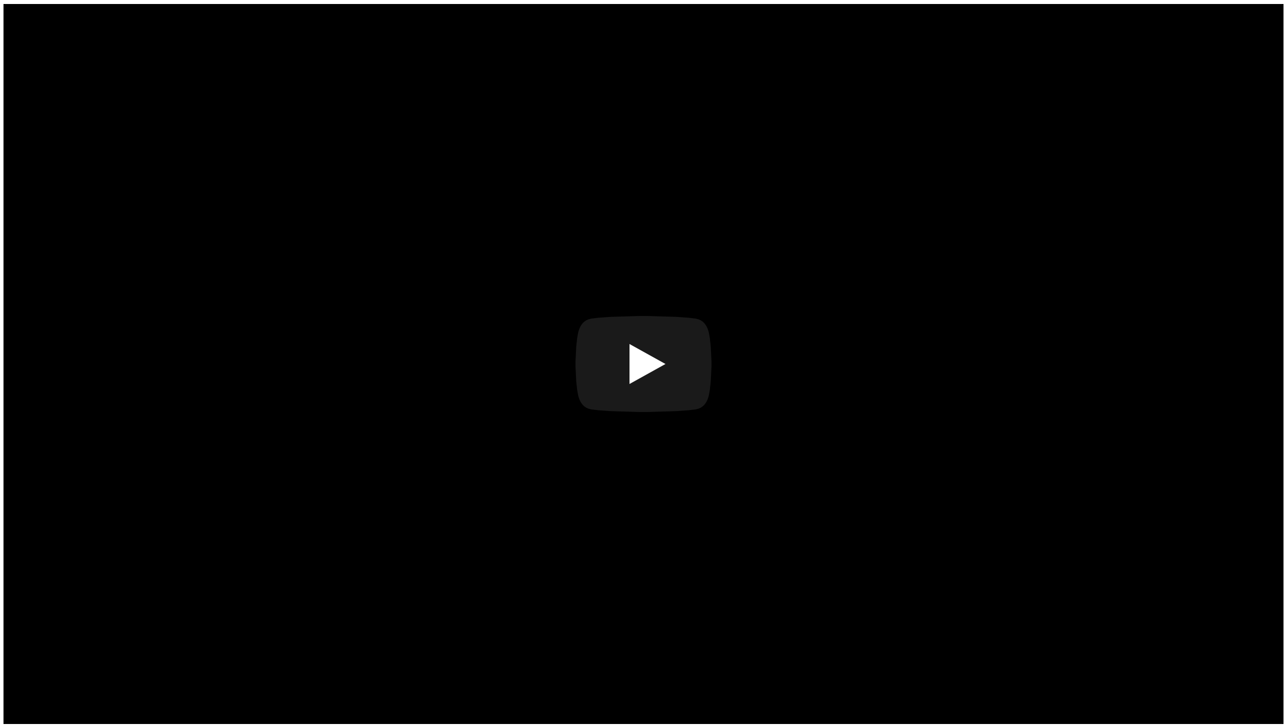


Figure 2.1.7: The complex chemistry around mid-ocean ridges.

In 1947 researchers started using an adaptation of SONAR to map a region in the middle of the Atlantic Ocean with poorly-understood topographic and thermal properties [9]. Using this information, Bruce Heezen and Marie Tharp created the first detailed map of the ocean floor to reveal the Mid-Atlantic Ridge [10], a basaltic mountain range that spanned the length of the Atlantic Ocean, with rock chemistry and dimensions unlike the mountains found on the continents. Initially, scientists thought the ridge was part of a mechanism that explained the expanding Earth or ocean-basin growth hypotheses [11; 12]. In 1959, Harry Hess proposed the hypothesis of seafloor spreading – that the mid-ocean ridges represented tectonic plate factories, where a new oceanic plate was issuing from these long volcanic ridges. Scientists later included transform faults perpendicular to the ridges to better account for varying rates of movement between the newly formed plates [13]. When earthquake epicenters were discovered along the ridges, the idea that earthquakes were linked to plate movement took hold [14].



Seafloor sediment, measured by dredging and drilling, provided another clue. Scientists once believed sediment accumulated on the ocean floors over a very long time in a static environment. When some studies showed less sediment than expected, these results were initially used to argue against the continental movement [15; 16]. With more time, researchers discovered these thinner sediment layers were located close to mid-ocean ridges, indicating the ridges were younger than the surrounding ocean floor. This finding supported the idea that the seafloor was not fixed in one place [17].

Paleomagnetism

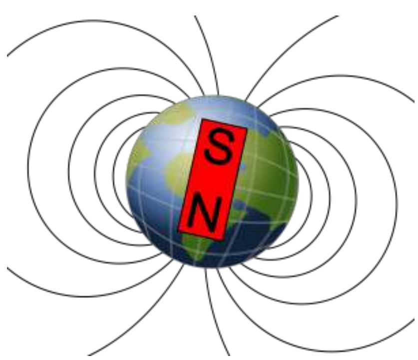


Figure 2.1.8: The magnetic field of Earth, simplified as a bar magnet.

The seafloor was also mapped magnetically. Scientists had long known of strange magnetic anomalies that formed a striped pattern of symmetrical rows on both sides of mid-oceanic ridges. What made these features unusual was the north and south magnetic poles within each stripe was reversed in alternating rows [18]. By 1963, Harry Hess and other scientists used these magnetic reversal patterns to support their model for seafloor spreading [19] (see also Lawrence W. Morley [20]).

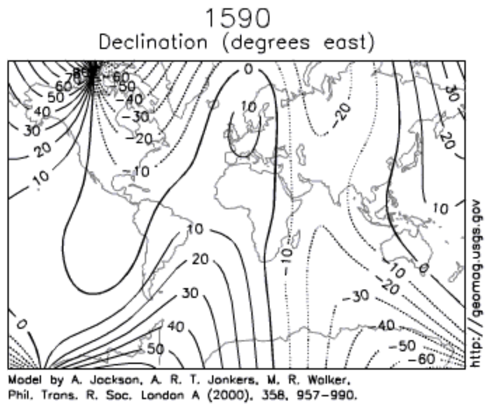


Figure 2.1.9: This animation shows how the magnetic poles have moved over 400 years.

Paleomagnetism is the study of magnetic fields frozen within rocks, basically a fossilized compass. In fact, the first hard evidence to support plate motion came from paleomagnetism.

Igneous rocks containing magnetic minerals like magnetite typically provide the most useful data. In their liquid state as magma or lava, the magnetic poles of the minerals align themselves with the Earth's magnetic field. When the rock cools and solidifies, this alignment is frozen into place, creating a permanent paleomagnetic record that includes magnetic inclination related to global latitude, and declination related to magnetic north.

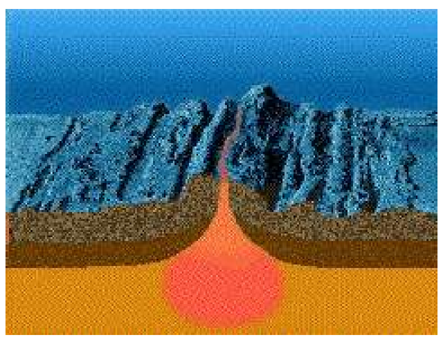


Figure 2.1.10: The iron in the solidifying rock preserves the current magnetic polarity as new oceanic plates form at mid-ocean ridges.

Scientists had noticed for some time the alignment of magnetic north in many rocks was nowhere close to the earth's current magnetic north. Some explained this as part of the normal movement of earth magnetic north pole. Eventually, scientists realized adding the idea of continental movement explained the data better than the pole movement alone [21].

Wadati-Benioff Zones

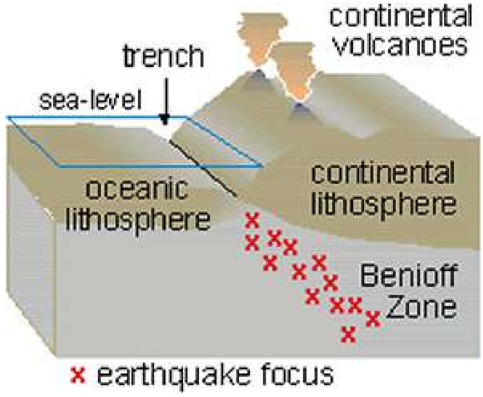


Figure 2.1.11: The Wadati-Benioff zone, showing earthquakes following the subducting slab down.

Around the same time mid-ocean ridges were being investigated, other scientists linked the creation of ocean trenches and island arcs to seismic activity and tectonic plate movement [22]. Several independent research groups recognized earthquake epicenters traced the shapes of oceanic plates sinking into the mantle. These deep earthquake zones congregated in planes that started near the surface around ocean trenches and angled beneath the continents and island arcs [23]. Today these earthquake zones called Wadati-Benioff zones.

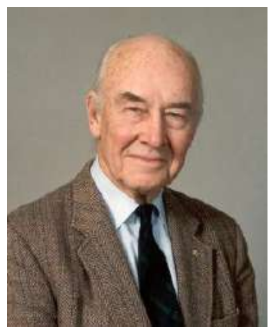


Figure 2.1.12: J. Tuzo Wilson

Based on the mounting evidence, the theory plate tectonics continued to take shape. J. Tuzo Wilson was the first scientist to put the entire picture together by proposing that the opening and closing of the ocean basins [24]. Before long, scientists proposed other models showing plates moving with respect to each other, with clear boundaries between them [25]. Others started piecing together complicated histories of tectonic plate movement [26]. The plate tectonic revolution had taken hold.

References

- 1. Fluegel, von H. W. Wegener-Ampferer-Schwinner. Ein Beitrag zur Geschichte der Geologie in Österreich. *Mitt. Oesterr. Geol. Ges.* **73**, 237–254 (1980).
- 3. Bacon, F. & Montagu, B. *The Works of Francis Bacon, Lord Chancellor of England: With a Life of the Author*. (Parry & McMillan, 1848).
- 4. Drake, E. T. Alfred Wegener's reconstruction of Pangea. *Geology* **4**, 41–44 (1976).
- 5. Mantovani, R. Les fractures de l'écorce terrestre et la théorie de Laplace. *Bull. Soc. Sc. et Arts Réunion* 41–53 (1889).
- 6. Wells, H. G., Huxley, J. & Wells, G. P. The Science of Life. *Philosophy* **6**, 506–507 (1931).
- 7. Scheidegger, A. E. Examination of the physics of theories of orogenesis. *Geol. Soc. Am. Bull.* **64**, 127–150 (1953).
- 8. Jacoby, W. R. Modern concepts of Earth dynamics anticipated by Alfred Wegener in 1912. *Geology* **9**, 25–27 (1981).
- 9. Tolstoy, I. & Ewing, M. North Atlantic hydrography and the Mid-Atlantic Ridge. *Geol. Soc. Am. Bull.* **60**, 1527–1540 (1949).
- 10. Heezen, B. C., Tharp, M. & Ewing, M. The Floors of the Oceans I. The North Atlantic. *Geological Society of America Special Papers* **65**, 1–126 (1959).
- 11. Heezen, B. C. The Rift in the Ocean Floor. *Sci. Am.* **203**, 98–110 (1960).
- 12. Dietz, R. S. Continent and ocean basin evolution by spreading of the seafloor. *Nature* **190**, 854–857 (1961).
- 13. Wilson, J. T. A new class of faults and their bearing on continental drift. *Nature* (1965).
- 14. Heezen, B. C. & Tharp, M. Tectonic Fabric of the Atlantic and Indian Oceans and Continental Drift. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **258**, 90–106 (1965).

- 15. Ewing, M., Ewing, J. I. & Talwani, M. Sediment distribution in the oceans: The Mid-Atlantic Ridge. *Geol. Soc. Am. Bull.* **75**, 17–36 (1964).
- 16. Saito, T., Ewing, M. & Burckle, L. H. Tertiary sediment from the mid-atlantic ridge. *Science* **151**, 1075–1079 (1966).
- 17. Ewing, M., Houtz, R. & Ewing, J. South Pacific sediment distribution. *J. Geophys. Res.* **74**, 2477–2493 (1969).
- 18. Mason, R. G. A magnetic survey off the west coast of the United-States between latitudes 32-degrees-N and 36-degrees-N longitudes 121-degrees-W and 128-degrees-W. *Geophysical Journal of the Royal Astronomical Society* **1**, 320 (1958).
- 19. Vine, F. J. & Matthews, D. H. Magnetic anomalies over oceanic ridges. *Nature* **199**, 947–949 (1963).
- 20. Frankel, H. The Development, Reception, and Acceptance of the Vine-Matthews-Morley Hypothesis. *Hist. Stud. Phys. Biol. Sci.* **13**, 1–39 (1982).
- 21. Irving, E. Palaeomagnetic and palaeoclimatological aspects of polar wandering. *Geofis. pura appl.* **33**, 23–41 (1956).
- 22. Coats, R. R. Magma type and crustal structure in the Aleutian Arc. in *The Crust of the Pacific Basin* 92–109 (American Geophysical Union, 1962). doi:10.1029/GM006p0092
- 23. Wadati, K. On the activity of deep-focus earthquakes in the Japan Islands and neighbourhoods. *Geophys. Mag.* **8**, 305–325 (1935).
- 24. Wilson, J. T. *Did the Atlantic close and then re-open?* (Nature, 1966).
- 25. McKenzie, D. P. & Parker, R. L. The North Pacific: an Example of Tectonics on a Sphere. *Nature* **216**, 1276–1280 (1967).
- 26. Atwater, T. Implications of Plate Tectonics for the Cenozoic Tectonic Evolution of Western North America. *Geol. Soc. Am. Bull.* **81**, 3513–3536 (1970).

Contributions and Attributions

2.2: Layers of the Earth

In order to understand the details of plate tectonics, it is essential to first understand the layers of the earth. Firsthand information about what is below the surface is very limited; most of what we know is pieced together from hypothetical models, and analyzing seismic wave data and meteorite materials. In general, the Earth can be divided into layers based on chemical composition and physical characteristics.

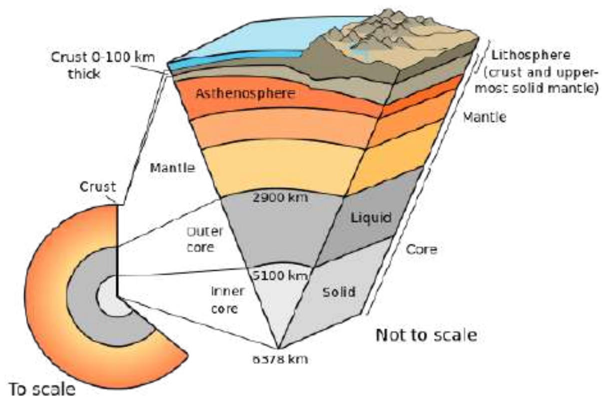


Figure 2.2.1: The layers of the Earth. Physical layers include the lithosphere and asthenosphere; chemical layers are crust, mantle, and core.

Chemical Layers

Certainly, the earth is composed of countless combinations of elements. Regardless of what elements are involved two major factors—temperature and pressure—are responsible for creating three distinct chemical layers.

Crust

The outermost chemical layer and the one we currently reside on is the crust. There are two types of crust. Continental crust has a relatively low density and composition similar to granite. Oceanic crust has a relatively high density, especially when cold and old, and composition similar to basalt. The surface levels of crust are relatively brittle. The deeper parts of the crust are subjected to higher temperatures and pressure, which makes them more ductile. Ductile materials are like soft plastics or putty, they move under force. Brittle materials are like solid glass or pottery, they break under force, especially when it is applied quickly. Earthquakes, generally occur in the upper crust and are caused by the rapid movement of relatively brittle materials.

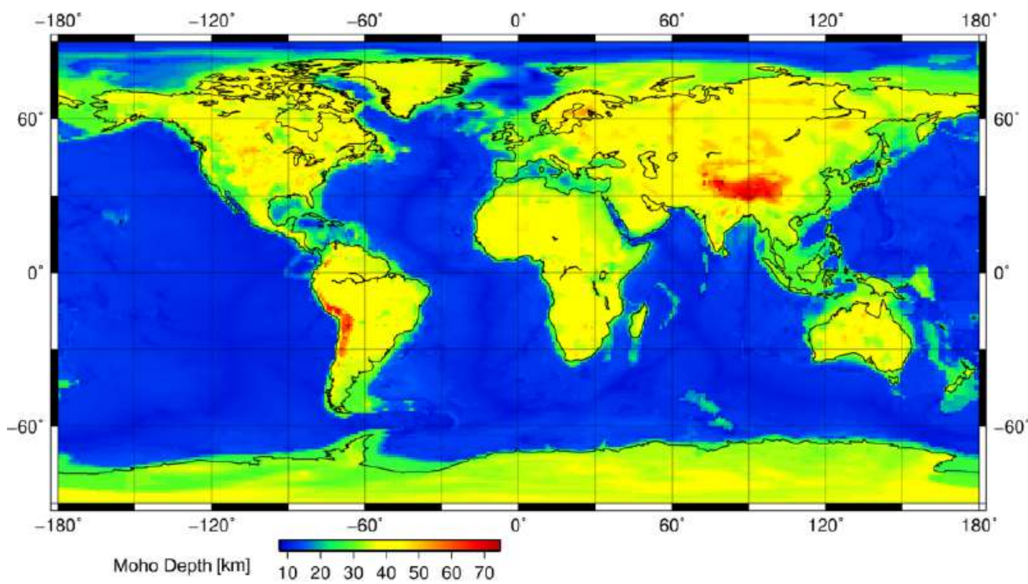


Figure 2.2.2: The global map of the depth of the moho.

The base of the crust is characterized by a large increase in seismic velocity, which measures how fast earthquake waves travel through solid matter. Called the Mohorovičić Discontinuity, or Moho for short, this zone was discovered by Andrija Mohorovičić (pronounced mo-ho-ro-vee-cheech; audio pronunciation) in 1909 after studying earthquake wave paths in his native Croatia [27]. The change in wave direction and speed is caused by dramatic chemical differences between the crust and mantle. Underneath the oceans, the Moho is found roughly 5 km below the ocean floor. Under the continents, it is located about 30-40 km below the surface. Near certain large mountain-building events known as orogenies, the continental Moho depth is doubled [28].

Mantle

Figure 2.2.3: This mantle peridotite xenolith containing olivine (green) is chemically weathering by hydrolysis and oxidation into the brown pseudo-mineral iddingsite, which is a complex of water, clay, and iron oxides. The more altered side of the rock has been exposed to the environment longer.

The mantle sits below the crust and above the core. It is the largest chemical layer by volume, extending from the base of the crust to a depth of about 2900 km [29]. Most of what we know about the mantle comes from seismic wave analysis, though the information is gathered by studying ophiolites and xenoliths. Ophiolites are pieces of the mantle that have risen through the crust until they are exposed as part of the ocean floor. Xenoliths are carried within magma and brought to the Earth's surface by volcanic eruptions. Most xenoliths are made of peridotite, an ultramafic class of igneous rock (see Chapter 4 for explanation). Because of this, scientists hypothesize most of the mantle is made of peridotite [30].

Core

Figure 2.2.4: A polished fragment of the iron-rich Toluca Meteorite, with octahedral Widmanstätten Pattern.

The core of the Earth, which has both liquid and solid layers, and consists mostly of iron, nickel, and possibly some oxygen [31]. Scientists looking at seismic data first discovered this innermost chemical layer in 1906 [32]. Through a union of hypothetical modeling, astronomical insight, and hard seismic data, they concluded the core is mostly metallic iron [33]. Scientists studying meteorites, which typically contain more iron than surface rocks, have proposed the earth was formed from meteoric material. They believe the liquid component of the core was created as the iron and nickel sank into the center of the planet, where it was liquefied by intense pressure [34].

Physical Layers

The Earth can also be broken down into five distinct physical layers based on how each layer responds to stress. While there is some overlap in the chemical and physical designations of layers, specifically the core-mantle boundary, there are significant differences between the two systems.

Lithosphere

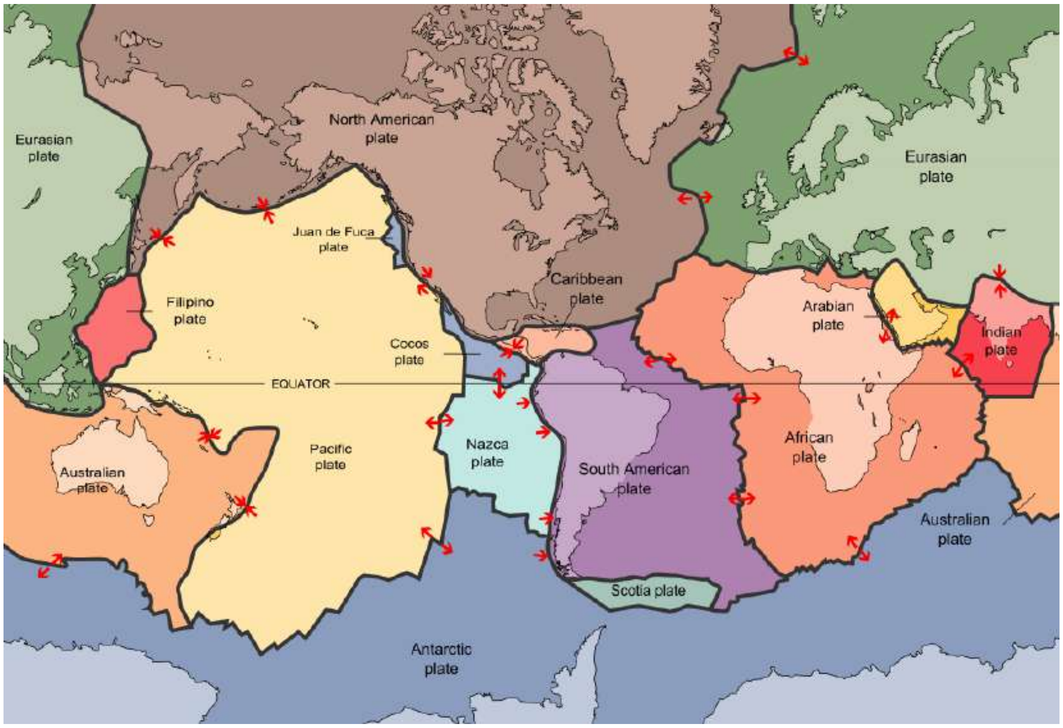


Figure 2.2.5: Map of the major plates and their motions along boundaries.

Lithos is Greek for stone, and the lithosphere is the outermost physical layer of the Earth. It is grouped into two types: oceanic and continental. Oceanic lithosphere is thin and relatively rigid. It ranges in thickness from nearly zero in new plates found around mid-ocean ridges, to an average of 140 km in most other locations. Continental lithosphere is generally thicker and considerably more plastic, especially at the deeper levels. Its thickness ranges from 40 to 280 km [35]. The lithosphere is not continuous. It is broken into segments called plates. A plate boundary is where two plates meet and move relative to each other. Plate boundaries are where we see plate tectonics in action—mountain building, triggering earthquakes, and generating volcanic activity.

Asthenosphere

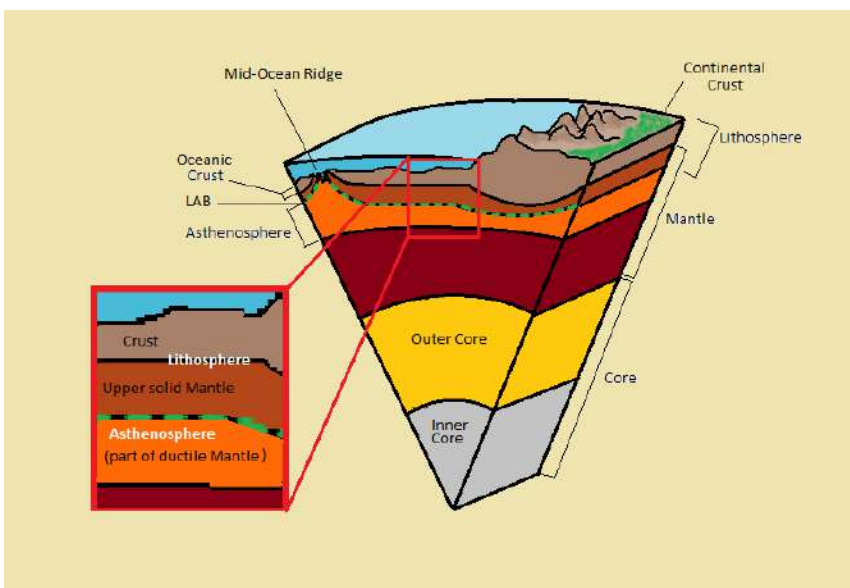


Figure 2.2.5: The lithosphere-asthenosphere boundary changes with certain tectonic situations.

The asthenosphere is the layer below the lithosphere. *Astheno-* means lacking strength, and the most distinctive property of the asthenosphere is movement. Because it is mechanically weak, this layer moves and flows due to convection currents created by heat coming from the earth's core cause [33]. Unlike the lithosphere that consists of multiple plates, the asthenosphere is relatively unbroken. Scientists have determined this by analyzing seismic waves that pass through the layer. The depth at which the asthenosphere is found is temperature-dependent [36]. It tends to lie closer to the earth's surface around mid-ocean ridges and much deeper underneath mountains and the centers of lithospheric plates.

Mesosphere

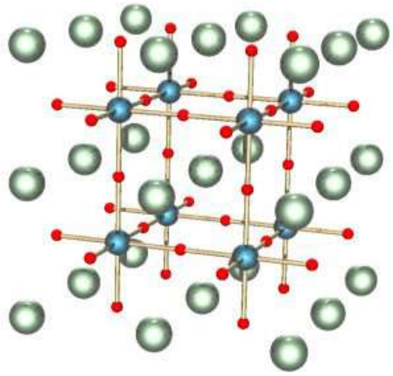


Figure 2.2.6: General perovskite structure. Perovskite silicates (e.g. Bridgmenite, $(Mg, Fe)SiO_3$) are thought to be the main component of the lower mantle, making it the most common mineral in or on Earth.

The mesosphere, sometimes known as the lower mantle, is more rigid and immobile than the asthenosphere. Located at a depth of approximately 410 and 660 km below the earth's surface, the mesosphere is subjected to very high pressures and temperatures. These extreme conditions create a transition zone in the upper mesosphere where minerals continuously change into various forms or pseudomorphs [37]. Scientists identify this zone by changes in seismic velocity and sometimes physical barriers to movement [38]. Below this transitional zone, the mesosphere is relatively uniform until it reaches the core.

Inner and Outer Core

Figure 2.2.7: Lehmann in 1932

The outer core is the only entirely liquid layer within the Earth. It starts at a depth of 2,890 km and extends to 5,150 km, making it about 2,300 km thick. In 1936, the Danish geophysicist Inge Lehmann analyzed seismic data and was the first to prove a solid inner core existed within a liquid outer core [39]. The solid inner core is about 1,220 km thick, and the outer core is about 2,300 km thick [40].

It seems like a contradiction that the hottest part of the Earth is solid, as the minerals making up the core should be liquified or vaporized at this temperature. Immense pressure keeps the minerals of the inner core in a solid phase [41]. The inner core grows slowly from the lower outer core solidifying as heat escapes the interior of the Earth and is dispersed to the outer layers [42].

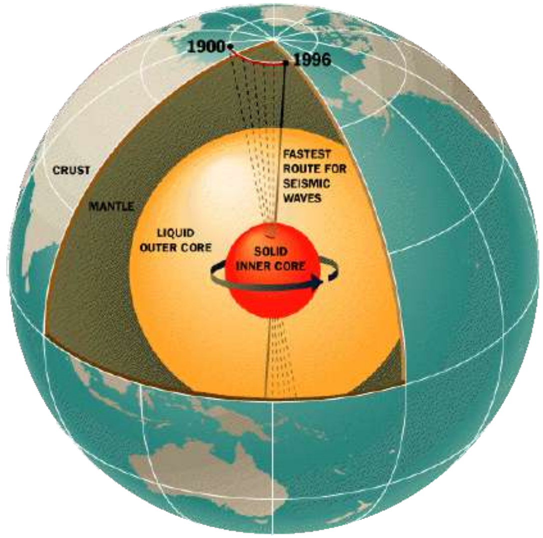


Figure 2.2.8: The outer core's spin most likely causes our protective magnetic field.

The earth's liquid outer core is critically important in maintaining a breathable atmosphere and other environmental conditions favorable for life. Scientists believe the earth's magnetic field is generated by the circulation of molten iron and nickel within the outer core [43]. If the outer core were to stop circulating or become solid, the loss of the magnetic field would result in Earth getting stripped of life-supporting gases and water. This is what happened, and continues to happen, on Mars [44].

Plate Tectonic Boundaries

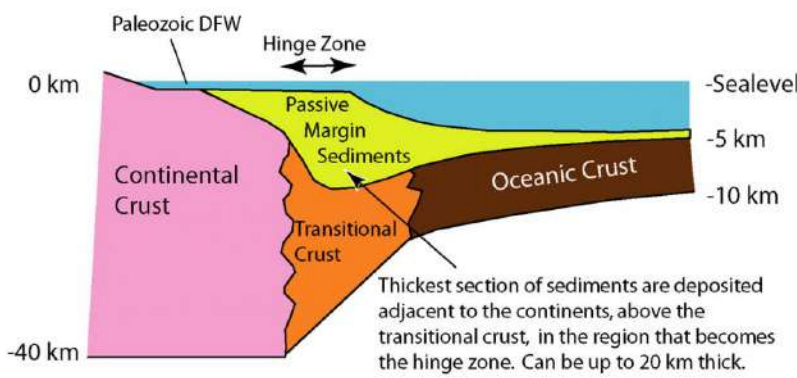


Figure 2.2.9: Passive Margin

At passive margins, the plates don't move—the continental lithosphere transitions into the oceanic lithosphere and forms plates made of both types. A tectonic plate may be made of both oceanic and continental lithosphere connected by a passive margin. North and South America's eastern coastlines are examples of passive margins. Active margins are places where the oceanic and continental lithospheric tectonic plates meet and move relative to each other, such as the western coasts of North and South America. This movement is caused by frictional drag created between the plates and differences in plate densities. The majority of mountain-building events, earthquake activity and active volcanism on the Earth's surface can be attributed to tectonic plate movement at active margins.

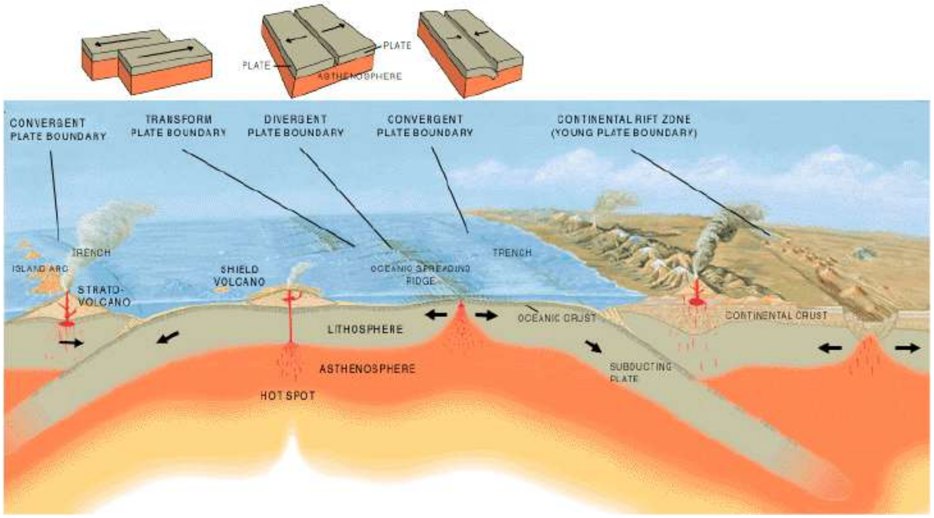


Figure 2.2.10: Schematic of plate boundary types.

In a simplified model, there are three categories of tectonic plate boundaries. Convergent boundaries are places where plates move toward each other. At divergent boundaries, the plates move apart. At transform boundaries, the plates slide past each other.

References

- 27. Herak, D. & Herak, M. Andrija Mohorovičić (1857-1936)—On the occasion of the 150th anniversary of his birth. *Seismol. Res. Lett.* **78**, 671–674 (2007).
- 28. Fernandez, L. M. & Careaga, J. The thickness of the crust in central United States and La Paz, Bolivia, from the spectrum of longitudinal seismic waves. *Bull. Seismol. Soc. Am.* **58**, 711–741 (1968).
- 29. Birch, F. Elasticity and constitution of the Earth's interior. *J. Geophys. Res.* **57**, 227–286 (1952).
- 30. Wyllie, P. J. Ultramafic rocks and the upper mantle. in *Fiftieth anniversary symposia: Mineralogy and petrology of the Upper Mantle; Sulfides; Mineralogy and geochemistry of non-marine evaporites* (ed. Morgan, B. A.) 3–32 (Mineralogical Society of America, 1970).
- 31. Alfe, D., Gillan, M. J. & Price, G. D. Composition and temperature of the Earth's core constrained by combining ab initio calculations and seismic data. *Earth Planet. Sci. Lett.* **195**, 91–98 (2002).
- 32. Oldham, R. D. The constitution of the interior of the Earth, as revealed by earthquakes. *Q. J. Geol. Soc. London* **62**, 456–475 (1906).
- 33. Birch, F. Density and composition of mantle and core. *J. Geophys. Res.* **69**, 4377–4388 (1964).
- 34. Hanks, T. C. & Anderson, D. L. The early thermal history of the earth. *Phys. Earth Planet. Inter.* **2**, 19–29 (1969).
- 35. Pasyanos, M. E. Lithospheric thickness modeled from long-period surface wave dispersion. *Tectonophysics* **481**, 38–50 (2010).
- 36. Forsyth, D. W. The Early Structural Evolution and Anisotropy of the Oceanic Upper Mantle. *Geophys. J. Int.* **43**, 103–162 (1975).
- 37. Ito, E. & Takahashi, E. Postspinel transformations in the system Mg2SiO4-Fe2SiO4 and some geophysical implications. *J. Geophys. Res. [Solid Earth]* **94**, 10637–10646 (1989).
- 38. Fukao, Y. & Obayashi, M. Subducted slabs stagnant above, penetrating through, and trapped below the 660 km discontinuity. *J. Geophys. Res. [Solid Earth]* **118**, 2013JB010466 (2013).
- 39. Lehmann, I. P', Publ. *Bur. Centr. Seism. Internat. Serie A* **14**, 87–115 (1936).
- 40. Engdahl, E. R., Flynn, E. A. & Masse, R. P. Differential PkiKP travel times and the radius of the core. *Geophysical J Royal Astro Soc* **40**, 457–463 (1974).
- 41. Aitta, A. Iron melting curve with a tricritical point. *J. Stat. Mech.* **2006**, P12015 (2006).
- 42. Waszek, L., Irving, J. & Deuss, A. Reconciling the hemispherical structure of Earth/'s inner core with its super-rotation. *Nat. Geosci.* **4**, 264–267 (2011).
- 43. de Wijs, G. A. *et al.* The viscosity of liquid iron at the physical conditions of the Earth's core. *Nature* **392**, 805–807 (1998).
- 44. Jakosky, B. M. *et al.* MAVEN observations of the response of Mars to an interplanetary coronal mass ejection. *Science* **350**, aad0210 (2015).

Contributions and Attributions

2.3: Convergent Boundaries

Convergent boundaries, also called destructive boundaries, are places where two or more plates move toward each other. Convergent boundary movement is divided into two types, subduction and collision, depending on the density of the involved plates. Continental lithosphere is of lower density and thus more buoyant than the underlying asthenosphere. Oceanic lithosphere is denser than continental lithosphere, and, when old and cold, may even be denser than asthenosphere.

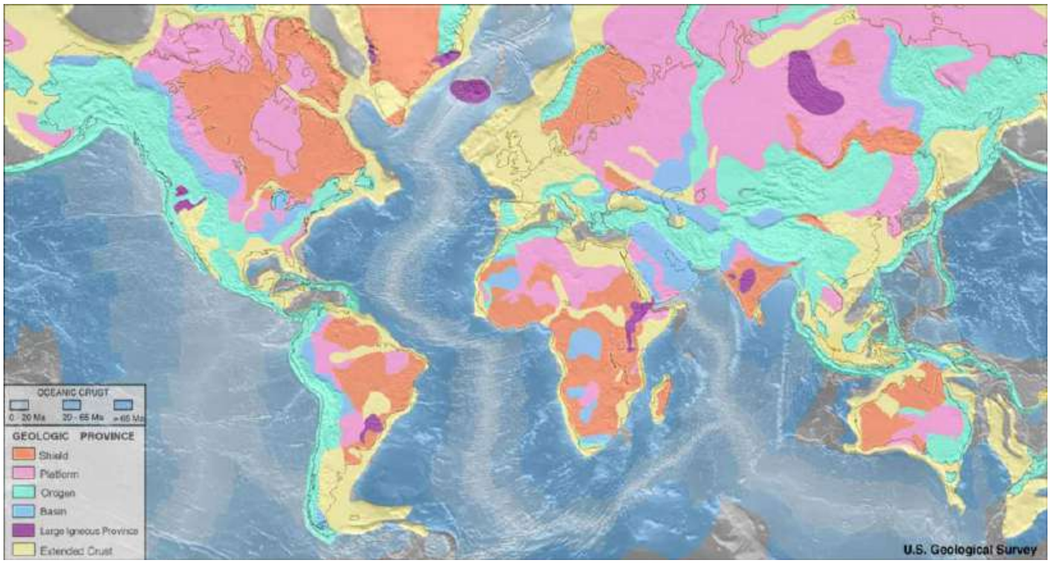


Figure 2.3.1: Geologic provinces of Earth. Orogenies are labeled light blue.

When plates of different densities converge, the higher density plate is pushed beneath the more buoyant plate in a process called subduction. When continental plates converge without subduction occurring, this process is called collision.

Subduction

Video showing continental-oceanic subduction, causing volcanism. By Tanya Atwater and John Iwerks.

Subduction occurs when a dense oceanic plate meets a more buoyant plate, like a continental plate or warmer/younger oceanic plate, and descends into the mantle [45]. The worldwide average rate of oceanic plate subduction is 25 miles per million years, about a half-inch per year [46]. As an oceanic plate descends, it pulls the ocean floor down into a trench. These trenches can be more than twice as deep as the average depth of the adjacent ocean basin, which is usually three to four km. The Mariana Trench, for example, approaches a staggering 11 km [47].

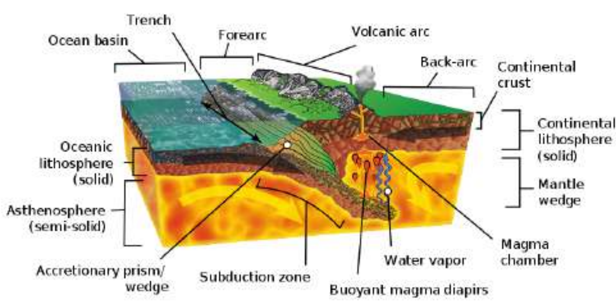


Figure 2.3.2: Diagram of ocean-continent subduction.

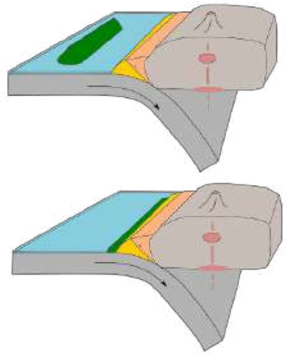


Figure 2.3.3: Microcontinents can become part of the accretionary prism of a subduction zone.

Within the trench, ocean floor sediments are scraped together and compressed between the subducting and overriding plates. This feature is called the accretionary wedge, mélange, or accretionary prism. Fragments of continental material, including microcontinents, riding atop the subducting plate may become sutured to the accretionary wedge and accumulate into a large area of land called a terrane [48]. Vast portions of California are comprised of accreted terranes [49].

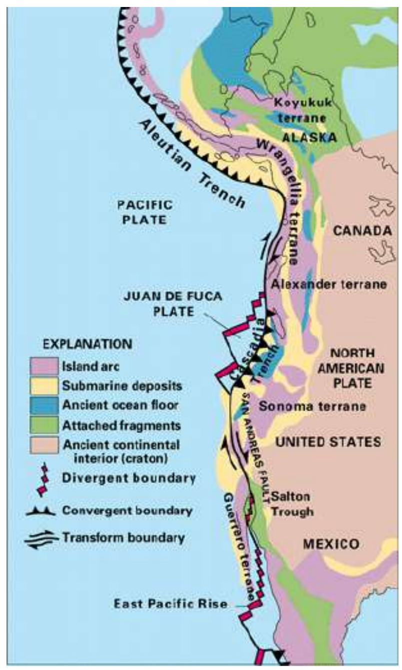


Figure 2.3.4: Accreted terranes of western North America. Everything that is not the "Ancient continental interior (craton)" has been smeared onto the side of the continent by accretion from subduction.

When the subducting oceanic plate, or slab, sinks into the mantle, the immense heat and pressure push volatile materials like water and carbon dioxide into an area below the continental plate and above the descending plate called the mantle wedge. The volatiles are released mostly by hydrated minerals that revert to non-hydrated minerals in these higher temperatures and pressure conditions. When mixed with asthenospheric material above the plate, the volatile lower the melting point of the mantle wedge, and through a process called flux melting it becomes liquid magma. The molten magma is more buoyant than the lithospheric plate above it and migrates to the Earth's surface where it emerges as volcanism. The resulting volcanoes frequently appear as curved mountain chains, volcanic arcs, due to the curvature of the earth. Both oceanic and continental plates can contain volcanic arcs.

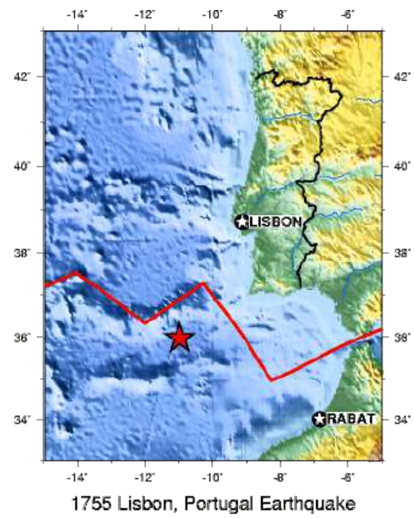


Figure 2.3.5: Location of the large (Mw 8.5-9.0) 1755 Lisbon Earthquake.

How subduction is initiated is still a matter of scientific debate [50]. It is generally accepted that subduction zones start as passive margins, where oceanic and continental plates come together, and then gravity initiates subduction and converts the passive margin into an active one [51]. One hypothesis is gravity pulls the denser oceanic plate down [52] or the plate can start to flow ductility at a low angle [53]. Scientists seeking to answer this question have collected evidence that suggests a new subduction zone is forming off the coast of Portugal [54]. Some scientists have proposed large earthquakes like the 1755 Lisbon earthquake may even have something to do with this process of creating a subduction zone [55], although the evidence is not definitive. Another hypothesis proposes subduction happens at transform boundaries involving plates of different densities [56].

Some plate boundaries look like they should be active, but show no evidence of subduction. The oceanic lithospheric plates on either side of the Atlantic Ocean for example, are denser than the underlying asthenosphere and are not subducting beneath the continental plates.

One hypothesis is the bond holding the oceanic and continental plates together is stronger than the downward force created by the difference in plate densities.

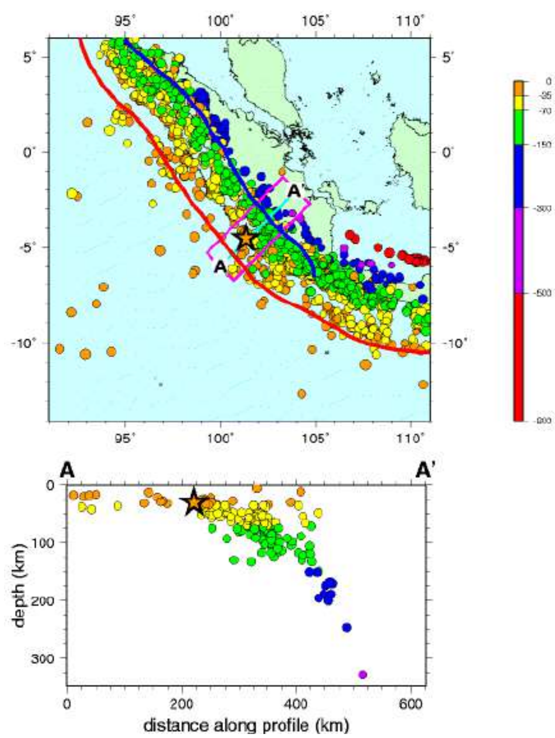


Figure 2.3.6: Earthquakes along the Sunda megathrust subduction zone, along the island of Sumatra, showing the 2006 Mw 9.1-9.3 Indian Ocean Earthquake as a star.

Subduction zones are known for having the largest earthquakes and tsunamis; they are the only places with fault surfaces large enough to create magnitude-9 earthquakes. These subduction-zone earthquakes not only are very large but also are very deep. When a subducting slab becomes stuck and cannot descend, a massive amount of energy builds up between the stuck plates. If this energy is not gradually dispersed, it may force the plates to suddenly release along several hundred kilometers of the subduction zone [57]. Because subduction-zone faults are located on the ocean floor, this massive amount of movement can generate giant tsunamis such as those that followed the 2004 Indian Ocean Earthquake and 2011 Tōhoku Earthquake in Japan.

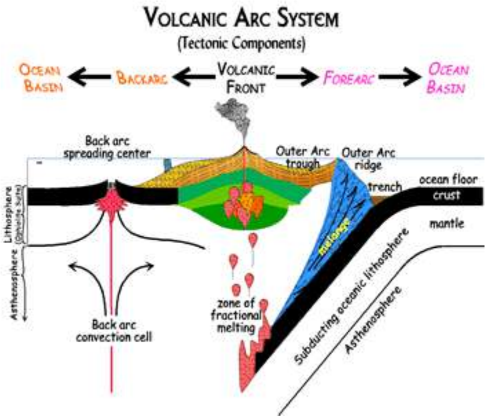


Figure 2.3.7: Various parts of a subduction zone. This subduction zone is ocean-ocean subduction, though the same features can apply to continent-ocean subduction.

All subduction zones have a forearc basin, a feature of the overriding plate found between the volcanic arc and oceanic trench. The forearc basin experiences a lot of faulting and deformation activity, particularly within the accretionary wedge [58].

In some subduction zones, tensional forces working on the continental plate create a backarc basin on the interior side of the volcanic arc. Some scientists have proposed a subduction mechanism called oceanic slab rollback creates extension faults in the overriding plates [59]. In this model, the descending oceanic slab does not slide directly under the overriding plate but instead rolls back, pulling the overlying plate seaward. The continental plate behind the volcanic arc gets stretched like pizza dough until the surface cracks and collapses to form a backarc basin. If the extension activity is extensive and deep enough, a backarc basin can develop into a continental rifting zone. These continental divergent boundaries may be less symmetrical than their mid-ocean ridge counterparts [60].

In places where numerous young buoyant oceanic plates are converging and subducting at a relatively high velocity, they may force the overlying continental plate to buckle and crack [61]. This is called back-arc faulting. Extensional back-arc faults pull rocks and chunks of plates apart. Compressional back-arc faults, also known as thrust faults, push them together.

The dual spines of the Andes Mountain range include a example of compressional thrust faulting. The western spine is part of a volcanic arc. Thrust faults have deformed the non-volcanic eastern spine, pushing rocks and pieces of a continental plate on top of each other.

There are two styles of thrust fault deformation: thin-skinned faults that occur in superficial rocks lying on top of the continental plate and thick-skinned faults that reach deeper into the crust. The Sevier Orogeny in the western U.S. is a notable thin-skinned type of deformation created during the Cretaceous Period. The Laramide Orogeny, a thick-skinned type of deformation, occurred near the end of and slightly after the Sevier Orogeny in the same region.

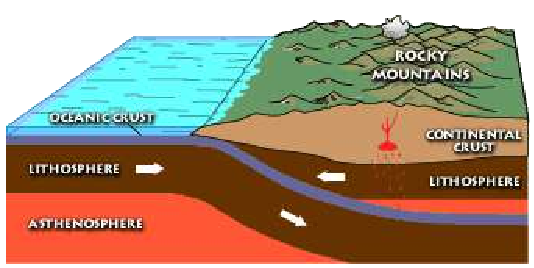


Figure 2.3.8: Shallow subduction during the Laramide Orogeny.

Flat-slab, or shallow, subduction caused the Laramide Orogeny. When the descending slab subducts at a low angle, there is more contact between the slab and the overlying continental plate than in a typical subduction zone. The shallowly-subducting slab pushes against the overriding plate and creates an area of deformation on the overriding plate many kilometers away from the subduction zone [62].

Oceanic-Continental Subduction

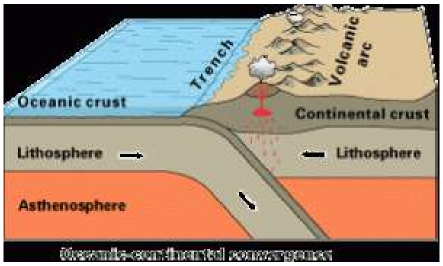


Figure 2.3.9: Subduction of an oceanic plate beneath a continental plate, forming a trench and volcanic arc.

Oceanic-continental subduction occurs when an oceanic plate dives below a continental plate. This convergent boundary has a trench and mantle wedge and frequently, a volcanic arc. Well-known examples of continental volcanic arcs are the Cascade Mountains in the Pacific Northwest [63] and the Western Andes Mountains in South America [64].

Oceanic-Oceanic Subduction

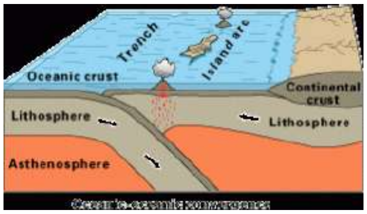


Figure 2.3.10: Subduction of an oceanic plate beneath another oceanic plate, forming a trench and an island arc.

The boundaries of oceanic-oceanic subduction zones show very different activity from those involving oceanic-continental plates. Since both plates are made of oceanic lithosphere, it is usually the older plate that subducts because it is colder and denser. The volcanism on the overlying oceanic plate may remain hidden underwater.. If the volcanoes rise high enough the reach the ocean surface, the chain of volcanism forms an island arc. Examples of these island arcs include the Aleutian Islands in the northern Pacific Ocean, Lesser Antilles in the Caribbean Sea, and numerous island chains scattered throughout the western Pacific Ocean [65].

Collisions

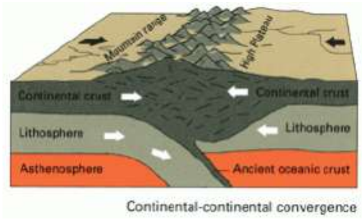


Figure 2.3.11: Two continental plates colliding.

When continental plates converge, during the closing of an ocean basin, for example, subduction is not possible between the equally buoyant plates. Instead of one plate descending beneath another, the two masses of continental lithosphere slam together in a process known as collision [66]. Without subduction, there is no magma formation and no volcanism. Collision zones are characterized by tall, non-volcanic mountains; a broad zone of frequent, large earthquakes; and very little volcanism.

When oceanic crust connected by a passive margin to continental crust completely subducts beneath a continent, an ocean basin closes, and continental collision begins. Eventually, as ocean basins close, continents join together to form a massive accumulation of continents called a supercontinent, a process that has taken place in ~500 million-year-old cycles over earth's history.

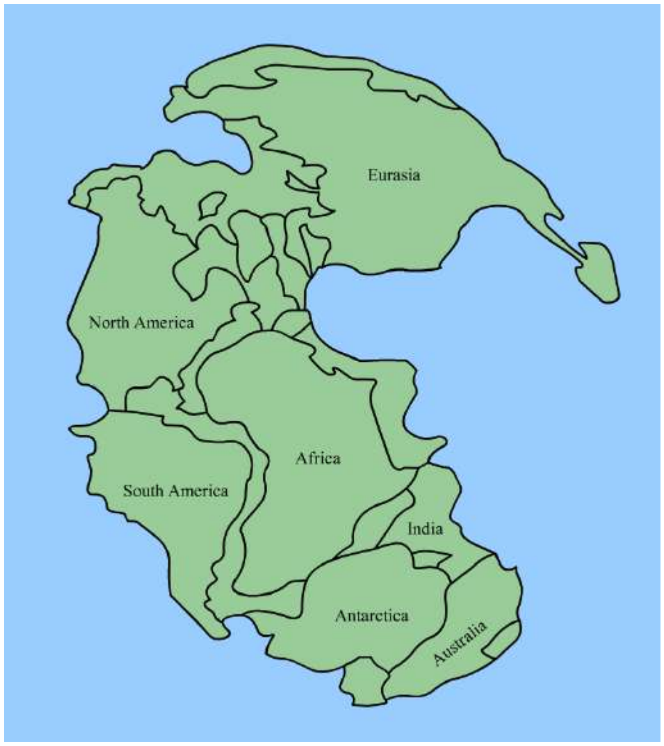


Figure 2.3.12: A reconstruction of the supercontinent Pangaea, showing approximate positions of modern continents.

The process of collision created Pangea, the supercontinent envisioned by Wegener as the key component of his continental drift hypothesis. Geologists now have evidence that continental plates have been continuously converging into supercontinents and splitting into smaller basin-separated continents throughout Earth's existence, calling this process the supercontinent cycle, a process that takes place in approximately 500 million years. For example, they estimate Pangea began separating 200 million years ago. Pangea was preceded by an earlier supercontinents, one of which being Rodinia, which existed 1.1 billion years ago and started breaking apart 800 million to 600 million years ago.

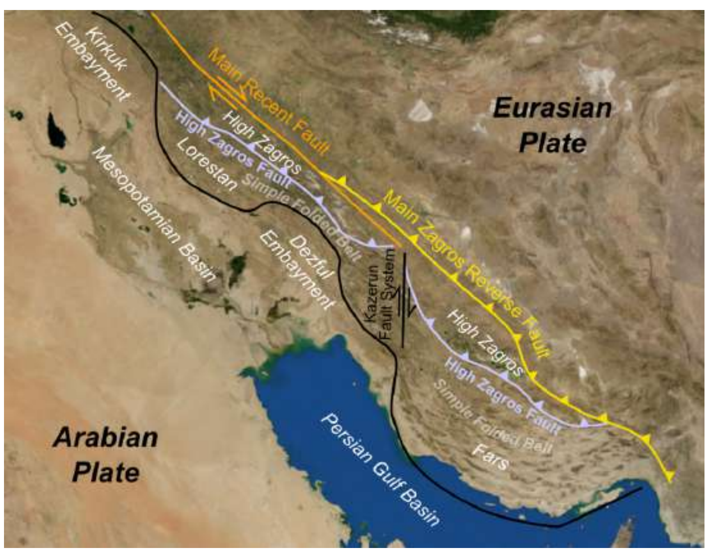


Figure 2.3.13: The tectonics of the Zagros Mountains. Note the Persian Gulf foreland basin.

A foreland basin is a feature that develops near mountain belts, as the combined mass of the mountains forms a depression in the lithospheric plate. While foreland basins may occur at subduction zones, they are most commonly found at collision boundaries. The Persian Gulf is possibly the best modern example, created entirely by the weight of the nearby Zagros Mountains.

Figure 2.3.14: Pillow lavas, which only form underwater, from an ophiolite in the Apennine Mountains of central Italy.

If continental and oceanic lithosphere are fused on the same plate, it can partially subduct but its buoyancy prevents it from fully descending. In very rare cases, part of a continental plate may become trapped beneath a descending oceanic plate in a process called obduction [67]. When a portion of the continental crust is driven down into the subduction zone, due to its buoyancy it returns to the surface relatively quickly.

As pieces of the continental lithosphere break loose and migrate upward through the obduction zone, they bring along bits of the mantle and ocean floor and amend them on top of the continental plate. Rocks composed of this mantle and ocean-floor material are called ophiolites and they provide valuable information about the composition of the mantle.

The area of collision-zone deformation and seismic activity usually covers a broader area because the continental lithosphere is plastic and malleable. Unlike subduction-zone earthquakes, which tend to be located along a narrow swath near the convergent boundary, collision-zone earthquakes may occur hundreds of kilometers from the boundary between the plates.

The Eurasian continent has many examples of collision-zone deformations covering vast areas. The Pyrenees mountains begin in the Iberian Peninsula and cross into France. Also, there are the Alps stretching from Italy to central Europe; the Zagros mountains from Arabia to Iran; and Himalaya mountains from the Indian subcontinent to central Asia.

Video *Animation of India crashing into Asia, by Tanya Atwater.*

References

- 45. Stern, R. J. A subduction primer for instructors of introductory geology courses and authors of introductory-geology textbooks. *J. Geosci. Educ.* **46**, 221 (1998).
- 46. Tatsumi, Y. The subduction factory: how it operates in the evolving Earth. *GSA Today* **15**, 4 (2005).
- 47. Todo, Y., Kitazato, H., Hashimoto, J. & Gooday, A. J. Simple foraminifera flourish at the ocean's deepest point. *Science* **307**, 689 (2005).
- 48. Schermer, E. R., Howell, D. G. & Jones, D. L. THE ORIGIN OF ALLOCHTHONOUS TERRANES: Perspectives on the Growth and shaping of Continents. *Annual Review of Earth and Planetary Sciences* **12**, 107 (1984).
- 49. Moores, E. M., Sloan, D. & Stout, D. L. *Classic Cordilleran Concepts: A View from California*. (Geological Society of America, 1999).
- 50. Stern, R. J. Subduction initiation: spontaneous and induced. *Earth Planet. Sci. Lett.* **226**, 275–292 (2004).
- 51. Mueller, S. & Phillips, R. J. On the initiation of subduction. *J. Geophys. Res. [Solid Earth]* **96**, 651–665 (1991).
- 52. Mart, Y. *et al.* Analogue modelling of the initiation of subduction. *Geophys. J. Int.* **160**, 1081–1091 (2005).
- 53. Goren, L., Aharonov, E., Mulugeta, G., Koyi, H. A. & Mart, Y. Ductile deformation of passive margins: A new mechanism for subduction initiation. *J. Geophys. Res.* **113**, B08411 (2008).
- 54. Duarte, J., Schellart, W. & Rosas, F. The future of Earth's oceans: consequences of subduction initiation in the Atlantic and implications for supercontinent formation. in **17**, (adsabs.harvard.edu, 2015).
- 55. Stich, D., Mancilla, F. de L., Pondrelli, S. & Morales, J. Source analysis of the February 12th 2007, Mw 6.0 Horseshoe earthquake: Implications for the 1755 Lisbon earthquake. *Geophys. Res. Lett.* **34**, L12308 (2007).
- 56. Casey, J. F. & Dewey, J. F. Initiation of subduction zones along transform and accreting plate boundaries, triple-junction evolution, and forearc spreading centres—implications for ophiolitic geology and obduction. *Geological Society, London, Special Publications* **13**, 269–290 (1984).
- 57. Satake, K. & Atwater, B. F. Long-term perspectives on giant earthquakes and tsunamis at subduction zones*. *Annu. Rev. Earth Planet. Sci.* **35**, 349–374 (2007).

- 58. Fuller, C. W., Willett, S. D. & Brandon, M. T. Formation of forearc basins and their influence on subduction zone earthquakes. *Geology* **34**, 65–68 (2006).
- 59. Uyeda, S. Subduction zones: Their diversity, mechanism and human impacts. *GeoJournal* **8**, 381–406 (1984).
- 60. Deschamps, A. & Fujiwara, T. Asymmetric accretion along the slow-spreading Mariana Ridge. *Geochem. Geophys. Geosyst.* (2003).
- 61. Molnar, P. & Atwater, T. Interarc spreading and Cordilleran tectonics as alternates related to the age of subducted oceanic lithosphere. *Earth Planet. Sci. Lett.* **41**, 330–340 (1978).
- 62. Li, Z.-X. & Li, X.-H. Formation of the 1300-km-wide intracontinental orogen and postorogenic magmatic province in Mesozoic South China: A flat-slab subduction model. *Geology* **35**, 179–182 (2007).
- 63. Geller, D. Cascadia subduction zone volcanism in British Columbia. *academic.emporia.edu* (2008).
- 64. De Silva, S. L. & Francis, P. W. *Volcanoes of the central Andes*. (Springer, 1991).
- 65. Smith, A. L. & Roobol, M. J. Mt. Pelée, Martinique; A Study of an Active Island-arc Volcano. *Geological Society of America Memoirs* **175**, 1–110 (1990).
- 66. O'Brien, P. J. Subduction followed by collision: Alpine and Himalayan examples. *Phys. Earth Planet. Inter.* **127**, 277–291 (2001).
- 67. Ernst, W. G. Preservation/exhumation of ultrahigh-pressure subduction complexes. *Lithos* **92**, 321–335 (2006).

Contributions and Attributions

2.4: Divergent Boundaries

At divergent boundaries, sometimes called constructive boundaries, lithospheric plates move away from each other. There are two types of divergent boundaries, categorized by where they occur: continental rift zones and mid-ocean ridges. Continental rift zones occur in weak spots in the continental lithospheric plate. A mid-ocean ridge usually originates in a continental plate as a rift zone that expands to the point of splitting the plate apart, with seawater filling in the gap. The separate pieces continue to drift apart and become individual continents. This process is known as rift-to-drift.

Continental Rifting

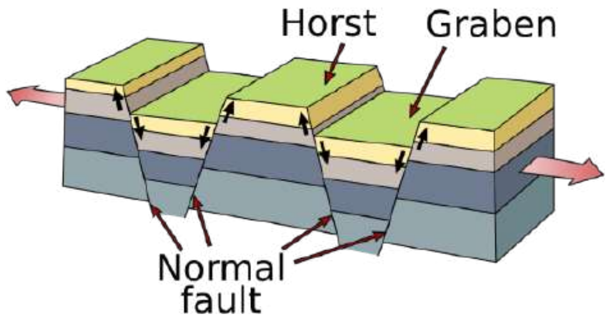


Figure 2.4.1: Faulting that occurs in divergent boundaries.

In places where the continental plates are very thick, they reflect so much heat back into the mantle it develops strong convection currents that push super-heated mantle material up against the overlying plate, softening it. Tensional forces created by this convective upwelling begin to pull the weakened plate apart. As it stretches, it becomes thinner and develops deep cracks called extension or normal faults. Eventually, plate sections located between large faults drop into deep depressions known as rift valleys, which often contain keystone-shaped blocks of down-dropped crust known as grabens. The shoulders of these grabens are called horsts. If only one side of a section drops, it is called a half-graben. Depending on the conditions, rifts can grow into very large lakes and even oceans.

Figure 2.4.2: The Afar Triangle (center) has the Red Sea ridge (center to upper left), Gulf of Aden ridge (center to right), and East African Rift (center to lower left) form a triple junction that are about 120° apart.

While seemingly occurring at random, rifting is dictated by two factors. Rifting does not occur in continents with older and more stable interiors, known as cratons. When continental rifting does occur, the break-up pattern resembles the seams of a soccer ball, also called a truncated icosahedron. This is the most common surface-fracture pattern to develop on an evenly expanding sphere because it uses the least amount of energy [68].

Using the soccer ball model, rifting tends to lengthen and expand along a particular seam while fizzling out in the other directions. These seams with little or no tectonic activity are called failed rift arms. A failed rift arm is still a weak spot in the continental plate; even without the presence of active extension faults, it may develop into a called an aulacogen. One example of a failed rift arm is the Mississippi Valley Embayment, a depression through which the upper end of the Mississippi River flows. Occasionally connected rift arms do develop concurrently, creating multiple boundaries of active rifting. In places where the rift arms do not fail, for example the Afar Triangle, three divergent boundaries can develop near each other forming a triple junction.

Figure 2.4.3: NASA image of the Basin and Range horsts and grabens across central Nevada.

Rifts come in two types: narrow and broad. Narrow rifts are characterized by a high density of highly active divergent boundaries. The East African Rift Zone, where the horn of Africa is pulling away from the mainland, is an excellent example of an active narrow rift. Lake Baikal in Russia is another. Broad rifts also have numerous fault zones, but they are distributed over wide areas of deformation. The Basin and Range region located in the western United States is a type of broad rift. The Wasatch Fault, which also created the Wasatch Mountain Range in the state of Utah, forms the eastern divergent boundary of this broad rift (Animation 1 and Animation 2).

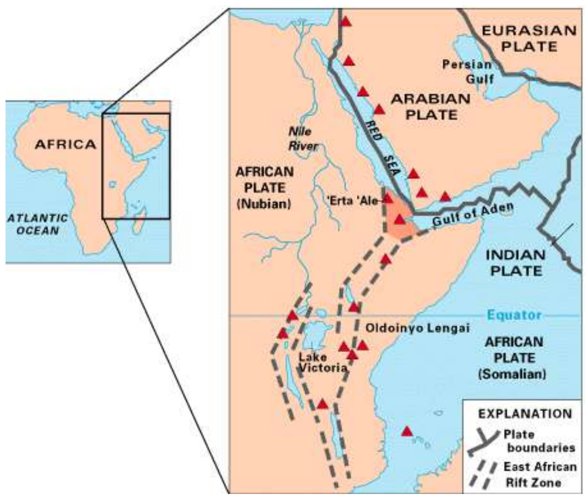


Figure 2.4.4: The narrow East African Rift.

Rifts have earthquakes, although not of the magnitude and frequency of other boundaries. They may also exhibit volcanism. Unlike the flux-melted magma found in subduction zones, rift-zone magma is created by decompression melting. As the continental plates are pulled apart, they create a region of low pressure that melts the lithosphere and draws it upwards. When this molten magma reaches the weakened and fault-riddled rift zone, it migrates to the surface by breaking through the plate or escaping via an open fault. Examples of young rift volcanoes dot the Basin and Range region in the United States. Rift-zone activity is responsible for generating some unique volcanism, such as the Ol Doinyo Lengai in Tanzania. This volcano erupts lava consisting largely of carbonatite, a relatively cold, liquid carbonate mineral [69].

South America and Africa rift, forming the Atlantic. Video by Tanya Atwater.

Mid-Ocean Ridges

Figure 2.4.5: Progression from rift to mid-ocean ridge.

As rifting and volcanic activity progress, the continental lithosphere becomes more mafic (see Chapter 4) and thinner, with the eventual result transforming the plate under the rifting area into the oceanic lithosphere. This is the process that gives birth to a new ocean, much like the narrow Red Sea emerged with the movement of Arabia away from Africa. As the oceanic lithosphere continues to diverge, a mid-ocean ridge is formed.

Mid-ocean ridges, also known as spreading centers, have several distinctive features. They are the only places on earth that create new oceanic lithosphere. Decompression melting in the rift zone changes asthenosphere material into the new lithosphere, which oozes up through cracks in the oceanic plate. The amount of new lithosphere being created at mid-ocean ridges is highly significant. These undersea rift volcanoes produce more lava than all other types of volcanism combined. Despite this, most mid-oceanic ridge volcanism remains unmapped because the volcanoes are located deep on the ocean floor.

In rare cases, such as a few locations in Iceland, rift zones display the type of volcanism, spreading, and ridge formation found on the ocean floor.

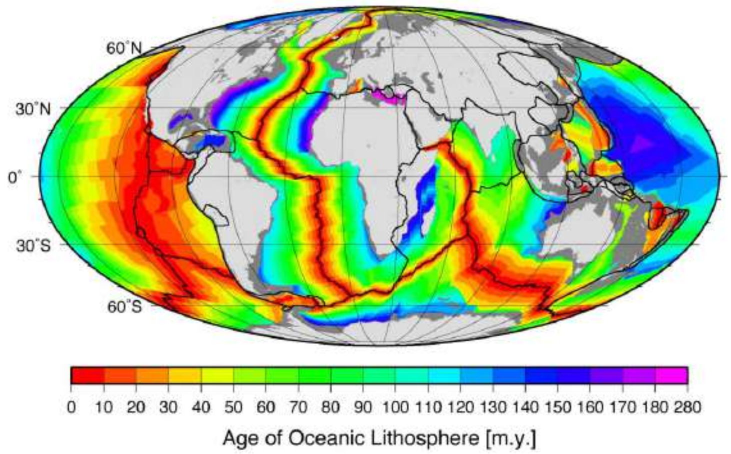


Figure 2.4.6: Age of oceanic lithosphere, in millions of years. Notice the differences in the Atlantic Ocean along the coasts of the continents.

The ridge feature is created by the accumulation of hot lithosphere material, which is lighter than the dense underlying asthenosphere. This chunk of isostatically buoyant lithosphere sits partially submerged and partially exposed to the asthenosphere, like an ice cube floating in a glass of water.

As the ridge continues to spread, the lithosphere material is pulled away from the area of volcanism and becomes colder and denser. As it continues to spread and cool, the lithosphere settles into wide swathes of relatively featureless topography called abyssal plains with lower topography [70].

This model of ridge formation suggests the sections of the lithosphere furthest away from the mid-ocean ridges will be the oldest. Scientists have tested this idea by comparing the age of rocks located in various locations on the ocean floor. Rocks found near ridges are younger than those found far away from any ridges. Sediment accumulation patterns also confirm the idea of sea-floor spreading. Sediment layers tend to be thinner near mid-ocean ridges, indicating it has had less time to build up.

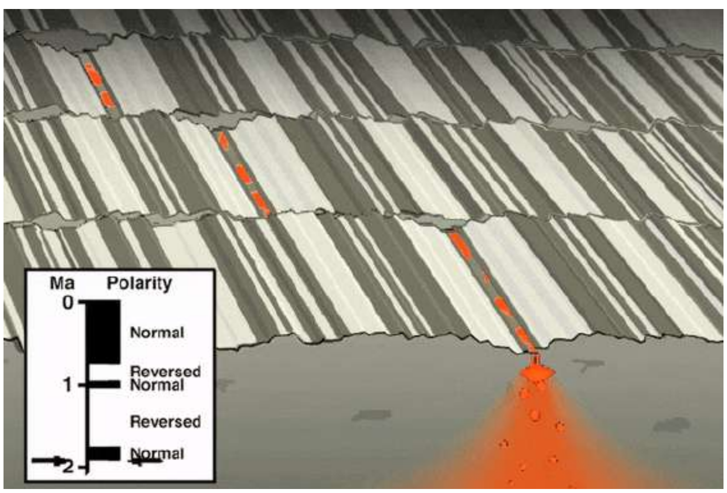


Figure 2.4.8: Spreading along several mid-ocean ridges, showing magnetic striping symmetry. By Tanya Atwater.

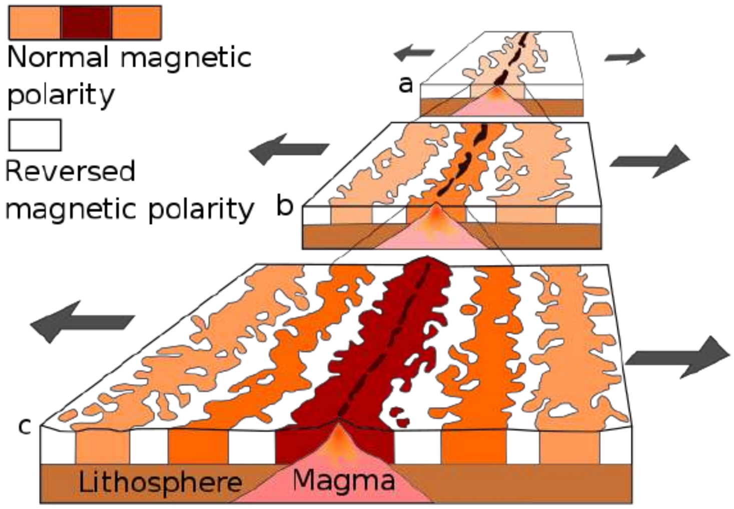
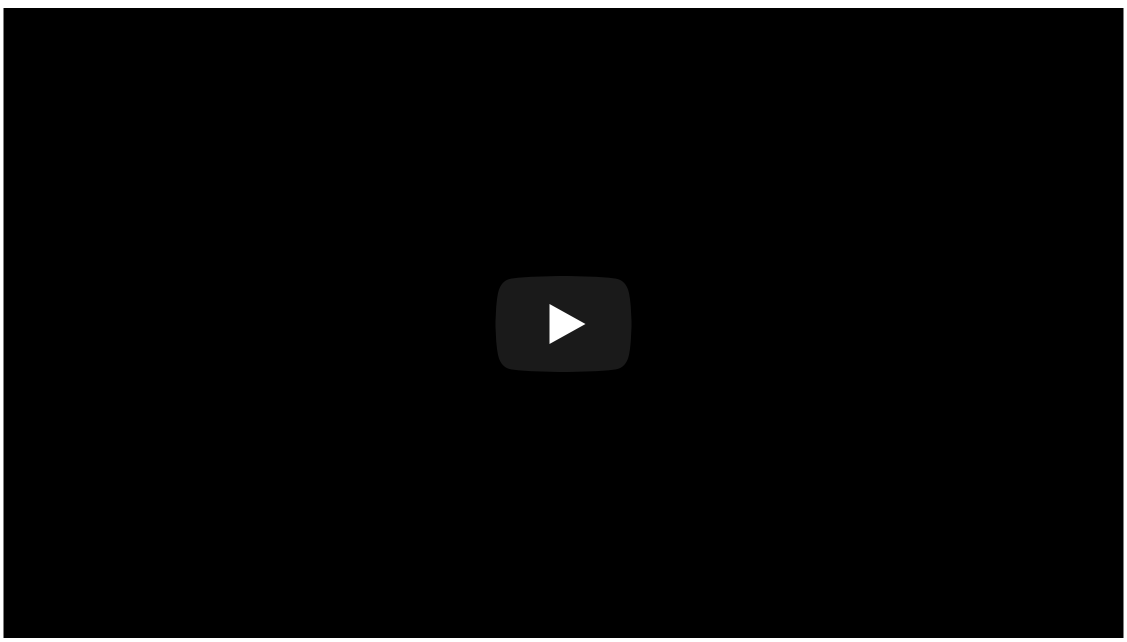
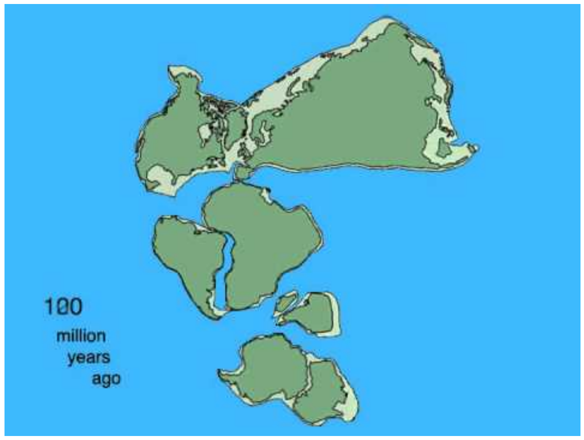


Figure 2.4.9: A time progression (with “a” being first and “c” being last) showing a spreading center getting wider while recording changes in the magnetic field of the Earth.

As mentioned in the section on paleomagnetism and the development of plate tectonic theory, scientists noticed mid-ocean ridges contained unique magnetic anomalies that show up as symmetrical striping on both sides of the ridge. The Vine-Matthews-Morley hypothesis [20] proposes these alternating reversals are created by the earth’s magnetic field being imprinted into magma after it emerges from the ridge [71]. Very hot magma has no magnetic field. As the oceanic plates get pulled apart, the magma cools below the Curie point, the temperature below which a magnetic field gets locked into magnetic minerals. The alternating magnetic reversals in the rocks reflect the periodic swapping of earth’s magnetic north and south poles. This paleomagnetic pattern provides a great historical record of ocean-floor movement, and is used to reconstruct past tectonic activity and determine rates of ridge spreading [72].





[Video](#) of the breakup of Pangea and formation of the northern Atlantic Ocean. By Tanya Atwater.

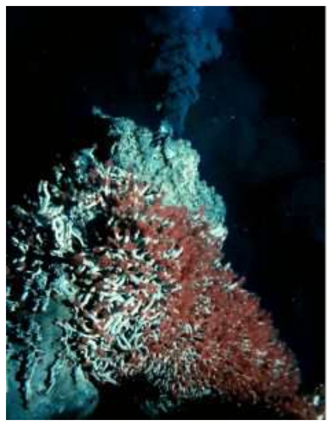


Figure 2.4.10: Black smoker hydrothermal vent with a colony of giant (6'+) tube worms.

Thanks to their distinctive geology, mid-ocean ridges are home to some of the most unique ecosystems ever discovered. The ridges are often studded with hydrothermal vents, deep fissures that allow seawater to circulate through the upper portions of the oceanic plate and interact with hot rock. The super-heated seawater rises back up to the surface of the plate, carrying dissolved gasses and minerals, and small particulates. The resulting emitted hydrothermal water looks like black underwater smoke.

Scientists had known about these geothermal areas on the ocean floor for some time. However, it was not until 1977, when scientists piloting a deep submergence vehicle, the Alvin, discovered a thriving community of organisms clustered around these hydrothermal vents [73]. These unique organisms, which include 10-foot-long tube worms taller than people, live in the complete darkness of the ocean floor deprived of oxygen and sunlight. They use the geothermal energy provided by the vents and a process called bacterial chemosynthesis to feed on sulfur compounds. Before this discovery, scientists believed life on earth could not exist without photosynthesis, a process that requires sunlight. Some scientists suggest this type of environment could have been the origin of life on Earth [74], and perhaps even extraterrestrial life elsewhere in the galaxy, such as on Jupiter's moon Europa [75].

References

- 20. Frankel, H. The Development, Reception, and Acceptance of the Vine-Matthews-Morley Hypothesis. *Hist. Stud. Phys. Biol. Sci.* **13**, 1–39 (1982).
- 68. Sears, J. W., St. George, G. M. & Winne, J. C. Continental rift systems and anorogenic magmatism. *Lithos* **80**, 147–154 (2005).
- 69. Dawson, J. B., Pinkerton, H., Norton, G. E. & Pyle, D. M. Physicochemical properties of alkali carbonatite lavas:Data from the 1988 eruption of Oldoinyo Lengai, Tanzania. *Geology* **18**, 260–263 (1990).
- 70. Davis, E. E. & Lister, C. R. B. Fundamentals of ridge crest topography. *Earth Planet. Sci. Lett.* **21**, 405–413 (1974).
- 71. Cox, A., Doell, R. R. & Dalrymple, G. B. Reversals of the Earth's Magnetic Field. *Science* **144**, 1537–1543 (1964).
- 72. Atwater, T. & Molnar, P. Relative motion of the Pacific and North American plates deduced from sea-floor spreading in the Atlantic, Indian, and South Pacific Oceans. in **13**, 136–148 (Stanford University Publication of …, 1973).
- 73. Corliss, J. B. *et al.* Submarine thermal springs on the Galapagos Rift. *Science* **203**, 107321083 (1979).

- 74. Wächtershäuser, G. Evolution of the first metabolic cycles. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 200–204 (1990).
- 75. McCollom, T. M. Methanogenesis as a potential source of chemical energy for primary biomass production by autotrophic organisms in hydrothermal systems on Europa. *J. Geophys. Res.* **104**, 30729–30742 (1999).

Contributions and Attributions

2.5: Transform Boudaries

A transform boundary, sometimes called a strike-slip or conservative boundary, is where the lithospheric plates slide past each other in the horizontal plane. This movement is described based on the perspective of an observer standing on one of the plates, looking across the boundary at the opposing plate. Dextral, also known as right-lateral, movement describes the opposing plate moving to the right. Sinistral, also known as left lateral, movement describes the opposing plate moving to the left.

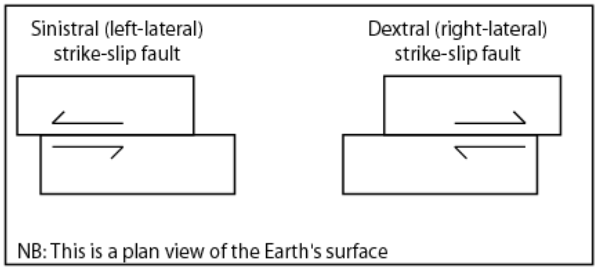


Figure 2.5.1: The two types of transform/strike-slip faults.

Most transform boundaries are found on the ocean floor, around mid-ocean ridges. These boundaries form aseismic fracture zones, filled with earthquake-free transform faults, to accommodate different rates of spreading occurring at the ridge.

Figure 2.5.2: Map of the San Andreas fault, showing relative motion.

Some transform boundaries produce significant seismic activity, primarily as earthquakes, with very little mountain-building or volcanism. This type of transform boundary may contain a single fault or series of faults, which develop in places where plate tectonic stresses are transferred to the surface. As with other types of active boundaries, if the plates are unable to shear past each other the tectonic forces will continue to build up. If the built-up energy between the plates is suddenly released, the result is an earthquake.

In the eyes of humanity, the most significant transform faults occur within continental plates and have a shearing motion that frequently produces moderate-to-large magnitude earthquakes. Notable examples include the San Andreas Fault in California, Northern and Eastern Anatolian Faults in Turkey, Altyn Tagh Fault in Central Asia, and Alpine Fault in New Zealand.

Transpression and Transtension

Bends along transform faults may create compressional or extensional forces that cause secondary faulting zones. Transpression occurs where there is a component of compression in addition to the shearing motion. These forces build up around the area of the bend, where the opposing plates are restricted from sliding past each other. As the forces continue to build up, they create mountains in the restraining bend around the fault. The Big Bend area, located in the southern part of the San Andreas Fault includes a large area of transpression where many mountains have been built, moved, and even rotated [76].

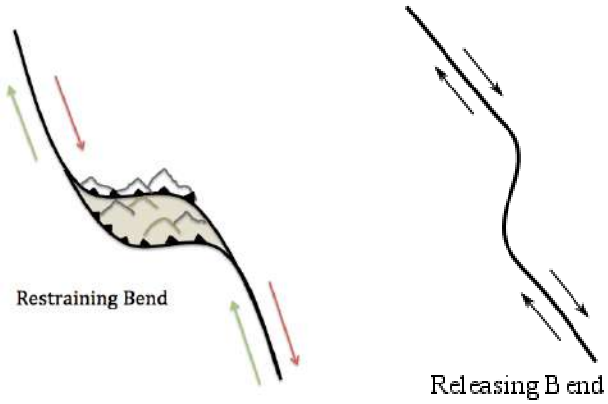


Figure 2.5.3: A transpressional strike-slip fault, causing uplift called a restraining bend. A transtensional strike-slip fault, resulting in a releasing bend.

Transtension zones require a fault that includes a releasing bend, where the plates are being pulled apart by extensional forces. Depressions and sometimes volcanism develop in the releasing bend, along the fault. The Dead Sea found between Israel and Jordan, and the Salton Sea of California are examples of basins formed by transtensional forces.

Piercing Points

Figure 2.5.4: Wallace (dry) Creek on the Carrizo Plain, California. Note as the creek flows from the northern mountainous part of the image, it takes a sharp right (as viewed from the flow of water), then a sharp left. This is caused by the San Andreas Fault cutting roughly perpendicular to the creek, and shifting the location of the creek over time. The fault can be seen about halfway down, trending left to right, as a change in the topography.

When a geological feature is cut by a fault, it is called a piercing point. Piercing points are very useful for recreating past fault movement, especially along transform boundaries. Transform faults are unique because their horizontal motion keeps a geological feature relatively intact, preserving the record of what happened. Other types of faults—normal and reverse —tend to be more destructive, obscuring or destroying these features. The best type of piercing point includes unique patterns that are used to match the parts of a geological feature separated by fault movement. Detailed studies of piercing points show the San Andreas Fault has experienced over 225 km of movement in the last 20 million years, and this movement occurred at three different fault traces.

References

- 76. Ingersoll, R. V. & Rumelhart, P. E. Three-stage evolution of the Los Angeles basin, southern California. *Geology* **27**, 593–596 (1999).

Contributions and Attributions

2.6: The Wilson Cycle

The Wilson Cycle is named for J. Tuzo Wilson who first described it in 1966 [24], and it outlines the ongoing origin and breakup of supercontinents, such as Pangea and Rodinia [77]. Scientists have determined this cycle has been operating for at least three billion years and possibly earlier.

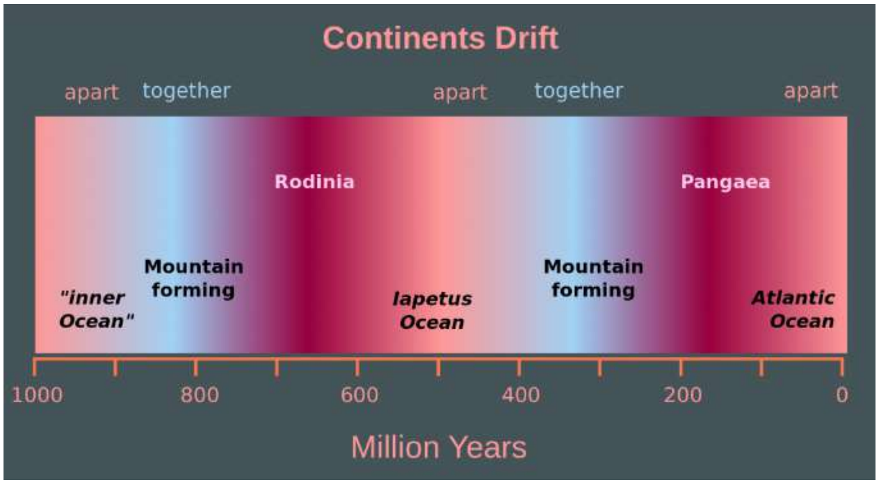


Figure 2.6.1: Diagram of the Wilson Cycle, showing rifting and collision phases.

There are a number of hypotheses about how the Wilson Cycle works. One mechanism proposes that rifting happens because continental plates reflect the heat much better than oceanic plates [78]. When continents congregate together, they reflect more of the Earth's heat back into the mantle, generating more vigorous convection currents that then start the continental rifting process [79].

Some geologists believe mantle plumes are remnants of these periods of increased mantle temperature and convection upwelling and study them for clues about the origin of continental rifting. The mechanism behind how supercontinents are created is still largely a mystery. There are three schools of thought about what continues to drive the continents further apart and eventually bring them together. The ridge-push hypothesis suggests after the initial rifting event, plates continue to be pushed apart by mid-ocean spreading centers and their underlying convection currents. Slab-pull proposes the plates are pulled apart by descending slabs in the subduction zones of the oceanic-continental margins [80]. A third idea, gravitational sliding, attributes the movement to gravitational forces pulling the lithospheric plates down from the elevated mid-ocean ridges and across the underlying asthenosphere [81]. Current evidence seems to support slab pull more than ridge push or gravitational sliding.

References

- 24. Wilson, J. T. *Did the Atlantic close and then re-open?* (Nature, 1966).
- 77. Zhao, G., Cawood, P. A., Wilde, S. A. & Sun, M. Review of global 2.1–1.8 Ga orogens: implications for a pre-Rodinia supercontinent. *Earth-Sci. Rev.* **59**, 125–162 (2002).
- 78. Pollack, H. N., Hurter, S. J. & Johnson, J. R. Heat flow from the Earth's interior: Analysis of the global data set. *Rev. Geophys.* **31**, 267–280 (1993).
- 79. Li, Z. X., Li, X. H., Kinny, P. D. & Wang, J. The breakup of Rodinia: did it start with a mantle plume beneath South China? *Earth Planet. Sci. Lett.* **173**, 171–181 (1999).
- 80. Conrad, C. P. & Lithgow-Bertelloni, C. How mantle slabs drive plate tectonics. *Science* **298**, 207–209 (2002).
- 81. Spence, W. Slab pull and the seismotectonics of subducting lithosphere. *Rev. Geophys.* **25**, 55–69 (1987).

Contributions and Attributions

2.7: Hotspots

The Wilson Cycle provides a broad overview of the tectonic plate movement. To analyze plate movement more precisely, scientists study hotspots. First postulated by J. Tuzo Wilson in 1963, a hotspot is an area in the lithospheric plate where molten magma breaks through and creates a volcanic center, islands in the ocean and mountains on land [82]. As the plate moves across the hotspot, the volcano center becomes extinct because it is no longer over an active magma source. Instead, the magma emerges through another area in the plate to create a new active volcano. Over time, the combination of a moving plate and a stationary hotspot creates a chain of islands or mountains. The classic definition of hotspots states they do not move, although recent evidence suggests that there may be exceptions [83].

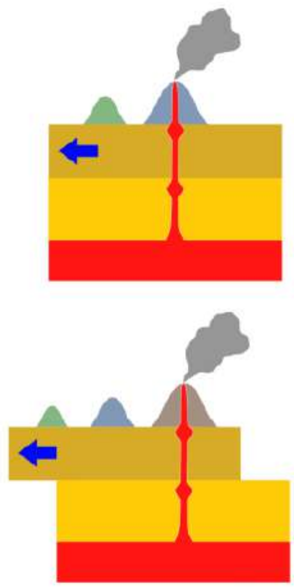


Figure 2.7.1: Diagram showing a non-moving source of magma (mantle plume) and a moving overriding plate.

Hotspots are the only types of volcanism not associated with subduction or rifting zones at plate boundaries; they seem totally disconnected from any plate tectonics processes, such as earthquakes. However, there are relationships between hotspots and plate tectonics. There are several hotspots, current and former, that are believed to have begun at the time of rifting. Also, scientists use the age of volcanic eruptions and shape of the chain to quantify the rate and direction of plate movements relative to the hotspot.

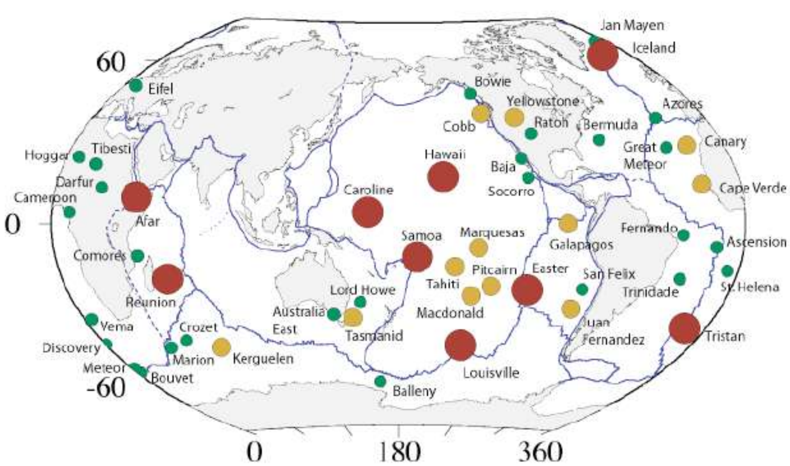


Figure 2.7.2: Map of world hotspots. Larger circles indicate more active hotspots.

Scientists are divided over how magma is generated in hotspots. Some suggest that hotspots originate from super-heated material from as deep as the core that reaches the Earth's crust as a mantle plume [84]. Others argue the molten material that feeds hotspots is sourced from the mantle [85]. Of course, it is difficult to collect data from these deep-Earth features due to the extremely high pressure and temperature [86].

How hotspots are initiated is another highly debated subject. The prevailing mechanism has hotspots starting in divergent boundaries during supercontinent rifting [87]. Scientists have identified a number of current and past hotspots believed to have begun this way. Subducting slabs have also been named as causing mantle plumes and hot-spot volcanism [88]. Some geologists have suggested another geological process not involving plate tectonics may be involved, such as large space objects crashing into the earth [89]. Regardless of how they are formed, dozens are on the Earth. Some well-known examples include the Tahiti Islands, Afar Triangle, Easter Island, Iceland, Galapagos Islands, and the Samoan Islands. The United States is home to two of the largest and best-studied hotspots: Hawaii and Yellowstone.

Hawaiian Hotspot

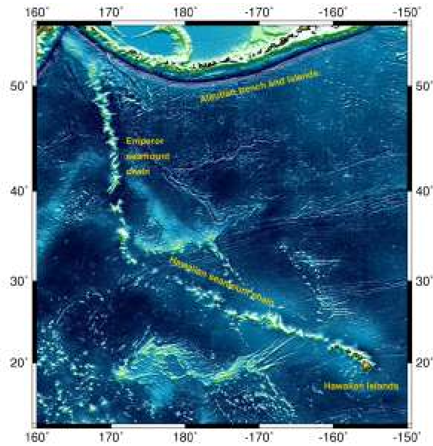


Figure 2.7.3: The Hawaii-Emperor seamount and island chain.

The active volcanoes at the end of the represent one of the most active hotspot sites on earth. Scientific evidence indicates the Hawaiian hotspot is at least 80 million years old [90]. Geologists believe it is actually much older. However, any rocks with proof of this have been subducted under the ocean floor. The big island of Hawaii sits atop a large mantle plume that marks the active hotspot. The Kilauea volcano is the main vent for this hotspot and has been actively erupting since 1983.

This enormous volcanic island chain, much of which is underwater, stretches across the Pacific for almost 6,000 km. The seamount chain's most striking feature is a sharp 60-degree bend located at the midpoint, which marks a significant change in plate movement direction that occurred 50 million years ago. The change in direction has been more often linked to a plate reconfiguration [91], but also to other things like plume migration [83].

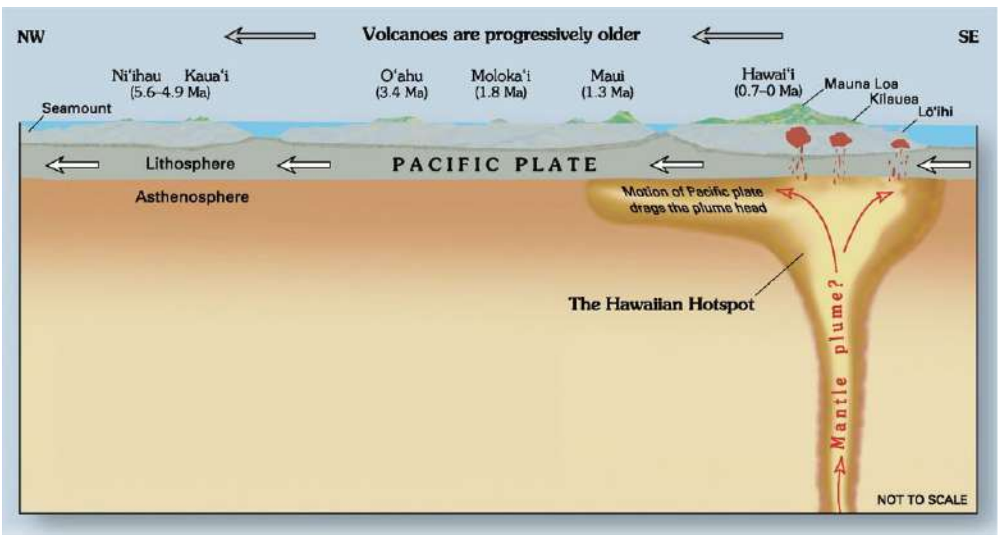


Figure 2.7.4:: Diagram of the Hawaiian hotspot and islands that it formed.

In an attempt to map the Hawaiian mantle plume as far down as the lower mantle [92], scientists have used tomography, a type of three-dimensional seismic imaging. This information—along with other evidence gathered from rock ages, vegetation types, and island size—indicates the oldest islands in the chain are located the furthest away from the active hotspot.

Yellowstone Hotspot

Like the Hawaiian version, the Yellowstone hotspot is formed by magma rising through the lithosphere. What makes it different is this hotspot is located under a thick, continental plate. Hawaii sits on a thin oceanic plate, which is easily breached by magma coming to the surface. At Yellowstone, the thick continental plate presents a much more difficult barrier for magma to penetrate. When it does emerge, the eruptions are generally much more violent. Thankfully they are also less frequent.

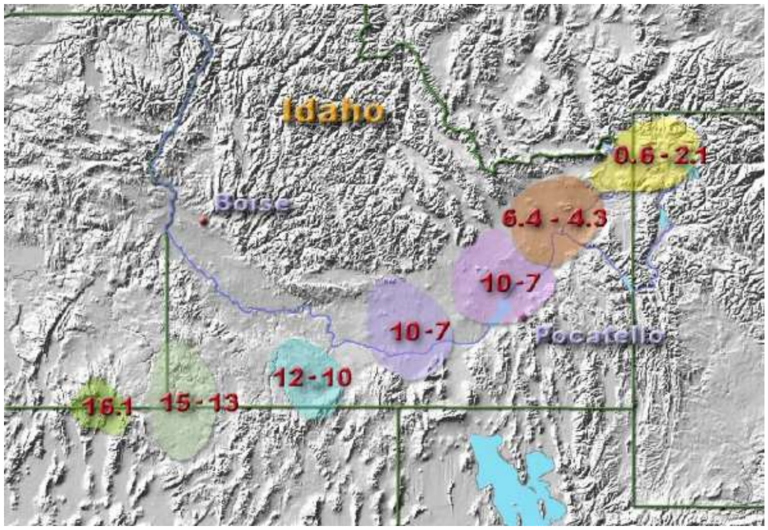


Figure 2.7.5:: The track of the Yellowstone hotspot which shows the age of different eruptions millions of years ago.

Over 15 million years of eruptions by this hotspot have carved a curved path across the western United States. It has been suggested the Yellowstone hotspot is connected to the much older Columbia River flood basalts [93] and even to 70 million-year-old volcanism found in the Yukon region of Canada [94].

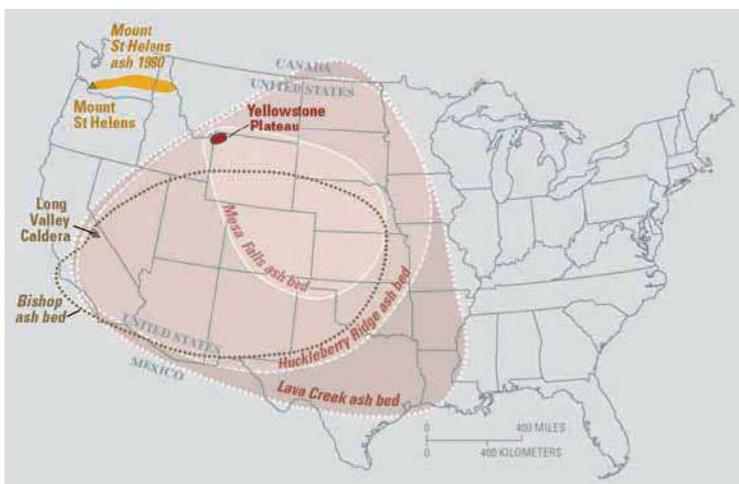


Figure 2.7.6:: Several prominent ash beds found in North America, including three Yellowstone eruptions which are shaded pink (Mesa Falls, Huckleberry Ridge, and Lava Creek), the Bisho Tuff ash bed (brown dashed line), and the modern May 18th, 1980 ash fall (yellow).

The most recent major eruption of this hotspot created the Yellowstone Caldera and Lava Creek tuff formation approximately 631,000 years ago [95]. The eruption threw 1,000 cubic kilometers of ash and magma into the atmosphere, some of which was found as far away as Mississippi. Should the hotspot erupt again, scientists predict it will be another massive event. This would be a calamity reaching far beyond the western United States. These super volcanic eruptions fill the earth's atmosphere with so much gas and ash, they block sunlight from reaching the earth. Not only would this drastically alter climates and environments around the globe, but it could also affect worldwide food production.

References

- 82. Wilson, J. T. A possible origin of the Hawaiian Islands. *Can. J. Phys.* **41**, 863–870 (1963).
- 83. Tarduno, J. A. *et al.* The Emperor Seamounts: southward motion of the Hawaiian hotspot plume in Earth's mantle. *Science* **301**, 1064–1069 (2003).
- 84. Morgan, W. J. Convection Plumes in the Lower Mantle. **230**, 42–43 (1971).
- 85. Courtillot, V., Davaille, A., Besse, J. & Stock, J. Three distinct types of hotspots in the Earth's mantle. *Earth Planet. Sci. Lett.* **205**, 295–308 (2003).
- 86. Foulger, G. R. *et al.* Caveats on tomographic images. *Terra Nova* **25**, 259–281 (2013).
- 87. Kerr, R. A. Rising plumes in Earth's mantle: phantom or real? *Science* **313**, 1726 (2006).
- 88. James, D. E., Fouch, M. J., Carlson, R. W. & Roth, J. B. Slab fragmentation, edge flow and the origin of the Yellowstone hotspot track. *Earth Planet. Sci. Lett.* **311**, 124–135 (2011).
- 89. Hagstrum, J. T. Antipodal hotspots and bipolar catastrophes: Were oceanic large-body impacts the cause? *Earth Planet. Sci. Lett.* **236**, 13–27 (2005).
- 90. Dalrymple, G. B., Silver, E. A. & Jackson, E. D. Origin of the Hawaiian Islands: Recent studies indicate that the Hawaiian volcanic chain is a result of relative motion between the Pacific plate and a melting spot in the Earth's mantle. *Am. Sci.* **61**, 294–308 (1973).
- 91. Whittaker, J. M. *et al.* Major Australian-Antarctic plate reorganization at Hawaiian-Emperor bend time. *Science* **318**, 83–86 (2007).
- 92. Ji, Y. & Nataf, H.-C. Detection of mantle plumes in the lower mantle by diffraction tomography: Hawaii. *Earth Planet. Sci. Lett.* **159**, 99–115 (1998).
- 93. Geist, D. & Richards, M. Origin of the Columbia Plateau and Snake River plain: Deflection of the Yellowstone plume. *Geology* **21**, 789–792 (1993).
- 94. Johnston, S. T. *et al.* Yellowstone in Yukon: The Late Cretaceous Carmacks Group. *Geology* **24**, 997–1000 (1996).
- 95. Matthews, N. E., Vazquez, J. A. & Calvert, A. T. Age of the Lava Creek supereruption and magma chamber assembly at Yellowstone based on 40Ar/39Ar and U-Pb dating of sanidine and zircon crystals. *Geochem. Geophys. Geosyst.* **16**, 2508–2528 (2015).

Contributions and Attributions

CHAPTER OVERVIEW

3: MINERALS

Learning Objectives

Define mineral.
Describe the basic structure of the atom.
Derive basic atomic information from the Periodic Table of Elements.
Describe chemical bonding related to minerals.
Describe the main ways minerals form.
Describe the silicon-oxygen tetrahedron and how it forms common silicate minerals.
List common non-silicate minerals in oxide, sulfide, sulfate, and carbonate groups.
Identify minerals using physical properties and identification tables.

The International Mineralogical Association in 1985 defined “A mineral is an element or chemical compound that is normally crystalline and that has been formed as a result of geological processes.” This means that the calcite in the shell of a clam is not considered a mineral. But once that clamshell undergoes burial, diagenesis, or other geological processes, then the calcite is considered a mineral. Typically, substances like coal, pearl, opal, or obsidian that do not fit the definition of a mineral are called mineraloids. A **rock** is a substance that contains one or more minerals or mineraloids. As is discussed in later chapters, there are three types of rocks composed of minerals: igneous (rocks crystallizing from molten material), sedimentary (rocks made of products of mechanical weathering (sand, gravel, etc.), chemical weathering (things precipitated from solution), and metamorphic (rocks produced by alteration of other rocks by heat and pressure.

3.1: Prelude to Minerals

The term “minerals” as used in nutrition labels and pharmaceutical products is not the same as a mineral in a geological sense. In geology, the classic definition of a **mineral** is:

1. Naturally occurring
2. Inorganic
3. Solid at room temperature
4. Regular crystal structure
5. Defined chemical composition

Some natural substances technically should not be considered minerals, but are included by exception. For example, water and mercury are liquid at room temperature. Both are considered minerals because they were classified before the room-temperature rule was accepted as part of the definition. Calcite is quite often formed by organic processes but is considered a mineral because it is widely found and geologically important. Because of these discrepancies, the International Mineralogical Association in 1985 amended the definition to: “A mineral is an element or chemical compound that is normally crystalline and that has been formed as a result of geological processes.” This means that the calcite in the shell of a clam is not considered a mineral. But once that clamshell undergoes burial, diagenesis, or other geological processes, then the calcite is considered a mineral. Typically, substances like coal, pearl, opal, or obsidian that do not fit the definition of a mineral are called mineraloids.

Figure 3.1.1: These selenite (gypsum) crystals, found in The Cave of the Crystals in Naica, Mexico, has some of the largest minerals ever found. The largest crystal found here is 39 feet (12 meters) and 55 tones.

A **rock** is a substance that contains one or more minerals or mineraloids. As is discussed in later chapters, there are three types of rocks composed of minerals: igneous (rocks crystallizing from molten material), sedimentary (rocks made of products of mechanical weathering (sand, gravel, etc.), chemical weathering (things precipitated from solution), and metamorphic (rocks produced by alteration of other rocks by heat and pressure).

3.2: Chemistry of Minerals

Rocks are composed of minerals that have a specific chemical composition. To understand mineral chemistry, it is essential to examine the fundamental unit of all matter, the atom.

The Atom

Matter is made of atoms. Atoms consist of subatomic particles—**protons**, **neutrons**, and **electrons**. A simple model of the atom has a central nucleus composed of protons, which have positive charges, and neutrons which have no charge. A cloud of negatively charged electrons surrounds the nucleus, the number of electrons equaling the number of protons thus balancing the positive charge of the protons for a neutral atom. Protons and neutrons each have a mass number of 1. The mass of an electron is less than 1/1000th that of a proton or neutron, meaning most of the atom's mass is in the nucleus.

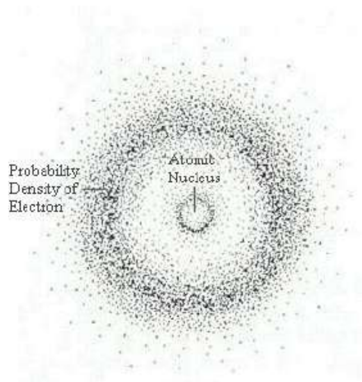


Figure 3.2.1: Electron cloud model of the atom

Periodic Table of the Elements

Matter is composed of elements which are atoms that have a specific number of protons in the nucleus. This number of protons is called the **Atomic Number** for the element. For example, an oxygen atom has 8 protons and an iron atom has 26 protons. An element cannot be broken down chemically into a simpler form and retains unique chemical and physical properties. Each element behaves in a unique manner in nature. This uniqueness led scientists to develop a periodic table of the elements, a tabular arrangement of all known elements listed in order of their atomic number.

PERIODIC TABLE OF ELEMENTS

1 H Atomic Number Symbol

1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	*	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	**	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn	113 Nh	114 Fl	115 Mc	116 Lv	117 Ts	118 Og
		*	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu
		**	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr

Figure 3.2.2: The Periodic Table of the Elements

The first arrangement of elements into a periodic table was done by Dmitri Mendeleev in 1869 using the elements known at the time [1]. In the periodic table, each element has a chemical symbol, name, atomic number, and atomic mass. The chemical symbol is an abbreviation for the element, often derived from a Latin or Greek name for the substance [2]. The atomic number is the number of protons in the nucleus. The atomic mass is the number of protons and neutrons in the nucleus, each with a mass number of one. Since the mass of electrons is so much less than the protons and neutrons, the atomic mass is effectively the number of protons plus neutrons.

$$^{14}_{7}N + ^{1}_{0}n \rightarrow ^{14}_{6}C + ^{1}_{1}p$$

Figure 3.2.3: Formation of Carbon 14 from Nitrogen 14

The atomic mass of natural elements represents an average mass of the atoms comprising that substance in nature and is usually not a whole number as seen on the periodic table, meaning that an element exists in nature with atoms having different numbers of neutrons. The differing number of neutrons affects the mass of an element in nature and the atomic mass number represents this average. This gives rise to the concept of 'isotope'. **Isotopes** are forms of an element with the same number of protons but different numbers of neutrons. There are usually several isotopes for a particular element. For example, 98.9% of carbon atoms have 6 protons and 6 neutrons. This isotope of carbon is called carbon-12 (^{12}C). A few carbon atoms, carbon-13 (^{13}C), have 6 protons and 7 neutrons. A trace amount of carbon atoms, carbon-14 (^{14}C), has 6 protons and 8 neutrons. Among the 118 known elements, the heaviest are fleeting human creations known only in high energy particle accelerators, and they decay rapidly. The heaviest naturally occurring element is uranium, atomic number 92. The eight most abundant elements in Earth's continental crust are shown in Table 3.2.1: [3; 4]. These elements are found in the most common rock-forming minerals.

Table 3.2.1:: Eight Most Abundant Elements in the Earth's Continental Crust % by weight (Source: USGS). All other elements are less than 1%..

Element	Symbol	Abundance %
Oxygen	O	47%
Silicon	Si	28%
Aluminum	Al	8%
Iron	Fe	5%
Calcium	Ca	4%
Sodium	Na	3%
Potassium	K	3%
Magnesium	Mg	2%

Chemical Bonding

Most substances on Earth are compounds containing multiple elements. Chemical bonding describes how these atoms attach with each other to form compounds, such as sodium and chlorine combining to form NaCl, common table salt. Compounds that are held together by chemical bonds are called molecules. Water is a compound of hydrogen and oxygen in which two hydrogen atoms are covalently bonded with one oxygen making the water molecule. The oxygen we breathe is formed when one oxygen atom covalently bonds with another oxygen atom to make the molecule O_2. The subscript 2 in the chemical formula indicates the molecule contains two atoms of oxygen.

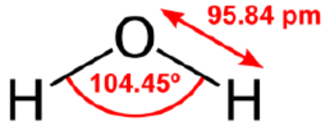


Figure 3.2.4: A model of a water molecule, showing the bonds between the hydrogen and oxygen.

Most minerals are also compounds of more than one element. The common mineral calcite has the chemical formula $CaCO_3$ indicating the molecule consists of one calcium, one carbon, and three oxygen atoms. In calcite, one carbon and three oxygen atoms are held together by covalent bonds to form a **molecular ion**, called carbonate, which has a negative charge. Calcium as an **ion** has a positive charge of plus two. The two oppositely charged ions attract each other and combine to form the mineral calcite, $CaCO_3$. The name of the chemical compound is calcium carbonate, where calcium is Ca and carbonate refers to the molecular ion CO_3^{-2}.

The mineral olivine has the chemical formula $(Mg,Fe)_2SiO_4$, in which one silicon and four oxygen atoms are bonded with two atoms of either magnesium or iron. The comma between iron (Fe) and magnesium (Mg) indicates the two elements can occupy the same location in the crystal structure and substitute for one another.

Valence and Charge

The electrons around the atom's nucleus are located in shells representing different energy levels. The outermost shell is called the **valence shell**. Electrons in the valence shell are involved in chemical bonding. In 1913, Niels Bohr proposed a simple model of the atom that states atoms are more stable when their outermost shell is full [5; 6]. Atoms of most elements thus tend to gain or lose electrons so the outermost or valence shell is full. In Bohr's model, the innermost shell can have a maximum of two electrons and the second and third shells can have a maximum of eight electrons. When the innermost shell is the valence shell, as in the case of hydrogen and helium, it obeys the octet rule when it is full with two electrons. For elements in higher rows, the octet rule of eight electrons in the valence shell applies.

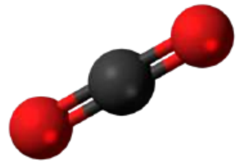


Figure 3.2.5: The carbon dioxide molecule. Since Oxygen is -2 and Carbon is +4, the two oxygens

The rows in the periodic table present the elements in order of atomic number and the columns organize elements with similar characteristics, such as the same number of electrons in their valence shells. Columns are often labeled from left to right with Roman numerals I to VIII, and Arabic numerals 1 through 18. The elements in columns I and II have 1 and 2 electrons in their respective valence shells and the elements in columns VI and VII have 6 and 7 electrons in their respective valence shells.

In row 3 and column I, sodium (Na) has 11 protons in the nucleus and 11 electrons in three shells—2 electrons in the inner shell, 8 electrons in the second shell, and 1 electron in the valence shell. To maintain a full outer shell of 8 electrons per the octet rule, sodium readily gives up that 1 electron so there are 10 total electrons. With 11 positively charged protons in the nucleus and 10 negatively charged electrons in two shells, sodium when forming chemical bonds is an ion with an overall net charge of +1.

All elements in column I have a single electron in their valence shell and a valence of 1. These other columns I elements also readily give up this single valence electron and thus become ions with a +1 charge. Elements in column II readily give up 2 electrons and end up as ions with a charge of +2. Note that elements in columns I and II which readily give up their valence electrons, often form bonds with elements in columns VI and VII which readily take up these electrons. Elements in columns 3 through 15 are usually involved in covalent bonding. The last column 18 (VIII) contains the **noble gases**. These elements are chemically inert because the valence shell is already full with 8 electrons, so they do not gain or lose electrons. An example is the noble gas helium which has 2 valence electrons in the first shell. Its valence shell is therefore full. All elements in column VIII possess full valence shells and do not form bonds with other elements.

As seen above, an atom with a net positive or negative charge as a result of gaining or losing electrons is called an **ion**. In general, the elements on the left side of the table lose electrons and become positive ions, called cations because they are attracted to the cathode in an electrical device. The elements on the right side tend to gain electrons. These are called anions because they are attracted to the anode in an electrical device. The elements in the center of the periodic table, columns 3 through 15, do not consistently follow the octet rule. These are called transition elements. A common example is iron, which has a +2 or +3 charge depending on the oxidation state of the element. Oxidized Fe^{+3} carries a +3 charge and reduced Fe^{+2} is +2. These two different oxidation states of iron often impart dramatic colors to rocks containing their minerals—the oxidized form producing red colors and the reduced form producing green.

Ionic Bonding

Ionic bonds, also called electron-transfer bonds, are formed by the electrostatic attraction between atoms having opposite charges. Atoms of two opposite charges attract each other electrostatically and form an **ionic bond** in which the positive ion transfers its electron (or electrons) to the negative ion which takes them up. Through this transfers both atoms thus achieve a full valence shell. For example one atom of sodium (Na^{+1}) and one atom of chlorine (Cl^{-1}) form an ionic bond to make the compound sodium chloride (NaCl). This is also known as the mineral halite or common table salt. Another example is calcium (Ca^{+2}) and chlorine (Cl^{-1}) combining to make the compound calcium chloride ($CaCl_2$). The subscript 2 indicates two atoms of chlorine are ionically bonded to one atom of calcium.

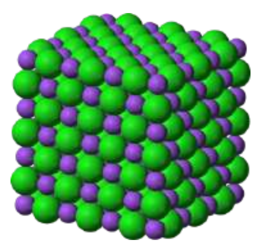


Figure 3.2.6: Cubic arrangement of Na and Cl ions in Halite

Covalent Bonding

Ionic bonds are usually formed between a **metal** and a **nonmetal**. Another type, called a covalent or electron-sharing bond, commonly occurs between nonmetals. Covalent bonds share electrons between ions to complete their valence shells. For example, oxygen (atomic number 8) has 8 electrons—2 in the inner shell and 6 in the valence shell. Gases like oxygen often form diatomic molecules by sharing valence electrons. In the case of oxygen, two atoms attach to each other and share 2 electrons to fill their valence shells to become the common oxygen molecule we breathe (O_2). Methane (CH_4) is another covalently bonded gas. The carbon atom needs 4 electrons and each hydrogen needs 1. Each hydrogen shares its electron with the carbon to form a molecule as shown in the figure.

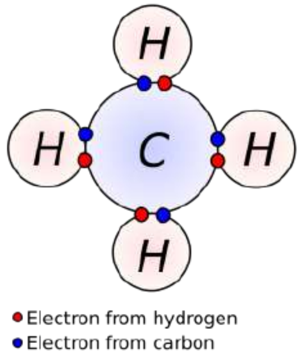


Figure 3.2.7: Methane molecule

References

- Mendeleev, D. The relation between the properties and atomic weights of the elements. *Journal of the Russian Chemical Society* **1**, 60–77 (1869).
- Scerri, E. R. *The Periodic Table: Its Story and Its Significance*. (Oxford University Press, USA, 2007).
- Clarke, F. W. H. S. W. *The Composition of the Earth's Crust*. **127**, (United States Geological Survey, 1927).
- Hans Wedepohl, K. The composition of the continental crust. *Geochim. Cosmochim. Acta* **59**, 1217–1232 (1995).
- *Science literacy: concepts, contexts, and consequences*. (National Academies Press (US), 2016).
- Bohr, N. On the constitution of atoms and molecules. *Philos. Mag.* **26**, 1–24 (1913).

Contributions and Attributions

3.3: Formation of Minerals

Minerals form when atoms bond together in a crystalline arrangement. Three main ways this occurs in nature are:

1. Precipitation directly from an aqueous (water) solution with a temperature change
2. Crystallization from a magma with a temperature change
3. Biological precipitation by the action of organisms

Precipitation from Aqueous Solution

Solutions consist of ions or molecules, known as solutes, dissolved in a medium or solvent. In nature, this solvent is usually water. Many minerals can be dissolved in water, such as halite or table salt, which has the composition sodium chloride, NaCl. The Na^{+1} and Cl^{-1} ions separate and disperse into the solution.

Figure 3.3.1: Calcium carbonate deposits from hard water on a faucet

Precipitation is the reverse process, in which ions in solution come together to form solid minerals. Precipitation is dependent on the concentration of ions in solution and other factors such as temperature and pressure. The point at which a solvent cannot hold any more solute is called saturation. Precipitation can occur when the temperature of the solution falls, when the solute evaporates, or with changing chemical conditions in the solution. An example of precipitation in our homes is when water evaporates and leaves behind a rind of minerals on faucets, showerheads, and drinking glasses.

In nature, changes in environmental conditions may cause the minerals dissolved in water to form bonds and grow into crystals or cement grains of sediment together. In Utah, deposits of tufa formed from mineral-rich springs that emerged into the ice age Lake Bonneville. Now exposed in dry valleys, this porous tufa was a natural insulation used by pioneers to build their homes with a natural protection against summer heat and winter cold. The travertine terraces at Mammoth Hot Springs in Yellowstone Park are another example formed by calcite precipitation at the edges of the shallow spring-fed ponds.

Figure 3.3.2: The Bonneville Salt Flats of Utah

Another example of precipitation occurs in the Great Salt Lake, Utah, where the concentration of sodium chloride and other salts is nearly eight times greater than in the world's oceans [7]. Streams carry salt ions into the lake from the surrounding mountains. With no other outlet, the water in the lake evaporates and the concentration of salt increases until saturation is reached and the minerals precipitate out as sediments. Similar salt deposits include halite and other precipitates, and occur in other lakes like Mono Lake in California and the Dead Sea.

Crystallization from Magma

Heat is energy that causes atoms in substances to vibrate. Temperature is a measure of the intensity of the vibration. If the vibrations are violent enough, chemical bonds are broken and the crystals melt releasing the ions into the melt. Magma is molten rock with freely moving ions. When magma is emplaced at depth or extruded onto the surface (then called lava), it starts to cool and mineral crystals can form.

Figure 3.3.3: Lava, magma at the earth's surface

Precipitation by Organisms

Many organisms build bones, shells, and body coverings by extracting ions from water and precipitating minerals biologically. The most common mineral precipitated by organisms is calcite, or calcium carbonate (CaCO3). Calcite is often precipitated by organisms as a polymorph called aragonite. **Polymorphs** are crystals with the same chemical formula but different crystal structures. Marine invertebrates such as corals and clams precipitate aragonite or calcite for their shells and structures. Upon death, their hard parts accumulate on the ocean floor as sediments and eventually may become the sedimentary rock limestone. Though limestone can form inorganically, the vast majority is formed by this biological process.

Figure 3.3.4: Ammonite shell made of calcium carbonate

Another example is marine organisms called radiolaria, which are zooplankton that precipitates silica for their microscopic external shells. When the organisms die, the shells accumulate on the ocean floor and can form the sedimentary rock chert. An example of biologic precipitation from the vertebrate world is bone, which is composed mostly of a type of apatite, a mineral in the phosphate group. The apatite found in bones contains calcium and water in its structure and is called **hydroxycarbonate apatite**, $\mathrm{Ca_5(PO_4)_3(OH)}$. As mentioned above, such substances are not technically minerals until the organism dies and these hard parts become fossils.

References

- 7. Utah Geological Survey. Commonly Asked Questions About Utah's Great Salt Lake & Lake Bonneville. (2017). Available at: http://geology.utah.gov/popular/general-geology/great-salt-lake/commonly-asked-questions-about-utahs-great-salt-lake-lake-bonneville/. (Accessed: 26th February 2017)

Contributions and Attributions

3.4: Silicate Minerals

Minerals are categorized based on their composition and structure. Silicate minerals are built around a molecular ion called the **silicon-oxygen tetrahedron**. A tetrahedron has a pyramid-like shape with four sides and four corners. Silicate minerals form the largest group of minerals on Earth, comprising the vast majority of the Earth's mantle and crust. Of the nearly four thousand known minerals on Earth, most are rare. There are only a few that make up most of the rocks likely to be encountered by surface dwelling creatures like us. These are generally called the **rock-forming minerals**.

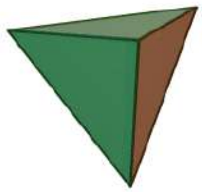


Figure 3.4.1: Rotating animation of a tetrahedron

The silicon-oxygen tetrahedron (SiO_4) consists of a single silicon atom at the center and four oxygen atoms located at the four corners of the tetrahedron. Each oxygen ion has a -2 charge and the silicon ion has a +4 charge. The silicon ion shares one of its four valence electrons with each of the four oxygen ions in a covalent bond to create a symmetrical geometric four-sided pyramid figure. Only half of the oxygen's valence electrons are shared, giving the silicon-oxygen tetrahedron an ionic charge of -4. This silicon-oxygen tetrahedron forms bonds with many other combinations of ions to form the large group of silicate minerals.

Figure 3.4.2: Ping pong ball model of a tetrahedron: balls are oxygen, small space in the center is silicon

The silicon ion is much smaller than the oxygen ions (see the figures) and fits into a small space in the center of the four large oxygen ions, see if the top ball is removed (as shown in the figure to the right). Because only one of the valence electrons of the corner oxygens is shared, the silicon-oxygen tetrahedron has chemically active corners available to form bonds with other silica tetrahedra or other positively charged ions such as Al_{+3}, $Fe_{+2,+3}$, Mg_{+2}, K_{+1}, Na_{+1}, and Ca_{+2}. Depending on many factors, such as the original magma chemistry, silica-oxygen tetrahedra can combine with other tetrahedra in several different configurations. For example, tetrahedra can be isolated, attached in chains, sheets, or three-dimensional structures. These combinations and others create the chemical structure in which positively charged ions can be inserted for unique chemical compositions forming silicate mineral groups.

Figure 3.4.3: The silicon atom in the center of the tetrahedron (with the top oxygen removed) represented by a metal ball

The Dark Ferromagnesian Silicates

Figure 3.4.4: Green olivine in basalt

The Olivine Family

Olivine is the primary mineral component in mantle rock such as peridotite and basalt. It is characteristically green when not weathered. The chemical formula is $(Fe,Mg)_2SiO_4$. As previously described, the comma between iron (Fe) and magnesium (Mg) indicates these two elements occur in a solid solution. Not to be confused with a liquid solution, a solid solution occurs when two or more elements have similar properties and can freely substitute for each other in the same location in the crystal structure.

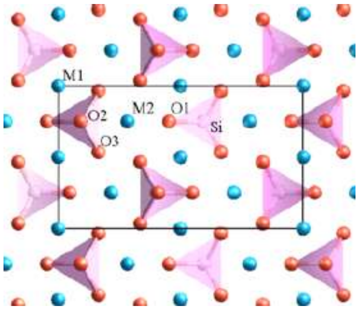


Figure 3.4.5: Tetrahedral structure of olivine

Olivine is referred to as a mineral family because of the ability of iron and magnesium to substitute for each other. Iron and magnesium in the olivine family indicate a solid solution forming a compositional series within the mineral group which can form crystals of all iron as one end member and all mixtures of iron and magnesium in between to all magnesium at the other end member. Different mineral names are applied to compositions between these end members. In the olivine series of minerals, the iron and magnesium ions in the solid solution are about the same size and charge, so either atom can fit into the same location in the growing crystals. Within the cooling magma, the mineral crystals continue to grow until they solidify into igneous rock. The relative amounts of iron and magnesium in the parent magma determine which minerals in the series form. Other rarer elements with similar properties to iron or magnesium, like manganese (Mn), can substitute into the olivine crystalline structure in small amounts. Such ionic substitutions in mineral crystals give rise to the great variety of minerals and are often responsible for differences in color and other properties within a group or family of minerals. Olivine has a pure iron end-member (called fayalite) and a pure magnesium end-member (called forsterite). Chemically, olivine is mostly silica, iron, and magnesium and therefore is grouped among the dark-colored ferromagnesian (iron=ferro, magnesium=magnesian) or **mafic** minerals, a contraction of their chemical symbols Ma and Fe. Mafic minerals are also referred to as dark-colored ferromagnesian minerals. *Ferro* means iron and *magnesian* refers to magnesium.

The crystal structure of olivine is built from independent silica tetrahedra. Minerals with independent tetrahedral structures are called neosilicates (or orthosilicates). In addition to olivine, other common neosilicate minerals include garnet, topaz, kyanite, and zircon.

Two other similar arrangements of tetrahedra are close in structure to the neosilicates and grade toward the next group of minerals, the pyroxenes. In a variation on independent tetrahedra called sorosilicates, there are minerals that share one oxygen between two tetrahedra and include minerals like pistachio-green epidote, a gemstone. Another variation are the cyclosilicates, which as the name suggests, consist of tetrahedral rings, and include gemstones such as beryl, emerald, aquamarine, and tourmaline

Pyroxene Family

Figure 3.4.6: Crystals of diopside, a member of the pyroxene family

Pyroxene is another family of dark ferromagnesian minerals, typically black or dark green in color. Members of the pyroxene family have a complex chemical composition that includes iron, magnesium, aluminum, and other elements bonded to polymerized silica tetrahedra. **Polymers** are chains, sheets, or three-dimensional structures, and are formed by multiple tetrahedra covalently bonded via their corner oxygen atoms. Pyroxenes are commonly found in mafic igneous rocks such as peridotite, basalt, and gabbro, as well as metamorphic rocks like eclogite and blue-schist.

Figure 3.4.7: Single chain

Pyroxenes are built from long, single chains of polymerized silica tetrahedra in which tetrahedra share two corner oxygens. The silica chains are bonded together into the crystal structures by metal cations. A common member of the pyroxene family is augite, itself containing several solid solution series with a complex chemical formula $(Ca,Na)(Mg,Fe,Al,Ti)(Si,Al)_2O_6$ that gives rise to a number of individual mineral names.

This single-chain crystalline structure bonds with many elements, which can also freely substitute for each other. The generalized chemical composition for pyroxene is $XZ(Al,Si)_2O_6$. X represents the ions Na, Ca, Mg, or Fe, and Z represents Mg, Fe, or Al. These ions have similar ionic sizes, which allows many possible substitutions among them. Although the cations may freely substitute for each other in the crystal, they carry different ionic charges that must be balanced out in the final crystalline structure. For example, Na has a charge of +1, but Ca has a charge of +2. If a Na^{+} ion substitutes for a Ca^{+2} ion, it creates an unequal charge that must be balanced by other ionic substitutions elsewhere in the crystal. Note that ionic size is more important than ionic charge for substitutions to occur in solid solution series in crystals.

Amphibole Family

Figure 3.4.1: Hornblende crystals

Figure 3.4.8: Elongated crystals of hornblende in orthoclase

Amphibole minerals are built from polymerized double silica chains and they are also referred to as inosilicates. Imagine two pyroxene chains that connect together by sharing the third oxygen on each tetrahedron. Amphiboles are usually found in igneous and metamorphic rocks and typically have a long-bladed **crystal habit**. The most common amphibole, hornblende, is usually black; however, they come in a variety of colors depending on their chemical composition. The metamorphic rock, amphibolite, is primarily composed of amphibole minerals.

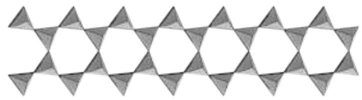


Figure 3.4.9: Double chain structure

Amphiboles are composed of iron, magnesium, aluminum, and other cations bonded with silica tetrahedra. These dark ferromagnesian minerals are commonly found in gabbro, basalt, diorite, and often form the black specks in granite. Their chemical formula is very complex and generally written as $(RSi_4O_{11})_2$, where R represents many different cations. For example, it can also be written more exactly as $AX_2Z_5((Si,Al,Ti)_8O_{22})(OH,F,Cl,O)_2$. In this formula A may be Ca, Na, K, Pb, or blank; X equals Li, Na, Mg, Fe, Mn, or Ca; and Z is Li, Na, Mg, Fe, Mn, Zn, Co, Ni, Al, Cr, Mn, V, Ti, or Zr. The substitutions create a wide variety of colors such as green, black, colorless, white, yellow, blue, or brown. Amphibole crystals can also include hydroxide ions (OH^-), which occurs from an interaction between the growing minerals and water dissolved in the magma.

Sheet Silicates

Figure 3.4.10: (left) Sheet crystals of biotite mica. (right) A stack of sheets of muscovite mica

Sheet silicates are built from tetrahedra which share all three of their bottom corner oxygens thus forming sheets of tetrahedra with their top corners available for bonding with other atoms. Micas and clays are common types of sheet silicates, also known as phyllosilicates. Mica minerals are usually found in igneous and metamorphic rocks, while clay minerals are more often found in sedimentary rocks. Two frequently found micas are dark-colored biotite, frequently found in granite, and light-colored muscovite, found in the metamorphic rock called schist.

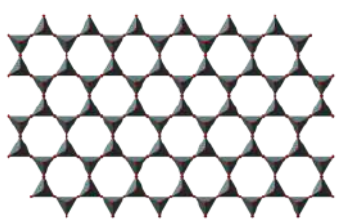


Figure 3.4.11: Sheet structure of mica, view perpendicular to the sheets

Chemically, sheet silicates usually contain silicon and oxygen in a 2:5 ratio (Si_4O_{10}). Micas contain mostly silica, aluminum, and potassium. Biotite mica has more iron and magnesium and is considered a ferromagnesian silicate mineral. Muscovite micas belong to the felsic silicate minerals. Felsic is a contraction formed from feldspar, the dominant mineral in felsic rocks.

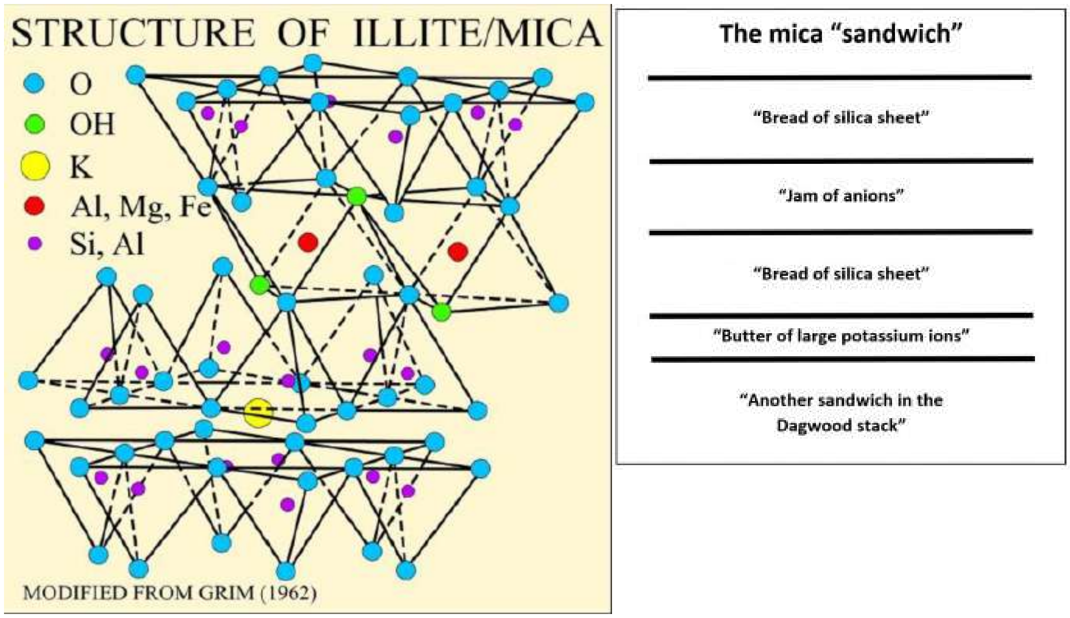


Figure 3.4.12: (left) Crystal structure of a mica, view parallel to the sheets. (right) Mica "silica sandwich" structure related to layers in illite structure.

The illustration of the crystalline structure of mica shows the corner O atoms bonded with K, Al, Mg, Fe, and Si atoms, forming polymerized sheets of linked tetrahedra, with an octahedral layer of Fe, Mg, or Al, between them. The yellow potassium ions form Van der Waals bonds (attraction and repulsion between atoms, molecules, and surfaces) and hold the sheets together. Van der Waals bonds differ from covalent and ionic bonds, and exist here between the sandwiches, holding them together into a stack of sandwiches. The Van der Waals bonds are weak compared to the bonds within the sheets, allowing the sandwiches to be separated along the potassium layers. This gives mica its characteristic property of easily cleaving into sheets.

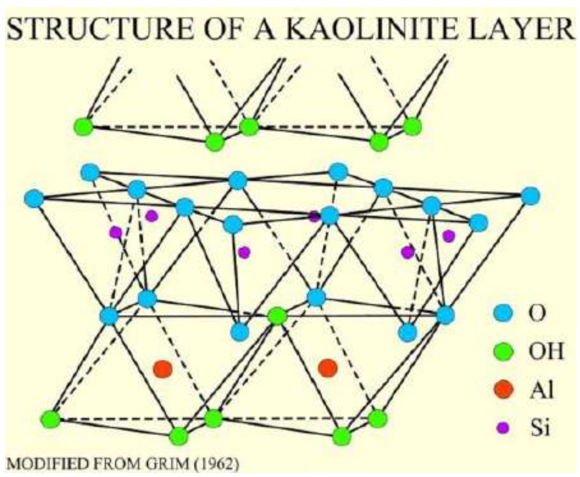


Figure 3.4.13: Structure of kaolinite

Clays minerals occur in sediments formed by the weathering of rocks and are another family of silicate minerals with a tetrahedral sheet structure. Clay minerals form a complex family and are an important component of many sedimentary rocks. Other sheet silicates include serpentine and chlorite, found in metamorphic rocks.

Clay minerals are composed of hydrous aluminum silicates. One type of clay, kaolinite, has a structure like an open-faced sandwich, with the bread being a single layer of silicon-oxygen tetrahedra and a layer of aluminum as the spread in an octahedral configuration with the top oxygens of the sheets.

Framework Silicates

Figure 3.4.14: Quartz crystals

Quartz and feldspar are the two most abundant minerals in the continental crust. In fact, feldspar itself is the single most abundant mineral in the Earth’s crust. There are two types of feldspar, one containing potassium and abundant in felsic rocks of the continental crust, and the other with sodium and calcium abundant in the mafic rocks of oceanic crust. Together with quartz, these minerals are classified as framework silicates. They are built with a three-dimensional framework of silica tetrahedra in which all four corner oxygens are shared with adjacent tetrahedra. Within these frameworks in feldspar are holes and spaces into which other ions like aluminum, potassium, sodium, and calcium can fit giving rise to a variety of mineral compositions and mineral names. They are usually found in igneous rocks, such as granite, rhyolite, and basalt as well as metamorphic rocks and detrital sedimentary rocks. Detrital sedimentary rocks are composed of mechanically weathered rock particles, like sand and gravel. Quartz is especially abundant in detrital sedimentary rocks because it is very resistant to disintegration by weathering.

Figure 3.4.15: Pink orthoclase crystals

Quartz is composed of pure silica, SiO_2 with the tetrahedra arranged in a three-dimensional framework. Impurities consisting of atoms within this framework give rise to many varieties of quartz among which are gemstones like amethyst, rose quartz, and citrine. Feldspars are mostly silica with aluminum, potassium, sodium, and calcium. Orthoclase feldspar ($KAlSi_3O_8$), also called potassium feldspar or K-spar, is made of silica, aluminum, and potassium. Quartz and orthoclase feldspar are felsic minerals. Felsic is the compositional term applied to continental igneous minerals and rocks that contain an abundance of orthoclase feldspar. Another feldspar is plagioclase with the formula $(Ca,Na)AlSi_3O_8$, the solid solution (Ca,Na) indicating a series of minerals, one end of the series with calcium $CaAl_2Si_2O_8$, called anorthite, and the other end with sodium $NaAlSi_3O_8$, called albite. Note how the mineral accommodates the substitution of Ca^{++} and Na^{+}. Minerals in this solid solution series have different mineral names.

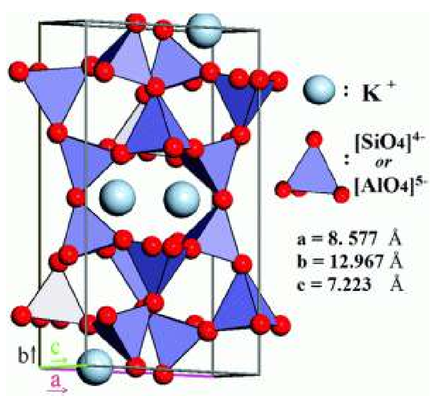


Figure 3.4.16: Crystal structure of feldspar

Note that aluminum, which has a similar ionic size to silicon, can substitute for silicon inside the tetrahedra (see figure). Because potassium ions are so much larger than sodium and calcium ions, which are very similar in size, the inability of the crystal lattice to accommodate both potassium and sodium/calcium gives rise to the two families of feldspar: orthoclase and plagioclase respectively. Framework silicates are called tectosilicates and include the alkali metal-rich feldspathoids and zeolites.

Contributions and Attributions

3.5: Non-Silicate Minerals

The crystal structure of non-silicate minerals (see table) does not contain silica-oxygen tetrahedra. Many non-silicate minerals are economically important and provide metallic resources such as copper, lead, and iron. They also include valuable non-metallic products such as salt, construction materials, and fertilizer.

Table 3.5.1: Common non-silicate mineral groups.

Mineral Group	Examples	Formula	Uses
Native elements	gold, silver, copper	Au, Ag, Cu	Jewelry, coins, industry
Carbonates	calcite, dolomite	$CaCO_3$, $CaMg(CO_3)_2$	Lime, Portland cement
Oxides	hematite, magnetite, bauxite	Fe_2O_3, Fe_3O_4, a mixture of aluminum oxides	Ores of iron & aluminum, pigments
Halides	halite, sylvite	NaCl, KCl	Table salt, fertilizer
Sulfides	galena, chalcopyrite, cinnabar	PbS, $CuFeS_2$, HgS	Ores of lead, copper, mercury
Sulphates	gypsum, epsom salts	$CaSo_4 \cdot 2H_2O$, $MgSO_4 \cdot 7H_2O$	Sheetrock, therapeutic soak
Phosphates	apatite	$Ca_5(PO_4)_3(F,Cl,OH)$	Fertilizer, teeth, bones

Carbonates

Figure 3.5.1: Calcite crystal in shape of rhomb. Note the double-refracted word “calcite” in the center of the figure due to birefringence.

Figure 3.5.2: Limestone full of small fossils

Calcite ($CaCO_3$) and dolomite ($CaMg(CO_3)_2$) are the two most frequently occurring carbonate minerals, and usually occur in sedimentary rocks, such as limestone and dolostone rocks, respectively. Some carbonate rocks, such as calcite and dolomite, are formed via evaporation and precipitation. However, most carbonate-rich rocks, such as limestone, are created by the lithification of fossilized marine organisms. These organisms, including those we can see and many microscopic organisms, have shells or exoskeletons consisting of calcium carbonate ($CaCO_3$). When these organisms die, their remains accumulate on the floor of the water body in which they live and the soft body parts decompose and dissolve away. The calcium carbonate hard parts become included in the sediments, eventually becoming the sedimentary rock called limestone. While limestone may contain large, easy to see fossils, most limestones contain the remains of microscopic creatures and thus originate from biological processes.

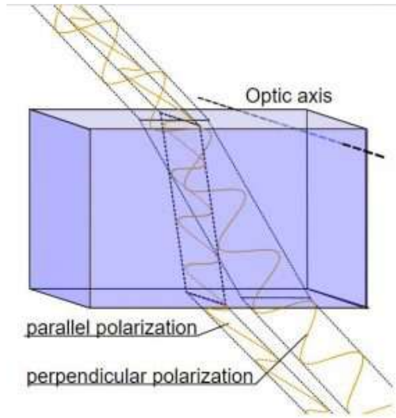


Figure 3.5.3: Birefringence in calcite crystals

Calcite crystals show an interesting property called **birefringence**, meaning they polarize light into two wave components vibrating at right angles to each other. As the two light waves pass through the crystal, they travel at different velocities and are separated by refraction into two different travel paths. In other words, the crystal produces a double image of objects viewed through it. Because they polarize light, calcite crystals are used in special petrographic microscopes for studying minerals and rocks.

Many non-silicate minerals are referred to as salts. The term **salts** used here refers to compounds made by replacing the hydrogen in natural acids. The most abundant natural acid is carbonic acid that forms by the solution of carbon dioxide in water. Carbonate minerals are salts built around the carbonate ion ($CO3^{-2}$) where calcium and/or magnesium replace the hydrogen in carbonic acid (H_2CO_3). Calcite and closely-related polymorph aragonite are secreted by organisms to form shells and physical structures like corals. Many such creatures draw both calcium and carbonate from dissolved bicarbonate ions (HCO_3^-) in ocean water. As seen in the mineral identification section below, calcite is easily dissolved in acid and thus effervesces in dilute hydrochloric acid (HCl). Small dropper bottles of dilute hydrochloric acid are often carried by geologists in the field as well as used in mineral identification labs.

Other salts include halite (NaCl) in which sodium replaces the hydrogen in hydrochloric acid and gypsum ($Ca[SO_4] \cdot 2\ H_2O$) in which calcium replaces the hydrogen in sulfuric acid. Note that some water molecules are also included in the gypsum crystal. Salts are often formed by evaporation and are called evaporite minerals.

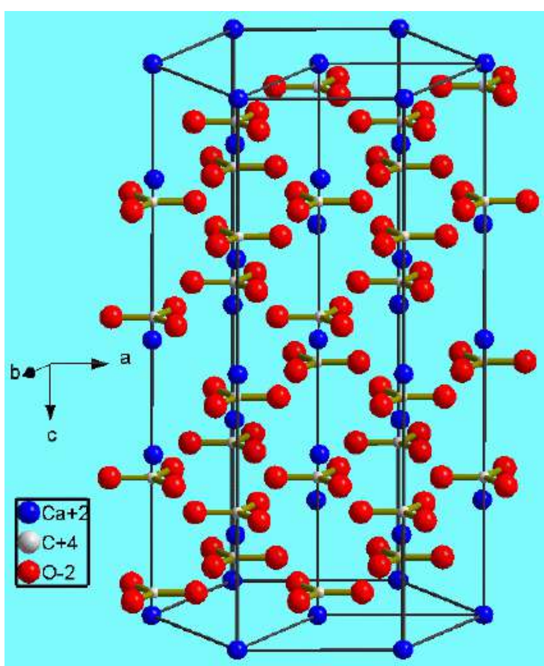


Figure 3.5.4: Crystal structure of calcite

The figure shows the crystal structure of calcite ($CaCO_3$). Like silicon, carbon has four valence electrons. The carbonate unit consists of carbon atoms (tiny white dots) covalently bonded to three oxygen atoms (red), one oxygen sharing two valence electrons with the carbon and the other two sharing one valence electron each with the carbon, thus creating triangular units with a charge of -2. The negatively charged carbonate unit forms an ionic bond with the Ca ion (blue), which as a charge of +2.

Oxides, Halides, and Sulfides

Figure 3.5.5: Limonite, hydrated oxide of iron

After carbonates, the next most common non-silicate minerals are the oxides, halides, and sulfides.

Oxides consist of metal ions covalently bonded with oxygen. The most familiar oxide is rust, which is a combination of iron oxides (Fe_2O_3) and hydrated oxides. Hydrated oxides form when the iron is exposed to oxygen and water. Iron oxides are important for producing metallic iron. When iron oxide or ore is smelted, it produces carbon dioxide (CO_2) and metallic iron.

The red color in rocks is usually due to the presence of iron oxides. For example, the red sandstone cliffs in Zion National Park and throughout Southern Utah consist of white or colorless grains of quartz coated with iron oxide which serve as cementing agents holding the grains together.

Figure 3.5.6: Oolitic hematite

Other iron oxides include limonite, magnetite, and hematite. Hematite occurs in many different crystal forms. The massive form shows no external structure. Botryoidal hematite shows large concentric blobs. Specular hematite looks like a mass of shiny metallic crystals. Oolitic hematite looks like a mass of dull red fish eggs. These different forms of hematite are polymorphs and all have the same formula, Fe_2O_3.

Other common oxide minerals include:

- Ice (H_2O), an oxide of hydrogen
- Bauxite ($Al_2H_2O_4$), hydrated oxides of aluminum, an ore for producing metallic aluminum
- Corundum (Al_2O_3), which includes ruby and sapphire gemstones.

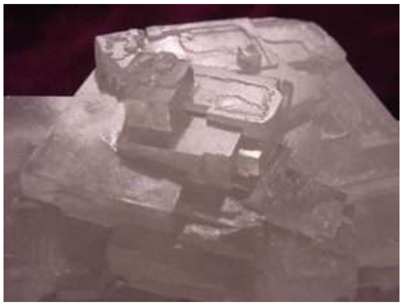


Figure 3.5.7: Halite crystal showing cubic habit

The **halides** consist of halogens in column VII, usually fluorine or chlorine, ionically bonded with sodium or other cations. These include halite or sodium chloride ($NaCl$), common table salt; sylvite or potassium chloride (KCl); and fluorite or calcium fluoride (CaF_2).

Figure 3.5.8: Salt crystals at the Bonneville Salt Flats

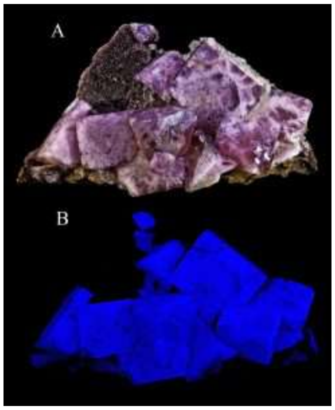


Figure 3.5.8: Fluorite. B shows fluorescence of fluorite under UV light

Halide minerals usually form from the evaporation of seawater or other isolated bodies of water. A well-known example of halide mineral deposits created by evaporation is the Bonneville Salt Flats, located west of the Great Salt Lake in Utah (see figure).

Figure 3.5.9: Cubic crystals of pyrite

Many important metal ores are **sulfides,** in which metals are bonded to sulfur. Significant examples include galena (lead sulfide), sphalerite (zinc sulfide), pyrite (iron sulfide, sometimes called “fool’s gold”), and chalcopyrite (iron-copper sulfide). Sulfides are well known for being important ore minerals. For example, galena is the main source of lead, sphalerite is the main source of zinc, and chalcopyrite is the main copper ore mineral mined in porphyry deposits like the Bingham mine (see chapter 16). The largest sources of nickel, antimony, molybdenum, arsenic, and mercury are also sulfides.

Sulfates

Figure 3.5.10: Gypsum crystal

Sulfate minerals contain a metal ion, such as calcium, bonded to a sulfate ion. The sulfate ion is a combination of sulfur and oxygen (SO_4^{-2}). The sulfate mineral gypsum ($CaSO_4 \cdot 2H_2O$) is used in construction materials such as plaster and drywall. Gypsum is often formed from evaporating water and usually contains water molecules in its crystalline structure. The $\cdot 2H_2O$ in the formula indicates the water molecules are whole H_2O. This is different from minerals like amphibole, which contain a hydroxide ion (OH^-) that is derived from water but is missing a hydrogen ion (H^+). The calcium sulfate without water is a different mineral than gypsum called anhydrite ($CaSO_4$).

Phosphates

Figure 3.5.11: Apatite crystal

Phosphate minerals have a tetrahedral phosphate unit (PO_4^{-3}) combined with various anions and cations. In some cases arsenic or vanadium can substitute for phosphorus. Phosphates are an important ingredient of fertilizers as well as detergents, paint, and other products. The best known phosphate mineral is apatite, $Ca_5(PO_4)_3(F,Cl,OH)$, variations of which are found in teeth and bones. The gemstone turquoise [$CuAl_6(PO_4)_4(OH)_8 \cdot 4H2O$] is a copper-rich phosphate mineral that, like gypsum, contains water molecules.

Native Element Minerals

Native element minerals, usually metals, occur in nature in a pure or nearly pure state. Gold is an example of a native element mineral; it is not very reactive and rarely bonds with other elements so it is usually found in an isolated or pure state. The non-metallic and poorly-reactive mineral carbon is often found as a native element, such as graphite and diamonds. Mildly reactive metals like silver, copper, platinum, mercury, and sulfur sometimes occur as native element minerals. Reactive metals such as iron, lead, and aluminum almost always bond to other elements and are rarely found in a native state.

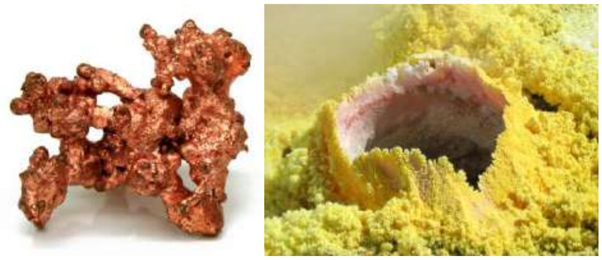


Figure 3.5.12: (left) Native copper. (right) Native sulfur deposited around a volcanic fumarole

Contributions and Attributions

3.6: Identifying Minerals

Geologists identify minerals by their physical properties. In the field, where geologists may have limited access to advanced technology and powerful machines, they can still identify minerals by testing several physical properties: luster and color, streak, hardness, crystal habit, cleavage and fracture, and some special properties. Only a few common minerals make up the majority of Earth's rocks and are usually seen as small grains in rocks. Of the several properties used for identifying minerals, it is good to consider which will be most useful for identifying them in small grains surrounded by other minerals.

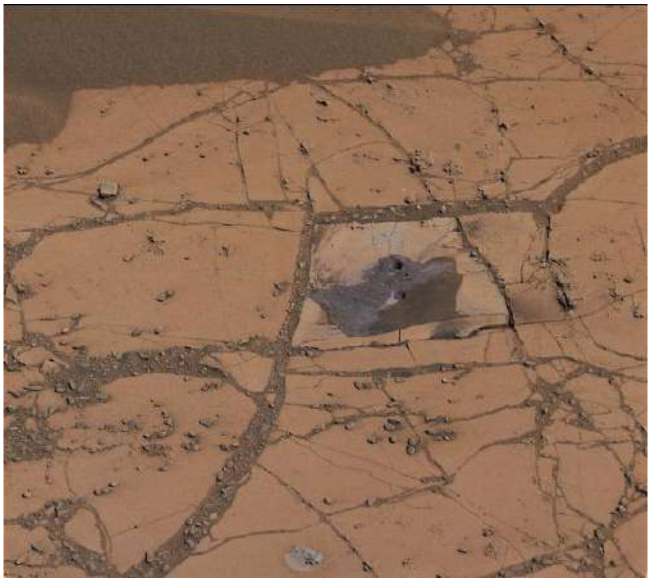


Figure 3.6.1: The rover Curiosity drilled a hole in this rock from Mars, and confirmed the mineral Hematite, as mapped from satellites.

Luster and Color

Figure 3.6.2: 15 mm metallic hexagonal molybdenite crystal from Quebec.

The first thing to notice about a mineral is its surface appearance, specifically luster and color. Luster describes how the mineral looks. Metallic luster looks like a shiny metal such as chrome, steel, silver, or gold. Submetallic luster has a duller appearance. Pewter, for example, shows submetallic luster.

Figure 3.6.3: Submetallic luster shown on an antique pewter plate.

Nonmetallic luster doesn't look like metal and may be described as vitreous (glassy), earthy, silky, pearly, and other surface qualities. Nonmetallic minerals may be shiny, although their vitreous shine is different from metallic luster. See the table for descriptions and examples of nonmetallic luster.

Luster	Image	Description
Vitreous/glassy	 *Quartz crystals*	Surface is shiny like glass
Earthy/dull	 *Kaolin specimen showing dull or earthy luster*	Dull, like dried mud or clay
Silky	 *Specimen showing silky luster*	Soft shine like silk fabric
Pearly	 *Specimen showing pearly luster*	Like the inside of a clam shell or mother-of-pearl
Submetallic	 *Submetallic luster on sphalerite*	Has the appearance of dull metal, like pewter. These minerals would usually still be considered metallic. Submetallic appearance can occur in metallic minerals because of weathering.

Figure 3.6.4: Azurite is ALWAYS a dark blue color and has been used for centuries for blue pigment.

Surface color may be helpful in identifying minerals, although it can be quite variable within the same mineral family. Mineral colors are affected by the main elements as well as impurities in the crystals. These impurities may be rare elements—like manganese, titanium, chromium, or lithium—even other molecules that are not normally part of the mineral formula. For example, the incorporation of water molecules gives quartz, which is normally clear, a milky color.

Some minerals predominantly show a single color. Malachite and azurite are green and blue, respectively, because of their copper content. Other minerals have a predictable range of colors due to elemental substitutions, usually via a solid solution. Feldspars, the most abundant minerals in the earth's crust, are complex, have solid solution series, and present several colors including pink, white, green, gray and others. Other minerals also come in several colors, influenced by trace amounts of several elements. The same element may show up as different colors, in different minerals. With notable exceptions, color is usually not a definitive property of minerals. For identifying many minerals. a more reliable indicator is a streak, which is the color of the powdered mineral.

Streak

Figure 3.6.5: Some minerals have different streaks than their visual color

Streak examines the color of a powdered mineral and can be seen when a mineral sample is scratched or scraped on an unglazed porcelain streak plate. A paper page in a field notebook may also be used for the streak of some minerals. Minerals that are harder than the streak plate will not show streak but will scratch the porcelain. For these minerals, a streak test can be obtained by powdering the mineral with a hammer and smearing the powder across a streak plate or notebook paper.

While mineral surface colors and appearances may vary, their streak colors can be diagnostically useful. An example of this property is seen in the iron-oxide mineral hematite. Hematite occurs in a variety of forms, colors, lusters (from shiny metallic silver to earthy red-brown), and different physical appearances. A hematite streak is consistently reddish-brown, no matter what the original specimen looks like. Iron sulfide, or pyrite, is a brassy metallic yellow. Commonly named fool's gold, pyrite has a characteristic black to greenish-black streak.

Hardness

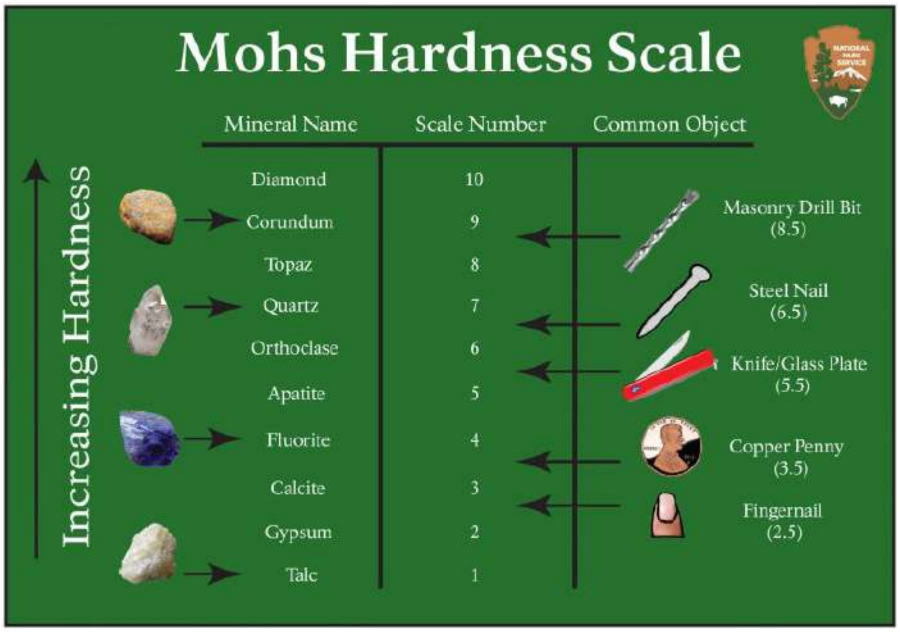


Figure 3.6.6: Mohs Hardness Scale

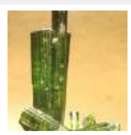
Hardness measures the ability of a mineral to scratch other substances. The Mohs Hardness Scale gives a number showing the relative scratch-resistance of minerals when compared to a standardized set of minerals of increasing hardness. The Mohs scale was developed by German geologist Fredrick Mohs in the early 20th century, although the idea of identifying minerals by hardness goes back thousands of years. Mohs hardness values are determined by the strength of a mineral's atomic bonds.

The figure shows the minerals associated with specific hardness values, together with some common items readily available for use in field testing and mineral identification. The hardness values run from 1 to 10, with 10 being the hardest; however, the scale is not linear. Diamond defines a hardness of 10 and is actually about four times harder than corundum, which is 9. A steel pocketknife blade, which has a hardness value of 5.5, separates between hard and soft minerals on many mineral identification keys.

Crystal Habit

Minerals can be identified by **crystal habit**, how their crystals grow and appear in rocks. Crystal shapes are determined by the arrangement of the atoms within the crystal structure. For example, a cubic arrangement of atoms gives rise to a cubic-shaped mineral crystal. Crystal habit refers to typically observed shapes and characteristics; however, they can be affected by other minerals crystallizing in the same rock. When minerals are constrained so they do not develop their typical crystal habit, they are called **anhedral**. **Subhedral** crystals are partially formed shapes. For some minerals, characteristic crystal habit is to grow crystal faces even when surrounded by other crystals in the rock. An example is garnet. Minerals grow freely where the crystals are unconstrained and can take characteristic shapes often form crystal faces. A **euhedral** crystal has a perfectly formed, unconstrained shape. Some minerals crystallize in such tiny crystals, they do not show a specific crystal habit to the naked eye. Other minerals, like pyrite, can have an array of different crystal habits, including cubic, dodecahedral, octahedral, and massive. The table lists typical crystal habits of various minerals.

Habit	Image	Examples
Bladed long and flat crystals	 *Bladed kyanite*	kyanite, amphibole, gypsum
Botryoidal/mammillary blobby, circular crystals	 *Malachite from the Congo*	hematite, malachite, smithsonite
Coating/laminae/druse crystals that are small and coat surfaces	 *Quartz (var. amethyst) in a geode*	quartz, calcite, malachite, azurite
Cubic cube-shaped crystals	 *Cubic crystals of galena*	pyrite, galena, halite
Dodecahedral 12-sided polygon shapes	 *Pyrite crystals with dodecahedral habit*	garnet, pyrite
Dendritic branching crystals	 *Manganese dendrites, scale in mm.*	Mn-oxides, copper, gold
Equant crystals that do not have a long direction		olivine, garnet, pyroxene

	 *Olivine crystal*	
Fibrous thin, very long crystals	 *Tremolite, a type of amphibole*	serpentine, amphibole, zeolite
Layered, sheets stacked, very thin, flat crystals	 *Sheet crystals of muscovite*	mica (biotite, muscovite, etc.)
Lenticular/platy crystals that are plate-like	 *Orange wulfenite on calcite*	selenite roses, wulfenite, calcite
Hexagonal crystals with six sides	 *Hexagonal hanksite*	quartz, hanksite, corundum
Massive/granular Crystals with no obvious shape, microscopic crystals	 *Limonite, a hydrated oxide of iron*	limonite, pyrite, azurite, bornite
Octahedral 4-sided double pyramid crystals	 *Octahedral fluorite*	diamond, fluorite, magnetite, pyrite
Prismatic/columnar very long, cylindrical crystals	 *Columnar tourmaline*	tourmaline, beryl, barite
Radiating crystals that grow from a point and fan-out	*Pyrophyllite*	pyrite "suns", pyrophyllite
Rhombohedral crystals shaped like slanted cubes	*Calcite crystal in shape of rhomb*	calcite, dolomite
Tabular/blocky/stubby sharp-sided crystals with no long direction	*Crystals of diopside, a member of the pyroxene family*	feldspar, pyroxene, calcite

Tetrahedral three-sided, pyramid-shaped crystals	magnetite, spinel, tetrahedrite

Tetrahedrite

Another crystal habit that may be used to identify minerals is striations, which are dark and light parallel lines on a crystal face. Twinning is another, which occurs when the crystal structure replicates in mirror images along certain directions in the crystal.

Figure 3.6.7: Twinned staurolite Gypsum with striations

Striations and twinning are related properties in some minerals including plagioclase feldspar. Striations are optical lines on a cleavage surface. Because of twinning in the crystal, striations show up on one of the two cleavage faces of the plagioclase crystal.

Figure 3.6.8: Striations on plagioclase

Cleavage and Fracture

Minerals often show characteristic patterns of breaking along specific cleavage planes or show characteristic fracture patterns. Cleavage planes are smooth, flat, parallel planes within the crystal. The cleavage planes may show as reflective surfaces on the crystal, as parallel cracks that penetrate into the crystal, or show on the edge or side of the crystal as a series of steps like rice terraces. Cleavage arises in crystals where the atomic bonds between atomic layers are weaker along some directions than others, meaning they will break preferentially along these planes. Because they develop on atomic surfaces in the crystal, cleavage planes are optically smooth and reflect light, although the actual break on the crystal may appear jagged or uneven. In such cleavages, the cleavage surface may appear like rice terraces on a mountainside that all reflect sunlight from a particular sun angle. Some minerals have a strong cleavage, some minerals only have weak cleavage or do not typically demonstrate cleavage.

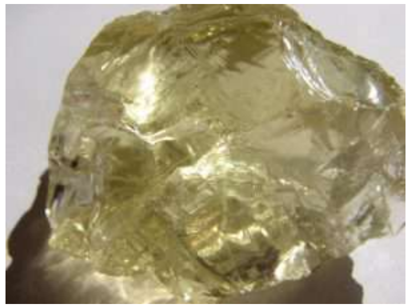


Figure 3.6.9: Citrine, a variety of quartz showing conchoidal fracture

For example, quartz and olivine rarely show cleavage and typically break into conchoidal fracture patterns.

Figure 3.6.10: Graphite showing layers of carbon atoms separated by a gap with weak bonds holding the layers together.

Graphite has its carbon atoms arranged into layers with relatively strong bonds within the layer and very weak bonds between the layers. Thus graphite cleaves readily between the layers and the layers slide easily over one another giving graphite its lubricating quality.

Mineral fracture surfaces may be rough, uneven, or show a conchoidal fracture. Uneven fracture patterns are described as irregular, splintery, fibrous. A conchoidal fracture has a smooth, curved surface like a shallow bowl or conch shell, often with curved ridges. Natural volcanic glass, called obsidian, breaks with this characteristic conchoidal pattern

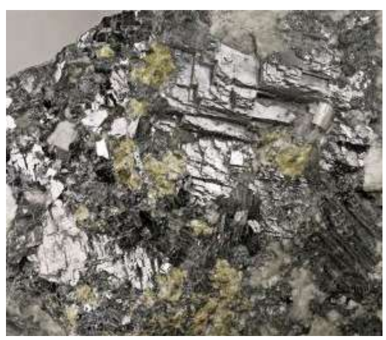


Figure 3.6.11: Cubic cleavage of galena; note how the cleavage surfaces show up as different but parallel layers in the crystal.

To work with cleavage, it is important to remember that cleavage is a result of bonds separating along planes of atoms in the crystal structure. On some minerals, **cleavage planes** may be confused with crystal faces. This will usually not be an issue for crystals of minerals that grew together within rocks. The act of breaking the rock to expose a fresh face will most likely break the crystals along cleavage planes. Some cleavage planes are parallel with crystal faces but many are not. Cleavage planes are smooth, flat, parallel planes within the crystal. The cleavage planes may show as parallel cracks that penetrate into the crystal (see amphibole below), or show on the edge or side of the crystal as a series of steps like rice terraces. For some minerals, characteristic crystal habit is to grow crystal faces even when surrounded by other crystals in rock. An example is garnet. Minerals grow freely where the crystals are unconstrained and can take characteristic shapes often form crystal faces (see quartz below).

Figure 3.6.12: Freely growing quartz crystals showing crystal faces

In some minerals, distinguishing cleavage planes from crystal faces may be challenging for the student. Understanding the nature of cleavage and referring to the number of cleavage planes and cleavage angles on identification keys should provide the student with enough information to distinguish cleavages from crystal faces. Cleavage planes may show as multiple parallel cracks or flat surfaces on the crystal. Cleavage planes may be expressed as a series of steps like terraced rice paddies. See the cleavage surfaces on galena above or plagioclase below. Cleavage planes arise from the tendency of mineral crystals to break along specific planes of weakness within the crystal favored by atomic arrangements. The number of cleavage planes, the quality of the cleavage surfaces, and the angles between them are diagnostic for many minerals and cleavage is one of the most useful properties for identifying minerals. Learning to recognize cleavage is an especially important and useful skill in studying minerals.

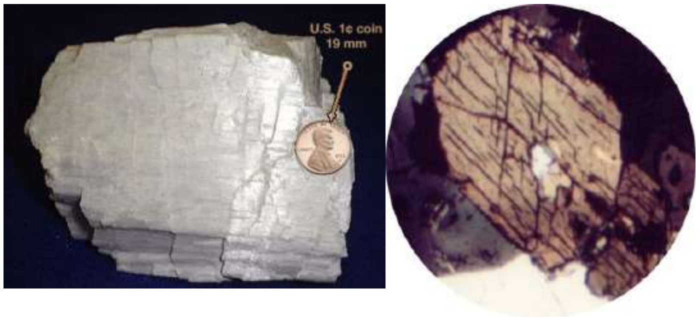


Figure 3.6.13: (left) Steps of cleavage along the same cleavage direction. (right) Photomicrograph showing 120/60 degree cleavage within a grain of amphibole

As an identification property of minerals, cleavage is usually given in terms of the quality of the cleavage (perfect, imperfect, or none), the number of cleavage surfaces, and the angles between the surfaces. The most common number of cleavage plane directions in the common rock-forming minerals are one perfect cleavage (as in mica), two cleavage planes (as in feldspar, pyroxene, and amphibole), and three cleavage planes (as in halite, calcite, and galena). One perfect cleavage (as in mica) develops on the top and bottom of the mineral specimen with many parallel cracks showing on the sides but no angle of intersection. Two cleavage planes intersect at an angle. Common cleavage angles are 60°, 75°, 90°, and 120°. Amphibole has two cleavage planes at 60° and 120°. Galena and halite have three cleavage planes at 90° (cubic cleavage). Calcite cleaves readily in three directions producing a cleavage figure called a rhomb that looks like a cube squashed over toward one corner giving rise to the approximately 75° cleavage angles. Pyroxene has an imperfect cleavage with two planes at 90°.

Cleavages on Common Rock-Forming Minerals

- Quartz—none (conchoidal fracture)
- Olivine—none (conchoidal fracture)
- Mica—1 perfect
- Feldspar—2 perfect at 90°
- Pyroxene—2 imperfect at 90°
- Amphibole—2 perfect at 60°/120°
- Calcite—3 perfect at approximately 75°
- Halite, galena, pyrite—3 perfect at 90°

Special Properties

Special properties are unique and identifiable characteristics used to identify minerals or that allow some minerals to be used for special purposes. Ulexite has a fiber-optic property that can project images through the crystal-like a high-definition television screen (see figure). A simple identifying special property is taste, such as the salty flavor of halite or common table salt (NaCl). Sylvite is potassium chloride (KCl) and has a more bitter taste.

Figure 3.6.14: A demonstration of ulexite's image projection

Another property geologists may use to identify minerals is a property related to density called **specific gravity**. Specific gravity measures the weight of a mineral specimen relative to the weight of an equal volume of water. The value is expressed as a ratio between the mineral and water weights. To measure specific gravity, a mineral specimen is first weighed in grams then submerged in a graduated cylinder filled with pure water at room temperature. The rise in water level is noted using the cylinder's graduated scale. Since the weight of water at room temperature is 1 gram per cubic centimeter, the ratio of the two-weight numbers gives the specific gravity. Specific gravity is easy to measure in the laboratory but is less useful for mineral identification in the field than other more easily observed properties, except in a few rare cases such as the very dense galena or native gold. The high density of these minerals gives rise to a qualitative property called "heft." Experienced geologists can roughly assess specific gravity by heft, a subjective quality of how heavy the specimen feels in one's hand relative to its size.

Figure 3.6.15: Native gold has one of the highest specific gravities.

A simple test for identifying calcite and dolomite is to drop a bit of dilute hydrochloric acid (10-15% HCl) on the specimen. If the acid drop effervesces or fizzes on the surface of the rock, the specimen is calcite. If it does not, the specimen is scratched to produce a small amount of powder and test with acid again. If the acid drop fizzes slowly on the powdered mineral, the specimen is dolomite. The difference between these two minerals can be seen in the video. Geologists who work with carbonate rocks carry a small dropper bottle of dilute HCl in their field kit. Vinegar, which contains acetic acid, can be used for this test and is used to distinguish non-calcite fossils from limestone. While acidic, vinegar produces less of a fizzing reaction because acetic acid is a weaker acid.

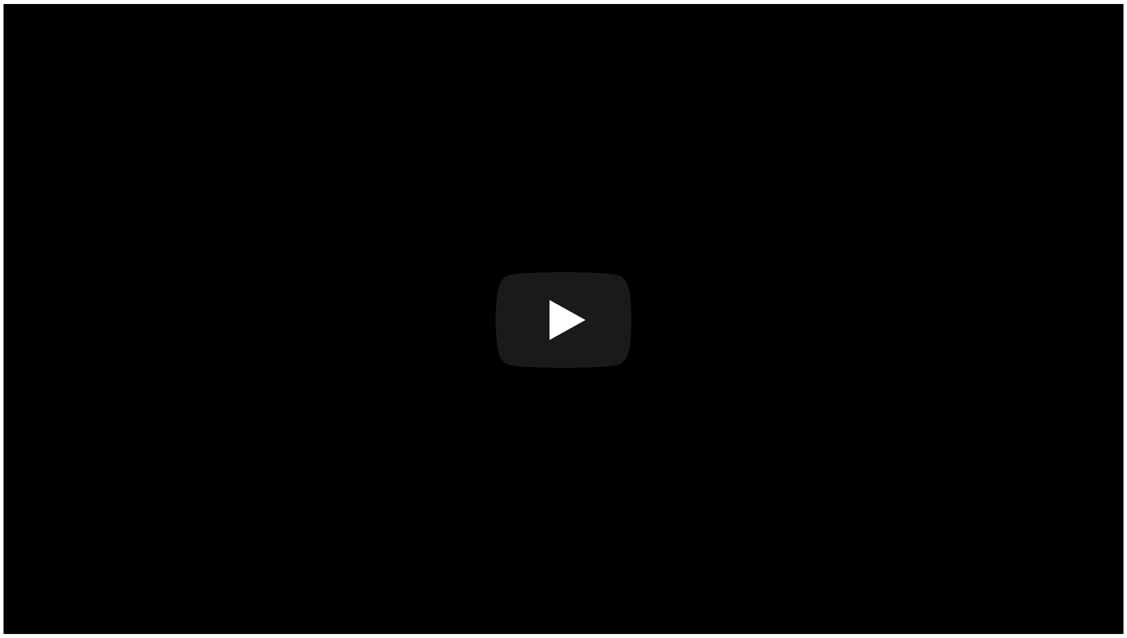


Figure 3.6.16: Paperclips attracted to lodestone (magnetite).

Some iron-oxide minerals are magnetic and are attracted to magnets. A common name for naturally magnetic iron oxide is **lodestone**. Others include magnetite ($Fe3O_4$) and ilmenite ($FeTiO_3$). Magnetite is strongly attracted to magnets and can be magnetized. Ilmenite and some types of hematite are weakly magnetic.

Figure 3.6.17: Iridescence on plagioclase; also showing striations on the cleavage surface

Some minerals and mineraloids scatter light via a phenomenon called **iridescence**. This property occurs in labradorite (a variety of plagioclase) and opal. It is also seen in biologically created substances like pearls and seashells. Cut diamonds show iridescence and the jeweler’s diamond cut is designed to maximize this property.

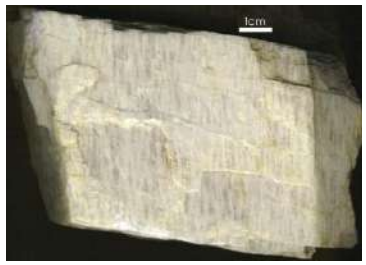


Figure 3.6.18: Exsolution lamellae or perthitic lineations within potassium feldspar

Striations on mineral cleavage faces are an optical property that can be used to separate plagioclase feldspar from potassium feldspar (K-spar). A process called twinning creates parallel zones in the crystal that are repeating mirror images. The actual cleavage angle in plagioclase is slightly different than 90^{o} and the alternating mirror images in these twinned zones produce a series of parallel lines on one of plagioclase’s two cleavage faces. Light reflects off these twinned lines at slightly different angles which then appear as light and dark lines called striations on the cleavage surface. Potassium feldspar does not exhibit twinning or striations but may show linear features called **exsolution lamellae**, also known as perthitic lineation or simply perthite. Because sodium and potassium do not fit into the same feldspar crystal structure, the lines are created by small amounts of sodium feldspar (albite) separating from the dominant potassium feldspar (K-spar) within the crystal structure. The two different feldspars crystallize out into roughly parallel zones within the crystal, which are seen as these linear markings.

Figure 3.6.19: Fluorite. Lower image shows fluorescence of fluorite under UV light

One of the most interesting special mineral properties is **fluorescence**. Certain minerals, or trace elements within them, give off visible light when exposed to ultraviolet radiation or black light. Many mineral exhibits have a fluorescence room equipped with black lights so this property can be observed. An even rarer optical property is phosphorescence. **Phosphorescent** minerals absorb light and then slowly release it, much like a glow-in-the-dark sticker.

Contributions and Attributions

3.S: Summary

Minerals are the building blocks of rocks and essential to understanding geology. Mineral properties are determined by their atomic bonds. Most minerals begin in a fluid, and either crystallize out of cooling magma or precipitate as ions and molecules out of a saturated solution. The silicates are largest group of minerals on Earth, by number of varieties and relative quantity, making up a large portion of the crust and mantle. Based on the silicon-oxygen tetrahedra, the crystal structure of silicates reflects the fact that silicon and oxygen are the top two of Earth's most abundant elements. Non-silicate minerals are also economically important, and providing many types of construction and manufacturing materials. Minerals are identified by their unique physical properties, including luster, color, streak, hardness, crystal habit, fracture, cleavage, and special properties.

CHAPTER OVERVIEW

4: IGNEOUS PROCESSES AND VOLCANOES

Learning Objectives

Explain the origin of magma it relates to plate tectonics
Describe how the Bowen's Reaction Series relates mineral crystallization and melting temperatures
Explain how cooling of magma leads to rock compositions and textures, and how these are used to classify igneous rocks
Analyze the features of common igneous landforms and how they relate to their origin
Explain partial melting and fractionation, and how they change magma compositions
Describe how silica content affects magma viscosity and eruptive style of volcanoes
Describe volcano types, eruptive styles, composition, and their plate tectonic settings
Describe volcanic hazards

Igneous rock is formed when liquid rock freezes into solid rock. This molten material is called **magma** when it is in the ground and **lava** when it is on the surface. Only the Earth's outer core is liquid; the Earth's mantle and crust are naturally solid. However, there are a few minor pockets of magma that form near the surface where geologic processes cause melting. It is this magma that becomes the source for volcanoes and igneous rocks. This chapter will describe the classification of igneous rocks, the unique processes that form magmas, types of volcanoes and volcanic processes, volcanic hazards, and igneous landforms.

Figure \PageIndex1: Lava flow in Hawaii

Lava cools quickly on the surface of the earth and forms tiny microscopic crystals. These are known as fine-grained **extrusive**, or volcanic, igneous rocks. Extrusive rocks are often **vesicular**, filled with holes from escaping gas bubbles. **Volcanism** is the process in which lava has erupted. Depending on the properties of the lava that is erupted, the volcanism can be drastically different, from smooth and gentle to dangerous and explosive. This leads to different types of volcanoes and different volcanic hazards.

Figure \PageIndex1: Half Dome, a mass of intrusive igneous rock in Yosemite National Park, now exposed by erosion.

In contrast, magma that cools slowly below the earth's surface forms larger crystals which can be seen with the naked eye. These are known as coarse-grained **intrusive**, or plutonic, igneous rocks. This relationship between cooling rates and grain sizes of the solidified minerals in igneous rocks is important for interpreting the rock's geologic history.

4.1: CLASSIFICATION OF IGNEOUS ROCKS

Igneous rocks are classified based on texture and composition. Texture describes the physical characteristics of the minerals, such as grain size. This relates to the cooling history of the molten magma from which it came. Composition refers to the rock's specific mineralogy and chemical composition. Cooling history is also related to changes that can occur to the composition of igneous rocks.

4.2: BOWEN'S REACTION SERIES

Bowen's Reaction Series describes the temperature at which minerals crystallize when cooled, or melt when heated. The low end of the temperature scale where all minerals crystallize into solid rock is approximately 700°C (158°F). The upper end of the range where all minerals exist in a molten state is approximately 1,250°C (2,282°F). These numbers reference minerals that crystallize at standard sea-level pressure, 1 bar.

4.3: MAGMA GENERATION

Magma and lava contain three components: melt, solids, and volatiles. The melt is made of ions from minerals that have liquefied. The solids are made of crystallized minerals floating in the liquid melt. These may be minerals that have already cooled Volatiles are gaseous components—such as water vapor, carbon dioxide, sulfur, and chlorine—dissolved in the magma. The presence and amount of these three components affect the physical behavior of the magma and will be discussed more in this page.

4.4: PARTIAL MELTING AND CRYSTALLIZATION

Even though all magmas originate from similar mantle rocks, other things, like partial melting and crystallization processes, can change the chemistry of the magma. This explains the wide variety of resulting igneous rocks that are found all over Earth. Because the mantle is composed of many different minerals, it does not melt uniformly. As minerals with lower melting points turn into liquid magma, those with higher melting points remain as solid crystals. This is known as partial melting.

4.5: VOLCANISM

When magma emerges onto the Earth's surface, the molten rock is called lava. A volcano is a type of land formation created when lava solidifies into rock. Volcanoes have been an important part of human society for centuries, though their understanding has greatly increased as our understanding of plate tectonics has made them less mysterious. This section describes volcano location, type, hazards, and monitoring.

4.1: Classification of Igneous Rocks

Igneous rocks are classified based on texture and composition. Texture describes the physical characteristics of the minerals, such as grain size. This relates to the cooling history of the molten magma from which it came. Composition refers to the rock's specific mineralogy and chemical composition. Cooling history is also related to changes that can occur to the composition of igneous rocks.

Texture

Figure 4.1.1: Granite is a classic coarse-grained (phaneritic) intrusive igneous rock. The different colors are unique minerals. The black colors are likely two or three different minerals.

If magma cools slowly, deep within the crust, the resulting rock is called intrusive or plutonic. The slow cooling process allows crystals to grow large, giving the intrusive igneous rock a coarse-grained or **phaneritic** texture. The individual crystals in phaneritic texture are readily visible to the unaided eye.

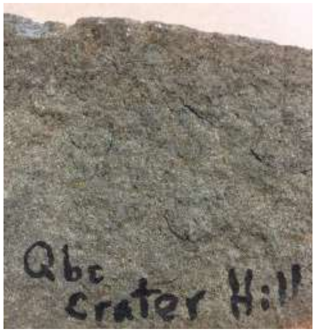


Figure 4.1.1: Basalt is a classic fine-grained (aphanitic) extrusive igneous rock. This sample is mostly fine groundmass with a few small green phenocrysts that are the mineral olivine.

When lava is extruded onto the surface, or intruded into shallow fissures near the surface and cools, the resulting igneous rock is called extrusive or volcanic. Extrusive igneous rocks have a fine-grained or **aphanitic** texture, in which the grains are too small to see with the unaided eye. The fine-grained texture indicates the quickly cooling lava did not have time to grow large crystals. These tiny crystals can be viewed under a petrographic microscope [1]. In some cases, extrusive lava cools so rapidly it does not develop crystals at all. This non-crystalline material is not classified as minerals but as volcanic glass. This is a common component of volcanic ash and rocks like obsidian.

Figure 4.1.1: Porphyritic texture

Some igneous rocks have a mix of coarse-grained minerals surrounded by a matrix of fine-grained material in a texture called **porphyritic**. The large crystals are called **phenocrysts** and the fine-grained matrix is called the **groundmass** or **matrix**.

Porphyritic texture indicates the magma body underwent a multi-stage cooling history, cooling slowly while deep under the surface and later rising to a shallower depth or the surface where it cooled more quickly.

Figure 4.1.1: Pegmatitic texture

Residual molten material expelled from igneous intrusions may form veins or masses containing very large crystals of minerals like feldspar, quartz, beryl, tourmaline, and mica. This texture, which indicates a very slow crystallization, is called **pegmatitic**. A rock that chiefly consists of pegmatitic texture is known as a **pegmatite**. To give an example of how large these crystals can get, transparent cleavage sheets of pegmatitic muscovite mica were used as windows during the Middle Ages.

Figure 4.1.1: Scoria, a vesicular extrusive igneous rock

All magmas contain gases dissolved in a solution called **volatiles**. As the magma rises to the surface, the drop in pressure causes the dissolved volatiles to come bubbling out of solution, like the fizz in an opened bottle of soda. The gas bubbles become trapped in the solidifying lava to create a **vesicular** texture, with the holes specifically called vesicles. The type of volcanic rock with common vesicles is called **scoria**.

Figure 4.1.1: Pumice

An extreme version of scoria occurs when volatile-rich lava is very quickly quenched and becomes a meringue-like froth of glass called **pumice**. Some pumice is so full of vesicles that the density of the rock drops low enough that it will float.

Figure 4.1.1: Obsidian (volcanic glass). Note conchoidal fracture.

Lava that cools extremely quickly may not form crystals at all, even microscopic ones. The resulting rock is called **volcanic glass**. **Obsidian** is a rock consisting of volcanic glass. Obsidian as a glassy rock shows an excellent example of conchoidal fracture similar to the mineral quartz (see Chapter 3).

Figure 4.1.1: This tuff has crystals, rock fragments, and bedrock mixed together.

When volcanoes erupt explosively, vast amounts of lava, rock, ash, and gases are thrown into the atmosphere. The solid parts, called tephra, settle back to earth and cool into rocks with **pyroclastic** textures. *Pyro,* meaning fire, refers to the igneous source of the tephra and *clastic* refers to the rock fragments. Tephra fragments are named based on size—**ash** (<2 mm), **lapilli** (2-64 mm), and **bombs or blocks** (>64 mm). Pyroclastic texture is usually recognized by the chaotic mix of crystals, angular glass shards, and rock fragments. Rock formed from large deposits of tephra fragments is called **tuff**. If the fragments accumulate while still hot, the heat may deform the crystals and weld the mass together, forming a welded tuff.

Composition

Composition refers to a rock's chemical and mineral make-up. For igneous rock, the composition is divided into four groups: **felsic**, **intermediate**, **mafic**, and **ultramafic**. These groups refer to differing amounts of silica, iron, and magnesium found in the minerals that make up the rocks. It is important to realize these groups do not have sharp boundaries in nature, but rather lie on a continuous spectrum with many transitional compositions and names that refer to specific quantities of minerals. As an example, granite is a commonly-used term but has a very specific definition which includes exact quantities of minerals like feldspar and quartz. Rocks labeled as 'granite' in laymen applications can be several other rocks, including syenite, tonalite, and monzonite. To avoid these complications, the following figure presents a simplified version of igneous rock nomenclature focusing on the four main groups, which is adequate for an introductory student.

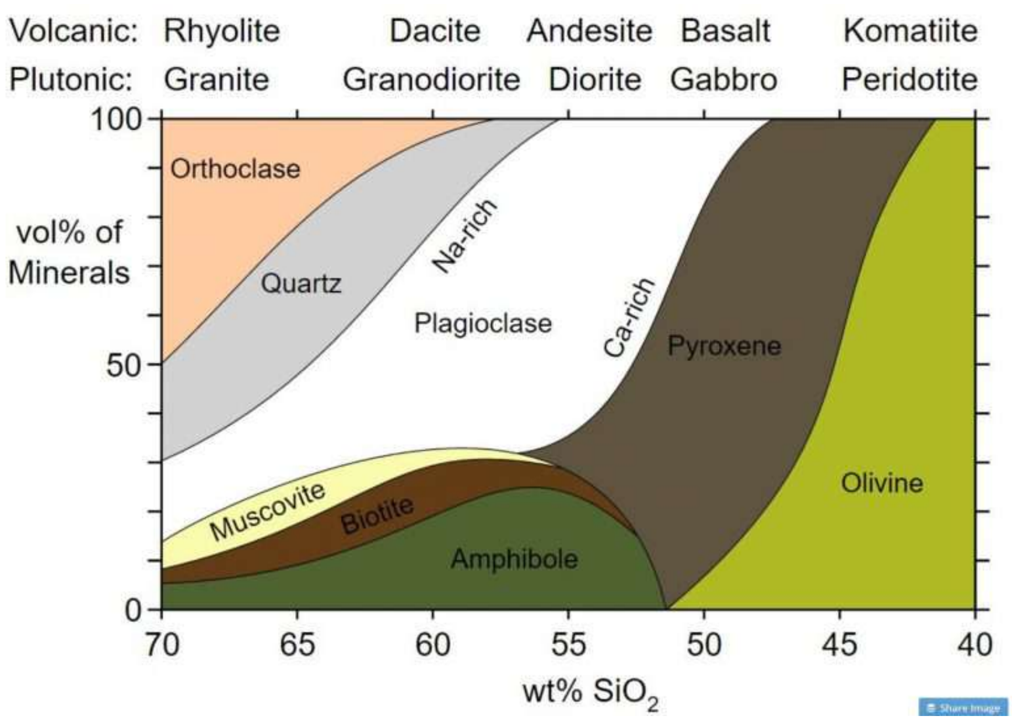


Figure 4.1.1: Mineral composition of common igneous rocks. The percentage of minerals is shown on the vertical axis. The percentage of silica is shown on the horizontal axis. Rock names at the top include a continuous spectrum of compositions grading from one into another.

- **Fel**sic refers to a predominance of the light-colored (felsic) minerals **fel**dspar and **si**lica in the form of quartz. These light-colored minerals have more silica as a proportion of their overall chemical formula. Minor amounts of dark-colored (mafic) minerals like amphibole and biotite mica may be present as well. Felsic igneous rocks are rich in silica (in the 65-75% range, meaning the rock would be 65-75% weight percent SiO^2) and poor in iron and magnesium.
- **Intermediate** is a composition between felsic and mafic. It usually contains roughly-equal amounts of light and dark minerals, including light grains of plagioclase feldspar and dark minerals like amphibole. It is intermediate in silica in the 55-60% range.
- **Maf**ic refers to an abundance of ferromagnesian minerals (with magnesium and iron, chemical symbols **M**g and **F**e) plus plagioclase feldspar. It is mostly made of dark minerals like pyroxene and olivine, which are rich in iron and magnesium and relatively poor in silica. Mafic rocks are low in silica, in the 45-50% range.
- **Ultramafic** refers to the extremely mafic rocks composed of mostly olivine and some pyroxene which have even more magnesium and iron and even less silica. These rocks are rare on the surface, but make up peridotite, the rock of the upper mantle. It is poor in silica, in the 40% or less range.

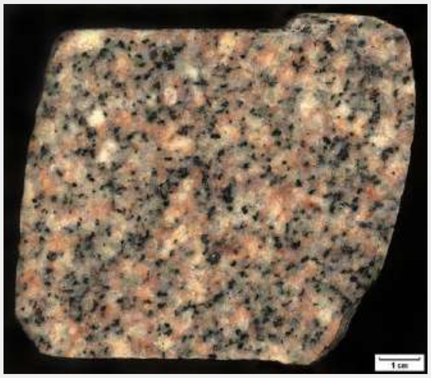
On the figure above, the top row has both plutonic and volcanic igneous rocks arranged in a continuous spectrum from felsic on the left to intermediate, mafic, and ultramafic toward the right. **Rhyolite** refers to the volcanic and felsic igneous rocks and **granite** refer to intrusive and felsic igneous rocks. **Andesite** and **diorite** likewise refer to extrusive and intrusive intermediate rocks (with dacite and granodiorite applying to those rocks with composition between felsic and intermediate).

Basalt and **gabbro** are the extrusive and intrusive names for mafic igneous rocks, and **peridotite** is ultramafic, with **komatiite** as the fine-grained extrusive equivalent. Komatiite is a rare rock because volcanic material that comes directly from the mantle is not common, although some examples can be found in ancient Archean rocks [2]. Nature rarely has sharp boundaries and the classification and naming of rocks often impose what appears to be sharp boundary names onto a continuous spectrum.

Figure 4.1.1: Igneous rock classification table with composition as vertical columns and texture as horizontal rows.

Aphanitic/Phaneritic Rock Types with Images

Felsic Composition



Granite from Cape Cod, Massachusetts.

Rhyolite (source: Michael C. Rygel via Wikimedia Commons)

Granite is a course-crystalline felsic intrusive rock. The presence of quartz is a good indicator of granite. Granite commonly has large amounts of salmon pink potassium feldspar and white plagioclase crystals that have visible cleavage planes. Granite is a good approximation for the continental crust, both in density and composition.

Rhyolite is a fine-crystalline felsic extrusive rock. Rhyolite is commonly pink and will often have glassy quartz phenocrysts. Because felsic lavas are less mobile, it is less common than granite. Examples of rhyolite include several lava flows in Yellowstone National Park and the altered rhyolite that makes up the Grand Canyon of the Yellowstone.

Intermediate Composition

Diorite

Andesite

Diorite is a coarse-crystalline intermediate intrusive igneous rock. Diorite is identifiable by it's Dalmatian-like appearance of black hornblende and biotite and white plagioclase feldspar. It is found in its namesake, the Andes Mountains as well as the Henry and Abajo mountains of Utah.

Andesite is a fine crystalline intermediate extrusive rock. It is commonly grey and porphyritic. It can be found in the Andes Mountains and in some island arcs (see Chapter 2). It is the fine grained compositional equivalent of diorite.

Mafic Composition

Gabbro	*Vesicular Basalt*
Gabbro is a coarse-grained mafic igneous rock, made with mainly mafic minerals like pyroxene and only minor plagioclase. Because mafic lava is more mobile, it is less common than basalt. Gabbro is a major component of the lower oceanic crust.	Basalt is a fine-grained mafic igneous rock. It is commonly vesicular and aphanitic. When porphyritic, it often has either olivine or plagioclase phenocrysts. Basalt is the main rock which is formed at mid-ocean ridges, and is therefore the most common rock on the Earth's surface, making up the entirety of the ocean floor (except where covered by sediment).

Igneous Rock Bodies

Igneous rocks are common in the geologic record, but surprisingly, it is the intrusive rocks that are more common. Extrusive rocks, because of their small crystals and glass, are less durable. Plus, they are, by definition, exposed to the elements of erosion immediately. Intrusive rocks, forming underground with larger, stronger crystals, are more likely to last. Therefore, most landforms and rock groups that owe their origin to igneous rocks are intrusive bodies. A significant exception to this is active volcanoes, which are discussed in a later section on volcanism. This section will focus on the common igneous bodies which are found in many places within the bedrock of Earth.

Figure 4.1.1: Dike of olivine gabbro cuts across Baffin Island in the Canadian Arctic

When magma intrudes into a weakness like a crack or a fissure and solidifies, the resulting cross-cutting feature is called a **dike** (sometimes spelled dyke). Because of this, dikes are often vertical or at an angle relative to the pre-existing rock layers that they intersect. Dikes are therefore discordant intrusions, not following any layering that was present. Dikes are important to geologists, not only for the study of igneous rocks themselves but also for dating rock sequences and interpreting the geologic history of an area. The dike is younger than the rocks it cuts across and, as discussed in the chapter on Geologic Time (Chapter 7), may be used to assign actual numeric ages to sedimentary sequences, which are notoriously difficult to age date.

Figure 4.1.1: Igneous sill intruding between Paleozoic strata in Nova Scotia

Sills are another type of intrusive structure. A sill is a concordant intrusion that runs parallel to the sedimentary layers in the country rock. They are formed when magma exploits a weakness between these layers, shouldering them apart and squeezing between them. As with dikes, sills are younger than the surrounding layers and may be radioactively dated to study the age of sedimentary strata.

Figure 4.1.1: Cottonwood Stock, a quartz monzonite pluton exposed at the mouth of Little Cottonwood Canyon, Utah

A magma chamber is a large underground reservoir of molten rock. The path of rising magma is called a **diapir**. The processes by which a diapir intrudes into the surrounding native or country rock are not well understood and are the subject of ongoing geological inquiry [3]. For example, it is not known what happens to the pre-existing country rock as the diapir intrudes. One theory is the overriding rock gets shouldered aside, displaced by the increased volume of magma. Another is the native rock is melted and consumed into the rising magma or broken into pieces that settle into the magma, a process known as **stoping**. It has also been proposed that diapirs are not a real phenomenon, but just a series of dikes that blend into each other. The dikes may be intruding over millions of years, but since they may be made of similar material, they would be appearing to be formed at the same time. Regardless, when a diapir cools, it forms a mass of intrusive rock called a **pluton**. Plutons can have irregular shapes, but can often be somewhat round.

Figure 4.1.1: Half Dome in Yosemite National Park, California, is a part of the Sierra Nevada batholith which is mostly made of granite.

When many plutons merge together in an extensive single feature, it is called a **batholith**. Batholiths are found in the cores of many mountain ranges, including the granite formations of Yosemite National Park in the Sierra Nevada of California. They are typically more than 100 km^2 in area, associated with subduction zones, and mostly felsic in composition. A **stock** is a type of pluton with less surface exposure than a batholith and may represent a narrower neck of material emerging from the top of a batholith. Batholiths and stocks are discordant intrusions that cut across and through surrounding country rock.

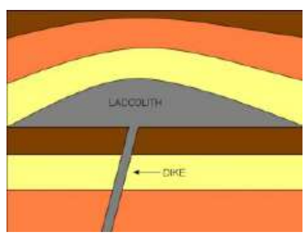
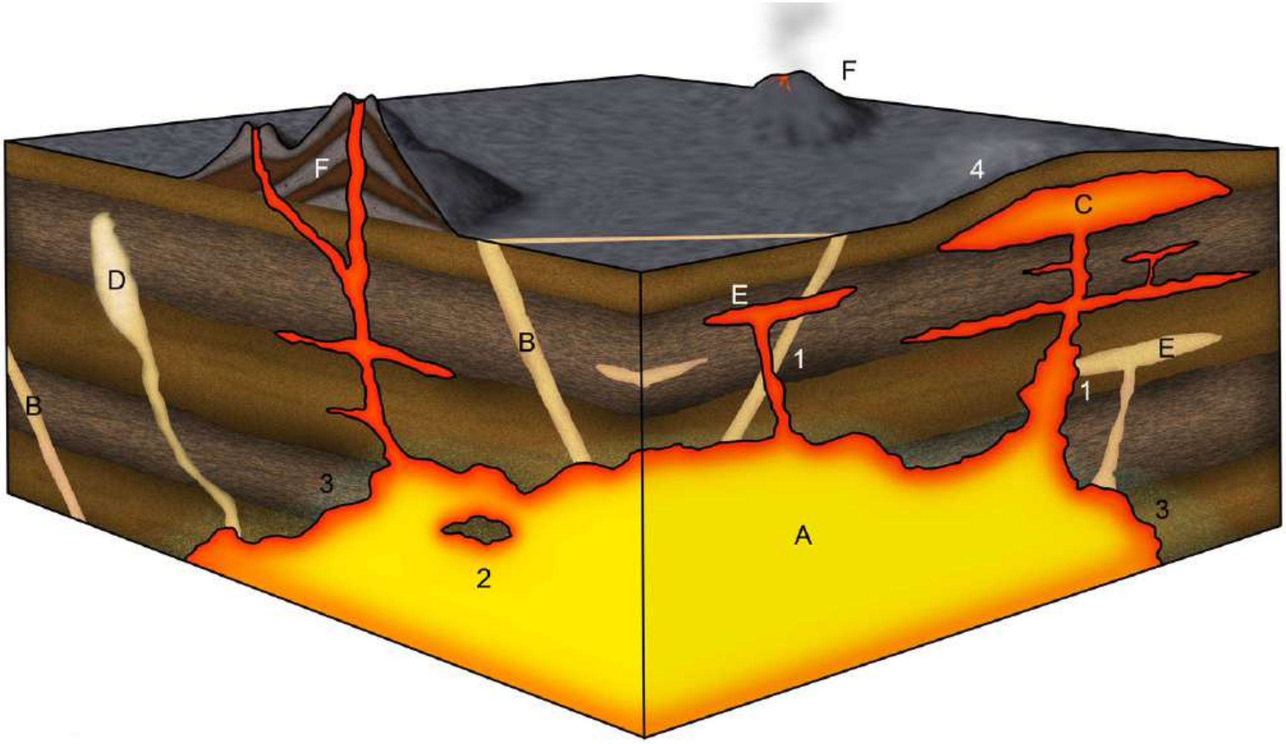


Figure 4.1.1: Laccolith forms as a blister in between sedimentary strata.

Figure 4.1.1: The Henry Mountains in Utah are interpreted to be a laccolith, exposed by erosion of the overlying layers.

Laccoliths are blister-like, concordant intrusions of magma that form between sedimentary layers. The Henry Mountains of Utah are a famous topographic landform formed by this process. Laccoliths bulge upwards; a similar downward-bulging intrusion is called a **lopolith**.



Guide for Igneous Structures Image (shown above)	
Number/Letter	Description
1	Young, emerging subvolcanic intrusion cutting through older one
2	Xenolith (solid rock of high melting temperature which has been transported within the magma from deep below) or roof pendant (fragment of the roof of the magma chamber that has detached from the roof and sunk into the melt)
3	Contact metamorphism in the country rock adjacent to the magma chamber (caused by the heat of the magma)
4	Uplift at the surface due to laccolith emplacement in the near sub-ground
A	Active magma chamber (called pluton when cooled and entirely crystallized; a batholith is a large rock body composed of several plutonic intrusions)
B	Old magmatic dykes/dikes
C	Emerging laccolith
D	Old pegmatite (late-magmatic dyke formed by aggressive and highly mobile residual melts of a magma chamber)
E	Old and emerging magmatic sills
F	Stratovolcano

References

1. Peacock, M. A. Classification of Igneous Rock Series. *The Journal of Geology* **39**, 54–67 (1931).

2. Arndt, N. T. Chapter 1 Archean Komatiites. in *Developments in Precambrian Geology* (ed. K.C. Condie) **11**, 11–44 (Elsevier, 1994).

3. Glazner, A. F., Bartley, J. M., Coleman, D. S., Gray, W. & Taylor, R. Z. Are plutons assembled over

Contributions and Attributions

4.2: Bowen's Reaction Series

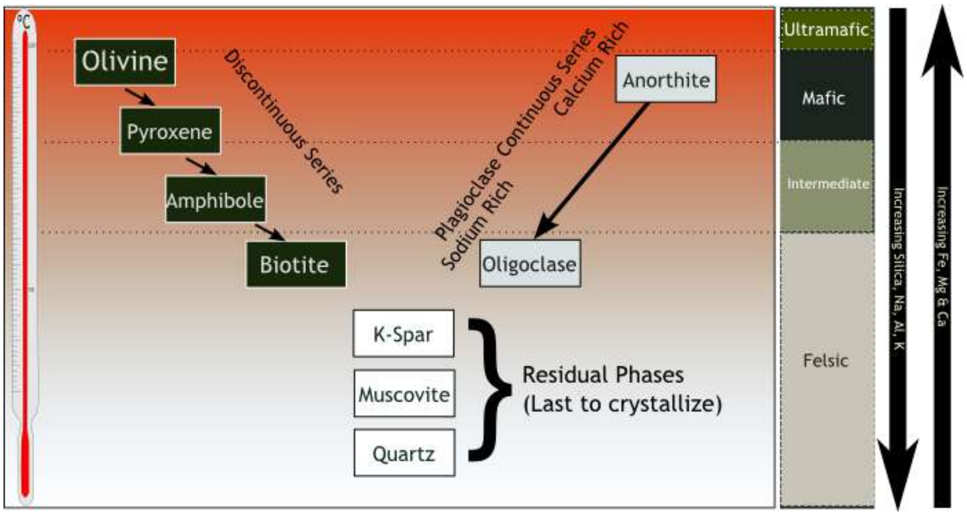


Figure 4.2.1: Bowen's Reaction Series. Minerals that crystallize at higher temperatures are at the top (olivine) and minerals that crystallize at lower temperatures are at the bottom (quartz). (Source Colivine, modified from Bowen, 1922)

Figure 4.2.1: Olivine, the first mineral to crystallize in a melt.

Bowen's Reaction Series describes the temperature at which minerals crystallize when cooled, or melt when heated. The low end of the temperature scale where all minerals crystallize into solid rock is approximately 700°C (158°F). The upper end of the range where all minerals exist in a molten state is approximately 1,250°C (2,282°F) [4]. These numbers reference minerals that crystallize at standard sea-level pressure, 1 bar. The values will be different for minerals located deep below the Earth's surface due to the increased pressure, which affects crystallization and melting temperatures (see Chapter 4.4). However, the order and relationships are maintained.

In the figure, the righthand column lists the four groups of igneous rock from top to bottom: ultramafic, mafic, intermediate, and felsic. The down-pointing arrow on the far right shows increasing amounts of silica, sodium, aluminum, and potassium as the mineral composition goes from ultramafic to felsic. The up-pointing arrow shows increasing ferromagnesian components, specifically iron, magnesium, and calcium. To the far left of the diagram is a temperature scale. Minerals near the top of the diagram, such as olivine and anorthite (a type of plagioclase), crystallize at higher temperatures. Minerals near the bottom, such as quartz and muscovite, crystalize at lower temperatures.

Figure 4.2.1: Normal L. Bowen

The most important aspect of Bowen's Reaction Series is to notice the relationships between minerals and temperature. Norman L. Bowen (1887-1956) was an early 20th Century geologist who studied igneous rocks. He noticed that in igneous rocks, certain minerals always occur together and these mineral assemblages exclude other minerals. Curious as to why, and with the hypothesis in mind that it had to do with the temperature at which the rocks cooled, he set about conducting experiments on igneous rocks in the early 1900s. He conducted experiments on igneous rock—grinding combinations of rocks into powder, sealing the powders into metal capsules, heating them to various temperatures, and then cooling them.

Figure 4.2.1: Norman L. Bowen working with his petrographic microscope

When he opened the quenched capsules, he found a glass surrounding mineral crystals that he could identify under his petrographic microscope. The results of many of these experiments, conducted at different temperatures over a period of several years, showed that the common igneous minerals crystallize from magma at different temperatures. He also saw that minerals occur together in rocks with others that crystallize within similar temperature ranges, and never crystallize with other minerals. This relationship can explain the main difference between mafic and felsic igneous rocks. Mafic igneous rocks contain more mafic minerals, and therefore, crystallize at higher temperatures than felsic igneous rocks. This is even seen in lava flows, with felsic lavas erupting hundreds of degrees cooler than their mafic counterparts. Bowen's work laid the foundation for understanding igneous **petrology** (the study of rocks) and resulted in his book, *The Evolution of the Igneous Rocks* in 1928 [5].

References

4. Boehler, R. Melting temperatures of the Earth's mantle and core: Earth's thermal structure. *Annual Review of Earth and Planetary Sciences* **24**, 15–40 (1996).

5. Bowen, N. L. *The evolution of the igneous rocks*. (Dover Publications, 1928).

Contributions and Attributions

4.3: Magma Generation

Magma and lava contain three components: melt, solids, and volatiles. The melt is made of ions from minerals that have liquefied. The solids are made of crystallized minerals floating in the liquid melt. These may be minerals that have already cooled **Volatiles** are gaseous components—such as water vapor, carbon dioxide, sulfur, and chlorine—dissolved in the magma [6]. The presence and amount of these three components affect the physical behavior of the magma and will be discussed more below.

Geothermal Gradient

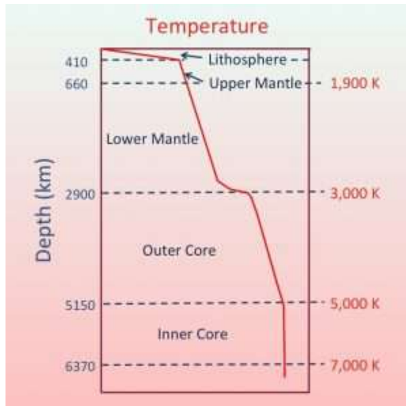


Figure 4.3.1: Geothermal gradient

Below the surface, the temperature of the Earth rises. This heat is caused by residual heat left from the formation of Earth and ongoing radioactive decay. The rate at which temperature increases with depth is called the **geothermal gradient**. The average geothermal gradient in the upper 100 km (62 mi) of the crust is about 25°C per kilometer of depth. So for every kilometer of depth, the temperature increases by about 25°C.

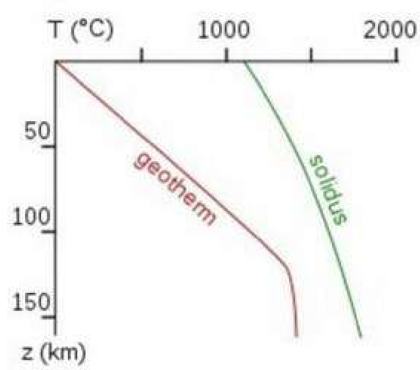
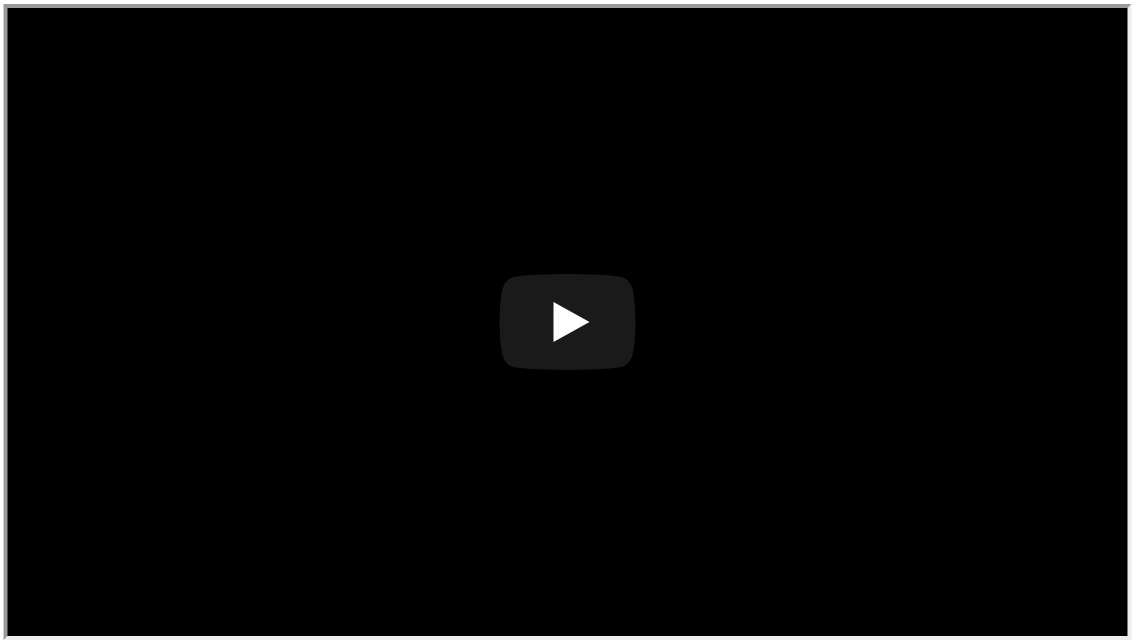


Figure 4.3.1: Pressure-temperature diagram showing temperature in degrees Celsius on the x-axis and depth below the surface in kilometers (km) on the y-axis. The red line is the geothermal gradient and the green solidus line represents the temperature and pressure regime at which melting begins. Rocks at pressures and temperatures left of the green line are solid. If pressure/temperature conditions change so that rocks pass to the right of the green line, then they will start to melt. (Source: Woudloper)

The depth-temperature graph (see figure) illustrates the relationship between the geothermal gradient (geotherm, red line) and the start of rock melting (solidus, green line). The geothermal gradient changes with depth (which has a direct relationship to pressure) through the crust into the upper mantle. The area to the left of the green line includes solid components; to the right is where liquid components start to form. The increasing temperature with depth makes the depth of about 125 kilometers (78 miles) where the natural geothermal gradient is closest to the solidus.

The temperature at 100 km (62 mi) deep is about 1,200°C (2,192°F). At bottom of the crust, 35 km (22 mi) deep, the pressure is about 10,000 bars [7]. A bar is a measure of pressure, with 1 bar being normal atmospheric pressure at sea level. At these pressures and temperatures, the crust and mantle are solid. To a depth of 150 km (93 mi), the geothermal gradient line stays to the left of the solidus line. This relationship continues through the mantle to the core-mantle boundary, at 2,880 km (1,790 mi).

The solidus line slopes to the right because the melting temperature of any substance depends on pressure. The higher pressure created at greater depth increases the temperature needed to melt rock. In another example, at sea level with an atmospheric pressure close to 1 bar, water boils at 100°C. But if the pressure is lowered, as shown in the video below, water boils at a much lower temperature.



There are three principal ways rock behavior crosses to the right of the green solidus line to create molten magma: 1) decompression melting caused by lowering the pressure, 2) flux melting caused by adding volatiles (see more below), and 3) heat-induced melting caused by increasing the temperature. The Bowen's Reaction Series shows that minerals melt at different temperatures. Since magma is a mixture of different minerals, the solidus boundary is more of a fuzzy zone rather than a well-defined line; some minerals are melted and some remain solid. This type of rock behavior is called **partial melting** and represents real-world magmas, which typically contain solid, liquid, and volatile components.

The figure below uses P-T diagrams to show how melting can occur at three different plate tectonic settings. The green line is called the **solidus**, the melting point temperature of the rock at that pressure. Setting A is a situation (called "normal") in the middle of a stable plate in which no magma is generated. In the other three situations, rock at a lettered location with a temperature at the geothermal gradient is moved to a new P-T situation on the diagram. This shift is indicated by the arrow and its temperature relative to the solidus is shown by the red line. Partial melting occurs where the red line temperature of the rock crosses the green solidus on the diagram. Setting B is at a mid-ocean ridge *(decompression melting)* where reduction of pressure carries the rock at its temperature across the solidus. Setting C is a hotspot where decompression melting plus the *addition of heat* carries the rock across the solidus, and setting D is a subduction zone where a process called *flux melting* takes place where the solidus (melting point) is actually shifted to below the temperature of the rock.

Graphs A-D below, along with the side view of the Earth's layers in various tectonic settings (see figure), show how melting occurs in different situations. Graph A illustrates a normal situation, located in the middle of a stable plate, where no melted rock can be found. The remaining three graphs illustrate rock behavior relative to shifts in the geothermal gradient or solidus lines. Partial melting occurs when the geothermal gradient line crosses the solidus line. Graph B illustrates the behavior of rock located at a mid-ocean ridge, labeled X in the graph and side view. Reduced pressure shifts the geotherm to the right of the solidus, causing decompression melting. Graph C and label Y illustrate a hotspot situation. Decompression melting, plus an addition of heat, shifts the geotherm across the solidus. Graph D and label Z show a subduction zone, where an addition of volatiles lowers the melting point, shifting the solidus to the left of the geothermal gradient. B, C, and D all show different ways the Earth produces intersections of the geothermal gradient and the solidus, which results in melting each time.

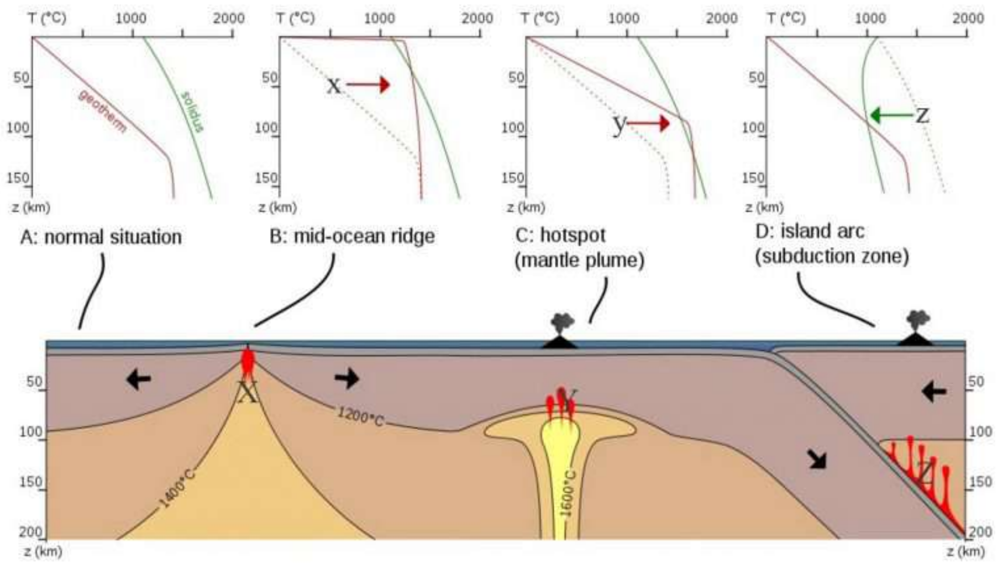


Figure 4.3.1: Four P-T diagrams showing the temperature in degrees Celsius on the x-axis and depth below the surface in kilometers (km) on the y-axis. The red line is the geothermal gradient and the green solidus line represents at temperature and pressure regime at which melting begins. Each of the four P-T diagrams is associated with a tectonic setting as shown by a side-view (cross-section) of the lithosphere and mantle.

Decompression Melting

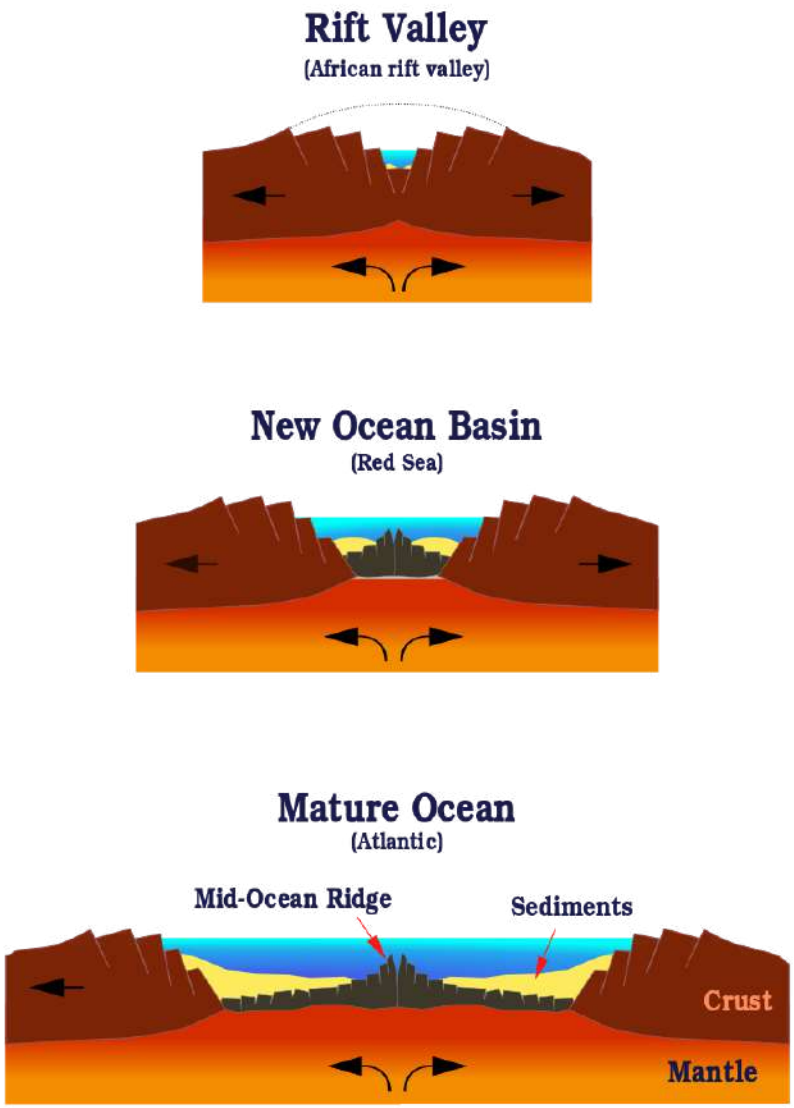


Figure 4.3.1: Progression from rift to the mid-ocean ridge, the divergent boundary types. Note the rising material in the center.

Magma is created at mid-ocean ridges via **decompression melting**. Strong convection currents cause the solid asthenosphere to slowly flow beneath the lithosphere. The upper part of the lithosphere (crust) is a poor heat conductor, so the temperature remains about the same throughout the underlying mantle material. Where the convection currents cause mantle material to rise, the pressure decreases, which causes the melting point to drop. In this situation, the rock at the temperature of the geothermal gradient is rising toward the surface, thus hotter rock is now shallower, at a lower pressure, and the rock, still at the temperature of the geothermal gradient at its old location, shifts past the its melting point (shown as the red line crossing over the solidus or green line in example B in previous figure) and partial melting starts. As this magma continues to rise, it cools and crystallizes to form new lithospheric crust.

Flux Melting

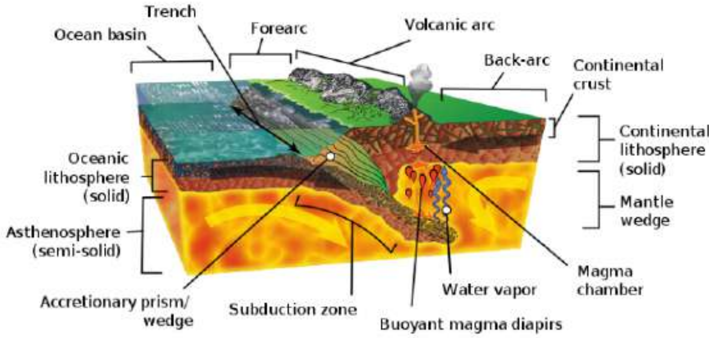


Figure 4.3.1: Diagram of ocean-continent subduction. Note water vapor driven out of hydrated minerals in the descending oceanic slab.

Flux melting or **fluid-induced melting** occurs in island arcs and subduction zones when volatile gases are added to mantle material (see figure: graph D, label Z). Flux-melted magma produces many of the volcanoes in the circum-Pacific subduction zones, also known as the Ring of Fire. The subducting slab contains oceanic lithosphere and hydrated minerals. As covered in Chapter 2, these hydrated forms are created when water ions bond with the crystal structure of silicate minerals. As the slab descends into the hot mantle, the increased temperature causes the hydrated minerals to emit water vapor and other volatile gases, which are expelled from the slab-like water being squeezed out of a sponge. The volatiles dissolve into the overlying asthenospheric mantle and decrease its melting point. In this situation the applied pressure and temperature have not changed, the mantle's melting point has been lowered by the addition of volatile substances. The previous figure (graph D) shows the green solidus line shifting to the left of and below the red geothermal gradient line, and melting begins. This is analogous to adding salt to an icy roadway. The salt lowers the freezing temperature of the solid ice so it turns into liquid water.

Heat-Induced Melting

Figure 4.3.1: Migmatite is a partially molten metamorphic rock. (Source: Peter Davis)

Heat-induced melting, transforming solid mantle into liquid magma by simply applying heat, is the least common process for generating magma (see figure: graph C, label Y). Heat-induced melting occurs at the mantle plumes or hotspots. The rock surrounding the plume is exposed to higher temperatures, the geothermal gradient crosses to the right of the green solidus line, and the rock begins to melt. The mantle plume includes rising mantle material, meaning some decompression melting is occurring as well. A small amount of magma is also generated by intense regional metamorphism (see Chapter 6). This magma becomes a hybrid metamorphic-igneous rock called migmatite.

References

6. Wallace, P. J. Volatiles in subduction zone magmas: concentrations and fluxes based on melt inclusion and volcanic gas data. *Journal of Volcanology and Geothermal Research* **140**, 217–240 (2005).

7. Petrini & Podladchikov. Lithospheric pressure–depth relationship in compressive regions of thickened crust. *Journal of Metamorphic Geology* **18**, 67–77 (2000).

Contributions and Attributions

4.4: Partial Melting and Crystallization

Even though all magmas originate from similar mantle rocks, and start out as similar magma, other things, like partial melting and crystallization processes like magmatic differentiation, can change the chemistry of the magma. This explains the wide variety of resulting igneous rocks that are found all over Earth.

Partial Melting

Because the mantle is composed of many different minerals, it does not melt uniformly. As minerals with lower melting points turn into liquid magma, those with higher melting points remain as solid crystals. This is known as partial melting. As magma slowly rises and cools into solid rock, it undergoes physical and chemical changes in a process called magmatic differentiation.

According to Bowen's Reaction Series (Section 4.2), each mineral has a unique melting and crystallization temperature. Since most rocks are made of many different minerals, when they start to melt, some minerals begin melting sooner than others. This is known as partial melting and creates magma with a different composition than the original mantle material.

The most important example occurs as magma is generated from mantle rocks (as discussed in Section 4.3). The chemistry of mantle rock (peridotite) is ultramafic, low in silicates and high in iron and magnesium. When peridotite begins to melt, the silica-rich portions melt first due to their lower melting point. If this continues, the magma becomes increasingly silica-rich, turning ultramafic mantle into mafic magma, and mafic mantle into intermediate magma. The magma rises to the surface because it is more buoyant than the mantle.

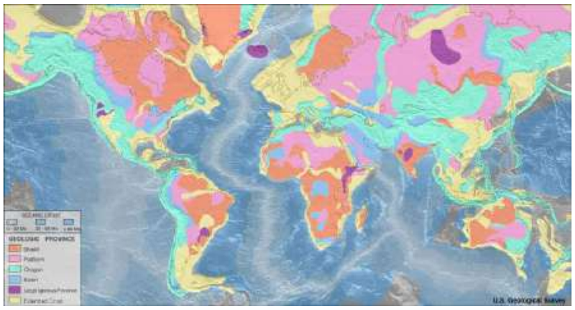


Figure 4.4.1: Geologic provinces with the Shield (orange) and Platform (pink) comprising the Craton, the stable interior of continents.

Partial melting also occurs as existing crustal rocks melt in the presence of heat from magmas. In this process, existing rocks melt, allowing the magma formed to be more felsic and less mafic than the pre-existing rock. Early in the Earth's history when the continents were forming, silica-rich magmas formed and rose to the surface and solidified into granitic continents. In the figure, the old granitic cores of the continents, called **shields**, are shown in orange.

Crystallization and Magmatic Differentiation

Liquid magma is less dense than the surrounding solid rock, so it rises through the mantle and crust. As magma begins to cool and crystallize, a process known as **magmatic differentiation** changes the chemistry of the resultant rock towards a more felsic composition. This happens via two main methods: assimilation and fractionation [8].

Figure 4.4.1: Xenoliths in Little Cottonwood Stock, Utah

During **assimilation**, pieces of country rock with a different, often more felsic, composition are added to the magma. These solid pieces may melt, which changes the composition of the original magma. At times, the solid fragments may remain intact within the cooling magma and only partially melt. The unmelted country rocks within an igneous rock mass are called **xenoliths**.

Xenoliths are also common in the processes of magma mixing and rejuvenation, two other processes that can contribute to magmatic differentiation. Magma mixing occurs when two different magmas come into contact and mix, though at times, the magmas can remain heterogeneous and create xenoliths, dikes, and other features. Magmatic rejuvenation happens when a cooled and crystallized body of rock is remelted and pieces of the original rock may remain as xenoliths.

Much of the continental lithosphere is felsic (i.e. granitic), and normally more buoyant than the underlying mafic/ultramafic mantle. When mafic magma rises through thick continental crust, it does so slowly, more slowly than when magma rises through oceanic plates. This gives the magma lots of time to react with the surrounding country rock. The mafic magma tends to assimilate felsic rock, becoming more silica-rich as it migrates through the lithosphere and changing into intermediate or felsic magma by the time it reaches the surface. This is why felsic magmas are much more common within continents.

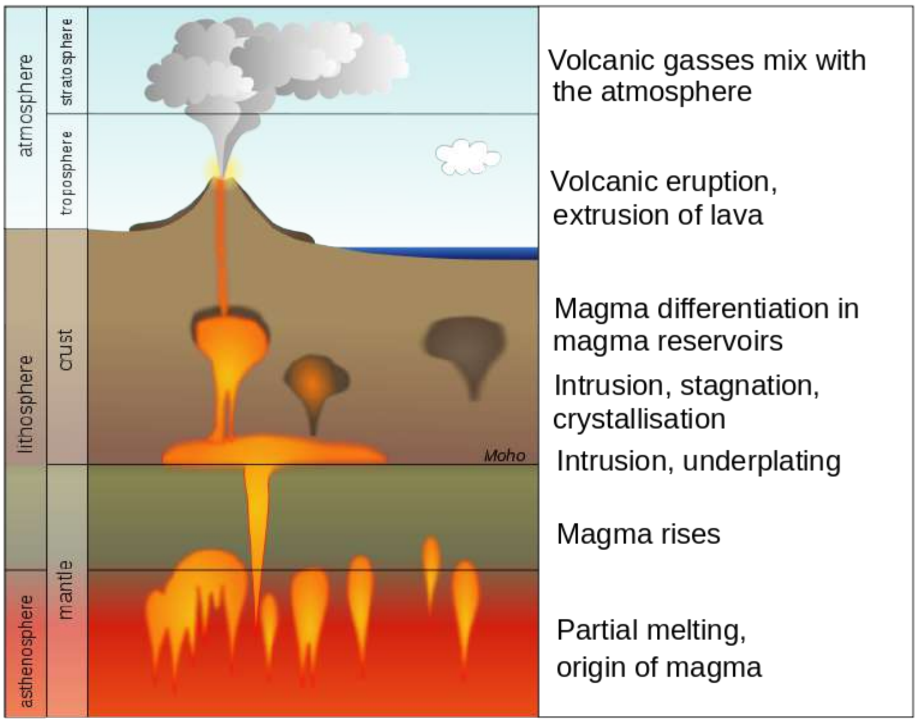


Figure 4.4.1: Rising magma diapirs in mantle and crust. Fractional crystallization occurs in the diapirs in the crust. (Source: Woudloper)

Fractionation or **fractional crystallization** is another process that increases the magma silica content, making it more felsic [9]. As the temperature drops within a magma diapir rising through the crust, some minerals will crystallize and settle to the bottom of the magma chamber, leaving the remaining melt depleted of those ions. Olivine is a mafic mineral at the top of the Bowen's Reaction series with a high melting point and a smaller percentage of silica versus other common igneous minerals. When ultramafic magma cools, the olivine crystallizes first and settles to the bottom of the magma chamber (see figure). This means the remaining melt becomes more silica-rich and felsic. As the mafic magma further cools, the next minerals on Bowen's Reaction Series (plagioclase and pyroxene) crystallize next, removing even more low-silica components from the magma, making it even more felsic. This crystal fractionation can occur in the oceanic lithosphere, but the formation of more differentiated, highly evolved felsic magmas is largely confined to continental regions where the longer time to the surface allows more fractionation to occur.

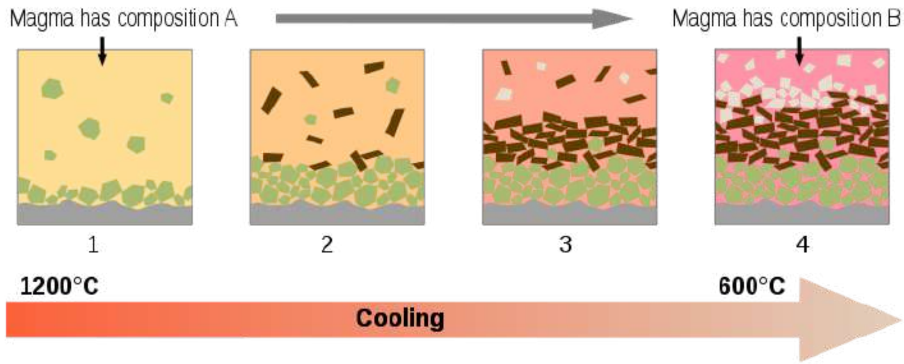


Figure 4.4.1: Schematic diagram illustrating fractional crystallization. If magma at composition A is ultramafic, as the magma cools it changes composition as different minerals crystallize from the melt and settle to the bottom of the magma chamber. In section 1, olivine crystallizes; section 2: olivine and pyroxene crystallize; section 3: pyroxene and plagioclase crystallize; and section 4: plagioclase crystallizes. The crystals are separated from the melt and the remaining magma (composition B) is more silica-rich. (Source: Woudloper)

References

8. Reid, J. B., Evans, O. C. & Fates, D. G. Magma mixing in granitic rocks of the central Sierra Nevada, California. *Earth and Planetary Science Letters* **66**, 243–261 (1983).

9. Bateman, P. C. & Chappell, B. W. Crystallization, fractionation, and solidification of the Tuolumne Intrusive Series, Yosemite National Park, California. *Geological Society of America Bulletin* **90**, 465–482 (1979).

Contributions and Attributions

4.5: Volcanism

When magma emerges onto the Earth's surface, the molten rock is called lava. A **volcano** is a type of land formation created when lava solidifies into rock. Volcanoes have been an important part of human society for centuries, though their understanding has greatly increased as our understanding of plate tectonics has made them less mysterious. This section describes volcano location, type, hazards, and monitoring.

Distribution and Tectonics

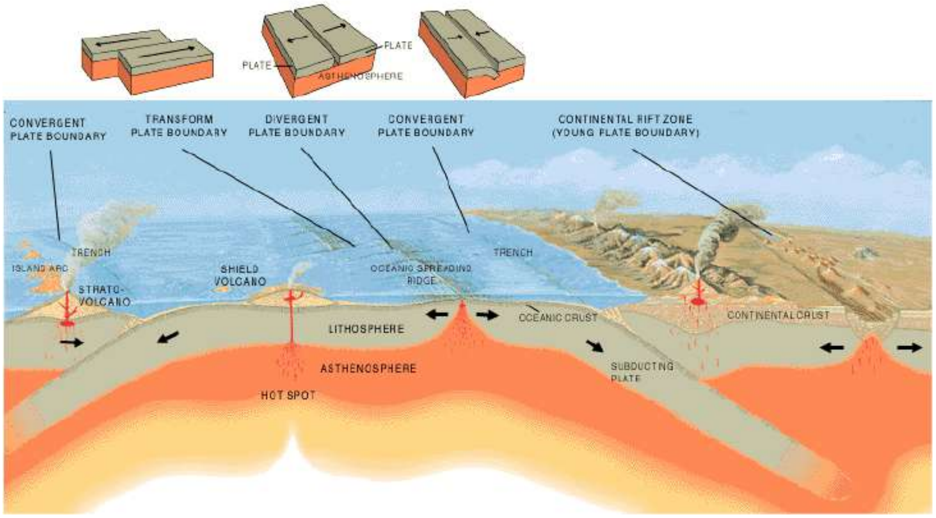


Figure 4.5.1: Association of volcanoes with plate boundaries

Most volcanoes are interplate volcanoes. Interplate volcanoes are located at active plate boundaries created by volcanism at mid-ocean ridges, subduction zones, and continental rifts. The prefix "*inter-*" means between. Some volcanoes are intraplate volcanoes. The prefix "*intra-*" means within, and intraplate volcanoes are located within tectonic plates, far removed from plate boundaries. Many intraplate volcanoes are formed by hotspots.

Volcanoes at Mid-Ocean Ridges

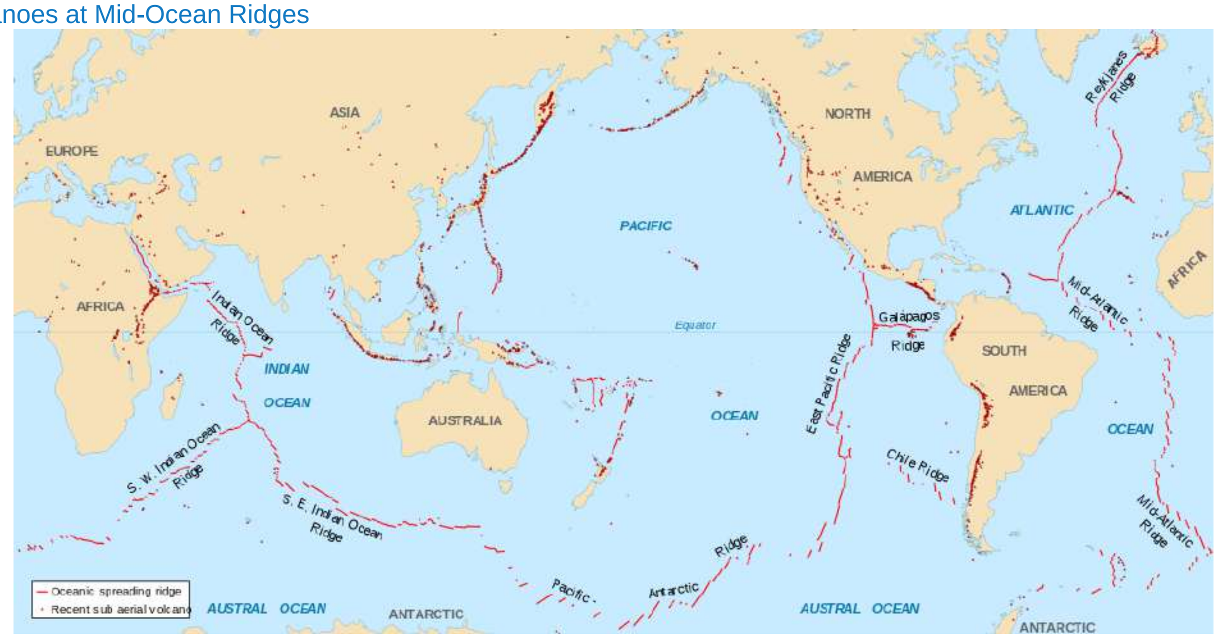


Figure 4.5.1: Map of mid-ocean ridges throughout the world.

Most volcanism on Earth occurs on the ocean floor along mid-ocean ridges, a type of divergent plate boundary (see Chapter 2). These interplate volcanoes are also the least observed and famous since most of them are located under 3,000-4,500 m (10,000-15,000 ft) of the ocean and the eruptions are slow, gentle, and oozing. One exception is the interplate volcanoes of Iceland. The diverging and thinning oceanic plates allow hot mantle rock to rise, releasing pressure and causing decompression melting. Ultramafic mantle rock, consisting largely of peridotite, partially melts and generates magma that is basaltic. Because of this, almost all volcanoes on the ocean floor are basaltic.

In fact, most oceanic lithosphere is basaltic near the surface, with phaneritic gabbro and ultramafic peridotite underneath [10].

Figure 4.5.1: Pillow basalt on the seafloor near Hawaii.

When basaltic lava erupts underwater it emerges in small explosions and/or forms pillow-shaped structures called pillow basalts. These seafloor eruptions enable entire underwater ecosystems to thrive in the deep ocean around mid-ocean ridges. This ecosystem exists around tall vents emitting black, hot mineral-rich water called deep-sea hydrothermal vents, also known as black smokers.

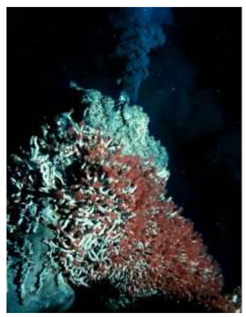


Figure 4.5.1: Black smoker hydrothermal vent with a colony of giant 2+ m (6'+) tube worms.

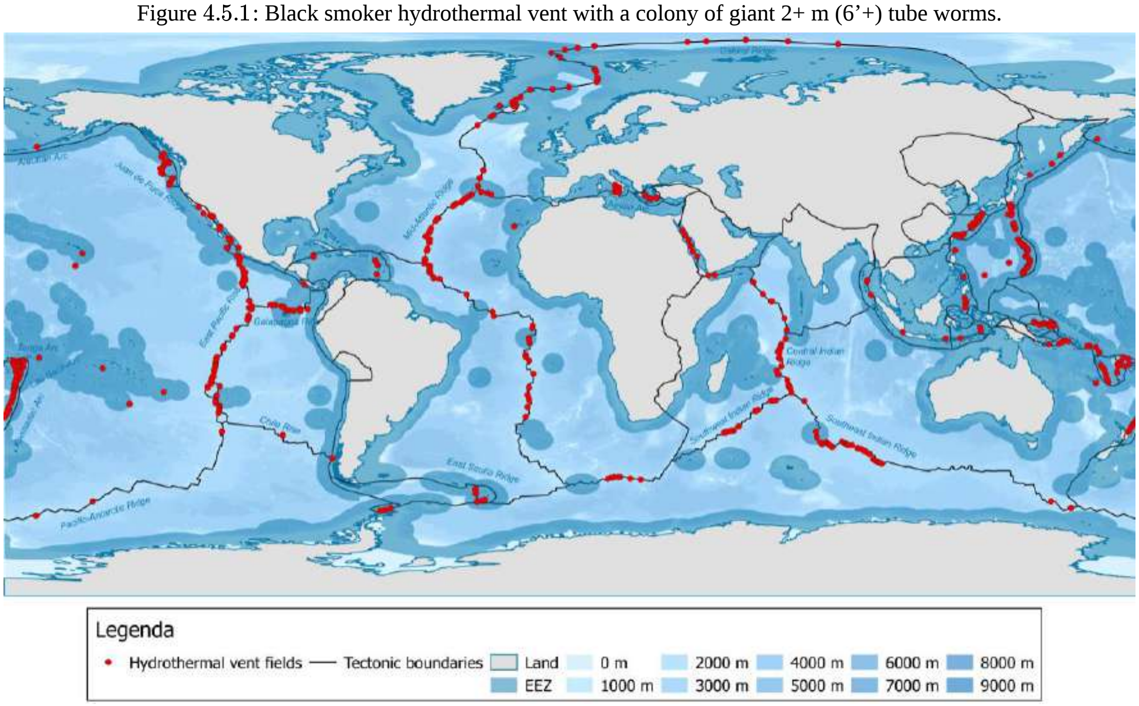
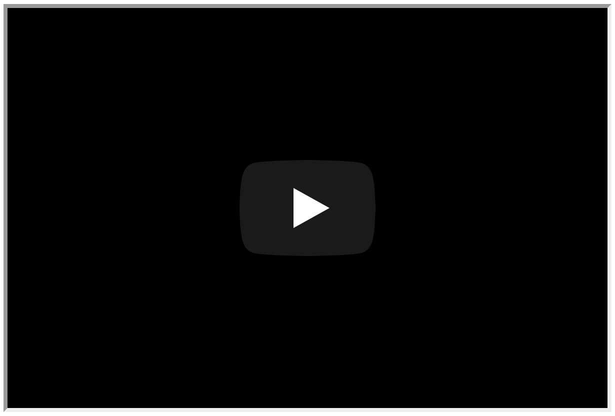
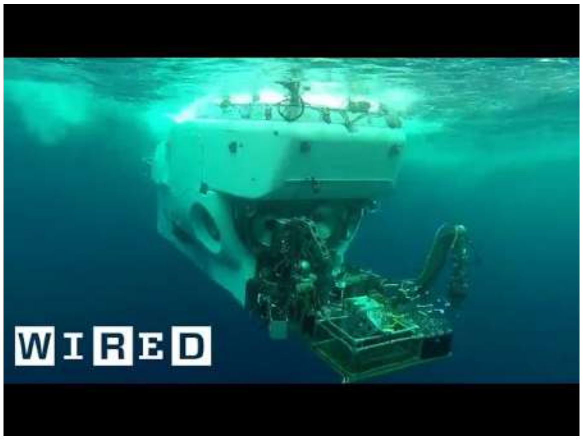
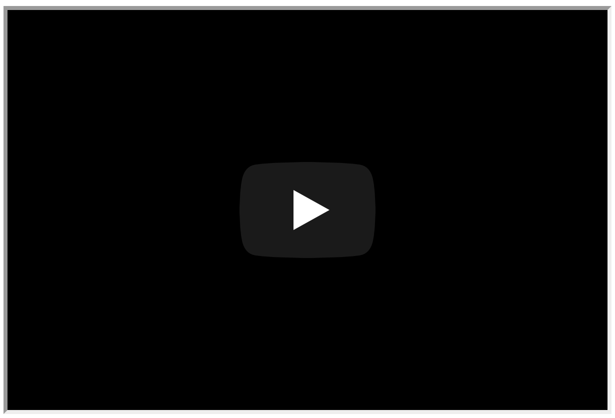


Figure 4.5.1: Distribution of hydrothermal vent fields.

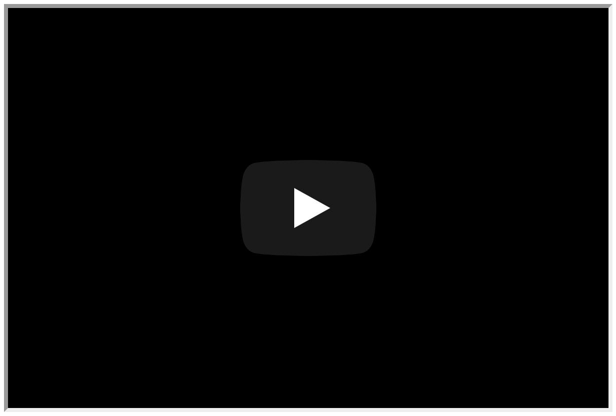
Without sunlight to support photosynthesis, these organisms instead utilize a process called **chemosynthesis**. Certain bacteria are able to turn hydrogen sulfide (H_2S), a gas that smells like rotten eggs, into life-supporting nutrients and water. Larger organisms may eat these bacteria or absorb nutrients and water produced by bacteria living symbiotically inside their bodies [11]. The three videos show some of the ecosystems found around deep-sea hydrothermal vents.




WIRED




WIRED



Volcanoes at Subduction Zones

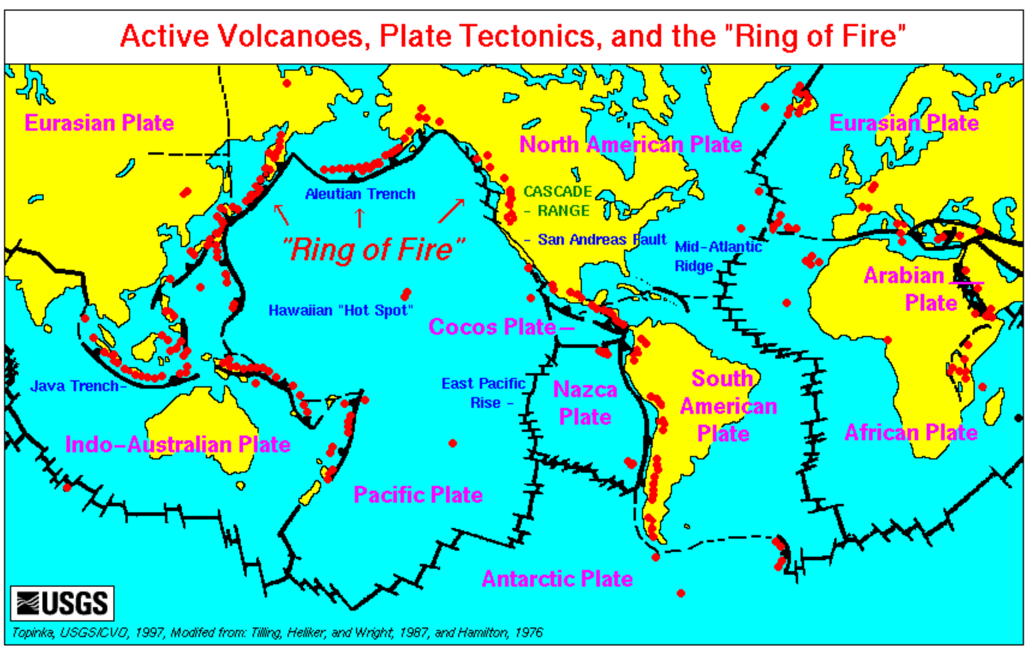


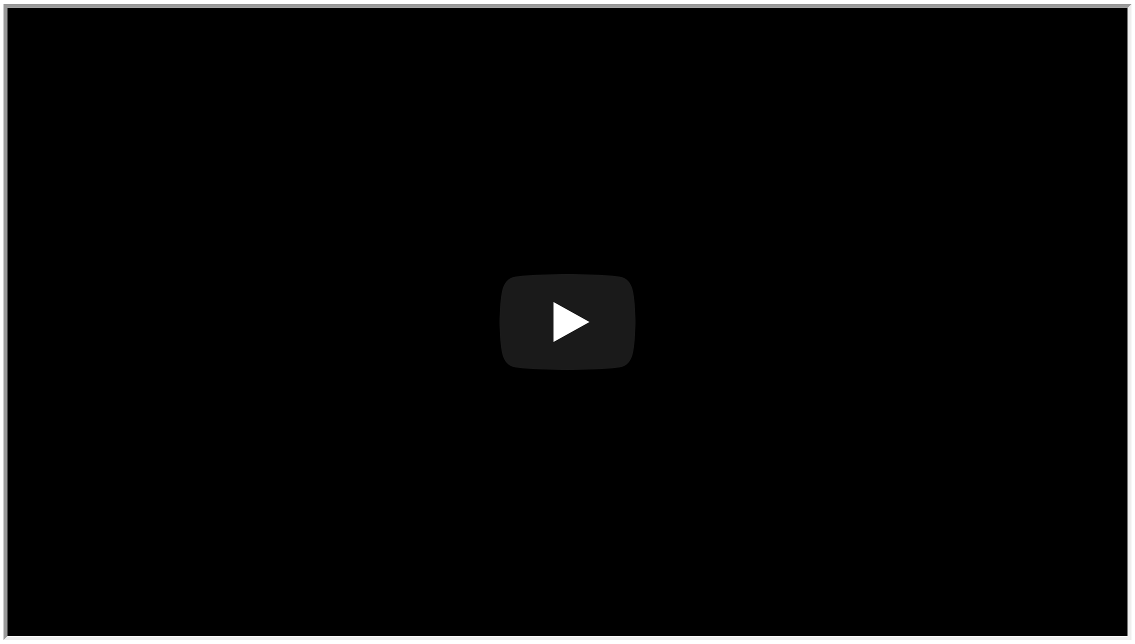
Figure 4.5.1: Distribution of volcanoes on the planet. Click here for an interactive map of volcano distributions.

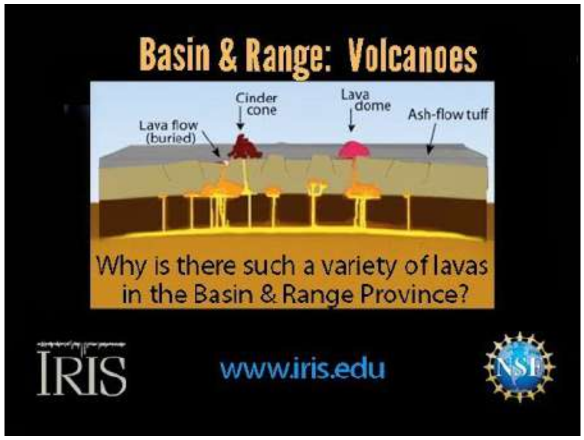
The second most commonly found location for volcanism is adjacent to subduction zones, a type of convergent plate boundary (see Chapter 2). The process of subduction expels water from hydrated minerals in the descending slab, which causes flux melting in the overlying mantle rock. Because subduction volcanism occurs in a volcanic arc, the thickened crust promotes partial melting and magma differentiation. These evolve the mafic magma from the mantle into more silica-rich magma. The Ring of Fire surrounding the Pacific Ocean, for example, is dominated by subduction-generated eruptions of mostly silica-rich lava; the volcanoes and plutons consist largely of intermediate-to-felsic rock such as andesite, rhyolite, pumice, and tuff.

Volcanoes at Continental Rifts

Figure 4.5.1: Basaltic cinder cones of the Black Rock Desert near Beaver, Utah.

Some volcanoes are created at continental rifts, where crustal thinning is caused by diverging lithospheric plates, such as the East African Rift Basin in Africa. Volcanism caused by crustal thinning without continental rifting is found in the Basin and Range Province in North America. In this location, volcanic activity is produced by rising magma that stretches the overlying crust (see figure). Lower crust or upper mantle material rises through the thinned crust, releases pressure, and undergoes decompression-induced partial melting. This magma is less dense than the surrounding rock and continues to rise through the crust to the surface, erupting as basaltic lava. These eruptions usually result in flood basalts, cinder cones, and basaltic lava flows (see video). Relatively young cinder cones of basaltic lava can be found in south-central Utah, in the Black Rock Desert Volcanic Field, which is part of the zone of Basin and Range crustal extension. These Utah cinder cones and lava flows started erupting around 6 million years ago, with the last eruption occurring 720 years ago [12].





Hotspots

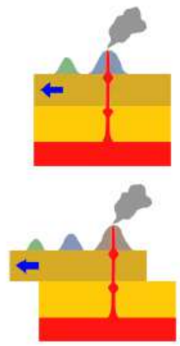


Figure 4.5.1: Diagram showing a non-moving source of magma (mantle plume) and a moving overriding plate.

Hotspots are the main source of intraplate volcanism. **Hotspots** occur when lithospheric plates glide over a hot mantle plume, an ascending column of solid heated rock originating from deep within the mantle. The mantle plume generates melts as material rises, with the magma rising even more. When the ascending magma reaches the lithospheric crust, it spreads out into a mushroom-shaped head that is tens to hundreds of kilometers across.

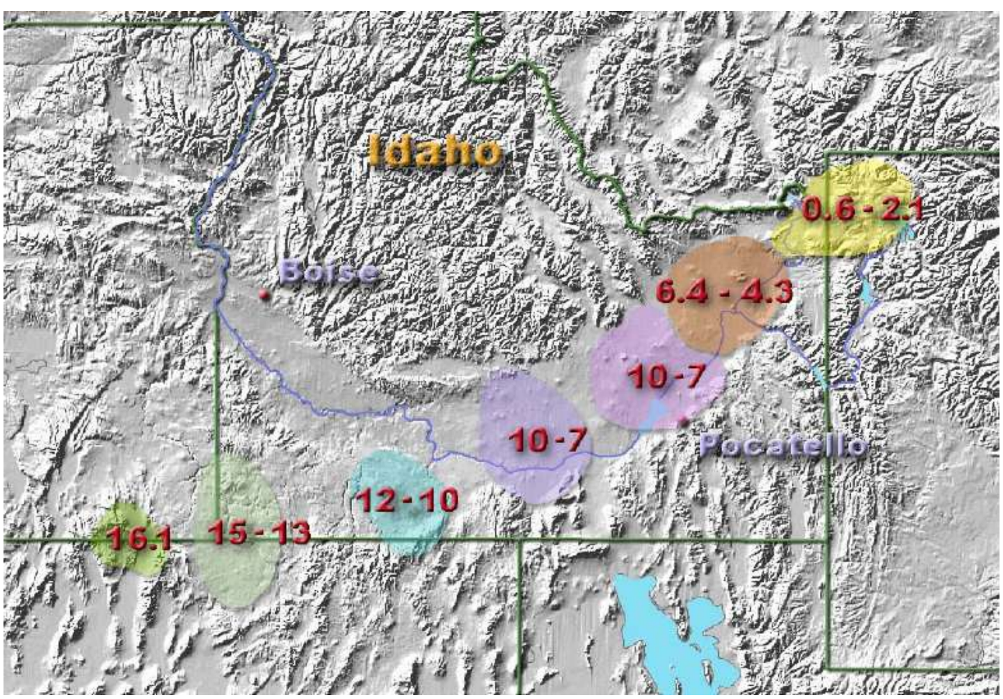


Figure 4.5.1: The track of the Yellowstone hotspot, which shows the age of different eruptions millions of years ago.

Since most mantle plumes are located beneath the oceanic lithosphere, the early stages of intraplate volcanism typically take place underwater. Over time, basaltic volcanoes may build up from the seafloor into islands, such as the Hawaiian Islands [13]. Where a hotspot is found under a continental plate, contact with the hot mafic magma may cause the overlying felsic rock to melt and mix with the mafic material below, forming intermediate magma. Or the felsic magma may continue to rise, and cool into a granitic batholith or erupt as a felsic volcano. The Yellowstone caldera is an example of hotspot volcanism that resulted in an explosive eruption.

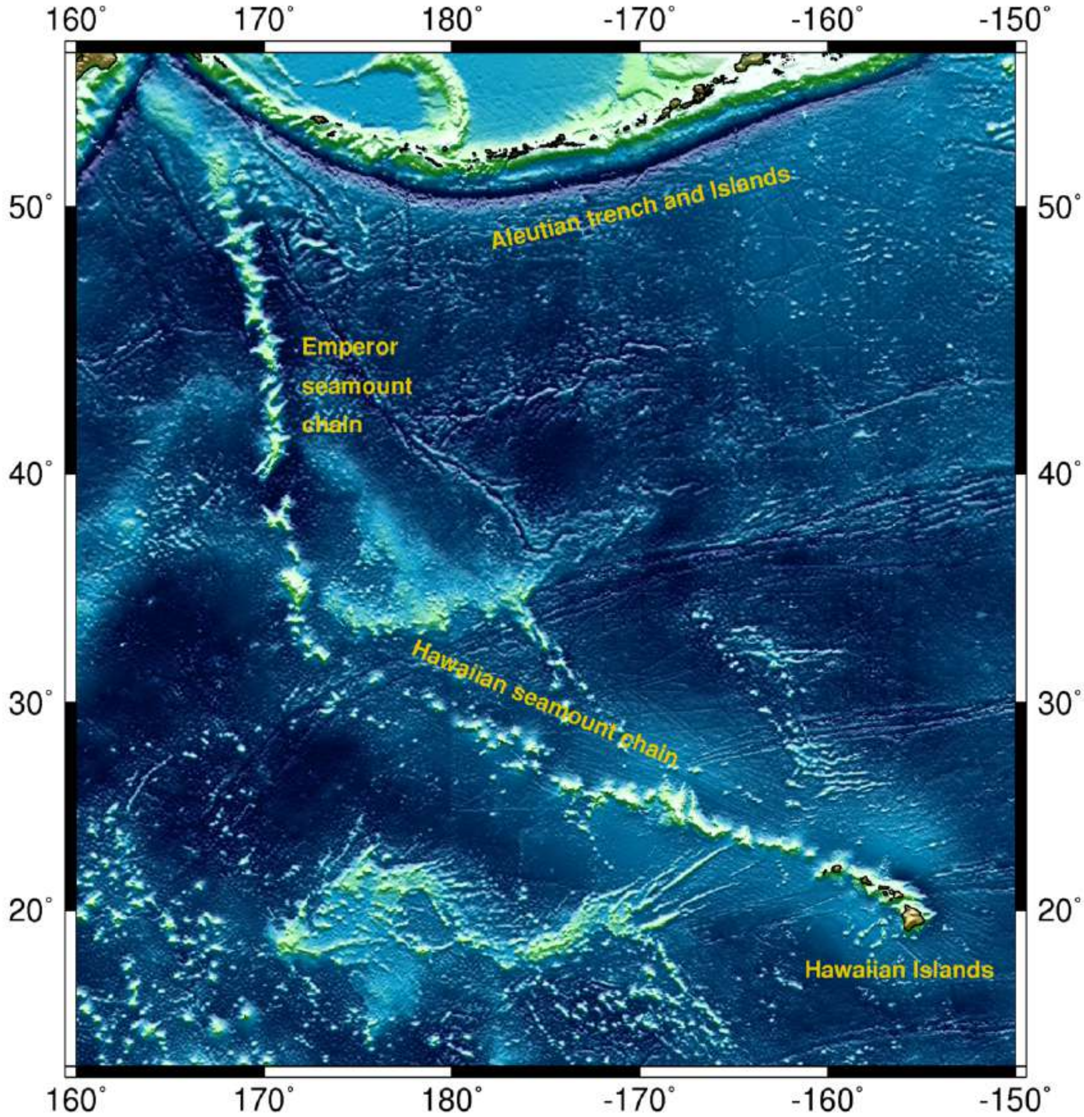
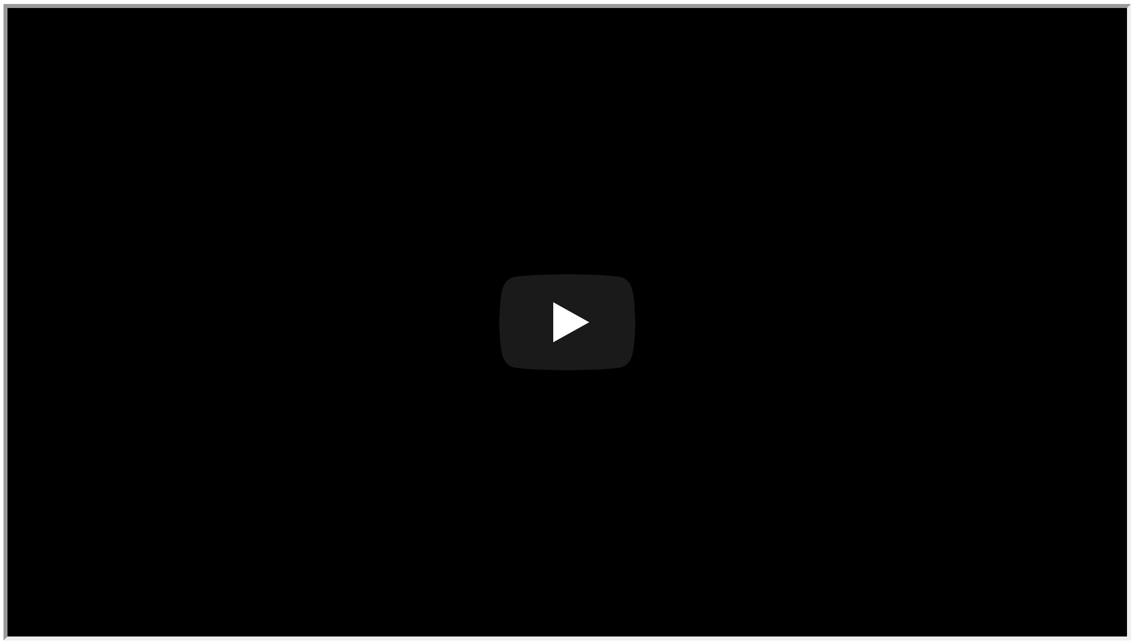
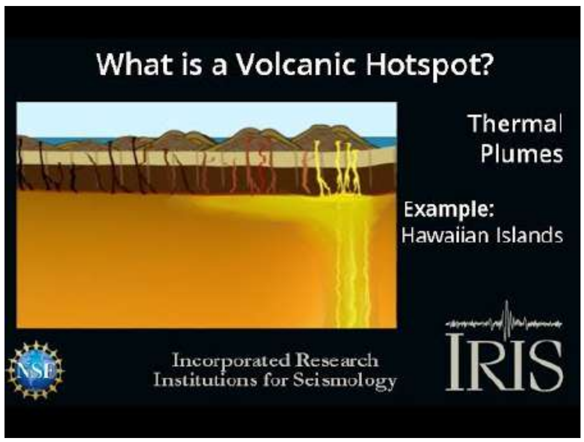


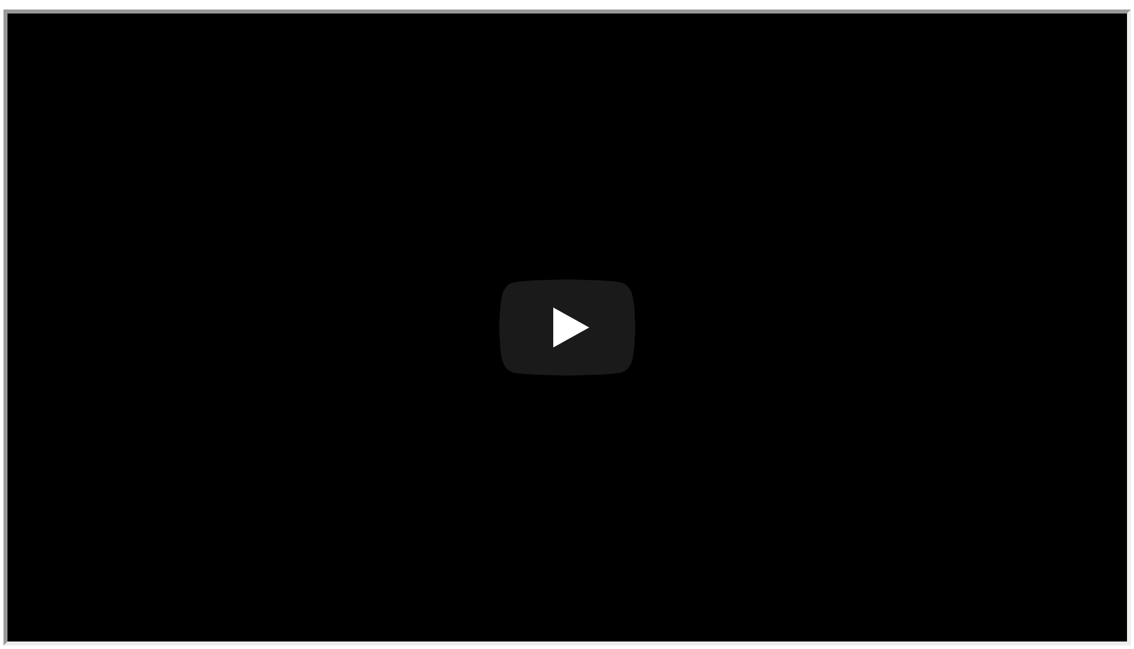
Figure 4.5.1: The Hawaii-Emperor seamount and island chain.

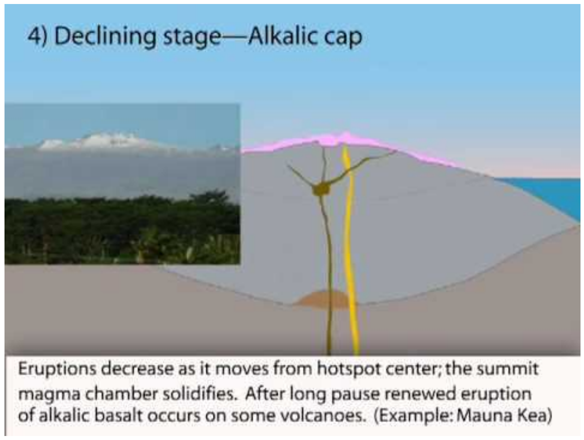
A zone of actively erupting volcanism connected to a chain of extinct volcanoes indicates intraplate volcanism located over a hotspot. These volcanic chains are created by the overriding oceanic plate slowly moving over a hotspot mantle plume. These chains are seen on the seafloor and continents and include volcanoes that have been inactive for millions of years. The Hawaiian Islands on the Pacific Oceanic plate are the active end of a long volcanic chain that extends from the northwest Pacific Ocean to the Emperor Seamounts, all the way to the subduction zone beneath the Kamchatka Peninsula. The overriding North American continental plate moved across a mantle plume hotspot for several million years, creating a chain of volcanic calderas that extends from Southwestern Idaho to the presently active Yellowstone caldera in Wyoming.

Two three-minute videos (below) illustrate hotspot volcanoes.



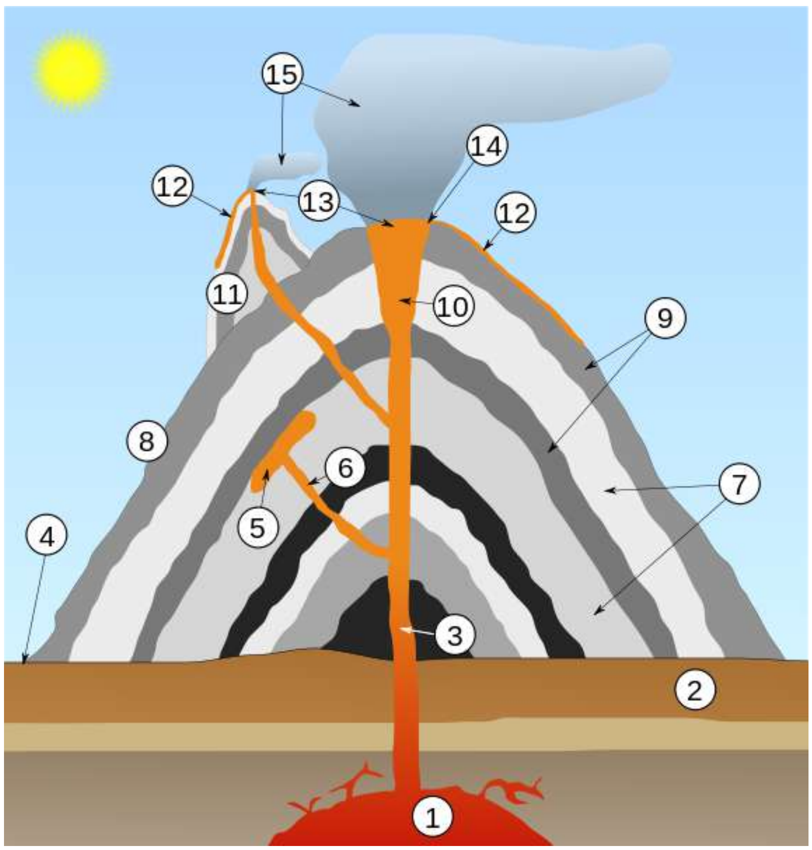

What is a Volcanic Hotspot?
Thermal Plumes
Example:
Hawaiian Islands
NSF
Incorporated Research Institutions for Seismology
IRIS





Volcano Features and Types

There are several different types of volcanoes based on their shape, eruption style, magmatic composition, and other aspects.



The figure shows the main features of a typical stratovolcano: 1) **magma chamber**, 2) upper layers of lithosphere, 3) the **conduit** or narrow pipe through which the lava erupts, 4) the base or edge of the volcano, 5) a **sill** of magma between layers of the volcano, 6) a **diapir** or feeder tube to the sill, 7) layers of **tephra** (ash) from previous eruptions, 8) & 9) layers of lava erupting from the vent and flowing down the sides of the volcano, 10) the **crater** at the top of the volcano, 11) layers of lava and tephra on 12), a parasitic cone. A **parasitic cone** is a small volcano located on the flank of a larger volcano such as Shastina on Mount Shasta. Kilauea sitting on the flank of Mauna Loa is not considered a parasitic cone because it has its own separate magma chamber [13]. 13) the **vents** of the parasite and the main volcano, 14) the rim of the crater, 15) clouds of ash blown into the sky by the eruption; this settles back onto the volcano and surrounding land.

Figure 4.5.1: Mt. Shasta in California with Shastina, its parasitic cone

Figure 4.5.1: Oregon’s Crater Lake was formed about 7700 years ago after the eruption of Mount Mazama.

The largest craters are called **calderas**, such as the Crater Lake Caldera in Oregon. Many volcanic features are produced by **viscosity**, a basic property of lava. Viscosity is the resistance to flowing by a fluid. Low viscosity magma flows easily more like syrup, the basaltic volcanism that occurs in Hawaii on shield volcanoes. High viscosity means a thick and sticky magma, typically felsic or intermediate, that flows slowly, similar to toothpaste.

Shield Volcano

Figure 4.5.1: Kilauea in Hawaii.

The largest volcanoes are **shield volcanoes**. They are characterized by broad low-angle flanks, small vents at the top, and mafic magma chambers. The name comes from the side view, which resembles a medieval warrior’s shield. They are typically associated with hotspots, mid-ocean ridges, or continental rifts with rising upper mantle material. The low-angle flanks are built up slowly from numerous low-viscosity basaltic lava flows that spread out over long distances. The basaltic lava erupts effusively, meaning the eruptions are small, localized, and predictable.

Figure 4.5.1: Eruption of Kilauea in 2018 produced relatively high viscosity a'a lava shown here crossing and destroying a road. This eruption caused much property damage in Hawaii.

Typically, shield volcano eruptions are not much of a hazard to human life—although non-explosive eruptions of Kilauea (Hawaii) in 2018 produced uncharacteristically large lavas that damaged roads and structures. Mauna Loa (see USGS page) and Kilauea (see USGS page) in Hawaii are examples of shield volcanoes. Shield volcanoes are also found in Iceland, the Galapagos Islands, Northern California, Oregon, and the East African Rift [14].

Figure 4.5.1: Olympus Mons, an enormous shield volcano on Mars, the largest volcano in the solar system.

The largest volcanic edifice in the Solar System is Olympus Mons on Mars. This (possibly extinct) shield volcano covers an area the size of the state of Arizona. This may indicate the volcano erupted over a hotspot for millions of years, which means Mars had little, if any, plate tectonic activity. [15] [16].

Figure 4.5.1: Ropey pahoehoe lava

Basaltic lava forms special landforms based on magma temperature, composition, and content of dissolved gases and water vapor. The two main types of basaltic volcanic rock have Hawaiian names—*pahoehoe* and *aa*. **Pahoehoe** might come from low-viscosity lava that flows easily into ropey strands.

Figure 4.5.1: Blocky a'a lava

Aa (sometimes spelled a'a or ʻaʻā and pronounced "ah-ah") is more viscous and has a crumbly blocky appearance [17]. The exact details of what forms the two types of flows are still up for debate. Felsic lavas have lower temperatures and more silica and thus are higher viscosity. These also form aa-style flows.

Figure 4.5.1: Volcanic fissure and flow, which could eventually form a lava tube.

Low-viscosity, fast-flowing basaltic lava tends to harden on the outside into a tube and continue to flow internally. Once lava flow subsides, the empty outer shell may be left as a lava tube. Lava tubes, with or without collapsed roofs, make famous caves in Hawaii, Northern California, the Columbia River Basalt Plateau of Washington and Oregon, El Malpais National Monument in New Mexico, and Craters of the Moon National Monument in Idaho.

Fissures are cracks that commonly originate from shield-style eruptions. Lava emerging from fissures is typically mafic and very fluid. The 2018 Kiluaea eruption included fissures associated with the lava flows. Some fissures are caused by the volcanic seismic activity rather than lava flows. Some fissures are influenced by plate tectonics, such as the common fissures located parallel to the divergent boundary in Iceland.

Figure 4.5.1: Devils Tower in Wyoming has columnar jointing.

Figure 4.5.1: Columnar jointing on Giant's Causeway in Ireland.

Cooling lava can contract into columns with semi-hexagonal cross sections called **columnar jointing**. This feature forms the famous Devils Tower in Wyoming, possibly an ancient volcanic vent from which the surrounding layers of lava and ash have been removed by erosion. Another well-known exposed example of columnar jointing is the Giant's Causeway in Ireland.

Stratovolcano

Figure 4.5.1: Mount Rainier towers over Tacoma, Washington.

A **stratovolcano**, also called a composite cone volcano, has steep flanks, a symmetrical cone shape, a distinct crater, and rises prominently above the surrounding landscape. The term composite refers to the alternating layers of pyroclastic fragments like ash and bombs, and solidified lava flows of varying composition. Examples include Mount Rainier in Washington state and Mount Fuji in Japan.

Figure 4.5.1: Mt. Fuji in Japan, a typical stratovolcano: symmetrical, increasing slope, and a visible crater at the top.

Stratovolcanoes usually have felsic to intermediate magma chambers, but can even produce mafic lavas. Stratovolcanoes have viscous lava flows and domes, punctuated by explosive eruptions. This produces volcanoes with steep flanks [14].

Lava Domes

Figure 4.5.1: Lava domes have started the rebuilding process at Mount St. Helens, Washington.

Lava domes are accumulations of silica-rich volcanic rock, such as rhyolite and obsidian. Too viscous to flow easily, the felsic lava tends to pile up near the vent in blocky masses. Lava domes often form in a vent within the collapsed crater of a stratovolcano and grow by internal expansion. As the dome expands, the outer surface cools, hardens, and shatters, and spills loose fragments down the sides. Mount Saint Helens has a good example of a lava dome inside of a collapsed stratovolcano crater. Examples of stand-alone lava domes are Chaiten in Chile and Mammoth Mountain in California [18][14].

Caldera

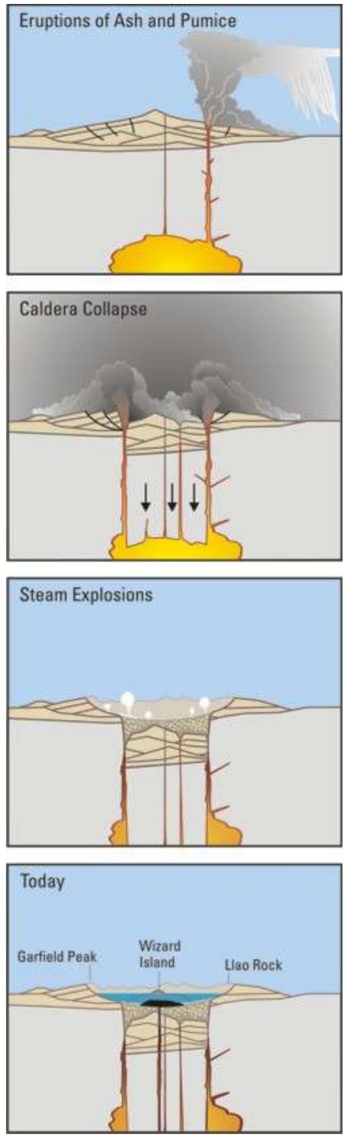


Figure 4.5.1: Timeline of events at Mount Mazama.


Figure 4.5.1: Wizard Island sits in the caldera at Crater Lake.

Calderas are steep-walled, basin-shaped depressions formed by the collapse of a volcanic edifice into an empty magma chamber. Calderas are generally very large, with diameters of up to 25 km (15.5 mi). The term caldera specifically refers to a volcanic vent but it is frequently used to describe a volcano type. Caldera volcanoes are typically formed by eruptions of high-viscosity felsic lava having high volatiles content.

Crater Lake, Yellowstone, and the Long Valley Caldera are good examples of this type of volcanism. The caldera at Crater Lake National Park in Oregon was created about 6,800 years ago when Mount Mazama, a composite volcano, erupted in a huge explosive blast. The volcano ejected large amounts of volcanic ash and rapidly drained the magma chamber, causing the top to collapse into a large depression that later filled with water. Wizard Island in the middle of the lake is a later resurgent lava dome that formed within the caldera basin [14].

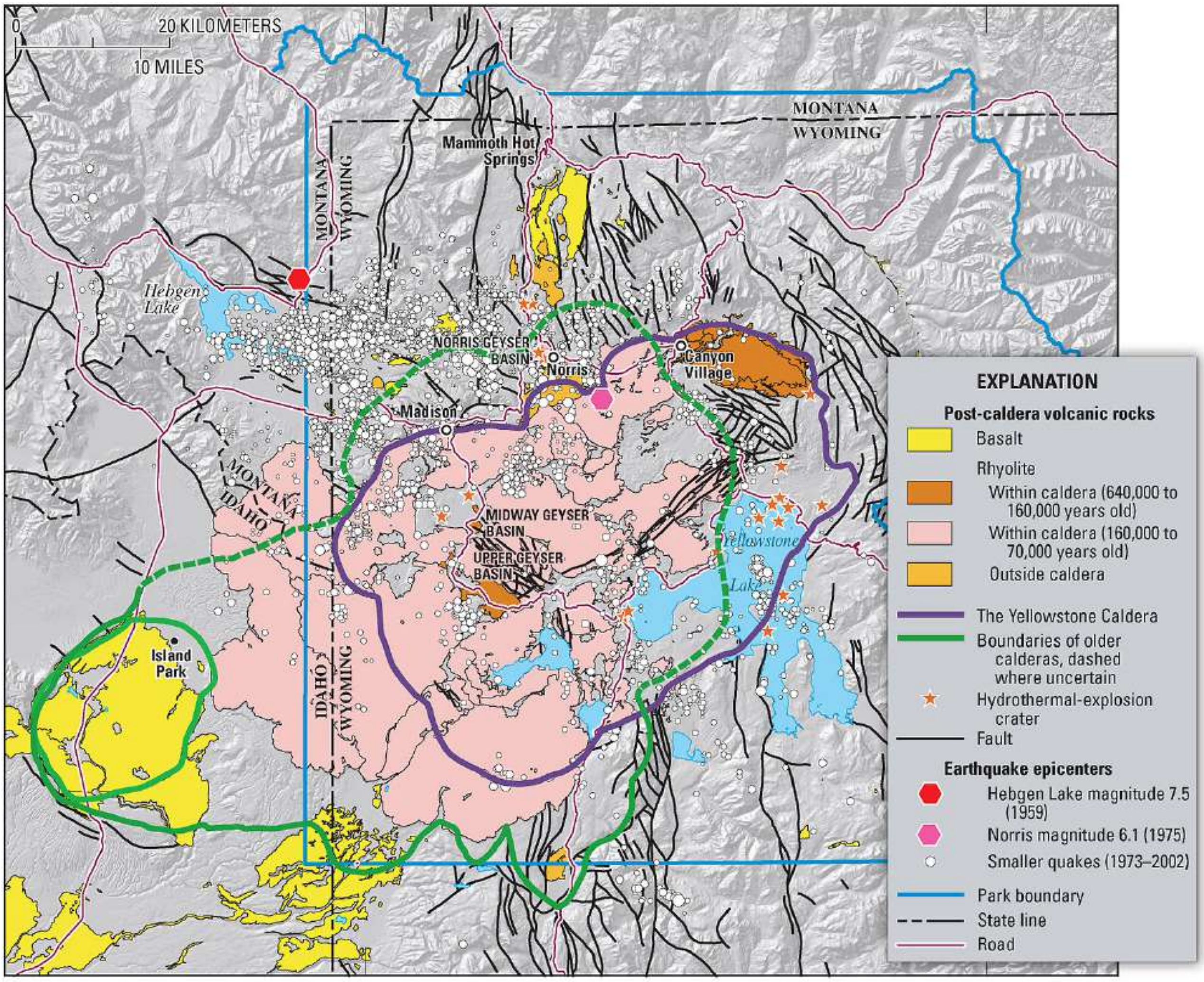


Figure 4.5.1: Map of calderas and related rocks around Yellowstone.

The Yellowstone volcanic system erupted three times in the recent geologic past—2.1, 1.3, and 0.64 million years ago—leaving behind three caldera basins. Each eruption created large rhyolite lava flows as well as pyroclastic flows that solidified into tuff formations. These extra-large eruptions rapidly emptied the magma chamber, causing the roof to collapse and form a caldera. The youngest of the three calderas contains most of Yellowstone National Park, as well as two resurgent lava domes. The calderas are difficult to see today due to the amount of time since their eruptions and subsequent erosion and glaciation.

Yellowstone volcanism started about 17-million years ago as a hotspot under the North American lithospheric plate near the Oregon/Nevada border. As the plate moved to the southwest over the stationary hotspot, it left behind a track of past volcanic activities. Idaho's Snake River Plain was created from volcanism that produced a series of calderas and lava flows. The plate eventually arrived at its current location in northwestern Wyoming, where hotspot volcanism formed the Yellowstone calderas [19].

Figure 4.5.1: Several prominent ash beds found in North America, including three Yellowstone eruptions shaded pink (Mesa Falls, Huckleberry Ridge, and Lava Creek), the Bisho Tuff ash bed (brown dashed line), and the modern May 18th, 1980 ash fall from Mt. St. Helens (yellow).

The Long Valley Caldera near Mammoth, California, is the result of a large volcanic eruption that occurred 760,000 years ago. The explosive eruption dumped enormous amounts of ash across the United States, in a manner similar to the Yellowstone eruptions. The Bishop Tuff deposit near Bishop, California, is made of ash from this eruption. The current caldera basin is 17 km by 32 km (10 mi by 20 mi), large enough to contain the town of Mammoth Lakes, major ski resort, airport, major highway, resurgent dome, and several hot springs [20].

Cinder Cone

Figure 4.5.1: Sunset Crater, Arizona is a cinder cone.

Cinder cones are small volcanoes with steep sides and made of pyroclastic fragments that have been ejected from a pronounced central vent. The small fragments are called **cinders** and the largest are **volcanic bombs**. The eruptions are usually short-lived events, typically consisting of mafic lavas with a high content of volatiles. Hot lava is ejected into the air, cooling and solidifying into fragments that accumulate on the flank of the volcano. Cinder cones are found throughout western North America [14].

Figure 4.5.1: Soon after the birth of Parícutin in 1943.

Figure 4.5.1: Lava from Parícutin covered the local church and destroyed the town of San Juan, Mexico

A recent and striking example of a cinder cone is the eruption near the village of Parícutin, Mexico that started in 1943 [21]. The cinder cone started explosively shooting cinders out of the vent in the middle of a farmer's field. The volcanism quickly built up the cone to a height of over 90 m (300 ft) within a week and 365 m (1,200 ft) within the first 8 months. After the initial explosive eruption of gases and cinders, basaltic lava poured out from the base of the cone. This is a common order of events for cinder cones: violent eruption, cone, and crater formation, low-viscosity lava flow from the base. The cinder cone is not strong enough to support a column of lava rising to the top of the crater, so the lava breaks through and emerges near the bottom of the volcano. During nine years of eruption activity, the ashfall covered about 260 km^2 (100 mi^2) and destroyed the nearby town of San Juan [14].

Flood Basalts

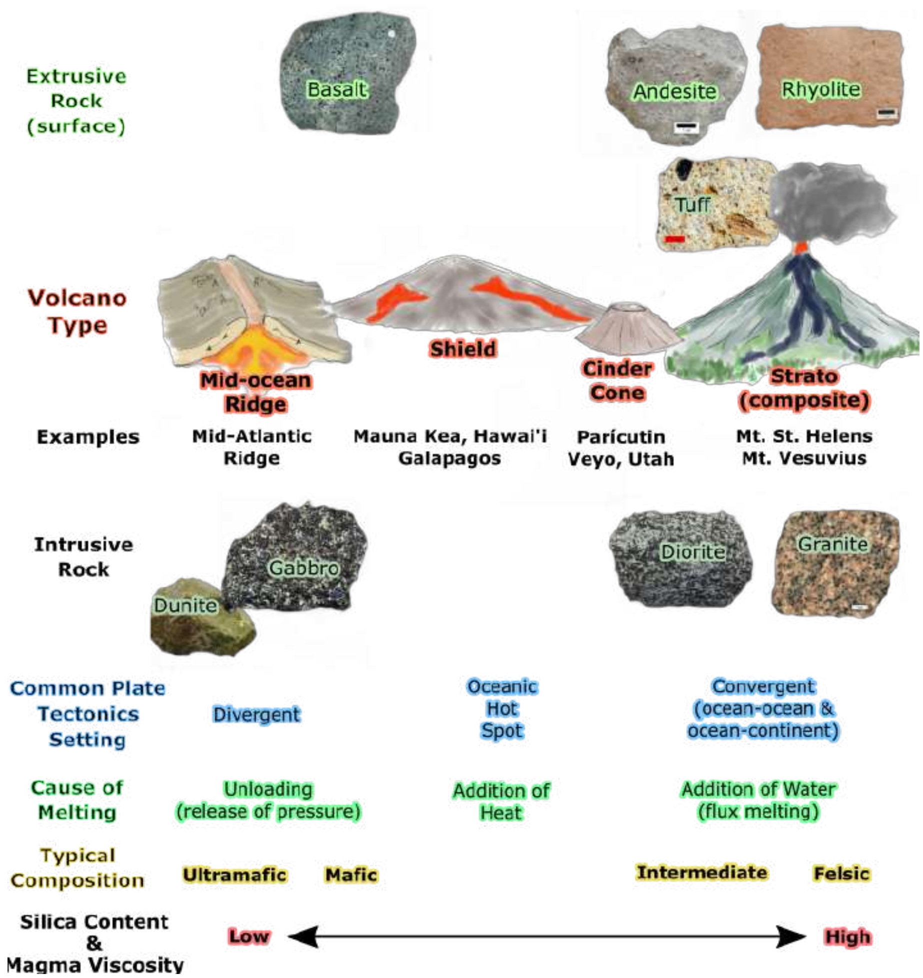
Figure 4.5.1: World map of flood basalts. Note the largest is the Siberian Traps

A rare volcanic eruption type, unobserved in modern times, is the **flood basalt**. Flood basalts are some of the largest and lowest viscosity types of eruptions known. They are not known from any eruption in human history, so the exact mechanisms of eruption are still mysterious. Some famous examples include the Columbia River Flood Basalts in Washington, Oregon, and Idaho, the Deccan Traps, which cover about 1/3 of the country of India, and the Siberian Traps, which may have been involved in the Earth's largest mass extinction (see chapter 8).

Carbonatites

Figure 4.5.1: Crater of Ol Doinyo Lengai in 2011. Note the white carbonatite in the walls of the crater.

Arguably the most unusual volcanic activity is **carbonatite** eruptions. Only one actively erupting carbonatite volcano exists on Earth today, Ol Doinyo Lengai, in the East African Rift Zone of Tanzania. While all other volcanism on Earth originates from silicate-based magma, carbonatites are a product of carbonate-based magma and produce volcanic rocks containing greater than 50% carbonate minerals. Carbonatite lavas are very low viscosity and relatively cold for lava. The erupting lava is black and solidifies to brown/grey rock that eventually turns white. These rocks are occasionally found in the geologic record and require special study to distinguish them from metamorphic marbles (see Chapter 6). They are mostly associated with continental rifting [22].



Igneous rock types and related volcano types. Mid-ocean ridges and shield volcanoes represent more mafic compositions, and strato (composite) volcanoes generally represent a more intermediate or felsic composition and a convergent plate tectonic boundary. Note that there are exceptions to this generalized layout of volcano types and igneous rock composition.

Volcanic Hazards and Monitoring

Figure 4.5.1: General diagram of volcanic hazards.

While the most obvious volcanic hazard is lava, the dangers posed by volcanoes go far beyond lava flows. For example, on May 18, 1980, Mount Saint Helens (Washington, United States) erupted with an explosion and landslide that removed the upper 400 m (1,300 ft) of the mountain. The initial explosion was immediately followed by a lateral blast, which produced a pyroclastic flow that covered nearly 600 km^2 (230 mi^2) of the forest with hot ash and debris [23]. The pyroclastic flow moved at speeds of 80-130 kph (50-80 mph), flattening trees and ejecting clouds of ash into the air. The USGS video provides an account of this explosive eruption that killed 57 people [24].

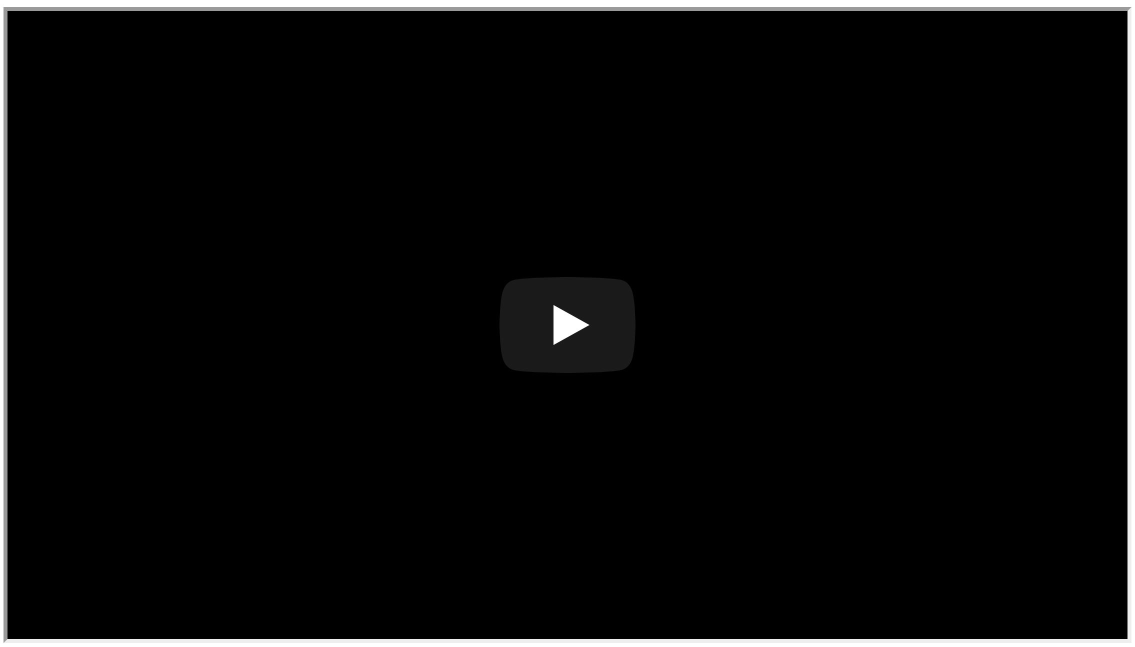


Figure 4.5.1: Human remains from the 79 CE eruption of Vesuvius.

In 79 AD, Mount Vesuvius, located near Naples, Italy, violently erupted sending a pyroclastic flow over the Roman countryside, including the cities of Herculaneum and Pompeii. The buried towns were discovered in an archeological expedition in the 18th century [25]. Pompeii famously contains the remains (casts) of people suffocated by ash and covered by 10 feet (3 m) of ash, pumice lapilli, and collapsed roofs [26].

Figure 4.5.1: Mount St. Helens, the day before the May 18th, 1980 eruption.

Figure 4.5.1: Picture of 4 months after the major eruption of Mount St. Helens.

Figure 4.5.1: Series of still images of the May 18, 1980, eruption of Mt. Saint Helens, Washington showing largest recorded landslide in history and subsequent eruption and pyroclastic flow (By The Associated Press via The Atlantic)

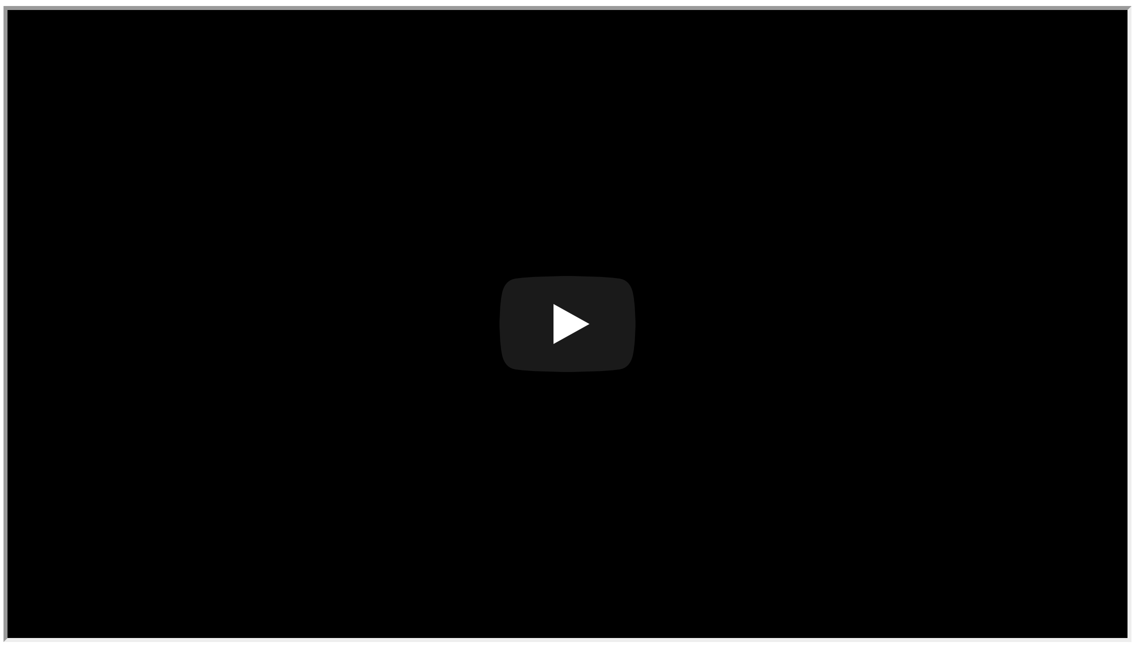
Pyroclastic Flows

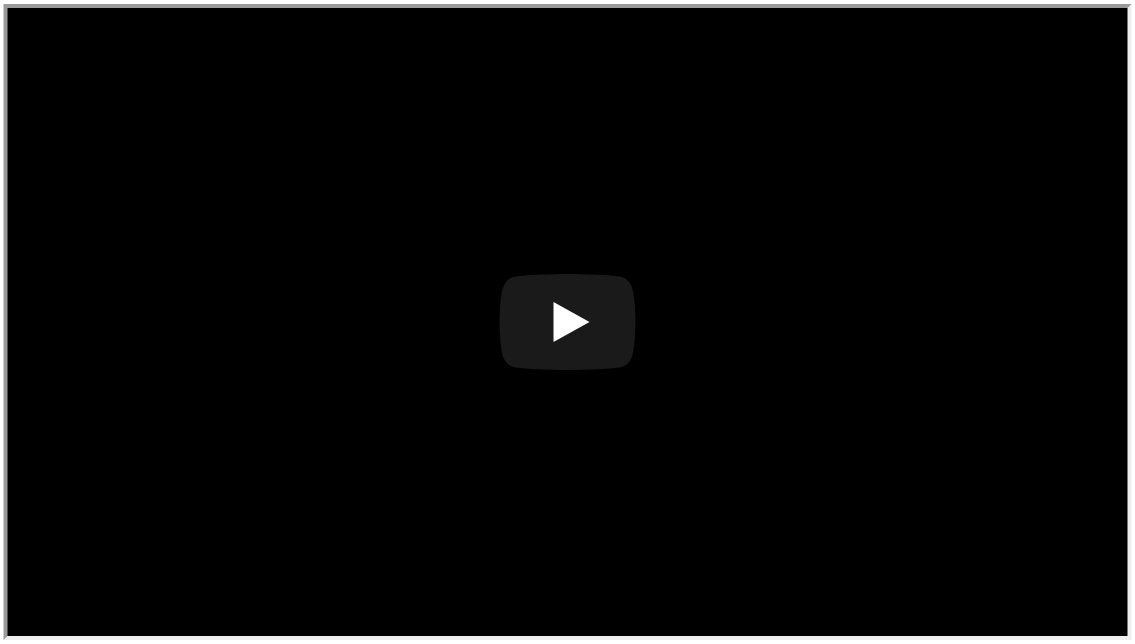
Figure 4.5.1: The material coming down from the eruption column onto the flanks is a pyroclastic flow.

The most dangerous volcanic hazard are **pyroclastic flows** (video). These flows are a mix of lava blocks, pumice, ash, and hot gases between 200°C-700°C (400°F-1,300°F). The turbulent cloud of ash and gas races down the steep flanks at high speeds up to 193 kph (120 mph) into the valleys around composite volcanoes [24]. Most explosive, silica-rich, high viscosity magma volcanoes such as composite cones usually have pyroclastic flows. The rock tuff and welded tuff is often formed from these pyroclastic flows.

Figure 4.5.1: The remains of St. Pierre.

There are numerous examples of deadly pyroclastic flows. In 2014, the Mount Ontake pyroclastic flow in Japan killed 47 people. The flow was caused by magma heating groundwater into steam, which then rapidly ejected with ash and volcanic bombs. Some were killed by inhalation of toxic gases and hot ash, while others were struck by volcanic bombs [27]. Two short videos below document an eye-witness video of pyroclastic flows. In the early 1990s, Mount Unzen erupted several times with pyroclastic flows. The pyroclastic flow shown in this famous short video killed 41 people. In 1902, on the Caribbean Island Martinique, Mount Pelee erupted with a violent pyroclastic flow that destroyed the entire town of St. Pierre and killing 28,000 people in moments [28].





Landslides and Landslide-Generated Tsunamis

Figure 4.5.1: Sequence of events for Mount St. Helens, May 18th, 1980. Note that an earthquake caused a landslide, which caused the "uncorking" of the mountain and started the eruption.

The steep and unstable flanks of a volcano can lead to slope failure and dangerous landslides. These landslides can be triggered by magma movement, explosive eruptions, large earthquakes, and/or heavy rainfall. During the 1980 Mount St. Helens eruption, the entire north flank of the volcano collapsed and released a huge landslide that moved at speeds of 160-290 kph (100-180 mph).

If enough landslide material reaches the ocean, it may cause a tsunami. In 1792, a landslide caused by the Mount Unzen eruption reached the Ariaka Sea, generating a tsunami that killed 15,000 people (see USGS page). When Mount Krakatau in Indonesia erupted in 1883, it generated ocean waves that towered 40 m (131 ft) above sea level. The tsunami killed 36,000 people and destroyed 165 villages [24].

Tephra

Figure 4.5.1: Aman sweeps ash from an eruption of Kelud, Indonesia.

Volcanoes, especially composite volcanoes, eject large amounts of **tephra** (ejected rock materials), most notably **ash** (tephra fragments less than 0.08 inches [2 mm]). Larger tephra is heavier and falls closer to the vent. Larger blocks and bombs pose hazards to those close to the eruption such as at the 2014 Mount Ontake disaster in Japan discussed earlier.

Figure 4.5.1: Micrograph of silica particle in volcanic ash. A cloud of these is capable of destroying an aircraft or automobile engine.

Hot ash poses an immediate danger to people, animals, plants, machines, roads, and buildings located close to the eruption. Ash is fine-grained (< 2mm) and can travel airborne long distances away from the eruption site. Heavy accumulations of ash can cause buildings to collapse. In people, it may cause respiratory issues like silicosis. Ash is destructive to aircraft and automobile engines, which can disrupt transportation and shipping services [24]. In 2010, the Eyjafjallajökull volcano in Iceland emitted a large ash cloud into the upper atmosphere, causing the largest air-travel disruption in northern Europe since World War II. No one was injured, but the service disruption was estimated to have cost the world economy billions of dollars [29].

Volcanic Ashes

As magma rises to the surface, the confining pressure decreases and allows dissolved gases to escape into the atmosphere. Even volcanoes that are not actively erupting may emit hazardous gases, such as carbon dioxide (CO_2), sulfur dioxide (SO_2), hydrogen sulfide (H_2S), and hydrogen halides (HF, HCl, or HBr).

Carbon dioxide tends to sink and accumulate in depressions and basins. In volcanic areas known to emit carbon dioxide, low-lying areas may trap hazardous concentrations of this colorless and odorless gas. The Mammoth Mountain Ski Resort in California, is located within the Long Valley Caldera, is one such area of carbon dioxide-producing volcanism. In 2006, three ski patrol members died of suffocation caused by carbon dioxide after falling into a snow depression near a fumarole (info).

In rare cases, volcanism may create a sudden emission of gases without warning. Limnic eruptions (*limne* is Greek for lake), occur in crater lakes associated with active volcanism. The water in these lakes is supercharged with high concentrations of dissolved gases. If the water is physically jolted by a landslide or earthquake, it may trigger an immediate and massive release of gases out of solution. An analogous example would be what happens to a vigorously shaken bottle of carbonated soda when the cap is opened. An infamous limnic eruption occurred in 1986 at Lake Nyos, Cameroon. Almost 2,000 people were killed by a massive release of carbon dioxide [24].

Lahars

Figure 4.5.1: Mud line shows the extent of lahars around Mount St. Helens

Lahar is an Indonesian word and is used to describe a volcanic mudflow that forms from rapidly melting snow or glaciers. Lahars are slurries resembling wet concrete and consist of water, ash, rock fragments, and other debris. These mudflows flow down the flanks of volcanoes or mountains covered with freshly-erupted ash and on steep slopes can reach speeds of up to 80 kph (50 mph).

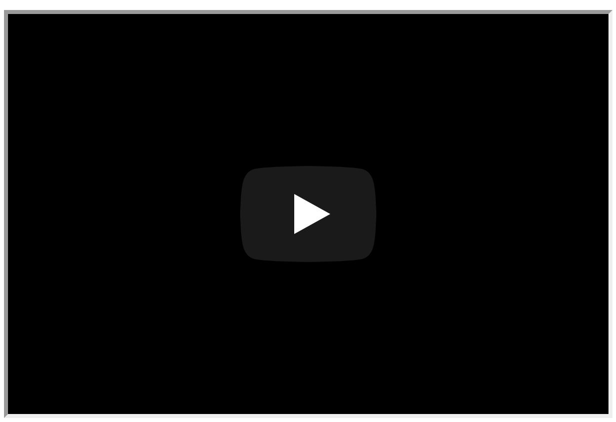
Figure 4.5.1: Old lahars around Tacoma, Washington.

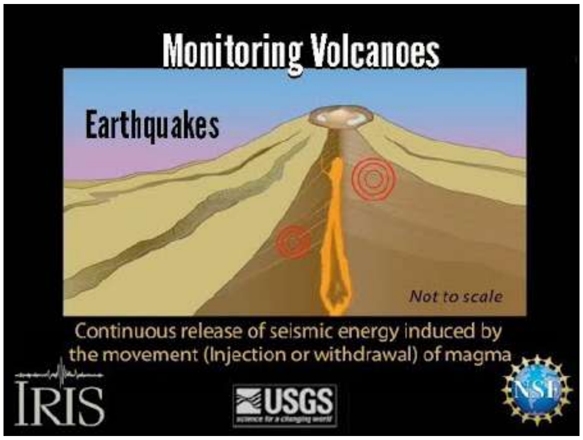
Several major cities, including Tacoma, are located on prehistoric lahar flows that extend for many kilometers across the flood plains surrounding Mount Rainier in Washington (see map). A map of Mount Baker in Oregon shows a similar potential hazard for lahar flows (see map) [24]. A tragic scenario played out recently, in 1985, when a lahar from the Nevado del Ruiz volcano in Colombia buried the town of Armero and killed an estimated 23,000 people.

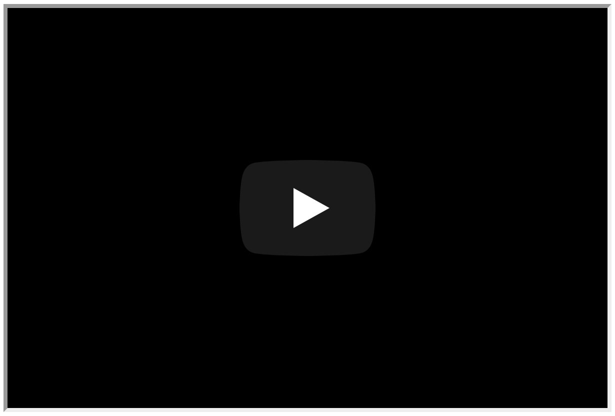
Monitoring

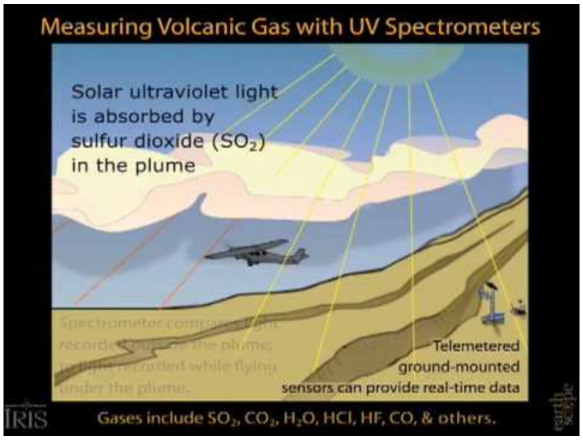
Geologists use various instruments to detect changes or indications that an eruption is imminent [30][31]. The three videos show different types of volcanic monitoring used to predict eruptions 1) earthquake activity, 2) increases in gas emission, and 3) changes in land surface orientation and elevation.

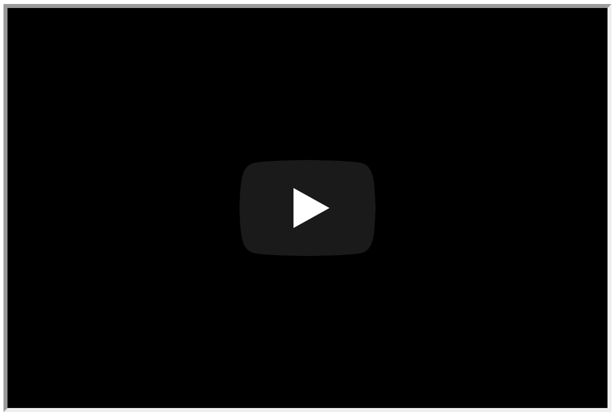
One video shows how monitoring earthquake frequency, especially special vibrational earthquakes called harmonic tremors, can detect magma movement and possible eruption. Another video shows how gas monitoring may be used to predict an eruption. A rapid increase of gas emission may indicate magma that is actively rising to surface and releasing dissolved gases out of solution, and that an eruption is imminent. The last video shows how a GPS unit and tiltmeter can detect land surface changes, indicating the magma is moving underneath it.

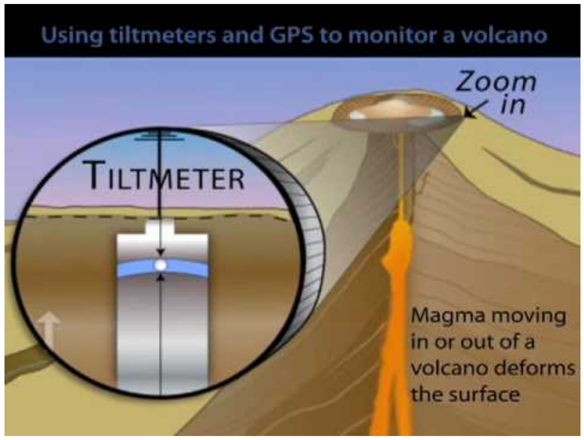



Monitoring Volcanoes
Earthquakes
Not to scale
Continuous release of seismic energy induced by the movement (Injection or withdrawal) of magma
IRIS
USGS
NSF




Measuring Volcanic Gas with UV Spectrometers
Solar ultraviolet light is absorbed by sulfur dioxide (SO_2) in the plume
Telemetered ground-mounted sensors can provide real-time data
IRIS
Gases include SO_2, CO_2, H_2O, HCl, HF, CO, & others.





References

10. Thorarinsson, S. The Lakagigar eruption of 1783. *Bull. Volcanol.* **33**, 910–929 (1969).

11. United States Geological Survey. Exploring the deep ocean floor. (1999). Available at: http://pubs.usgs.gov/gip/dynamic/exploring.html. (Accessed: 25th July 2016)

12. United States Geological Survey. Black Rock Desert Volcanic Field. (2012). Available at: http://volcanoes.usgs.gov/volcanoes/black_rock_desert/. (Accessed: 25th July 2016)

13. Rhodes, J. M. & Lockwood, J. P. Mauna Loa Revealed: Structure, Composition, History, and Hazards. *Washington DC American Geophysical Union Geophysical Monograph Series* **92**, (1995).

14. USGS. Volcanoes: Principal Types of Volcanoes. (2011). Available at: http://pubs.usgs.gov/gip/volc/types.html. (Accessed: 30th July 2016)

15. Carr, M. H. Geologic map of the Tharsis Quadrangle of Mars. *IMAP* (1975).

16. Frankel, C. *Worlds on Fire: Volcanoes on the Earth, the Moon, Mars, Venus and Io*. (Cambridge University Press, 2005).

17. Peterson, D. W. & Tilling, R. I. Transition of basaltic lava from pahoehoe to aa, Kilauea Volcano, Hawaii: Field observations and key factors - ScienceDirect. *J. Volcanol. Geotherm. Res.* **7**, 271–293 (1980).

18. USGS. Volcanoes General - What are the different types of volcanoes? (2016). Available at: https://www2.usgs.gov/faq/categories/9819/2730. (Accessed: 14th March 2017)

19. USGS. Yellowstone Volcano Observatory. (2012). Available at: https://volcanoes.usgs.gov/volcanoes/yellowstone/yellowstone_geo_hist_52.html. (Accessed: 30th July 2016)

20. United States Geological Survey. USGS: Volcano Hazards Program CalVO Long Valley Caldera. *Geologic history of the Long Valley, Mono Basin region* (2016). Available at: https://volcanoes.usgs.gov/volcanoes/long_valley/geo_hist_summary.html. (Accessed: 25th May 2017)

21. Luhr, J. F. & Cuasay, M. *Parícutin : The volcano born in a Mexican cornfield*. (U.S. Geoscience Press, 1993).

22. Bell, K. & Keller, J. *Carbonatite volcanism: Oldoinyo Lengai and the petrogenesis of natrocarbonatites*. (Springer Science & Business Media, 2012).

23. USGS. The Volcanoes of Lewis and Clark - Mount St. Helens. (2017). Available at: https://volcanoes.usgs.gov/observatories/cvo/Historical/LewisClark/Info/ summary_mount_st_helens.shtml. (Accessed: 14th March 2017)

24. USGS. USGS: Volcano Hazards Program. (2012). Available at: https://volcanoes.usgs.gov/vhp/hazards.html. (Accessed: 30th July 2016)

25. Scandone, R., Giacomelli, L. & Gasparini, P. Mount Vesuvius: 2000 years of volcanological observations. *Journal of Volcanology and Geothermal Research* **58**, 5–25 (1993).

26. Luongo, G. *et al.* Impact of the AD 79 explosive eruption on Pompeii, II. Causes of death of the inhabitants inferred by stratigraphic analysis and areal distribution of the human casualties. *J. Volcanol. Geotherm. Res.* **126**, 169–200 (2003).

27. EarthScope. Mount Ontake Volcanic Eruption. (2014). Available at: http://www.earthscope.org/science/geo-events/mount-ontake-volcanic-eruption. (Accessed: 30th July 2016)

28. USGS. Dual volcanic tragedies in the Caribbean led to founding of HVO. (2001). Available at: http://hvo.wr.usgs.gov/volcanowatch/archive/2001/01_05_03.html. (Accessed: 30th July 2016)

29. Perkins, S. 2010's Volcano-Induced Air Travel Shutdown Was Justified. (2011). Available at: http://www.sciencemag.org/news/2011/04/2010s-volcano-induced-air-travel-shutdown-was-justified. (Accessed: 30th July 2016)

30. Tilling, R. I. The critical role of volcano monitoring in risk reduction. *Adv. Geosci.* **14**, 3–11 (2008).

31. Stovall, W. K., Wilkins, A. M., Mandeville, C. W. & Driedger, C. L. *Fact Sheet*. (2016).

Contributions and Attributions

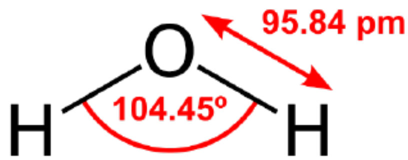
CHAPTER OVERVIEW

5: WEATHERING, EROSION, AND SEDIMENTARY ROCKS

Sedimentary rock and the processes that create it, which include weathering, erosion, and lithification, are an integral part of understanding Earth Science. This is because the majority of the Earth's surface is made up of sedimentary rocks and their common predecessor, sediments. Even though sedimentary rocks can form in drastically different ways, their origin and creation have one thing in common, water.

5.1: The Unique Properties of Water

Water plays a role in the formation of most sedimentary rocks. It is one of the main agents involved in creating the minerals in chemical sedimentary rock. It also is a weathering and erosion agent, producing the grains that become detrital sedimentary rock. Several special properties make water an especially unique substance, and integral to the production of sediments and sedimentary rock.



The water molecule consists of two hydrogen atoms covalently bonded to one oxygen atom arranged in a specific and important geometry. The two hydrogen atoms are separated by an angle of about 105 degrees, and both are located to one side of the oxygen atom [1]. This atomic arrangement, with the positively charged hydrogens on one side and negatively charged oxygen on the other side, gives the water molecule. This property is called **polarity**. Resembling a battery or a magnet, the molecule's positive-negative architecture leads to a whole suite of properties.

Figure 5.1.1: Dew on a spider's web.

Polarity allows water molecules to stick to other substances. This is called **adhesion**. Water is also attracted to itself, a property called **cohesion**, which leads to water's most common form in the air, a droplet. Cohesion is responsible for creating surface tension, which various insects use to walk on water by distributing their weight across the surface.

The fact that water is attracted to itself leads to another important property, one that is extremely rare in the natural world—the liquid form is denser than the solid form. The polarity of water creates a special type of weak bonding called **hydrogen bonds**. Hydrogen bonds allow the molecules in liquid water to sit close together. Water is densest at 4°C and is less dense above and below that temperature. As water solidifies into ice, the molecules must move apart in order to fit into the crystal lattice, causing water to expand and become less dense as it freezes. Because of this, ice floats and water sinks, which keeps the oceans liquid and prevents them from freezing solid from the bottom up. This unique property of water keeps Earth, the water planet, habitable.

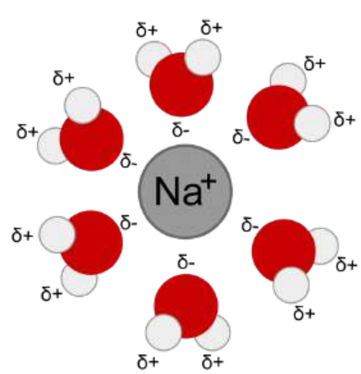


Figure 5.1.1: A sodium (Na) ion in solution.

Even more critical for supporting life, water remains liquid over a very large range of temperatures, which is also a result of cohesion. Hydrogen bonding allows liquid water can absorb high amounts of energy before turning into vapor or gas. The wide range across which water remains a liquid, 0°C-100°C (32°F-212°F), is rarely exhibited in other substances. Without this high boiling point, liquid water as we know it would be constricted to narrow temperature zones on Earth, instead of being found from pole to pole.

Water is a **universal solvent**, meaning it dissolves more substances than any other commonly found, naturally occurring liquid. The water molecules use polarity and hydrogen bonds to pry ions away from the crystal lattice. Water is such a powerful solvent, it can dissolve even the strongest rocks and minerals given enough time.

Reference

1. Hoy, A. R. & Bunker, P. R. A precise solution of the rotation bending Schrödinger equation for a triatomic molecule with application to the water molecule. *J. Mol. Spectrosc.* **74**, 1–8 (1979).

Contributions and Attributions

5.2: Weathering and Erosion

Bedrock refers to the solid crystalline rock that makes up the Earth's outer crust. Weathering is a process that turns **bedrock** into smaller particles, called **sediment** or **soil**. Mechanical weathering includes pressure expansion, frost wedging, root wedging, and salt expansion. Chemical **weathering** includes carbonic acid and hydrolysis, dissolution, and oxidation.

Erosion is a mechanical process, usually driven by water, wind, gravity, or ice, which transports sediment and soil from the place of weathering. Liquid water is the main agent of erosion. Gravity and mass wasting processes (see Chapter 10, Mass Wasting) move rocks and sediment to new locations. Gravity and ice, in the form of glaciers (see Chapter 14, Glaciers), move large rock fragments as well as fine sediment.

Erosion resistance is important in the creation of distinctive geological features. This is well-demonstrated in the cliffs of the Grand Canyon. The cliffs are made of rock left standing after less resistant materials have weathered and eroded away. Rocks with different levels of erosion resistance also create unique-looking features called hoodoos in Bryce Canyon National Park and Goblin Valley State Park in Utah.

Mechanical Weathering

Mechanical weathering physically breaks bedrock into smaller pieces. The usual agents of mechanical weathering are pressure, temperature, freezing/thawing cycle of water, plant or animal activity, and salt evaporation.

Pressure Expansion

Figure 5.2.1: The outer layer of this granite is fractured and eroding away, known as exfoliation.

Bedrock buried deep within the Earth is under high pressure and temperature. When uplift and erosion brings bedrock to the surface, its temperature drops slowly, while its pressure drops immediately. The sudden pressure drop causes the rock to rapidly expand and crack; this is called pressure expansion. Sheeting or **exfoliation** is when the rock surface spalls off in layers. **Spheroidal weathering** is a type of exfoliation that produces rounded features and is caused when chemical weathering moves along joints in the bedrock.

Frost Wedging

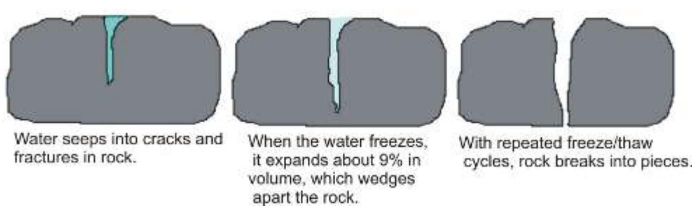


Figure 5.2.1: The process of frost wedging

Frost wedging, also called ice wedging, uses the power of expanding ice to break apart rocks. Water works its way into various cracks, voids, and crevices. As the water freezes, it expands with great force, exploiting any weaknesses. When ice melts, the liquid water moves further into the widened spaces. Repeated cycles of freezing and melting eventually pry the rocks apart. The cycles can occur daily when fluctuations of temperature between day and night go from freezing to melting.

Root Wedging

Figure 5.2.1: The roots of this tree are demonstrating the destructive power of root wedging. Though this picture is a man-made rock (asphalt), it works on a typical rock as well.

Like frost wedging, **root wedging** happens when plant roots work themselves into cracks, prying the bedrock apart as they grow. Occasionally these roots may become fossilized. **Rhizolith** is the term for these roots preserved in the rock record [2]. Tunneling organisms such as earthworms, termites, and ants are biological agents that induce weathering similar to root wedging.

Salt Expansion

Figure 5.2.1: Tafoni from Salt Point, California.

Salt expansion, which works similarly to frost wedging, occurs in areas of high evaporation or near-marine environments. Evaporation causes salts to precipitate out of solution and grow and expand into cracks in the rock. Salt expansion is one of the causes of **tafoni**, a series of holes in a rock. Tafoni, cracks, and holes are weak points that become susceptible to increased weathering. **Hopper crystal** describes a square-shaped crystal, commonly made of salt, preserved in rock.

Chemical Weathering

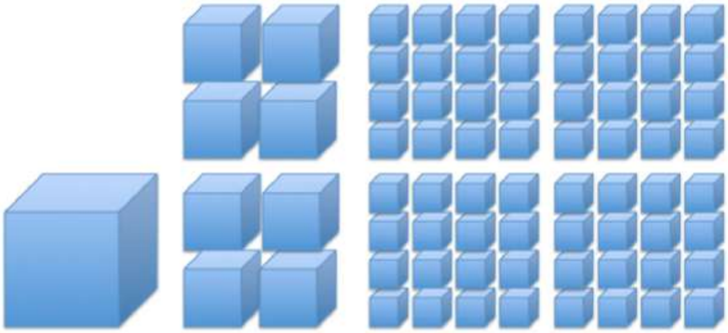


Figure 5.2.1: Each of these three groups of cubes has an equal volume. However, their surface areas are vastly different. On the left, the single cube has a length, width, and height of 4 units, giving it a surface area of $6(4 \times 4) = 48$ and a volume of $4^3 = 64$. The middle eight cubes have a length, width, and height of 2, meaning a surface area of $8(6(2 \times 2)) = 8 \times 24 = 96$. They also have a volume of $8(2^3) = 8 \times 8 = 64$. The 64 cubes on the right have a length, width, and height of 1, leading to a surface area of $64(6(1 \times 1)) = 64 \times 6 = 384$. The volume remains unchanged because $64(1^3) = 64 \times 1 = 64$. The surface area to volume ratio (SA:V), which is related to the amount of material available for reactions, changes for each as well. On the left, it is $\frac{48}{64} = 0.75$ or 3:4. The center has a SA/V of $\frac{96}{64} = 1.5$, or 3:2. On the right, the SA:V is $\frac{384}{64} = 6$, or 6:1.

Chemical weathering is the dominant weathering process in warm, humid environments. It happens when water, oxygen, and other reactants chemically degrade the mineral components of bedrock and turn them into water-soluble ions which can then be transported by water. Higher temperatures accelerate chemical weathering rates.

Chemical and mechanical weathering work hand-in-hand via a fundamental concept called surface-area-to-volume ratio. Chemical weathering only occurs on rock surfaces because water and reactants cannot penetrate the solid rock. Mechanical weathering penetrates the bedrock, breaking large rocks into smaller pieces and creating new rock surfaces. This exposes more surface area to chemical weathering, enhancing its effects. In other words, higher surface-area-to-volume ratios produce higher rates of overall weathering.

Carbonic Acid and Hydrolysis

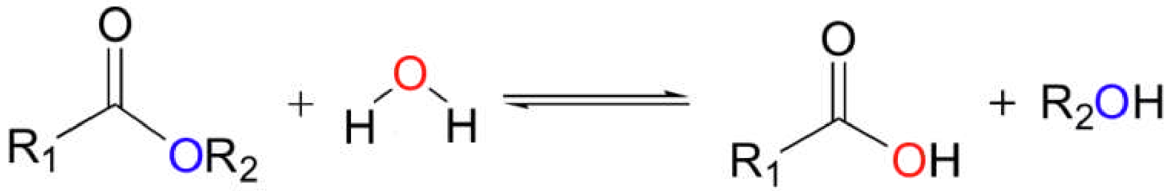


Figure 5.2.1: Generic hydrolysis diagram, where the bonds in mineral in question would represent the left side of the diagram.

Carbonic acid (H_2CO_3) forms when carbon dioxide, the fifth-most abundant gas in the atmosphere, dissolves in water. This happens naturally in clouds, which is why precipitation is normally slightly acidic. Carbonic acid is an important agent in two chemical weathering reactions, hydrolysis, and dissolution.

Hydrolysis occurs via two types of reactions. In one reaction, water molecules ionize into positively charged H^+ and OH^- ions and replace mineral cations in the crystal lattice. In another type of hydrolysis, carbonic acid molecules react directly with minerals, especially those containing silicon and aluminum (i.e. Feldspars), to form molecules of clay minerals.

Hydrolysis is the main process that breaks down silicate rock and creates clay minerals. The following is a hydrolysis reaction that occurs when silica-rich feldspar encounters carbonic acid to produce water-soluble clay and other ions:

> ***feldspar*** + ***carbonic acid*** *(in water) → clay + metal cations* (Fe^{++}, Mg^{++}, Ca^{++}, Na^+, *etc.) + bicarbonate anions* (HCO_3^-) + *silica* (SiO_2)

Clay minerals are platy silicates or phyllosilicates (see Chapter 3, Minerals) similar to micas, and are the main components of very fine-grained sediment. The dissolved substances may later precipitate into **chemical sedimentary rocks** like evaporite and limestone, as well as amorphous silica or chert nodules.

Dissolution

Figure 5.2.1: In this rock, a pyrite cube has dissolved (as seen with the negative “corner” impression in the rock), leaving behind small specks of gold.

Dissolution is a hydrolysis reaction that dissolves minerals in bedrock and leaves the ions in solution, usually in water. Some evaporites and carbonates, like salt and calcite, are more prone to this reaction; however, all minerals can be dissolved. Non-acidic water, having a neutral pH of 7, will dissolve any mineral, although it may happen very slowly. Water with higher levels of acid, natural or man-made, dissolves rocks at a higher rate. Liquid water is normally slightly acidic due to the presence of carbonic acid and free H+ ions. Natural rainwater can be highly acidic, with pH levels as low as 2 [3]. Dissolution can be enhanced by a biological agent, such as when organisms like lichen and bacteria release organic acids onto the rocks they are attached to. Regions with high humidity (airborne moisture) and precipitation experience more dissolution due to greater contact time between rocks and water.

Figure 5.2.1: This mantle xenolith containing olivine (green) is chemically weathering by hydrolysis and oxidation into the pseudo-mineral iddingsite, which is a complex of water, clay, and iron oxides. The more altered side of the rock has been exposed to the environment longer.

The **Goldich Dissolution Series** shows chemical weathering rates are associated with crystallization rankings in the Bowen's Reaction Series (see Chapter 4, Igneous Rock and Volcanic Processes) [4]. Minerals at the top of the Bowen series crystallize under high temperatures and pressures, and chemically weather at a faster rate than minerals ranked at the bottom. Quartz, a felsic mineral that crystallizes at 700°C, is very resistant to chemical weathering. High crystallization-point mafic minerals, such as olivine and pyroxene (1,250°C), weather relatively rapidly and more completely. Olivine and pyroxene are rarely found as end products of weathering because they tend to break down into elemental ions.

Figure 5.2.1: Eroded karst topography in Minevre, France.

Figure 5.2.1: A formation called The Great Heart of Timpanogos in Timpanogos Cave National Monument

Dissolution is also noteworthy for the special geological features it creates. In places with abundant carbonate bedrock, dissolution weathering can produce a **karst topography** characterized by sinkholes or caves (see Chapter 10, Mass Wasting).

Timpanogos Cave National Monument in Northern Utah is a well-known dissolution feature. The figure shows a cave formation created from dissolution followed by precipitation—groundwater saturated with calcite seeped into the cavern, where evaporation caused the dissolved minerals to precipitate out.

Oxidation

Figure 5.2.1: Pyrite cubes are oxidized, becoming new mineral goethite. In this case, goethite is a pseudomorph after pyrite, meaning it has taken the form of another mineral.

Oxidation, the chemical reaction that causes rust in metallic iron, occurs geologically when iron atoms in a mineral bond with oxygen. Any minerals containing iron can be oxidized. The resultant iron oxides may permeate a rock if it is rich in iron minerals. Oxides may also form a coating that covers rocks and grains of sediment, or lines rock cavities and fractures. If the oxides are more susceptible to weathering than the original bedrock, they may create void spaces inside the rock mass or hollows on exposed surfaces.

Three commonly found minerals are produced by iron-oxidation reactions: red or grey **hematite**, brown **goethite** (pronounced “GUR-tite”), and yellow **limonite**. These iron oxides coat and bind mineral grains together into sedimentary rocks in a process called cementation and often give these rocks a dominant color. They color the rock layers of the Colorado Plateau, as well as Zion, Arches, and Grand Canyon National Parks. These oxides can permeate a rock that is rich in iron-bearing minerals or can be a coating that forms in cavities or fractures. When the minerals replacing existing minerals in bedrock are resistant to weathering, iron concretions may occur in the rock. When bedrock is replaced by weaker oxides, this process commonly results in void spaces and weakness throughout the rock mass and often leaves hollows on exposed rock surfaces.

Erosion

Figure 5.2.1: A hoodoo near Moab, Utah. The more resistant cap has protected the less resistant underlying layers.

Erosion is a mechanical process, usually driven by water, gravity, (see Chapter 10), wind, or ice (see Chapter 14) that removes sediment from the place of weathering. Liquid water is the main agent of erosion.

Figure 5.2.1: Grand Canyon from Mather Point.

Erosion **resistance** is important in the creation of distinctive geological features. This is well demonstrated in the cliffs of the Grand Canyon. The cliffs are made of rock left standing after less resistant materials have weathered and eroded away. Rocks with different levels of erosion resistant also create a unique-looking features called hoodoos in Bryce Canyon National Park and Goblin Valley State Park in Utah.

Soil

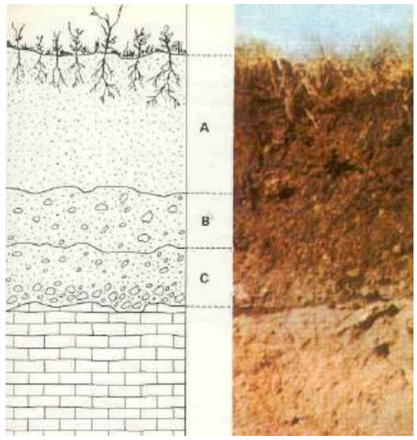


Figure 5.2.1: Sketch and picture of soil.

Soil is a combination of air, water, minerals, and organic matter that forms at the transition between the biosphere and geosphere. Soil is made when weathering breaks down the bedrock and turns it into sediment. If erosion does not remove the sediment significantly, organisms can access the mineral content of the sediments. These organisms turn minerals, water, and atmospheric gases into organic substances that contribute to the soil.

Soil is an important reservoir for organic components necessary for plants, animals, and microorganisms to live. The organic component of soil, called **humus**, is a rich source of bioavailable nitrogen. Nitrogen is the most common element in the atmosphere, but it exists in a form most life forms are unable to use. Special bacteria found only in the soil provide most nitrogen compounds that are usable, bioavailable, by life forms.

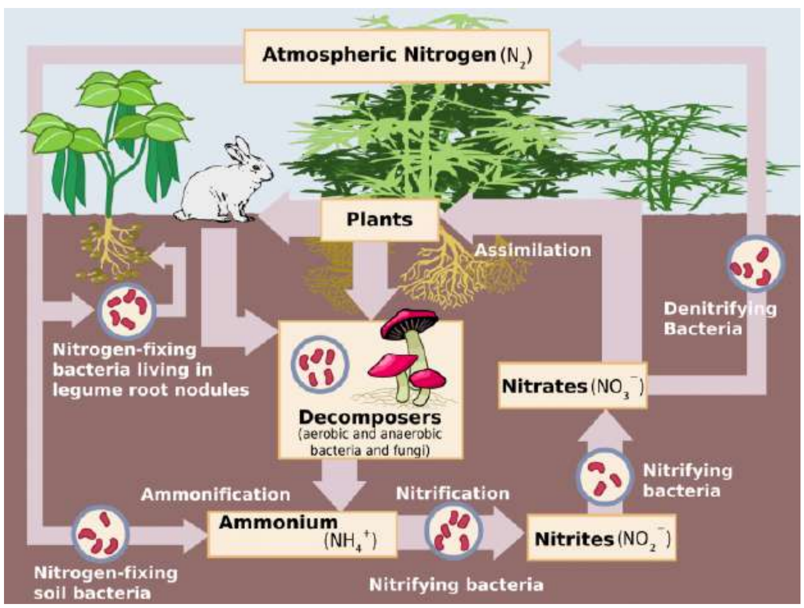


Figure 5.2.1: Schematic of the nitrogen cycle.

These nitrogen-fixing bacteria absorb nitrogen from the atmosphere and convert it into nitrogen compounds. These compounds are absorbed by plants and used to make DNA, amino acids, and enzymes. Animals obtain bioavailable nitrogen by eating plants, and this is the source of most of the nitrogen used by life. That nitrogen is an essential component of proteins and DNA. Soils range from poor to rich, depending on the amount of humus they contain. Soil productivity is determined by water and nutrient content. Freshly created volcanic soils, called andisols, and clay-rich soils that hold nutrients and water are examples of productive soils.

Figure 5.2.1: Agricultural terracing, as made by the Inca culture from the Andes, helps reduce erosion and promote soil formation, leading to better farming practices.

The nature of the soil, meaning its characteristics, is determined primarily by five components:

1. The mineralogy of the parent material
2. Topography
3. Weathering
4. Climate
5. The organisms that inhabit the soil.

For example, soil tends to erode more rapidly on steep slopes so soil layers in these areas may be thinner than in flood plains, where it tends to accumulate. The quantity and chemistry of organic matter of soil affect how much and what varieties of life it can sustain. Temperature and precipitation, two major weathering agents, are dependent on climate. Fungi and bacteria contribute organic matter and the ability of soil to sustain life, interacting with plant roots to exchange nitrogen and other nutrients [5].

In well-formed soils, there is a discernable arrangement of distinct layers called **soil horizons** [6]. These soil horizons can be seen in road cuts that expose the layers at the edge of the cut. Soil horizons make up the soil profile. Each soil horizon reflects climate, topography, and other soil-development factors, as well as its organic material and mineral sediment composition. The horizons are assigned names and letters. Differences in naming schemes depend on the area, soil type or research topic. The figure shows a simplified soil profile that uses commonly designated names and letters.

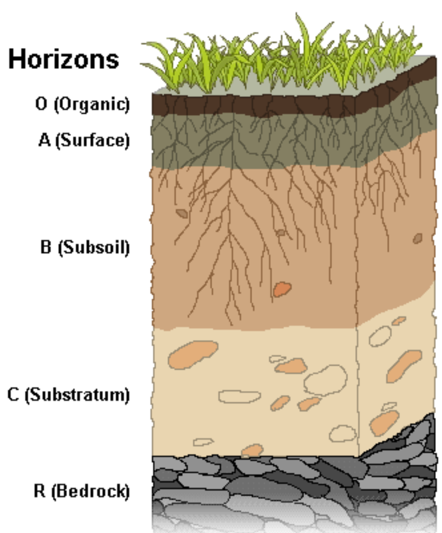


Figure 5.2.1: A simplified soil profile, showing labeled layers.

O Horizon: The top horizon is a thin layer of predominantly organic material, such as leaves, twigs, and other plant parts that are actively decaying into humus.

A Horizon: The next layer, called **topsoil**, consists of humus mixed with mineral sediment. As precipitation soaks down through this layer, it leaches out soluble chemicals. In wet climates with heavy precipitation, this leaching out produces a separate layer called horizon E, the leaching or eluviation zone.

B Horizon: Also called **subsoil**, this layer consists of sediment mixed with humus removed from the upper layers. The subsoil is where mineral sediment is chemically weathered. The amount of organic material and the degree of weathering decrease with depth. The upper subsoil zone, called **regolith**, is a porous mixture of humus and highly weathered sediment. In the lower zone, **saprolite**, scant organic material is mixed with largely unaltered parent rock.

C Horizon: This is substratum and is a zone of mechanical weathering. Here, bedrock fragments are physically broken but not chemically altered. This layer contains no organic material.

R Horizon: The final layer consists of unweathered, parent **bedrock** and fragments.

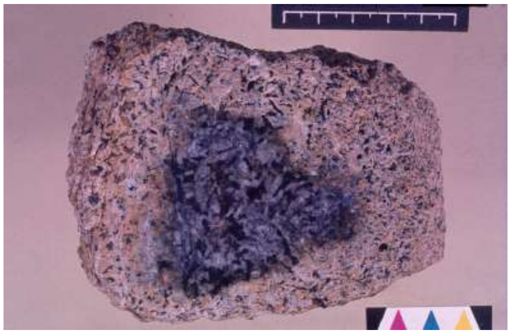


Figure 5.2.1: A sample of bauxite. Note the unweathered igneous rock in the center.

The United States governing body for agriculture, the USDA, uses a taxonomic classification to identify soil types, called soil orders. Xoxisols or laterite soils are nutrient-poor soils found in tropical regions. While poorly suited for growing crops, xosisols are home to most of the world's mineable aluminum ore (bauxite). Ardisol forms in dry climates and can develop layers of hardened calcite, called caliche. Andisols originate from volcanic ash deposits. Alfisols contain silicate clay minerals. These two soil orders are productive for farming due to their high content of mineral nutrients. In general, color can be an important factor in understanding soil conditions. Black soils tend to be anoxic, red oxygen-rich, and green oxygen-poor (i.e. reduced). This is true for many sedimentary rocks as well.

Figure 5.2.1: A dust storm approaches Stratford, Texas in 1935.

Not only is soil essential to terrestrial life in nature, but also human civilization via agriculture. Careless or uninformed human activity can seriously damage soil's life-supporting properties. A prime example is the famous Dust Bowl disaster of the 1930s, which affected the midwestern United States. The damage occurred because of large-scale attempts to develop prairie land in southern Kansas, Colorado, western Texas, and Oklahoma into farmland [7]. Poor understanding of the region's geology, ecology, and climate led to farming practices that ruined the soil profile.

The prairie soils and native plants are well adapted to a relatively dry climate. With government encouragement, settlers moved in to homestead the region. They plowed vast areas of prairie into long, straight rows and planted grain. The plowing broke up the stable soil profile and destroyed the natural grasses and plants, which had long roots that anchored the soil layers. The grains they planted had shallower root systems and were plowed up every year, which made the soil prone to erosion. The plowed furrows were aligned in straight rows running downhill, which favored erosion and loss of topsoil.

The local climate does not produce sufficient precipitation to support non-native grain crops, so the farmers drilled wells and over-pumped water from the underground aquifers. The grain crops failed due to lack of water, leaving bare soil that was stripped from the ground by the prairie winds. Particles of midwestern prairie soil were deposited along the east coast and as far away as Europe.

Huge dust storms called black blizzards made life unbearable, and the once-hopeful homesteaders left in droves. The setting for John Steinbeck's famous novel and John Ford's film, *The Grapes of Wrath*, is Oklahoma during this time. The lingering question is whether we have learned the lessons of the dust bowl, to avoid creating it again [8].

References

2. Klappa, C. F. Rhizoliths in terrestrial carbonates: classification, recognition, genesis and significance. *Sedimentology* **27**, 613–629 (1980).

3. Willey, J. D. *et al.* Effect of storm type on rainwater composition in southeastern North Carolina. *Environ. Sci. Technol.* **22**, 41–46 (1988).

4. Goldich, S. S. A study in rock-weathering. *J. Geol.* **46**, 17–58 (1938).

5. Jenny, H. *Factors of soil formation: a system of quantitative pedology*. (Courier Corporation, 1994).

6. Gerlach, A. C. Soil Profile Sections. *J. Geog.* **48**, 344–349 (1949).

7. Worster, D. *Dust bowl: the southern plains in the 1930s*. (Oxford University Press, 2004).

8. Eckholm, E. Losing ground. *Environment: Science and Policy for Sustainable Development* **18**, 6–11 (1976).

Contributions and Attributions

5.3: Sedimentary Rocks

Sedimentary rock is classified into two main categories: clastic and chemical. **Clastic** or **detrital** sedimentary rocks are made from pieces of bedrock, sediment, derived primarily by mechanical weathering. Clastic rocks may also include chemically weathered sediment. Clastic rocks are classified by **grain shape**, **grain size**, and **sorting**. **Chemical** sedimentary rocks are precipitated from water saturated with dissolved minerals. Chemical rocks are classified mainly by the composition of minerals in the rock.

Lithification and Diagenesis

Lithification turns loose sediment grains, created by weathering and transported by erosion, into clastic sedimentary rock via three interconnected steps. **Deposition** happens when friction and gravity overcome the forces driving sediment transport, allowing sediment to accumulate. **Compaction** occurs when material continues to accumulate on top of the sediment layer, squeezing the grains together and driving out the water. The mechanical compaction is aided by weak attractive forces between the smaller grains of sediment. Groundwater typically carries cementing agents into the sediment. These minerals, such as calcite, amorphous silica, or oxides, may have a different composition than the sediment grains. **Cementation** is the process of cementing minerals coating the sediment grains and gluing them together into a fused rock.

Figure 5.3.1: Permineralization in petrified wood

Diagenesis is an accompanying process of lithification and is a low-temperature form of rock metamorphism (see Chapter 6, Metamorphic Rock). During diagenesis, sediments are chemically altered by heat and pressure. A classic example is aragonite ($CaCO_3$), a form of calcium carbonate that makes up most organic shells. When lithified aragonite undergoes diagenesis, the aragonite reverts to calcite ($CaCO_3$), which has the same chemical formula but a different crystalline structure. In sedimentary rock containing calcite and magnesium (Mg), diagenesis may transform the two minerals into dolomite ($CaMg(CO_3)_2$). Diagenesis may also reduce the pore space, or open volume, between sedimentary rock grains. The processes of cementation, compaction, and ultimately lithification occur within the realm of diagenesis, which includes the processes that turn organic material into fossils.

Detrital Sedimentary Rocks (Clastic)

Detrital or **clastic** sedimentary rocks consist of preexisting sediment pieces that come from weathered bedrock. Most of this is mechanically weathered sediment, although some clasts may be pieces of chemical rocks. This creates some overlap between the two categories, since clastic sedimentary rocks may include chemical sediments. Detrital or clastic rocks are classified and named based on their grain size.

Grain Size

Detrital rock is classified according to sediment **grain size**, which is graded from large to small on the Wentworth scale (see figure). Grain size is the average diameter of sediment fragments in sediment or rock. Grain sizes are delineated using a logbase-2 scale [9; 10]. For example, the grain sizes in the pebble class are 2.52, 1.26, 0.63, 0.32, 0.16, and 0.08 inches, which correlate respectively to very coarse, coarse, medium, fine, and very fine granules. Large fragments, or clasts, include all grain sizes larger than 2 mm (5/64 in). These include boulders, cobbles, granules, and gravel. Sand has a grain size between 2 mm and 0.0625 mm, about the lower limit of the naked eye's resolution. Sediment grains smaller than sand are called silt. Silt is unique; the grains can be felt with a finger or as grit between your teeth, but are too small to see with the naked eye.

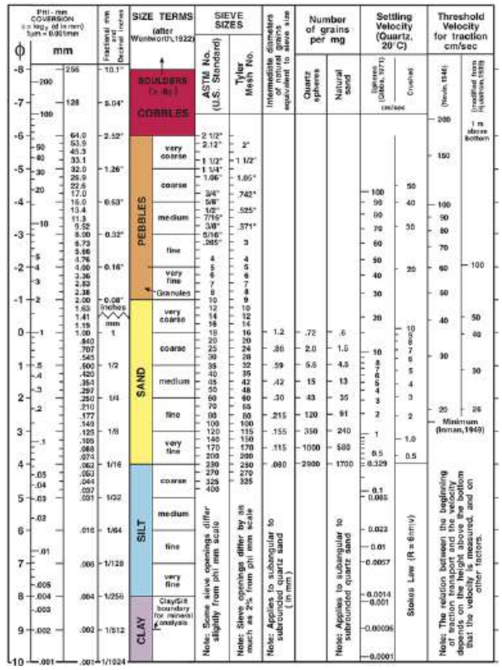


Figure 5.3.1: Size categories of sediments, known as the Wentworth scale.

Sorting and Rounding

Sorting describes the range of grain sizes within sediment or sedimentary rock. Geologists use the term "**well-sorted**" to describe a narrow range of grain sizes, and "poorly sorted" for a wide range of grain sizes (see figure) [11]. It is important to note that soil engineers use similar terms with opposite definitions; well-graded sediment consists of a variety of grain sizes, and poorly graded sediment has roughly the same grain sizes.

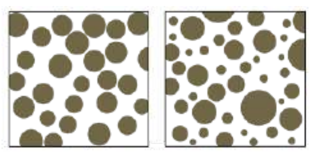


Figure 5.3.1: Well-sorted sediment (left) and Poorly-sorted sediment (right).

When reading the story told by rocks, geologists use sorting to interpret erosion or transport processes, as well as deposition energy. For example, wind-blown sands are typically extremely well sorted, while glacial deposits are typically poorly sorted. These characteristics help identify the type of erosion process that occurred. Coarse-grained sediment and poorly sorted rocks are usually found nearer to the source of sediment, while fine sediments are carried farther away. In a rapidly flowing mountain stream, you would expect to see boulders and pebbles. In a lake fed by the stream, there should be sand and silt deposits. If you also find large boulders in the lake, this may indicate the involvement of another sediment transport process, such as rockfall caused by ice- or root-wedging.

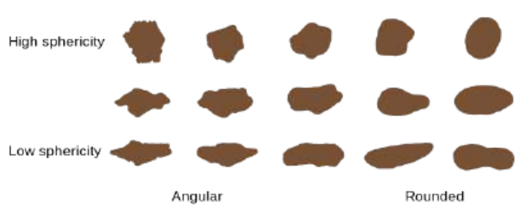


Figure 5.3.1: Degree of rounding in sediments. Sphericity refers to the spherical nature of an object, a completely different measurement unrelated to rounding.

Rounding is created when angular corners of rock fragments are removed from a piece of sediment due to abrasion during transport. Well-rounded sediment grains are defined as being free of all sharp edges. Very angular sediment retains the sharp corners. Most clast fragments start with some sharp edges due to the bedrock's crystalline structure, and those points are worn down during transport. More rounded grains imply a longer erosion time or transport distance, or more energetic erosional process. Mineral hardness is also a factor in rounding.

Composition and Provenance

Composition describes the mineral components found in sediment or sedimentary rock and may be influenced by local geology, like source rock and hydrology. Other than clay, most sediment components are easily determined by visual inspection (see Chapter 3, Minerals). The most commonly found sediment mineral is quartz because of its low chemical reactivity and high hardness, making it resistant to weathering, and its ubiquitous occurrence in continental bedrock. Other commonly found sediment grains include feldspar and lithic fragments. Lithic fragments are pieces of fine-grained bedrock [12], and include mud chips, volcanic clasts, or pieces of slate.

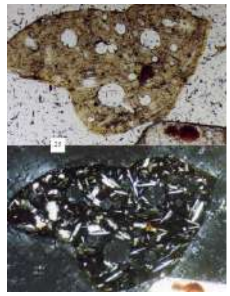


Figure 5.3.1: A sand grain made of basalt, known as a microlithic volcanic lithic fragment. The box is 0.25 mm. Top picture is plane-polarized light; bottom is cross-polarized light.

Weathering of volcanic rock produces Hawaii's famous black (basalt) and green (olivine) sand beaches, which are rare elsewhere on Earth. This is because the local rock is composed almost entirely of basalt and provides an abundant source of dark-colored clasts loaded with mafic minerals. According to the Goldich Dissolution Series, clasts high in mafic minerals are more easily destroyed compared to clasts composed of felsic minerals like quartz [13].

Figure 5.3.1: Hawaiian beach composed of green olivine sand from weathering of nearby basaltic rock.

Geologists use **provenance** to discern the original source of sediment or sedimentary rock. Provenance is determined by analyzing the mineral composition and types of fossils present, as well as textural features like sorting and rounding. Provenance is important for describing tectonic history [14], visualizing paleogeographic formations [15], unraveling an area's geologic history, or reconstructing past supercontinents [16].

In quartz sandstone, sometimes called quartz arenite (SiO_2), provenance may be determined using a rare, durable clast mineral called zircon ($ZrSiO_4$). Zircon, or zirconium silicate, contains traces of uranium, which can be used for age-dating the source bedrock that contributed sediment to the lithified sandstone rock (see Chapter 7, Geologic Time).

Classification of Clastic Rocks

Figure 5.3.1: Megabreccia in Titus Canyon, Death Valley National Park, California.

Clastic rocks are classified according to the grain size of their sediment [17]. Coarse-grained rocks contain clasts with a predominant grain size larger than sand. Typically, smaller sediment grains, collectively called groundmass or matrix, fill in much of the volume between the larger clasts, and hold the clasts together. **Conglomerates** are rocks containing coarse rounded clasts, and **breccias** contain angular clasts (see figure). Both conglomerates and breccias are usually poorly sorted.

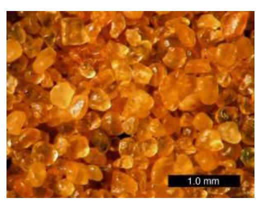


Figure 5.3.1: Enlarged image of frosted and rounded windblown sand grains

Medium-grained rocks composed mainly of sand are called **sandstone**, or sometimes **arenite** if well sorted. Sediment grains in sandstone can having a wide variety of mineral compositions, roundness, and sorting. Some sandstone names indicate the rock's mineral composition. Quartz sandstone contains predominantly quartz sediment grains. **Arkose** is sandstone with significant amounts of feldspar, usually greater than 25%. Sandstone that contains feldspar, which weathers more quickly than quartz, is useful for analyzing the local geologic history. **Greywacke** is a term with conflicting definitions [18]. Greywacke may refer to sandstone with a muddy matrix, or sandstone with many lithic fragments (small rock pieces).

Figure 5.3.1: The Rochester Shale, New York. Note the thin fissility in the layers.

Fine-grained rocks include mudstone, shale, siltstone, and claystone. **Mudstone** is a general term for rocks made of sediment grains smaller than sand (less than 2 mm). Rocks that are **fissile**, meaning they separate into thin sheets, are called shale. Rocks exclusively composed of silt or clay sediment, are called **siltstone** or **claystone**, respectively. These last two rock types are rarer than mudstone or shale.

Figure 5.3.1: Claystone laminations from Glacial Lake Missoula.

Rock types found as a mixture between the main classifications may be named using the less-common component as a descriptor. For example, a rock containing some silt but mostly rounded sand and gravel are called silty conglomerate. Sand-rich rock containing minor amounts of clay is called clayey sandstone.

Chemical, Biochemical, and Organic

Chemical sedimentary rocks are formed by processes that do not directly involve mechanical weathering and erosion. Chemical weathering may contribute to the dissolved materials in water that ultimately form these rocks. Biochemical and organic sediments are clastic in the sense that they are made from pieces of organic material that are deposited, buried, and lithified; however, they are usually classified as being chemically produced.

Inorganic chemical sedimentary rocks are made of minerals precipitated from ions dissolved in solution, and created without the aid of living organisms. Inorganic chemical sedimentary rocks form in environments where ion concentration, dissolved gasses, temperatures, or pressures are changing, which causes minerals to crystallize.

Biochemical sedimentary rocks are formed from shells and bodies of underwater organisms. The living organisms extract chemical components from the water and use them to build shells and other body parts. The components include aragonite, a mineral similar to and commonly replaced by calcite, and silica.

Organic sedimentary rocks come from organic material that has been deposited and lithified, usually underwater. The source materials are plant and animal remains that are transformed through burial and heat, and end up as coal, oil, and methane (natural gas).

Inorganic Chemical

Figure 5.3.1: Salt-covered plain known as the Bonneville Salt Flats, Utah.

Inorganic chemical sedimentary rocks are formed when minerals precipitate out of an aqueous solution, usually due to water evaporation. The precipitate minerals form various salts known as **evaporites**. For example, the Bonneville Salt Flats in Utah flood with winter rains and dry out every summer, leaving behind salts such as **gypsum** and **halite**. The deposition order of evaporites deposit is opposite to their solubility order, i.e. as water evaporates and increases the mineral concentration in solution, less soluble minerals precipitate out sooner than the highly soluble minerals. The deposition order and saturation percentages are depicted in the table, bearing in mind the process in nature may vary from laboratory-derived values [19].

Table after [20].

Mineral sequence	Percent Seawater remaining after
Calcite	50
Gypsum/anhydrite	20
Halite	10
Various potassium and magnesium salts	5

Figure 5.3.1: Ooids from Joulter's Cay, The Bahamas

Figure 5.3.1: Limestone tufa towers along the shores of Mono Lake, California.

Calcium carbonate-saturated water precipitates porous masses of calcite called **tufa**. Tufa can form near degassing water and in saline lakes. Waterfalls downstream of springs often precipitate tufa as the turbulent water enhances the degassing of carbon dioxide, which makes calcite less soluble and causes it to precipitate. Saline lakes concentrate calcium carbonate from a combination of wave action causing degassing, springs in the lakebed, and evaporation. In salty Mono Lake in California, tufa towers were exposed after water was diverted and lowered the lake levels.

Figure 5.3.1: Travertine terraces of Mammoth Hot Springs, Yellowstone National Park, USA

Cave deposits like stalactites and stalagmites are another form of chemical precipitation of calcite, in a form called **travertine**. Calcite slowly precipitates from water to form the travertine, which often shows banding. This process is similar to the mineral growth on faucets in your home sink or shower that comes from hard (mineral-rich) water. Travertine also forms at hot springs such as Mammoth Hot Spring in Yellowstone National Park.

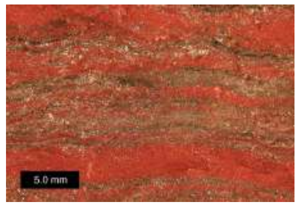


Figure 5.3.1: Alternating bands of iron-rich and silica-rich mud, formed as oxygen combined with dissolved iron.

Banded iron formation deposits commonly formed early in Earth's history, but this type of chemical sedimentary rock is no longer being created. Oxygenation of the atmosphere and oceans caused free iron ions, which are water-soluble, to become oxidized and precipitate out of solution. The iron oxide was deposited, usually in bands alternating with layers of chert.

Figure 5.3.1: A type of chert, flint, shown with a lighter weathered crust.

Chert, another commonly found chemical sedimentary rock, is usually produced from silica (SiO_2) precipitated from groundwater. Silica is highly insoluble on the surface of Earth, which is why quartz is so resistant to chemical weathering. Water deep underground is subjected to higher pressures and temperatures, which helps dissolve silica into an aqueous solution. As the groundwater rises toward or emerges at the surface the silica precipitates out, often as a cementing agent or into nodules. For example, the bases of the geysers in Yellowstone National Park are surrounded by silica deposits called geyserite or sinter. The silica is dissolved in water that is thermally heated by a relatively deep magma source. Chert can also form biochemically and is discussed in the Biochemical subsection. Chert has many synonyms, some of which may have gem value such as jasper, flint, onyx, and agate, due to subtle differences in colors, striping, etc., but chert is the more general term used by geologists for the entire group.

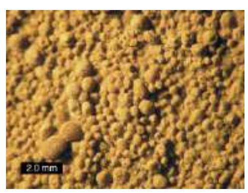


Figure 5.3.1: Ooids forming an oolite.

Oolites are among the few limestone forms created by an inorganic chemical process, similar to what happens in evaporite deposition. When water is oversaturated with calcite, the mineral precipitates out around a nucleus, a sand grain or shell fragment, and forms little spheres called ooids (see figure). As evaporation continues, the ooids continue building concentric layers of calcite as they roll around in gentle currents.

Biochemical

Biochemical sedimentary rocks are not that different from chemical sedimentary rocks; they are also formed from ions dissolved in solution. However, biochemical sedimentary rocks rely on biological processes to extract the dissolved materials out of the water. Most macroscopic marine organisms use dissolved minerals, primarily aragonite (calcium carbonate), to build hard parts such as shells. When organisms die the hard parts settle as sediment, which becomes buried, compacted, and cemented into rock.

Figure 5.3.1: Fossiliferous limestone (with brachiopods and bryozoans) from the Kope Formation of Ohio. Lower image is a section of the rock that has been etched with acid to emphasize the fossils.

This biochemical extraction and secretion is the main process for forming **limestone**, the most commonly occurring, non-clastic sedimentary rock. Limestone is mostly made of calcite ($CaCO_3$) and sometimes includes dolomite ($CaMgCO_3$), a close relative. Solid calcite reacts with hydrochloric acid by **effervescing** or fizzing. Dolomite only reacts to hydrochloric acid when ground into a powder, which can be done by scratching the rock surface (see Chapter 3, Minerals).

Figure 5.3.1: Close-up on coquina.

Limestone occurs in many forms, most of which originate from biological processes. Entire coral reefs and their ecosystems can be preserved in exquisite detail in limestone rock (see figure). **Fossiliferous limestone** contains many visible fossils. A type of limestone called **coquina** originates from beach sands made predominantly of shells that were then lithified. Coquina is composed of loosely-cemented shells and shell fragments. You can find beaches like this in modern tropical environments, such as the Bahamas. **Chalk** contains high concentrations of shells from a microorganism called a coccolithophore. **Micrite**, also known as microscopic calcite mud, is a very fine-grained limestone containing microfossils that can only be seen using a microscope.

Biogenetic chert forms on the deep ocean floor, created from biochemical sediment made of microscopic organic shells. This sediment, called ooze, may be calcareous (calcium carbonate-based) or siliceous (silica-based) depending on the type of shells deposited. For example, the shells of radiolarians (zooplankton) and diatoms (phytoplankton) are made of silica, so they produce siliceous ooze.

Organic

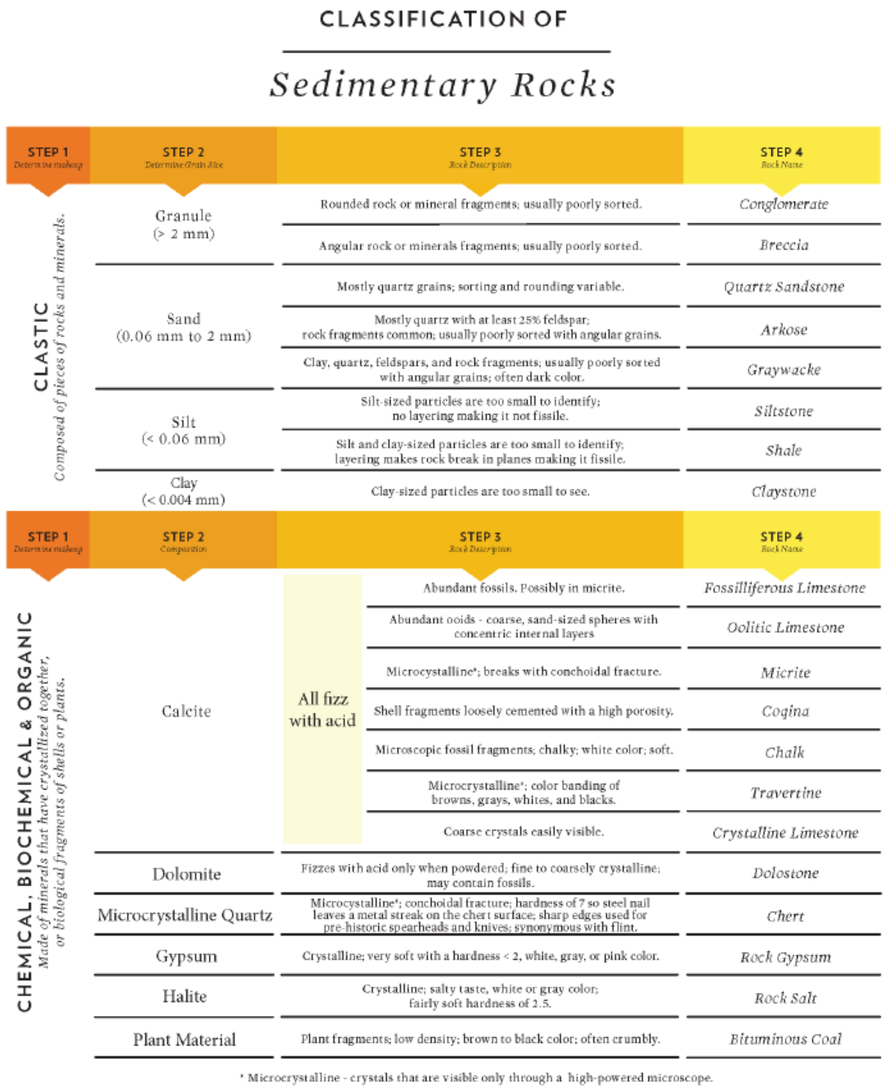
Under the right conditions, intact pieces of organic material or material derived from organic sources are preserved in the geologic record. Although not derived from sediment, this lithified organic material is associated with sedimentary strata and created by similar processes—burial, compaction, and diagenesis. Deposits of these fuels develop in areas where organic material collects in large quantities. Lush swamplands can create conditions conducive to the coal formation. Shallow-water, organic material-rich marine sediment can become highly productive petroleum and natural gas deposits. See Chapter 16, Energy and Mineral Resources, for a more in-depth look at these fossil-derived energy sources.

Figure 5.3.1: Anthracite coal, the highest grade of coal.

Classification of Chemical Sedimentary Rocks

Figure 5.3.1: Gyprock, a rock made of the mineral gypsum. From the Castile formation of New Mexico.

In contrast to detrital sediment, chemical, biochemical, and organic sedimentary rocks are classified based on mineral composition. Most of these are monomineralic, composed of a single mineral, so the rock name is usually associated with the identifying mineral. Chemical sedimentary rocks consisting of halite are called rock salt. Rocks made of Limestone (calcite) is an exception, having elaborate subclassifications and even two competing classification methods: Folk Classification and Dunham Classification [11; 21]. The Folk Classification deals with rock grains and usually requires a specialized, petrographic microscope. The Dunham Classification is based on rock texture, which is visible to the naked eye or using a hand lens and is easier for field applications. Most carbonate geologists use the Dunham system.



CLASSIFICATION OF

Sedimentary Rocks

STEP 1 Determine makeup	STEP 2 Determine Grain Size	STEP 3 Rock Description	STEP 4 Rock Name
CLASTIC *Composed of pieces of rocks and minerals.*	Granule (> 2 mm)	Rounded rock or mineral fragments; usually poorly sorted.	*Conglomerate*
		Angular rock or minerals fragments; usually poorly sorted.	*Breccia*
	Sand (0.06 mm to 2 mm)	Mostly quartz grains; sorting and rounding variable.	*Quartz Sandstone*
		Mostly quartz with at least 25% feldspar; rock fragments common; usually poorly sorted with angular grains.	*Arkose*
		Clay, quartz, feldspars, and rock fragments; usually poorly sorted with angular grains; often dark color.	*Graywacke*
	Silt (< 0.06 mm)	Silt-sized particles are too small to identify; no layering making it not fissile.	*Siltstone*
		Silt and clay-sized particles are too small to identify; layering makes rock break in planes making it fissile.	*Shale*
	Clay (< 0.004 mm)	Clay-sized particles are too small to see.	*Claystone*

STEP 1 Determine makeup	STEP 2 Composition		STEP 3 Rock Description	STEP 4 Rock Name
CHEMICAL, BIOCHEMICAL & ORGANIC *Made of minerals that have crystallized together, or biological fragments of shells or plants.*	Calcite	All fizz with acid	Abundant fossils. Possibly in micrite.	*Fossilliferous Limestone*
			Abundant ooids - coarse, sand-sized spheres with concentric internal layers	*Oolitic Limestone*
			Microcrystalline*; breaks with conchoidal fracture.	*Micrite*
			Shell fragments loosely cemented with a high porosity.	*Coqina*
			Microscopic fossil fragments; chalky; white color; soft.	*Chalk*
			Microcrystalline*; color banding of browns, grays, whites, and blacks.	*Travertine*
			Coarse crystals easily visible.	*Crystalline Limestone*
	Dolomite		Fizzes with acid only when powdered; fine to coarsely crystalline; may contain fossils.	*Dolostone*
	Microcrystalline Quartz		Microcystalline*; conchoidal fracture; hardness of 7 so steel nail leaves a metal streak on the chert surface; sharp edges used for pre-historic spearheads and knives; synonymous with flint.	*Chert*
	Gypsum		Crystalline; very soft with a hardness < 2, white, gray, or pink color.	*Rock Gypsum*
	Halite		Crystalline; salty taste, white or gray color; fairly soft hardness of 2.5.	*Rock Salt*
	Plant Material		Plant fragments; low density; brown to black color; often crumbly.	*Bituminous Coal*

* Microcrystalline - crystals that are visible only through a high-powered microscope.

Figure 5.3.1: Sedimentary rock identification chart

References

9. Udden, J. A. Mechanical composition of clastic sediments. *Geol. Soc. Am. Bull.* **25**, 655–744 (1914).

10. Wentworth, C. K. A scale of grade and class terms for clastic sediments. *J. Geol.* **30**, 377–392 (1922).

11. Folk, R. L. Petrography of sedimentary rocks. *Univ. Texas, Hemphill, Austin, Tex* **182**, (1974).

12. Dickinson, W. R. Interpreting detrital modes of graywacke and arkose. *J. Sediment. Res.* **40**, (1970).

13. Affolter, M. D. On the nature of volcanic lithic fragments: Definition source and evolution. (2004).

14. Dickinson, W. R. & Suczek, C. A. Plate tectonics and sandstone compositions. *AAPG Bull.* **63**, 2164–2182 (1979).

15. Johnson, C. L. *et al.* Sedimentary response to arc-continent collision, Permian, southern Mongolia. *Geological Society of America Special Papers* **436**, 363–390 (2008).

16. Cawood, P. A., Nemchin, A. A., Strachan, R., Prave, T. & Krabbendam, M. Sedimentary basin and detrital zircon record along East Laurentia and Baltica during assembly and breakup of Rodinia. *J. Geol. Soc. London* **164**, 257–275 (2007).

17. Grabau, A. W. *On the classification of sedimentary rocks*. (1904).

18. R. H. Dott, J. Wacke, Graywacke and Matrix–What Approach to Immature Sandstone Classification? *Journal of Sedimentary Research* **34**, (1964).

19. Borchert, H. & Muir, R. O. *Salt deposits: the origin, metamorphism and deformation of evaporites*. (Van Nostrand, 1964).

20. Boggs, S. J. *Principles of sedimentology and stratigraphy*. (Pearson, 2011).

21. Dunham, R. J. Classification of carbonate rocks according to depositional textures. (1962).

Contributions and Attributions

5.4: Sedimentary Structures

Sedimentary structures are visible textures or arrangements of sediments within a rock. Geologists use these structures to interpret the processes that made the rock and the environment in which it formed. They use uniformitarianism to usually compare sedimentary structures formed in modern environments to lithified counterparts in ancient rocks. Below is a summary discussion of common sedimentary structures that are useful for interpretations in the rock record.

Bedding Planes

Figure 5.4.1: Horizontal strata in southern Utah.

The most basic sedimentary structure is **bedding planes**, the planes that separate the layers or strata in sedimentary and some volcanic rocks. Visible in exposed outcroppings, each bedding plane indicates a change in sediment deposition conditions. This change may be subtle. For example, if a section of underlying sediment firms up, this may be enough to create a form or a layer that is dissimilar from the overlying sediment. Each layer is called a bed, or stratum, the most basic unit of **stratigraphy**, the study of sedimentary layering.

Figure 5.4.2: Students from the University of Wooster examine beds of Ordovician limestone in central Tennessee.

As would be expected, bed thickness can indicate sediment deposition quantity and timing. Technically, a bed is a bedding plane thicker than 1 cm (0.4 in) and the smallest mappable unit. A layer thinner than 1 cm (0.4 in) is called a **lamina** [22]. **Varves** are bedding planes created when laminae and beds are deposited in repetitive cycles, typically daily or seasonally [23]. Varves are valuable geologic records of climatic histories, especially those found in lakes and glacial deposits.

Graded Bedding

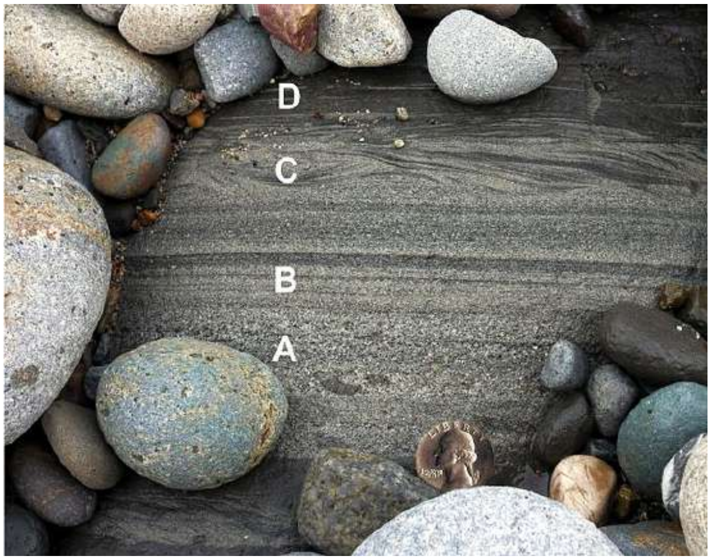


Figure 5.4.3: Image of the classic Bouma sequence. A=coarse- to fine-grained sandstone, possibly with an erosive base. B=laminated medium- to fine-grained sandstone. C=rippled fine-grained sandstone. D=laminated siltstone grading to mudstone.

Graded bedding refers to a sequence of increasingly coarse- or fine-grained sediment layers. Graded bedding often develops when sediment deposition occurs in an environment of decreasing energy. A **Bouma sequence** is graded bedding observed in a clastic rock called turbidite [24]. Bouma sequence beds are formed by offshore sediment gravity flows, which are underwater flows of sediment. These subsea density flows begin when sediment is stirred up by an energetic process and becomes a dense slurry of mixed grains. The sediment flow courses downward through submarine channels and canyons due to gravity acting on the density difference between the denser slurry and less dense surrounding seawater. As the flow reaches deeper ocean basins it slows down, loses energy, and deposits sediment in a Bouma sequence of coarse grains first, followed by increasingly finer grains (see figure).

Flow Regime and Bedforms

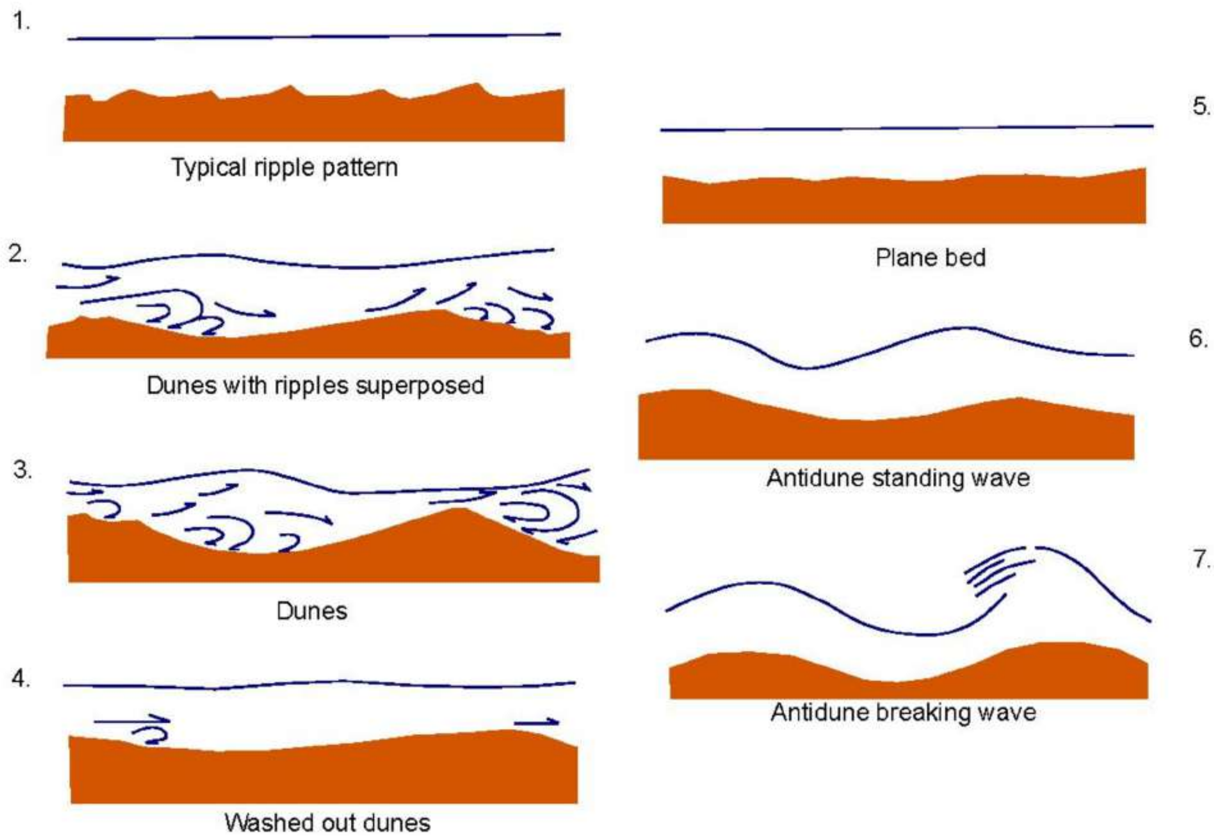


Figure 5.4.4: Bedforms from under increasing flow velocities.

In fluid systems, such as moving water or wind, sand is the most easily transported and deposited sediment grain. Smaller particles like silt and clay are less movable by fluid systems because the tiny grains are chemically attracted to each other and stick to the underlying sediment. Under higher flow rates, the fine silt and clay sediment tend to stay in place and the larger sand grains get picked up and moved.

Bedforms are sedimentary structures created by fluid systems working on sandy sediment [25]. Grain size, flow velocity, and **flow regime** or pattern interact to produce bedforms having unique, identifiable physical characteristics. Flow regimes are divided into upper and lower regimes, which are further divided into uppermost, upper, lower, and lowermost parts. The table below shows bedforms and their associated flow regimes. For example, the dunes bedform is created in the upper part of the lower flow regime.

Flow Regime (part)	Bedform	Description
Lower (lowest)	Plane bed	Lower plane bed, flat laminations
Lower (lower)	Ripples	Small (with respect to flow) inclined layers dipping downflow
Lower (upper)	Dunes	Larger inclined cross beds, ±ripples, dipping downflow
Upper (lower)	Plane bed	Flat layers, can include lined-up grains (parting lineations)
Upper (upper)	Antidunes	Hard to preserve reverse dunes dipping shallowly upflow
Upper (uppermost)	Chutes/pools *(rare)*	Erosional, not really a bedform; rarely found preserved

Plane Beds

Figure 5.4.5: Subtle lines across this sandstone (trending from the lower left to upper right) are parting lineations.

Plane beds created in the lower flow regime are like bedding planes, on a smaller scale. The flat, parallel layers form as sandy sediment piles and move on top of layers below. Even non-flowing fluid systems, such as lakes, can produce sediment plane beds. Plane beds in the upper flow regime are created by fast-flowing fluids. They may look identical to lower-flow-regime beds; however, they typically show **parting lineations**, slight alignments of grains in rows and swaths, caused by high sediment transport rates that only occur in upper flow regimes.

Ripples

Figure 5.4.6: Modern current ripple in sand from the Netherlands. The flow creates a steep side down current. In this image, the flow is from right to left.

Ripples are known by several names: ripple marks, ripple cross-beds, or ripple cross laminations. The ridges or undulations in the bed are created as sediment grains pile up on top of the plane bed. With the exception of dunes, the scale of these beds is typically measured in centimeters. Occasionally, large flows like glacial lake outbursts can produce ripples as tall as 20 m (66 ft).

Figure 5.4.7: A bidirectional flow creates this symmetrical wave ripple. From rocks in Nomgon, Mongolia. Note the crests of the ripples have been eroded away by subsequent flows in places.

First scientifically described by Hertha Ayrton [26], ripple shapes are determined by flow type and can be straight-crested, sinuous, or complex. Asymmetrical ripples form in a unidirectional flow. Symmetrical ripples are the result of an oscillating back-and-forth flow typical of intertidal swash zones. Climbing ripples are created from high sedimentation rates and appear as overlapping layers of ripple shapes (see figure).

Figure 5.4.8: Climbing ripple deposit from India.

Dunes

Figure 5.4.9: Lithified cross-bedded dunes from the high country of Zion National Park, Utah. The complexity of bedding planes results from the three-dimensional network of ancient dune flows.

Dunes are very large and prominent versions of ripples and typical examples of large cross-bedding [27]. Cross bedding happens when ripples or dunes pile atop one another, interrupting, and/or cutting into the underlying layers. Desert sand dunes are probably the first image conjured up by this category of bedform.

British geologist Agnold (1941) considered only Barchan and linear Seif dunes as the only true dune forms. Other workers have recognized transverse and star dunes as well as parabolic and linear dunes anchored by plants that are common in coastal areas as other types of dunes.

Figure 5.4.10: The modern sand dune in Morocco.

Dunes are the most common sedimentary structure found within channelized flows of air or water. The biggest difference between river dunes and air-formed (desert) dunes is the depth of the fluid system. Since the atmosphere's depth is immense when compared to a river channel, desert dunes are much taller than those found in rivers. Some famous air-formed dune landscapes include the Sahara Desert, Death Valley, and the Gobi Desert [28].

As airflow moves sediment along, the grains accumulate on the dune's windward surface (facing the wind). The angle of the windward side is typically shallower than the leeward (downwind) side, which has grains falling down over it. This difference in slopes can be seen in a bed cross-section and indicates the direction of the flow in the past. There are typically two styles of dune beds: the more common trough cross-beds with curved windward surfaces, and rarer planar cross-beds with flat windward surfaces.

In tidal locations with strong in-and-out flows, dunes can develop in opposite directions. This produces a feature called herringbone cross-bedding.

Figure 5.4.11: Herringbone cross-bedding from the Mazomanie Formation, upper Cambrian of Minnesota.

Figure 5.4.12: Hummocky-cross stratification, seen as wavy lines throughout the middle of this rock face. The best example is just above the pencil in the center.

Another dune formation variant occurs when very strong, hurricane-strength, winds agitate parts of the usually undisturbed seafloor. These beds are called **hummocky cross-stratification** and have a 3D architecture of hills and valleys, with inclined and declined layering that matches the dune shapes.

Antidunes

Figure 5.4.13: Antidunes forming in Urdaibai, Spain.

Antidunes are so named because they share similar characteristics with dunes, but are formed by a different, opposing process [29]. While dunes form in lower flow regimes, antidunes come from fast-flowing upper flow regimes. In certain conditions of high flow rates, the sediment accumulates upstream of a subtle dip instead of traveling downstream (see figure). Antidunes form in phase with the flow; in rivers, they are marked by rapids in the current. Antidunes are rarely preserved in the rock record because the high flow rates needed to produce the beds also accelerate erosion.

Bioturbation

Figure 5.4.14: Bioturbated dolomitic siltstone from Kentucky.

Bioturbation is the result of organisms burrowing through soft sediment, which disrupts the bedding layers. These tunnels are backfilled and eventually preserved when the sediment becomes rock. Bioturbation happens most commonly in shallow, marine environments, and can be used to indicate water depth [30].

Mudcracks

Figure 5.4.15: Lithified mud cracks from Maryland.

Mudcracks occur in clay-rich sediment that is submerged underwater and later dries out. Water fills voids in the clay's crystalline structure, causing the sediment grains to swell. When this waterlogged sediment begins to dry out, the clay grains shrink. The sediment layer forms deep polygonal cracks with tapered openings toward the surface [31], which can be seen in profile. The cracks fill with new sediment and become visible veins running through the lithified rock. These dried-out clay beds are a major source of **mud chips**, small fragments of mud or shale, which commonly become inclusions in sandstone and conglomerate. What makes this sedimentary structure so important to geologists is that they only form in certain depositional environments—such as tidal flats that form underwater and are later exposed to air. Syneresis cracks are similar in appearance to mud cracks but much rarer; they are formed when subaqueous (underwater) clay sediment shrinks [32].

Sole Marks

Figure 5.4.16: This flute cast shows a flow direction toward the upper right of the image, as seen by the bulge sticking down out of the layer above. The flute cast would have been scoured into a rock layer below that has been removed by erosion, leaving the sandy layer above to fill in the flute cast.

Sole marks are small features typically found in river deposits. They form at the base of a bed, the sole, and on top of the underlying bed. They can indicate several things about the deposition conditions, such as flow direction or stratigraphic up-direction (see Geopetal Structures section). **Flute casts** or scour marks are grooves carved out by the forces of fluid flow and sediment loads. The upstream part of the flow creates steep grooves and downstream the grooves are shallower. The grooves subsequently become filled by overlying sediment, creating a cast of the original hollow [33].

Figure 5.4.17: Groove casts at the base of a turbidite deposit in Italy.

Formed similarly to flute casts but with a more regular and aligned shape, **groove casts** are produced by larger clasts or debris carried along in the water that scrape across the sediment layer. Tool marks come from objects like sticks carried in the fluid downstream or embossed into the sediment layer, leaving a depression that later fills with new sediment.

Figure 5.4.18: A drill core showing a loaded cast showing light-colored sand sticking down into dark mud.

Load casts, an example of **soft-sediment deformation**, are small indentations made by an overlying layer of coarse sediment grains or clasts intruding into a softer, finer-grained sediment layer [34].

Raindrop Impressions

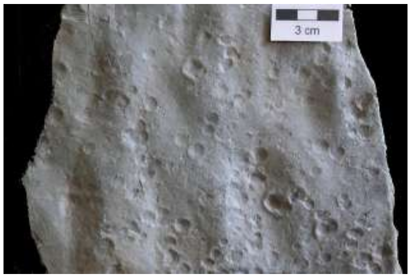


Figure 5.4.19: Mississippian raindrop impressions over wave ripples from Nova Scotia.

Like their name implies, **raindrop impressions** are small pits or bumps found in soft sediment. While they are generally believed to be created by rainfall, they may be caused by other agents such as escaping gas bubbles [35].

Imbrication

Figure 5.4.20: Cobbles in this conglomerate are positioned in a way that they are stacked on each other, which occurred as flow went from left to right.

Imbrication is a stack of large and usually flat clasts—cobbles, gravels, mud chips, etc—that are aligned in the direction of fluid flow [36]. The clasts may be stacked in rows, with their edges dipping down and flat surfaces aligned to face the flow (see figure). Or their flat surfaces may be parallel to the layer and long axes aligned with the flow. Imbrications are useful for analyzing **paleocurrents**, or currents found in the geologic past, especially in alluvial deposits.

Geopetal Structures

Figure 5.4.21: This bivalve (clam) fossil was partially filled with tan sediment, partially empty. Later fluids filled in the fossil with white calcite minerals. The line between the sediment and the later calcite is paleo-horizontal.

Figure 5.4.22: Eubrontes trace fossil from Utah, showing the geopetal direction is into the image.

Geopetal structures [37], also called up-direction indicators, are used to identify which way was up when the sedimentary rock layers were originally formed. This is especially important in places where the rock layers have been deformed, tilted, or overturned. Well-preserved mud cracks, sole marks, and raindrop impressions can be used to determine up direction. Other useful geopetal structures include:

- Vugs: Small voids in the rock that usually become filled during diagenesis. If the void is partially filled or filled in stages, it serves as a permanent record of a level bubble, frozen in time.
- Cross bedding – In places where ripples or dunes pile on top of one another, where one cross bed interrupts and/or cuts another below, this shows a cross-cutting relationship that indicates up direction.
- Ripples, dunes: Sometimes the ripples are preserved well enough to differentiate between the crests (top) and troughs (bottom).
- Fossils: Body fossils in life position, meaning the body parts are not scattered or broken, and trace fossils like footprints (see figure) can provide an up direction. Intact fossilized coral reefs are excellent up indicators because of their large size and easily distinguishable top and bottom. Index fossils, such as ammonites, can be used to age date strata and determine up direction based on relative rock ages.
- Vesicles – Lava flows eliminate gas upwards. An increase of vesicles toward the top of the flow indicates up.

References

- 22. Mckee, E. D. & Weir, G. W. Terminology for stratification and cross-stratification in sedimentary rocks. *Geol. Soc. Am. Bull.* **64**, 381–390 (1953).
- 23. de Geer, G. Geochronology of the last 12,000 years. in *Milestones in Geosciences* (eds. Dullo, P. D. W.-C. & e.V., G. V.) 100–110 (Springer Berlin Heidelberg, 2003).
- 24. Bouma, A. H., Kuenen, P. H. & Shepard, F. P. *Sedimentology of some flysch deposits: a graphic approach to facies interpretation*. **168**, (Elsevier Amsterdam, 1962).
- 25. Blatt, H., Middleton, G. V. & Murray, R. *Origin of Sedimentary Rocks*. (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, USA, 1980).
- 26. Ayrton, H. The origin and growth of ripple-mark. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **84**, 285–310 (1910).
- 27. Ashley, G. M. Classification of large-scale subaqueous bedforms: a new look at an old problem-SEPM bedforms and bedding structures. *J. Sediment. Res.* **60**, (1990).
- 28. Dingus, L. & Loope, D. Death in the Dunes. *Nat. Hist.* **109**, 50–55 (2000).
- 29. Gilbert, G. K. & Murphy, E. C. *The Transportation of Débris by Running Water*. (U.S. Government Printing Office, 1914).
- 30. Ekdale, A. A. Classification of Trace Fossils. (1984).
- 31. Plummer, P. S. & Gostin, V. A. Shrinkage cracks: desiccation or synaeresis? *J. Sediment. Res.* **51**, (1981).
- 32. Burst, J. F. Subaqueously formed shrinkage cracks in clay. *J. Sediment. Res.* **35**, (1965).
- 33. Yin, D. *et al.* Bedform genesis in bedrock substrates: Insights into formative processes from a new experimental approach and the importance of suspension-dominated abrasion. *Geomorphology* **255**, 26–38 (2016).
- 34. Pettijohn, F. J. & Potter, P. E. Atlas and glossary of primary sedimentary structures. (2012).
- 35. Metz, R. Why not raindrop impressions? *J. Sediment. Res.* **51**, (1981).

- 36. Karátson, D., Sztanó, O. & Telbisz, T. Preferred clast orientation in volcaniclastic mass-flow deposits: application of a new photo-statistical method. *J. Sediment. Res.* **72**, 823–835 (2002).
- 37. Sander, B. *Contributions to the study of depositional fabrics: rhythmically deposited Triassic limestones and dolomites*. (American Association of Petroleum Geologists, 1951).

Contributions and Attributions

5.5: Depositional Environments

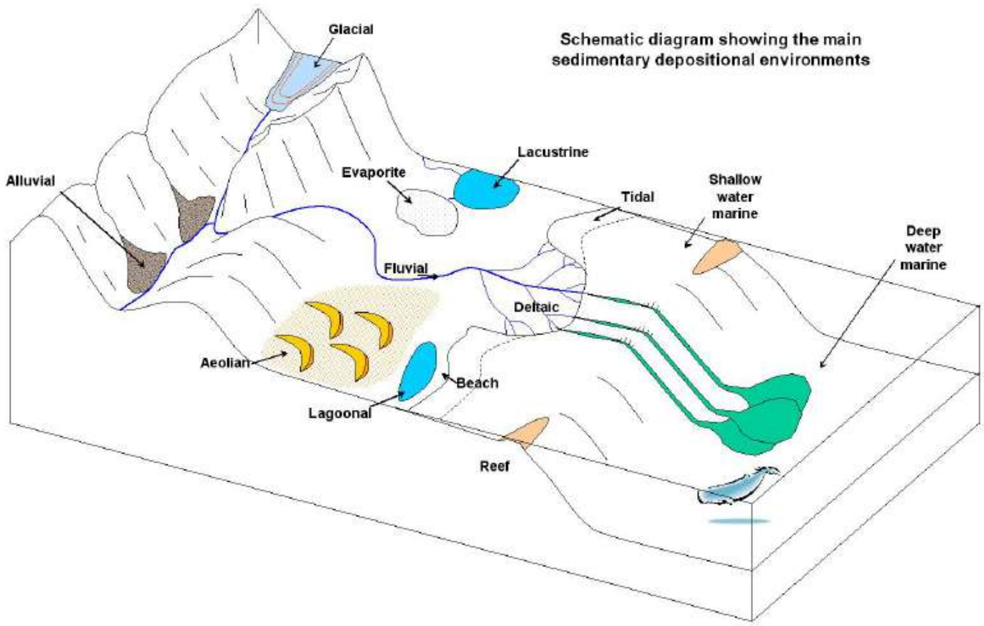


Figure 5.5.1: A representation of common depositional environments.

The ultimate goal of many stratigraphy studies is to understand the original **depositional environment**. Knowing where and how a particular sedimentary rock was formed can help geologists paint a picture of past environments—such as a mountain glacier, gentle floodplain, dry desert, or deep-sea ocean floor. The study of depositional environments is a complex endeavor; the table shows a simplified version of what to look for in the rock record.

Location	Sediment	Common Rock Types	Typical Fossils	Sedimentary Structures
Abyssal	very fine muds and oozes, diatomaceous Earth	chert	diatoms	few
Submarine fan	graded Bouma sequences, alternating sand/mud	clastic rocks	rare	channels, fan shape
Continental slope	mud, possible sand, countourites	shale, siltstone, limestone	rare	swaths
Lower shoreface	laminated sand	sandstone	bioturbation	hummocky cross beds
Upper shoreface	planar sand	sandstone	bioturbation	plane beds, cross beds
Littoral (beach)	very well sorted sand	sandstone	bioturbation	few
Tidal Flat	mud and sand with channels	shale, mudstone,siltstone	bioturbation	mudcracks, symmetric ripples
Reef	lime mud with coral	limestone	many, commonly coral	few
Lagoon	laminated mud	shale	many, bioturbation	laminations
Delta	channelized sand with mud, ±swamp	clastic rocks	many to few	cross beds
Fluvial (river)	sand and mud, can have larger sediments	sandstone, conglomerate	bone beds (rare)	cross beds, channels, asymmetric ripples
Alluvial	mud to boulders, poorly sorted	clastic rocks	rare	channels, mud cracks

Location	Sediment	Common Rock Types	Typical Fossils	Sedimentary Structures
Lacustrine (lake)	fine-grained laminations	shale	invertebrates, rare (deep) bone beds	laminations
Paludal (swamp)	plant material	coal	plant debris	rare
Aeolian (dunes)	very well-sorted sand and silt	sandstone	rare	cross beds (large)
Glacial	mud to boulders, poorly sorted	conglomerate (tillite)		striations, drop stones

Marine

Marine depositional environments are completely and constantly submerged in seawater. Their depositional characteristics are largely dependent on the depth of water with two notable exceptions: submarine fans and turbidites.

Abyssal

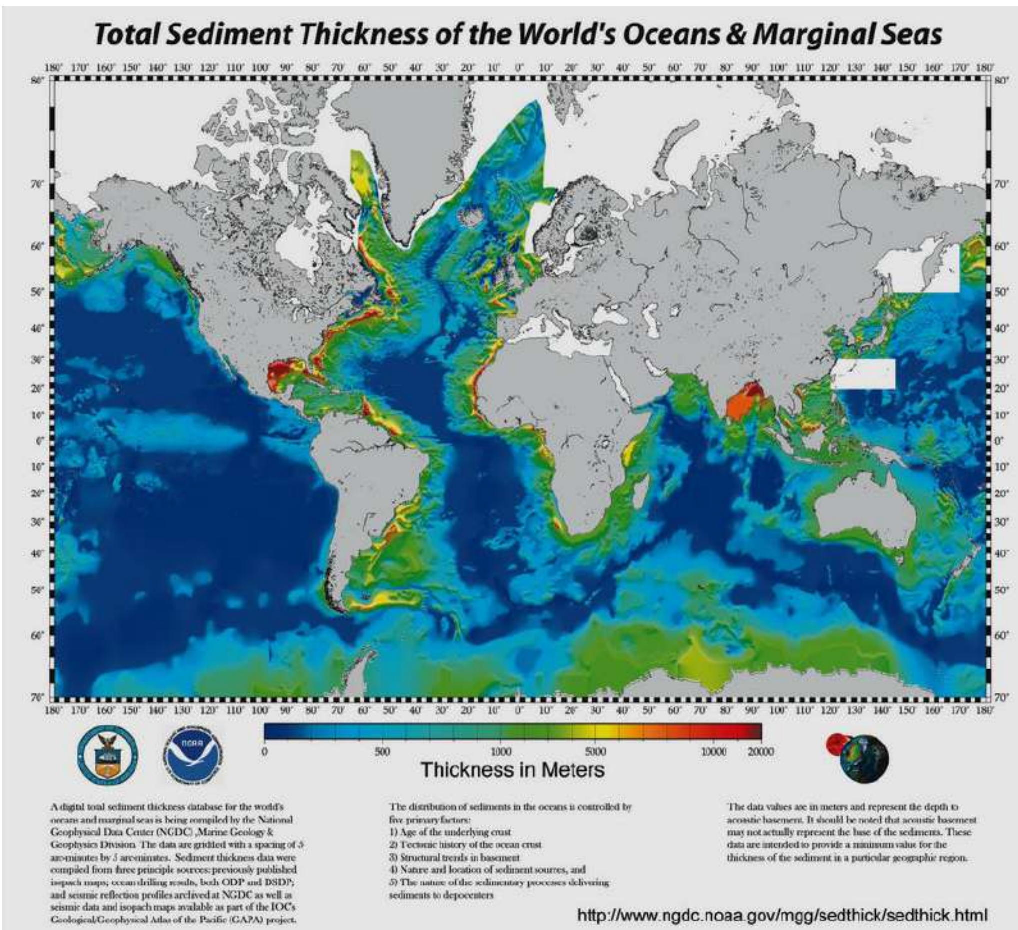


Figure 5.5.2: Marine sediment thickness. Note the lack of sediment away from the continents.

Abyssal sedimentary rocks form on the **abyssal plain**. The plain encompasses a relatively flat ocean floor with some minor topographical features, called abyssal hills. These small seafloor mounts range from 100 m to 20 km in diameter and are possibly created by extension [38]. Most abyssal plains do not experience significant fluid movement, so sedimentary rocks formed there are very fine-grained [39].

There are three categories of abyssal sediment. Calcareous oozes consist of calcite-rich plankton shells that have fallen to the ocean floor. An example of this type of sediment is chalk. Siliceous oozes are also made of plankton debris, but these organisms build their shells using silica or hydrated silica. In some cases such as with diatomaceous earth, sediment is deposited below the **calcite compensation depth**, a depth where calcite solubility increases. Any calcite-based shells are dissolved, leaving only silica-based shells. Chert is another common rock formed from these types of sediment. These two types of abyssal sediment are also classified as biochemical in origin. (see BIOCHEMICAL section).

Figure 5.5.3: Diatomaceous earth

The third sediment type is pelagic clay. Very fine-grained clay particles, typically brown or red, descend through the water column very slowly. Pelagic clay deposition occurs in areas of remote open ocean, where there is little plankton accumulation.

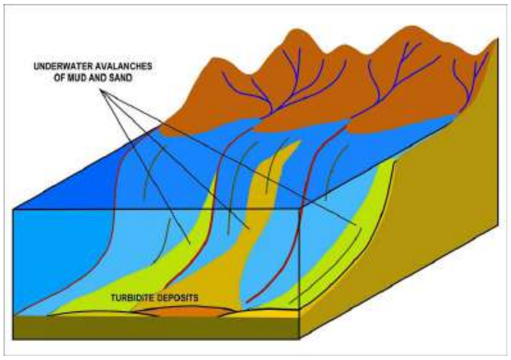


Figure 5.5.4: Turbidites inter-deposited within submarine fans.

Two notable exceptions to the fine-grained nature of abyssal sediment are **submarine fan** and **turbidite** deposits [40]. Submarine fans occur offshore at the base of large river systems. They are initiated during times of low sea level, as strong river currents carve submarine canyons into the continental shelf. When sea levels rise, sediments accumulate on the shelf typically forming large, fan-shaped floodplains called deltas. Periodically, the sediment is disturbed creating dense slurries that flush down the underwater canyons in large gravity-induced events called turbidites. The submarine fan is formed by a network of turbidites that deposit their sediment loads as the slope decreases, much like what happens above-water at alluvial fans and deltas. This sudden flushing transports coarser sediment to the ocean floor where they are otherwise uncommon. Turbidites are also the typical origin of graded Bouma sequences. (see Chapter 5, Weathering, Erosion, and Sedimentary Rock).

Continental Slope

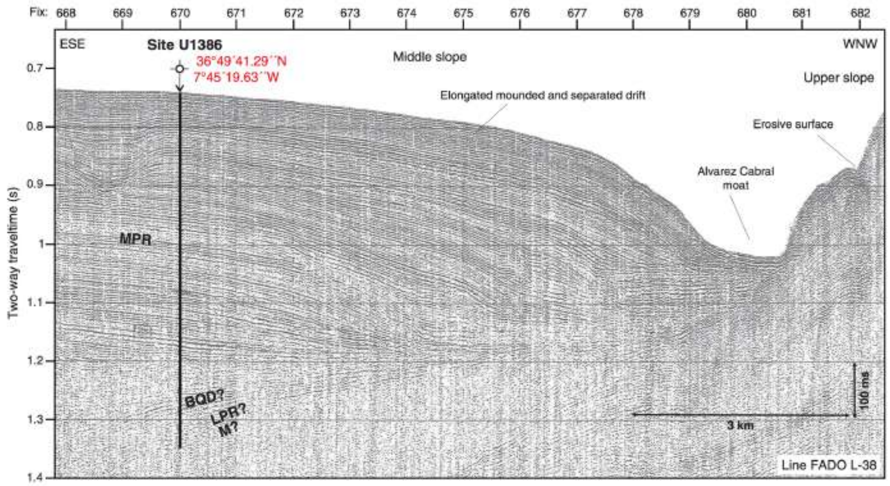


Figure 5.5.5: Contourite drift deposit imaged with seismic waves.

Continental slope deposits are not common in the rock record. The most notable type of continental slope deposits are contourites [41]. Contourites form on the slope between the continental shelf and the deep ocean floor. Deep-water ocean currents deposit sediment into smooth drifts of various architectures, sometimes interwoven with turbidites.

Lower Shoreface

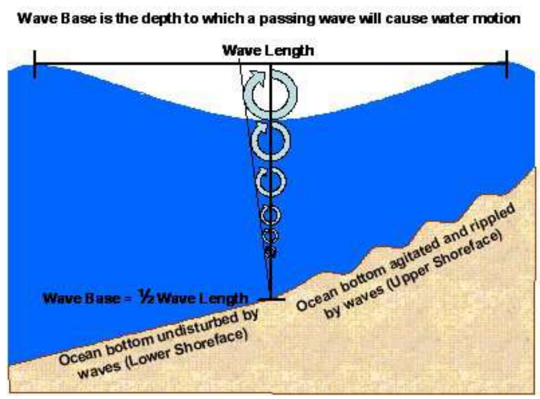


Figure 5.5.6: Diagram describing wavebase.

The **lower shoreface** lies below the normal depth of wave agitation, so the sediment is not subject to daily winnowing and deposition. These sediment layers are typically finely laminated and may contain hummocky cross-stratification. Lower shoreface beds are affected by larger waves, such those as generated by hurricanes and other large storms [42].

Upper Shoreface

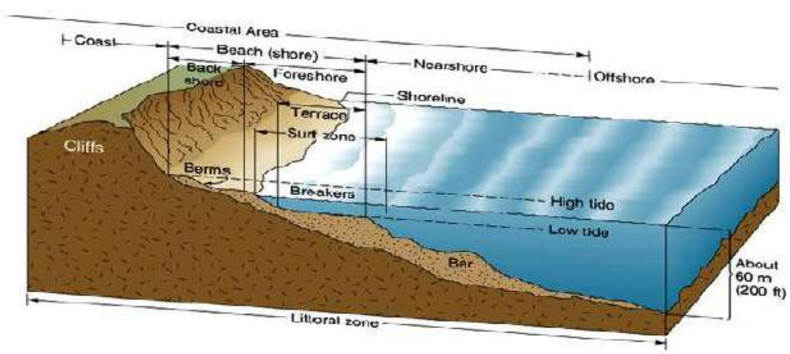


Figure 5.5.7: Diagram of zones of the shoreline.

The **upper shoreface** contains sediments within the zone of normal wave action but still submerged below the beach environment. These sediments usually consist of very well sorted, fine sand. The main sedimentary structure is planar bedding consistent with the lower part of the upper flow regime, but it can also contain cross-bedding created by longshore currents [43].

Transitional Coastline Environments

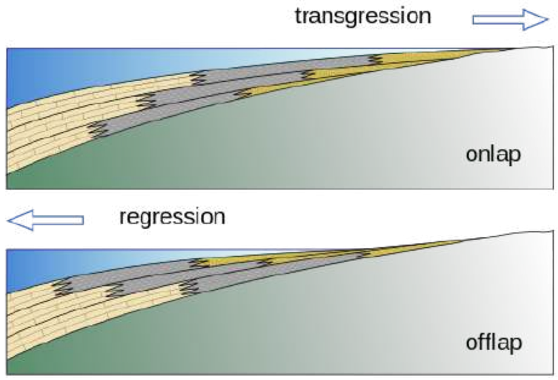


Figure 5.5.8: The rising sea levels of transgressions create onlapping sediments, regressions create offlapping.

Transitional environments, more often called shoreline or **coastline environments**, are zones of complex interactions caused by ocean water hitting land. The sediment preservation potential is very high in these environments because deposition often occurs on the **continental shelf** and underwater. Shoreline environments are an important source of hydrocarbon deposits (petroleum, natural gas).

The study of shoreline depositional environments is called **sequence stratigraphy**. Sequence stratigraphy examines depositional changes and 3D architectures associated with rising and falling sea levels, which is the main force at work in shoreline deposits. These sea-level fluctuations come from the daily tides, as well as climate changes and plate tectonics. A steady rise in sea level relative to the shoreline is called **transgression**. **Regression** is the opposite, a relative drop in sea level. Some common components of shoreline environments are littoral zones, tidal flats, reefs, lagoons, and deltas. For a more in-depth look at these environments, see Chapter 12, Coastlines.

Littoral

Figure 5.5.9: Lithified heavy mineral sand (dark layers) from a beach deposit in India.

The **littoral** zone, better known as the beach, consists of highly weathered, homogeneous, well-sorted sand grains made mostly of quartz. There are black sand and other types of sand beaches, but they tend to be unique exceptions rather than the rule. Because beach sands, past or present, are so highly evolved, the amount of grain weathering can be discerned using the minerals zircon, tourmaline, and rutile. This tool is called the ZTR (zircon, tourmaline, rutile) index [44]. The ZTR index is higher in more weathered beaches because these relatively rare and weather-resistant minerals become concentrated in older beaches. In some beaches, the ZTR index is so high the sand can be harvested as an economically viable source of these minerals. The beach environment has no sedimentary structures, due to the constant bombardment of wave energy delivered by surf action. Beach sediment is moved around via multiple processes. Some beaches with high sediment supplies develop dunes nearby.

Tidal Flats

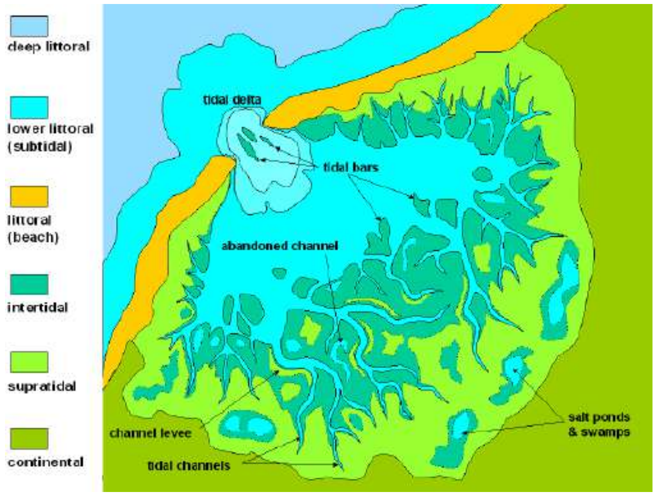


Figure 5.5.10: General diagram of a tidal flat and associated features.

Tidal flats, or mudflats, are sedimentary environments that are regularly flooded and drained by ocean tides. Tidal flats have large areas of fine-grained sediment but may also contain coarser sands. Tidal flat deposits typically contain gradational sediments and may include multi-directional ripple marks. Mudcracks are also commonly seen due to the sediment being regularly exposed to air during low tides; the combination of mud cracks and ripple marks is distinctive to tidal flats [45].

Tidal water carries in sediment, sometimes focusing the flow through a narrow opening called a tidal inlet. Tidal channels, creek channels influenced by tides, can also focus on tidally-induced flow. Areas of higher flow like inlets and tidal channels feature coarser grain sizes and larger ripples, which in some cases can develop into dunes.

Reefs

Figure 5.5.11: Waterpocket fold, Capitol Reef National Park, Utah

Reefs, which most people would immediately associate with tropical coral reefs found in the oceans, are not only made by living things. Natural buildups of sand or rock can also create reefs, similar to barrier islands. Geologically speaking, a **reef** is any topographically-elevated feature on the continental shelf, located oceanward of and separate from the beach. The term reef can also be applied to terrestrial (atop the continental crust) features. Capitol Reef National Park in Utah contains a topographic barrier, a reef, called the Waterpocket Fold.

Figure 5.5.12: A modern coral reef.

Most reefs, now and in the geologic past, originate from the biological processes of living organisms [46]. The growth habits of coral reefs provide geologists important information about the past. The hard structures in coral reefs are built by soft-bodied marine organisms, which continually add new material and enlarge the reef over time. Under certain conditions, when the land beneath a reef is subsiding, the coral reef may grow around and through existing sediment, holding the sediment in place, and thus preserving the record of environmental and geological conditions around it.

Figure 5.5.13: The light blue reef is fringing the island of Vanatinai. As the island erodes away, only the reef will remain, forming a reef-bound seamount.

Sediment found in coral reefs is typically fine-grained, mostly carbonate, and tends to deposit between the intact coral skeletons. Water with high levels of silt or clay particles can inhibit the reef growth because coral organisms require sunlight to thrive; they host symbiotic algae called zooxanthellae that provide the coral with nourishment via photosynthesis. Inorganic reef structures have much more variable compositions. Reefs have a big impact on sediment deposition in lagoon environments since they are natural storm breaks, wave and storm buffers, which allows fine grains to settle and accumulate.

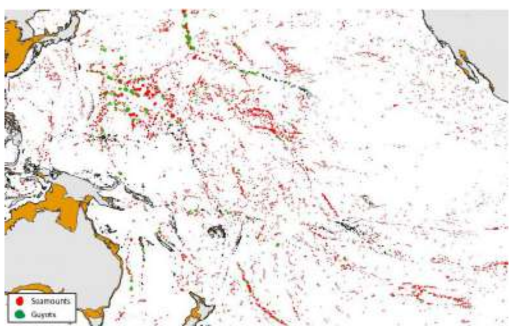


Figure 5.5.14: Seamounts and guyots in the North Pacific.

Reefs are found around shorelines and islands; coral reefs are particularly common in tropical locations. Reefs are also found around features known as **seamounts**, which is the base of an ocean island left standing underwater after the upper part is eroded away by waves. Examples include the Emperor Seamounts, formed millions of years ago over the Hawaiian Hotspot. Reefs live and grow along the upper edge of these flat-topped seamounts. If the reef builds up above sea level and completely encircles the top of the seamount, it is called a coral-ringed atoll. If the reef is submerged, due to erosion, subsidence, or sea-level rise, the seamount-reef structure is called a guyot.

Lagoon

Figure 5.5.15: Kara-Bogaz Gol lagoon, Turkmenistan.

Lagoons are small bodies of seawater located inland from the shore or isolated by another geographic feature, such as a reef or barrier island. Because they are protected from the action of tides, currents, and waves, lagoon environments typically have very fine-grained sediments [47]. Lagoons, as well as estuaries, are ecosystems with high biological productivity. Rocks from these environments often include bioturbation marks or coal deposits. Around lagoons where evaporation exceeds water inflow, salt flats, also known as sabkhas, and sand dune fields may develop at or above the high tide line.

Deltas

Figure 5.5.16: The Nile delta, in Egypt.

Figure 5.5.17: Birdfoot river-dominated delta of the Mississippi River.

Deltas form where rivers enter lakes or oceans and are of three basic shapes: river-dominated deltas, wave-dominated deltas, and tide-dominated deltas. The name delta comes from the Greek letter Δ (delta, uppercase) [48], which resembles the triangular shape of the Nile River delta. The velocity of water flow is dependent on riverbed slope or gradient, which becomes shallower as the river descends from the mountains. At the point where a river enters an ocean or lake, its slope angle drops to zero degrees (0°). The flow velocity quickly drops as well, and sediment is deposited, from coarse clasts to fine sand, and mud to form the delta. As one part of the delta becomes overwhelmed by sediment, the slow-moving flow gets diverted back and forth, over and over, and forms a spread out network of smaller distributary channels.

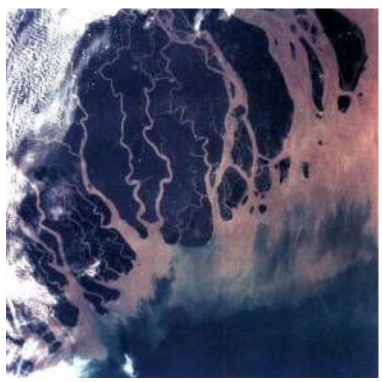


Figure 5.5.18: Tidal delta of the Ganges River.

Deltas are organized by the dominant process that controls their shape: tide-dominated, wave-dominated, or river-dominated. Wave-dominated deltas generally have smooth coastlines and beach-ridges on the land that represent previous shorelines. The Nile River delta is a wave-dominated type. (see figure).

The Mississippi River delta is a river-dominated delta, shaped by levees along the river and its distributaries that confine the flow forming a shape called a bird-foot delta. Other times the tides or the waves can be a bigger factor and can reshape the delta in various ways.

A tide-dominated delta is dominated by tidal currents. During flood stages when rivers have lots of water available, it develops distributaries that are separated by sand bars and sand ridges. The tidal delta of the Ganges River is the largest delta in the world.

Terrestrial

Terrestrial depositional environments are diverse. Water is a major factor in these environments, in liquid or frozen states, or even when it is lacking (arid conditions).

Fluvial

Figure 5.5.19: The Cauto River in Cuba. Note the sinuosity in the river, which is meandering.

Fluvial (river) systems are formed by water flowing in channels over the land. They generally come in two main varieties: meandering or braided. In meandering streams, the flow carries sediment grains via a single channel that wanders back and forth across the floodplain. The floodplain sediment away from the channel is mostly fine grained material that only gets deposited during floods.

Figure 5.5.20: The braided Waimakariri river in New Zealand.

Braided fluvial systems generally contain coarser sediment grains, and form a complicated series of intertwined channels that flow around gravel and sand bars [49] (see Chapter 11, Water).

Alluvial

Figure 5.5.21: An alluvial fan spreads out into a broad alluvial plain. From Red Rock Canyon State Park, California.

A distinctive characteristic of **alluvial** systems is the intermittent flow of water. Alluvial deposits are common in arid places with little soil development. Lithified alluvial beds are the primary basin-filling rock found throughout the Basin and Range region of the western United States. The most distinctive alluvial sedimentary deposit is the alluvial fan, a large cone of sediment formed by streams flowing out of dry mountain valleys into a wider and more open dry area. Alluvial sediments are typically poorly sorted and coarse-grained, and often found near playa lakes or aeolian deposits [50] (see Chapter 13, Deserts).

Lacustrine

Figure 5.5.22: Oregon's Crater Lake was formed about 7700 years ago after the eruption of Mount Mazama.

Lake systems and deposits, called **lacustrine**, form via processes somewhat similar to marine deposits, but on a much smaller scale. Lacustrine deposits are found in lakes in a wide variety of locations. Lake Baikal in southeast Siberia (Russia) is in a tectonic basin. Crater Lake (Oregon) sits in a volcanic caldera. The Great Lakes (northern United States) came from glacially carved and deposited sediment. Ancient Lake Bonneville (Utah) formed in a pluvial setting during a climate that was relatively wetter and cooler than that of modern Utah. Oxbow lakes, named for their curved shape, originated in fluvial floodplains.

Lacustrine sediment tends to be very fine-grained and thinly laminated, with only minor contributions from wind-blown, current, and tidal deposits [51]. When lakes dry out or evaporation outpaces precipitation, playas form. **Playa** deposits resemble those of normal lake deposits but contain more evaporite minerals. Certain tidal flats can have playa-type deposits as well.

Paludal

Paludal systems include bogs, marshes, swamps, or other wetlands, and usually contain lots of organic matter. Paludal systems typically develop in coastal environments but are common in humid, low-lying, low-latitude, warm zones with large volumes of flowing water. A characteristic paludal deposit is a peat bog, a deposit rich in organic matter that can be converted into coal when lithified. Paludal environments may be associated with tidal, deltaic, lacustrine, and/or fluvial deposition.

Aeolian

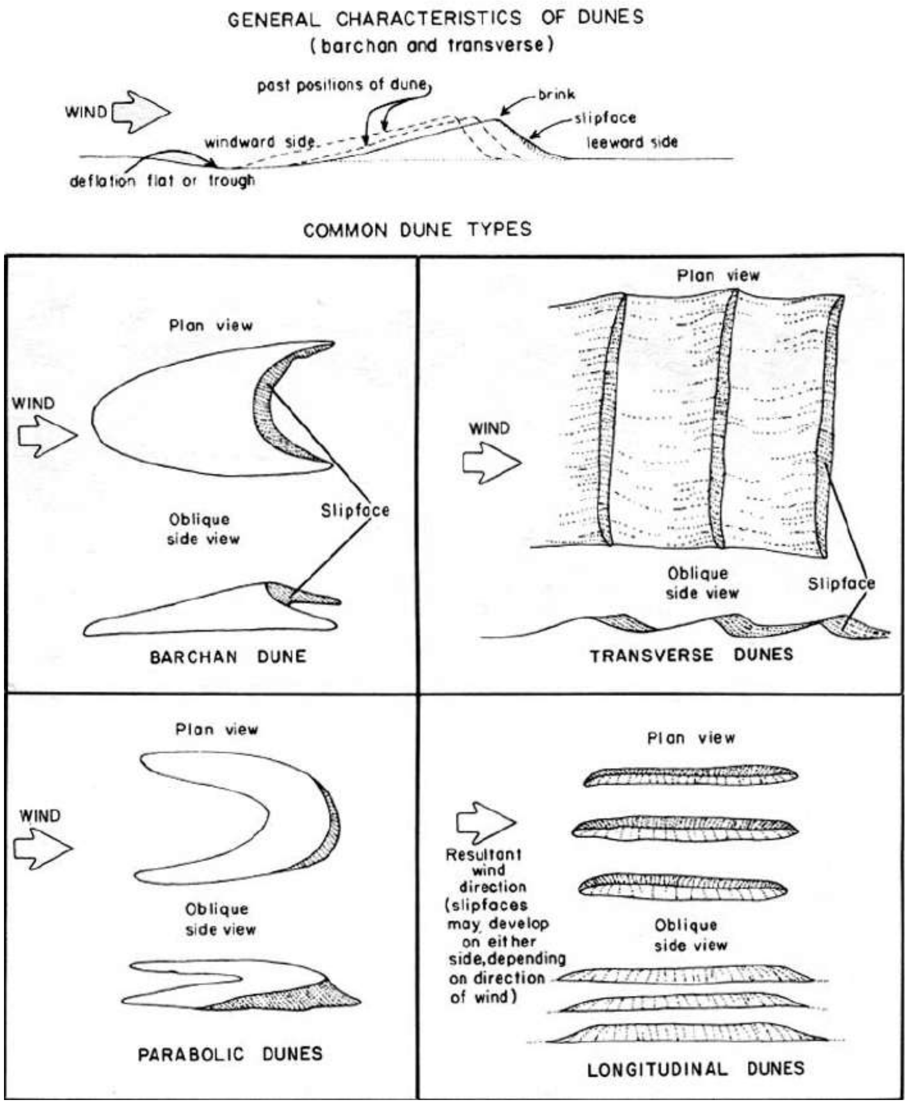


Figure 5.5.23: Formation and types of dunes.

Aeolian, sometimes spelled eolian or œolian, are deposits of windblown sediments. Since wind has a much lower carrying capacity than water, aeolian deposits typically consist of clast sizes from fine dust to sand [52]. Fine silt and clay can cross very long distances, even entire oceans suspended in the air.

With sufficient sediment influx, aeolian systems can potentially form large dunes in dry or wet conditions. The figure shows dune features and various types. British geologist Ralph A. Bagnold (1896-1990) considered only Barchan and linear Seif dunes as the only true dune forms. Other scientists recognize transverse, star, parabolic, and linear dune types. Parabolic and linear dunes grow from sand anchored by plants and are common in coastal areas.

Figure 5.5.24: Loess Plateau in China. The loess is so highly compacted that buildings and homes have been carved in it.

Compacted layers of wind-blown sediment are known as **loess**. Loess commonly starts as finely ground-up rock flour created by glaciers. Such deposits cover thousands of square miles in the Midwestern United States. Loess may also form in desert regions (see Chapter 13). Silt for the Loess Plateau in China came from the Gobi Desert in China and Mongolia.

Glacial

Figure 5.5.25: Wide range of sediments near Athabaska Glacier, Jasper National Park, Alberta, Canada.

Glacial sedimentation is very diverse and generally consists of the most poorly-sorted sediment deposits found in nature. The main clast type is called diamictite, which literally means two sizes, referring to the unsorted mix of large and small rock fragments found in glacial deposits [53]. Many glacial tills, glacially derived diamictites, include very finely-pulverized rock flour along with giant erratic boulders. The surfaces of larger clasts typically have striations from the rubbing, scraping, and polishing of surfaces by abrasion during the movement of glacial ice. Glacial systems are so large and produce so much sediment, they frequently create multiple, individualized depositional environments, such as fluvial, deltaic, lacustrine, pluvial, alluvial, and/or aeolian (see Chapter 14, Glaciers).

Facies

In addition to mineral composition and lithification process, geologists also classify sedimentary rock by its depositional characteristics, collectively called facies or lithofacies. Sedimentary facies consist of physical, chemical, and/or biological properties, including relative changes in these properties in adjacent beds of the same layer or geological age. Geologists analyze sedimentary rock facies to interpret the original deposition environment, as well as disruptive geological events that may have occurred after the rock layers were established.

It boggles the imagination to think of all the sedimentary deposition environments working next to each other, at the same time, in any particular region on Earth. The resulting sediment beds develop characteristics reflecting contemporaneous conditions at the time of deposition, which later may become preserved into the rock record. For example, in the Grand Canyon, rock strata of the same geologic age include many different depositional environments: beach sand, tidal flat silt, offshore mud, and farther offshore limestone. In other words, each sedimentary or stratigraphic facies presents recognizable characteristics that reflect specific, and different, depositional environments that were present at the same time.

Facies may also reflect depositional changes in the same location over time. During periods of rising sea level, called marine transgression, the shoreline moves inland as seawater covers what was originally dry land and creates new offshore depositional environments. When these sediment beds turn into sedimentary rock, the vertical stratigraphy sequence reveals beach lithofacies buried by offshore lithofacies.

Biological facies are remnants (coal, diatomaceous earth) or evidence (fossils) of living organisms. Index fossils, fossilized life forms specific to a particular environment and/or geologic time period, are an example of biological facies. The horizontal assemblage and vertical distribution of fossils are particularly useful for studying species evolution because transgression, deposition, burial, and compaction processes happen over a considerable geologic time range.

Fossil assemblages that show evolutionary changes greatly enhance our interpretation of Earth's ancient history by illustrating the correlation between stratigraphic sequence and geologic time scale. During the Middle Cambrian period (see Chapter 7, Geologic Time), regions around the Grand Canyon experienced marine transgression in a southeasterly direction (relative to current maps). This shift of the shoreline is reflected in the Tapeats Sandstone beach facies, Bright Angle Shale near-offshore mud facies, and Muav Limestone far-offshore facies. Marine organisms had plenty of time to evolve and adapt to their slowly changing environment; these changes are reflected in the biological facies, which show older life forms in the western regions of the canyon and younger life forms in the east.

References

- 38. Hill, M. N. *The Sea, Ideas and Observations on Progress in the Study of the Seas: Physical oceanography*. **1**, (Wiley-Interscience, 1962).
- 39. Stow, D. A. V. & Piper, D. J. W. Deep-water fine-grained sediments: facies models. *Geological Society, London, Special Publications* **15**, 611–646 (1984).
- 40. Normark, W. R. Fan valleys, channels, and depositional lobes on modern submarine fans: characters for recognition of sandy turbidite environments. *AAPG Bull.* **62**, 912–931 (1978).
- 41. Stow, D. A. V., Faugères, J.-C., Viana, A. & Gonthier, E. Fossil contourites: a critical review. *Sediment. Geol.* **115**, 3–31 (1998).
- 42. McCubbin, D. G. Barrier-island and strand-plain facies. (1982).
- 43. Reinson, G. E. Barrier-island and associated strand-plain systems. in *Facies Models* (ed. Walker, R. G.) 119–140 (Geoscience Canada Reprint Series 1, 1984).
- 44. Hubert, J. F. A zircon-tourmaline-rutile maturity index and the interdependence of the composition of heavy mineral assemblages with the gross composition and texture of sandstones. *J. Sediment. Res.* **32**, (1962).
- 45. Eisma, D. *Intertidal deposits: River mouths, tidal flats, and coastal lagoons*. (Taylor & Francis, 1998).
- 46. Longman, M. W. A process approach to recognizing facies of reef complexes. (1981).
- 47. Nichols, M. M., Biggs, R. B. & Davies, R. A. J. Estuaries. in *Coastal Sedimentary Environments* 77–173 (Springer-Verlag: New York, 1985).
- 48. Elliott, T. Deltas. *Sedimentary environments and facies* **2**, 113–154 (1986).
- 49. Cant, D. J. Fluvial facies models and their application. (1982).
- 50. Stanistreet, I. G. & McCarthy, T. S. The Okavango Fan and the classification of subaerial fan systems. *Sediment. Geol.* **85**, 115–133 (1993).
- 551. Sly, P. G. Sedimentary processes in lakes. in *Lakes* (ed. Lerman, A.) 65–89 (Springer New York, 1978). doi:10.1007/978-1-4757-1152-3_3
- 52. Bagnold, R. A. The physics of blown sand and desert dunes. *Methum, London, UK* 265 (1941).
- 53. Eyles, N., Eyles, C. & Miall, A. D. Glacial facies models. *Geology* **16**, 374–375 (1988).

Contributions and Attributions

5.S: Weathering, Erosion, and Sedimentary Rocks (Summary)

Sedimentary rocks are grouped into two main categories: clastic (detrital) and chemical. Clastic (detrital) rocks are made of mineral clasts or sediment that lithifies into solid material. Sediment is produced by the mechanical or chemical weathering of bedrock and transported away from the source via erosion. Sediment that is deposited, buried, compacted, and sometimes cemented becomes clastic rock. Clastic rocks are classified by grain size; for example sandstone is made of sand-sized particles. Chemical sedimentary rocks comes from minerals precipitated out an aqueous solution and is classified according to mineral composition. The chemical sedimentary rock limestone is made of calcium carbonate. Sedimentary structures have textures and shapes that give insight on depositional histories. Depositional environments depend mainly on fluid transport systems and encompass a wide variety of underwater and above ground conditions. Geologists analyze depositional conditions, sedimentary structures, and rock records to interpret the paleogeographic history of a region.

CHAPTER OVERVIEW

6: METAMORPHIC ROCKS

Metamorphic rocks is one of the three rock categories in the rock cycle. Metamorphic rock material has been changed by temperature, pressure, and/or fluids. The rock cycle shows that both igneous and sedimentary rocks can become metamorphic rocks. And metamorphic rocks themselves can be re-metamorphosed. Because metamorphism is caused by plate tectonic motion, metamorphic rock provides geologists with a history book of how past tectonic processes shaped our planet

6.1: Prelude to Metamorphic Rocks

Figure 6.1.1: Painted Wall of Black Canyon of the Gunnison National Park, Colorado, made of 1.7 billion-year-old gneiss intruded by younger pegmatites.

Metamorphic rocks, *meta-* meaning change and *–morphos* meaning form, is one of the three rock categories in the rock cycle (see Chapter 1). Metamorphic rock material has been changed by temperature, pressure, and/or fluids. The rock cycle shows that both igneous and sedimentary rocks can become metamorphic rocks. And metamorphic rocks themselves can be re-metamorphosed. Because metamorphism is caused by plate tectonic motion, metamorphic rock provides geologists with a history book of how past tectonic processes shaped our planet [1].

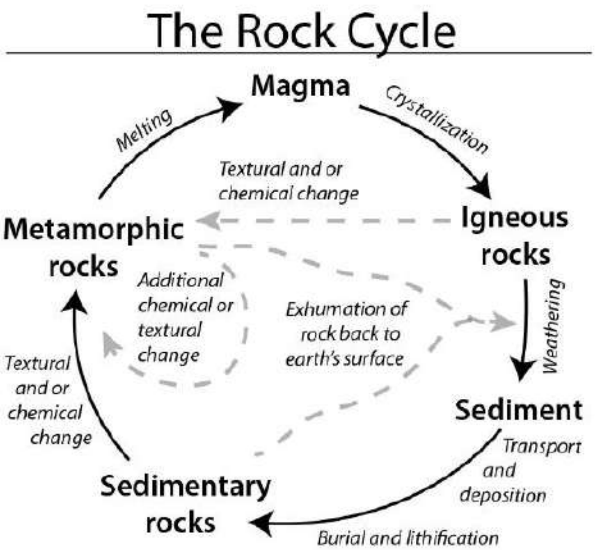


Figure 6.1.1: Rock cycle showing the five materials (such as igneous rocks and sediment) and the processes by which one changes into another (such as weathering). (Source: Peter Davis)

Reference

1. Yardley BWD (1989) An introduction to metamorphic petrology. Prentice Hall

6.2: Metamorphic Processes

Metamorphism occurs when solid rock changes in composition and/or texture without the mineral crystals melting, which is how igneous rock is generated. Metamorphic source rocks, the rocks that experience the metamorphism, are called the parent rock or **protolith**, from *proto*– meaning first, and *lithos*- meaning rock. Most metamorphic processes take place deep underground, inside the earth's crust. During metamorphism, protolith chemistry is mildly changed by increased temperature (heat), a type of pressure called confining pressure, and/or chemically reactive fluids. Rock texture is changed by heat, confining pressure, and a type of pressure called directed stress.

Temperature (Heat)

Temperature measures a substance's energy—an increase in temperature represents an increase in energy [2]. Temperature changes affect the chemical equilibrium or cation balance in minerals. At high temperatures, atoms may vibrate so vigorously they jump from one position to another within the crystal lattice, which remains intact. In other words, this atom swapping can happen while the rock is still solid.

The temperatures of metamorphic rock lie in between surficial processes (as in sedimentary rock) and magma in the rock cycle. Heat-driven metamorphism begins at temperatures as cold as 200˚C and can continue to occur at temperatures as high as 700°C-1,100°C [3; 4; 5]. Higher temperatures would create magma, and thus, would no longer be a metamorphic process. Temperature increases with increasing depth in the Earth along a geothermal gradient (see Chapter 4) and metamorphic rock records these depth-related temperature changes.

Pressure

Pressure is the force exerted over a unit area on a material. Like heat, pressure can affect the chemical equilibrium of minerals in a rock. The pressure that affects metamorphic rocks can be grouped into confining pressure and directed stress. **Stress** is a scientific term indicating a force. **Strain** is the result of this stress, including metamorphic changes within minerals.

Confining Pressure

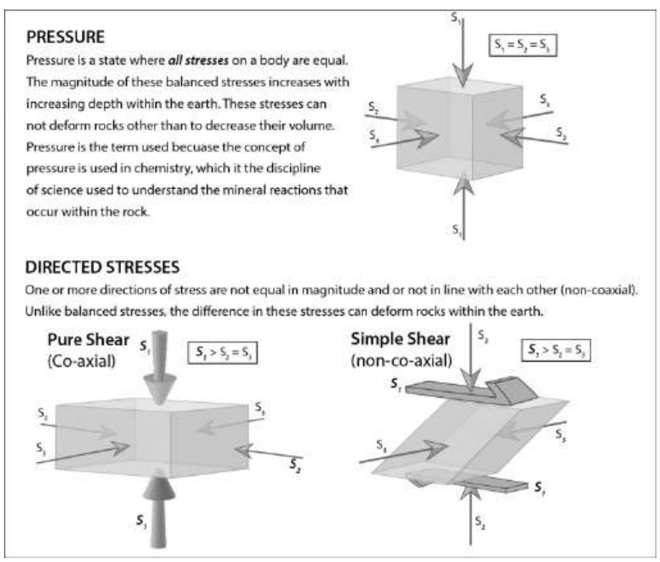


Figure 6.2.1: Difference between pressure and stress and how they deform rocks. Pressure (or confining pressure) has equal stress (forces) in all directions and increases with depth under the Earth's surface. Under directed stress, some stress directions (forces) are stronger than others, and this can deform rocks. Source: Peter Davis

Pressure exerted on rocks under the surface is due to the simple fact that rocks lie on top of one another. When pressure is exerted from rocks above, it is balanced from below and sides and is called **confining** or **lithostatic pressure**. Confining pressure has equal pressure on all sides (see figure) and is responsible for causing chemical reactions to occur, just like heat. These chemical reactions will cause new minerals to form.

Confining pressure is measured in bars and ranges from 1 bar at sea level to around 10,000 bars at the base of the crust [6]. For metamorphic rocks, pressures range from a relatively low-pressure of 3,000 bars [5] around 50,000 bars [5], which occurs around 15-35 kilometers below the surface.

Directed Stress

Figure 6.2.2: Pebbles (that used to be spherical or close to spherical) in quartzite deformed by directed stress

Directed stress, also called differential or tectonic stress, is an unequal balance of forces on a rock in one or more directions (see the previous figure). Directed stresses are generated by the movement of lithospheric plates. Stress indicates a type of force acting on the rock. Strain describes the resultant processes caused by stress and includes metamorphic changes in the minerals. In contrast to confining pressure, directed stress occurs at much lower pressures and does not generate chemical reactions that change the mineral composition and atomic structure [3]. Instead, directed stress modifies the parent rock at a mechanical level, changing the arrangement, size, and/or shape of the mineral crystals. These crystalline changes create identifying textures, which is shown in the figure below comparing the phaneritic texture of igneous granite with the foliated texture of metamorphic gneiss.

Figure 6.2.3: An igneous rock granite (left) and foliated high-temperature and high-pressure metamorphic rock gneiss (right) illustrating a metamorphic texture. (Source: Peter Davis)

Directed stresses produce rock textures in many ways. Crystals are rotated, changing their orientation in space. Crystals can get fractured, reducing their grain size. Conversely, they may grow larger as atoms migrate. Crystal shapes also become deformed. These mechanical changes occur via **recrystallization**, which is when minerals dissolve from an area of rock experiencing high stress and precipitate or regrow in a location having lower stress. For example, recrystallization increases grain size much like adjacent soap bubbles coalesce to form larger ones. Recrystallization rearranges mineral crystals without fracturing the rock structure, deforming the rock-like silly putty; these changes provide important clues to understanding the creation and movement of deep underground rock faults.

Fluids

A third metamorphic agent is chemically reactive fluids that are expelled by crystallizing magma and created by metamorphic reactions. These reactive fluids are made of mostly water (H_2O) and carbon dioxide (CO_2), and smaller amounts of potassium (K), sodium (Na), iron (Fe), magnesium (Mg), calcium (Ca), and aluminum (Al). These fluids react with minerals in the protolith, changing its chemical equilibrium and mineral composition, in a process similar to the reactions driven by heat and pressure. In addition to using elements found in the protolith, the chemical reaction may incorporate substances contributed by the fluids to create new minerals. In general, this style of metamorphism, in which fluids play an important role, is called **hydrothermal metamorphism** or hydrothermal alteration. Water actively participates in chemical reactions and allows extra mobility of the components in hydrothermal alteration.

Fluids-activated metamorphism is frequently involved in creating economically important mineral deposits that are located next to igneous intrusions of magma bodies. For example, the mining districts in the Cottonwood Canyons and Mineral Basin of northern

Utah produce valuable ores such as argentite (silver sulfide), galena (lead sulfide), and chalcopyrite (copper-iron sulfide), as well as the native element gold [2]. These mineral deposits were created from the interaction between a granitic intrusion called the Little Cottonwood Stock and country-rock consisting of mostly limestone and dolostone. Hot, circulating fluids expelled by the crystallizing granite reacted with and dissolved the surrounding limestone and dolostone, precipitating out new minerals created by the chemical reaction. Hydrothermal alteration of mafic mantle rock, such as olivine and basalt, creates the metamorphic rock **serpentinite**, a member of the serpentine subgroup of minerals. This metamorphic process happens at mid-ocean spreading centers where newly formed oceanic crust interacts with seawater.

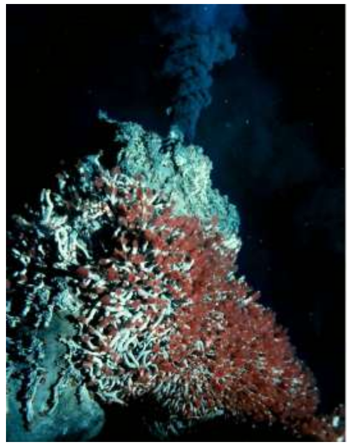


Figure 6.2.4: Black smoker hydrothermal vent with a colony of giant (6’+) tube worms.

Some hydrothermal alterations remove elements from the parent rock rather than deposit them. This happens when seawater circulates down through fractures in the fresh, still-hot basalt, reacting with and removing mineral ions from it. The dissolved minerals are usually ions that do not fit snugly in the silicate crystal structure, such as copper. The mineral-laden water emerges from the seafloor via hydrothermal vents called **black smokers**, named after the dark-colored precipitates produced when the hot vent water meets cold seawater. (see Chapter 4, Igneous Rock and Volcanic Processes) Ancient black smokers were an important source of copper ore for the inhabitants of Cyprus (Cypriots) as early as 4,000 BCE, and later by the Romans [8].

References

- 2. Jeong I-K, Heffner RH, Graf MJ, Billinge SJL (2003) Lattice dynamics and correlated atomic motion from the atomic pair distribution function. Phys Rev B Condens Matter 67:104301
- 3. Winter JD (2009) Principles of igneous and metamorphic petrology, 2 edition. Pearson, New York
- 4. Boehler R (1996) Melting temperatures of the Earth’s mantle and core: Earth’s thermal structure. Annual Review of Earth and Planetary Sciences 24:15–40. https://doi.org/10.1146/annurev.earth.24.1.15
- 5. Issitt ML (2012) Ultra-High-Pressure Metamorphism. Earth Science: Earth Materials & Resources 523
- 6. Petrini, Podladchikov (2000) Lithospheric pressure–depth relationship in compressive regions of thickened crust. Journal of Metamorphic Geology 18:67–77. https://doi.org/10.1046/j.1525-1314.2000.00240.x
- 8. (2007) Timeline of Art History. Reference Reviews 21:45–45

Contributions and Attributions

6.3: Metamorphic Textures

Metamorphic texture is the description of the shape and orientation of mineral grains in a metamorphic rock. Metamorphic rock textures are foliated, non-foliated, or lineated are described below.

METAMORPHIC ROCK IDENTIFICATION

1 IDENTIFY ROCK'S FOLIATION		2 TEXTURAL FEATURES	3 MINERAL COMPOSITION	4 ROCK NAME	5 PARENT ROCK
FOLIATED *layered texture*	Fine-grained or no visible grains	Flat, slaty cleavage is well developed. Dense, microscopic grains, may exhibit slight sheen (or dull luster). Clanky sound when struck. Breaks into hard, flat sheets.	Fine, microscopic clay or mica.	SLATE	Shale
		Finely crystalline; micas hardly discernible, but impart a sheen or luster. Breaks along wavy surfaces.	Dark silicates and micas.	PHYLLITE	Siltstone or shale
	Medium- to coarse-grained	Schistose texture. Foliation formed by alignment of visible crystals. Rock breaks along scaly foliation surfaces. Medium to fine-grained. Sparkling appearance.	Common minerals include chlorite, biotite, muscovite, garnet and hornblende. Recognizable minerals used as part of rock name. Porphyroblasts common.	MICA SCHIST	Siltstone or shale
				GARNET SCHIST	
		Gneissic banding. Coarse-grained. Foliation present as minerals arranged into alternating light and dark layers giving the rock a banded texture in side view. Crystalline texture. No cleavage.	Light-colored quartz and feldspar; dark ferromagnesian minerals.	GNEISS	Shale or granitic rocks
NON-FOLIATED *no layered texture*	Fine-grained or no visible grains	Medium- to coarse-grained crystalline texture.	Crystals of amphibole (hornblende) in blade-like crystals.	AMPHIBOLITE	Basalt, gabbro, or ultramafic igneous rocks
		Microcrystalline texture. Glassy black sheen. Conchoidal fracture. Low density.	Fine, tar-like, organic makeup.	ANTHRACITE COAL	Coal
		Dense and dark-colored. Fine or microcrystalline texture. Very hard. Color can range from gray, gray-green to black.	Microscopic dark silicates.	HORNFELS	Many rock types
		Microcrystalline or no visible grains with smooth, wavy surfaces. May be dull or glossy. Usually shades of green.	Serpentine. May have fibrous asbestos visible.	SERPENTINITE	Ultramafic igneous rocks or peridotite
	Fine- to coarse-grained	Microcrystalline or no visible grains. Can be scratched with a fingernail. Shades of green, gray, brown or white. Soapy feel.	Talc.	SOAPSTONE OR TALC SCHIST	Ultramafic igneous rocks
		Crystalline. Hard (scratches glass). Breaks across grains. Sandy or sugary texture. Color variable; can be white, pink, buff, brown, red, purple.	Quartz grains fused together. Grains will not rub off like sandstone.	QUARTZITE	Quartz sandstone
		Finely crystalline (resembling a sugar cube) to medium or coarse texture. Color variable; white, pink, gray, among others. Fossils in some varieties.	Calcite or dolomite crystals tightly fused together. Calcite effervesces with HCl; dolomite effervesces only when powdered.	MARBLE	Limestone or dolostone
		Texture of conglomerate, but breaks across clasts as easily as around them. Pebbles may be stretched (lineated) or cut by rock cleavage	Granules or pebbles are commonly granitic or jasper, chert, quartz or quartzite.	META-CONGLOMERATE	Conglomerate

Created by Belinda C. Madsen for Salt Lake Community College

Metamorphic rock identification table. (Source: Belinda Madsen)

Foliation and Lineation

Foliation is a term used that describes minerals lined up in planes. Certain minerals, most notably the mica group, are mostly thin and planar by default. Foliated rocks typically appear as if the minerals are stacked like pages of a book, thus the use of the term 'folia', like a leaf. Other minerals, with hornblende being a good example, are longer in one direction, linear like a pencil or a needle, rather than a planar-shaped book. These linear objects can also be aligned within a rock. This is referred to as a **lineation**. Linear crystals, such as hornblende, tourmaline, or stretched quartz grains, can be arranged as part of a foliation, a lineation, or foliation/lineation together. If they lie on a plane with mica, but with no common or preferred direction, this is foliation. If the minerals line up and point in a common direction, but with no planar fabric, this is lineation. When minerals lie on a plane AND point in a common direction; this is both foliation and lineation.

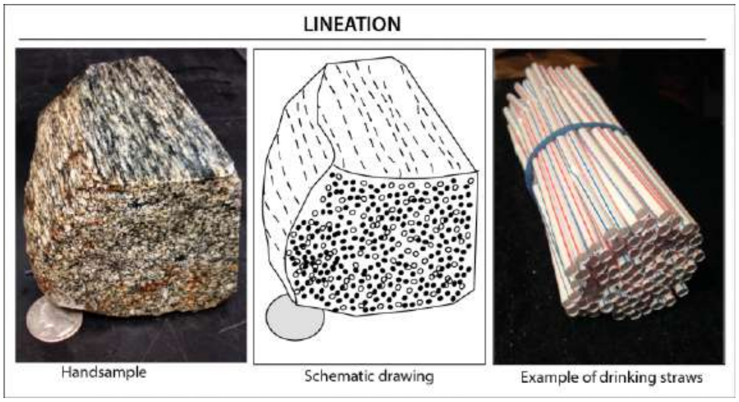


Figure 6.3.1: Example of lineation where minerals are aligned like a stack of straws or pencils. (Source: Peter Davis)

Figure 6.3.2: An example of foliation WITH lineation. (Source: Peter Davis)

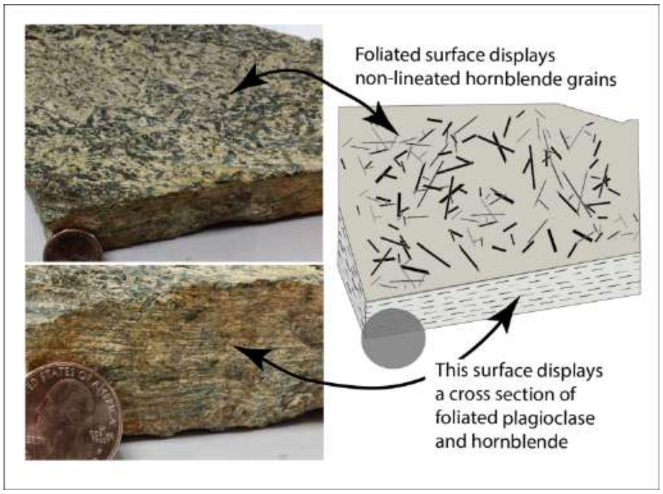


Figure 6.3.3: An example of foliation WITHOUT lineation. (Source: Peter Davis)

Foliated metamorphic rocks are named based on the style of their foliations. Each rock name has a specific texture that defines and distinguishes it, with their descriptions listed below.

Slate is a fine-grained metamorphic rock that exhibits a foliation called **slaty cleavage** that is the flat orientation of the small platy crystals of mica and chlorite forming perpendicular to the direction of stress. The minerals in slate are too small to see with the unaided eye. The thin layers in slate may resemble sedimentary bedding, but they are a result of directed stress and may lie at angles to the original strata. In fact, original sedimentary layering may be partially or completely obscured by the foliation. Thin slabs of slate are often used as a building material for roofs and tiles.

Figure 6.3.4: Slate mine in Germany cleavage.

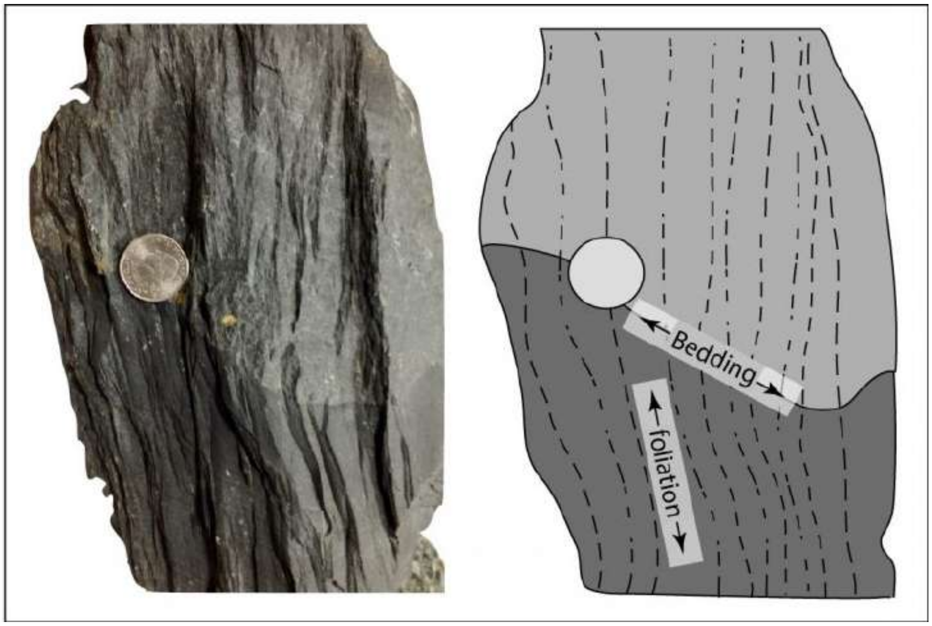


Figure 6.3.5: Foliation vs. bedding. Foliation is caused by metamorphism. Bedding is a result of sedimentary processes. They do not have to align. (Source: Peter Davis)

Figure 6.3.6: Phyllite with a small fold. (Source: Peter Davis)

Phyllite is a foliated metamorphic rock in which platy minerals have grown larger and the surface of the foliation shows a sheen from light reflecting from the grains, perhaps even a wavy appearance, called crenulations. Similar to phyllite but with even larger grains is the foliated metamorphic rock **schist**, which has large platy grains visible as individual crystals. Common minerals are muscovite, biotite, and porphyroblasts of garnets. A porphyroblast is a large crystal of a particular mineral surrounded by small grains. **Schistosity** is a textural description of foliation created by the parallel alignment of platy visible grains. Some schists are named for their minerals such as mica schist (mostly micas), garnet schist (mica schist with garnets), and staurolite schist (mica schists with staurolite).

Figure 6.3.7: Schist

Figure 6.3.8: Garnet staurolite muscovite schist.

Figure 6.3.9: Gneiss

Gneissic banding is a metamorphic foliation in which visible silicate minerals separate into dark and light bands or lineations. These grains tend to be coarse and often folded. A rock with this texture is called **gneiss**. Since gneisses form at the highest temperatures and pressures, some partial melting may occur. This partially melted rock is a transition between metamorphic and igneous rocks called a **migmatite** [9].

Figure 6.3.10: Migmatite

Migmatites appear as dark and light banded gneiss that may be swirled or twisted some since some minerals started to melt. Thin accumulations of light-colored rock layers can occur in a darker rock that is parallel to each other or even cut across the gneissic foliation. The lighter colored layers are interpreted to be the result of the separation of a felsic igneous melt from the adjacent highly metamorphosed darker layers, or injection of a felsic melt from some distance away.

Non-foliated

Non-foliated textures do not have lineations, foliations, or other alignments of mineral grains. Non-foliated metamorphic rocks are typically composed of just one mineral and, therefore, usually show the effects of metamorphism with recrystallization in which crystals grow together, but with no preferred direction. The two most common examples of non-foliated rocks are quartzite and marble. **Quartzite** is a metamorphic rock from the protolith sandstone. In quartzites, the quartz grains from the original sandstone are enlarged and interlocked by recrystallization. A defining characteristic for distinguishing quartzite from sandstone is that when broken with a rock hammer, the quartz crystals break across the grains. In a sandstone, only a thin mineral cement holds the grains together, meaning that a broken piece of sandstone will leave the grains intact. Because most sandstones are rich in quartz, and quartz is a mechanically and chemically durable substance, quartzite is very hard and resistant to weathering.

Figure 6.3.11: Marble Baraboo Quartzite

Marble is metamorphosed limestone (or dolostone) composed of calcite (or dolomite). Recrystallization typically generates larger interlocking crystals of calcite or dolomite. Marble and quartzite often look similar, but these minerals are considerably softer than quartz. Another way to distinguish marble from a quartzite is with a drop of dilute hydrochloric acid. Marble will effervesce (fizz) if it is made of calcite.

A third non-foliated rock is **hornfels** identified by its dense, fine-grained, hard, blocky or splintery texture composed of several silicate minerals. Crystals in hornfels grow smaller with metamorphism and become so small that specialized study is required to identify them. These are common around intrusive igneous bodies and are hard to identify. The protolith of hornfels can be even harder to distinguish, which can be anything from mudstone to basalt.

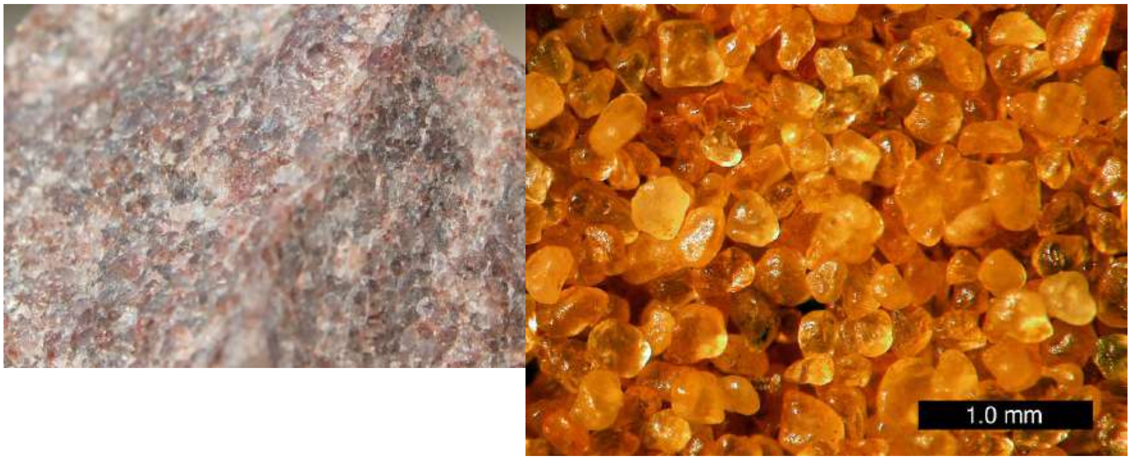


Figure 6.3.12: (left) Macro view of quartzite. Note the interconnectedness of the grains. (right) Unmetamorphosed, unconsolidated sand grains have space between the grains

References

- 9. Sawyer EW (2008) Atlas of migmatites. NRC Research Press and Mineralogical Association of Canada, Canada

Contributions and Attributions

6.4: Metamorphic Grade

Metamorphic grade refers to the range of metamorphic change a rock undergoes, progressing from low (little metamorphic change) grade to high (significant metamorphic change) grade. Low-grade metamorphism begins at temperatures and pressures just above sedimentary rock conditions. The sequence **slate** → **phyllite** → **schist** → **gneiss** illustrates an increasing metamorphic grade.

Figure 6.4.1: Garnet schist.

Geologists use **index minerals** that form at certain temperatures and pressures to identify metamorphic grade. These index minerals also provide important clues to a rock's sedimentary protolith and the metamorphic conditions that created it. Chlorite, muscovite, biotite, garnet, and staurolite are index minerals representing a respective sequence of low-to-high grade rock. The figure shows a **phase diagram** of three index minerals—sillimanite, kyanite, and andalusite—with the same chemical formula (Al_2SiO_5) but having different crystal structures (**polymorphism**) created by different pressure and temperature conditions.

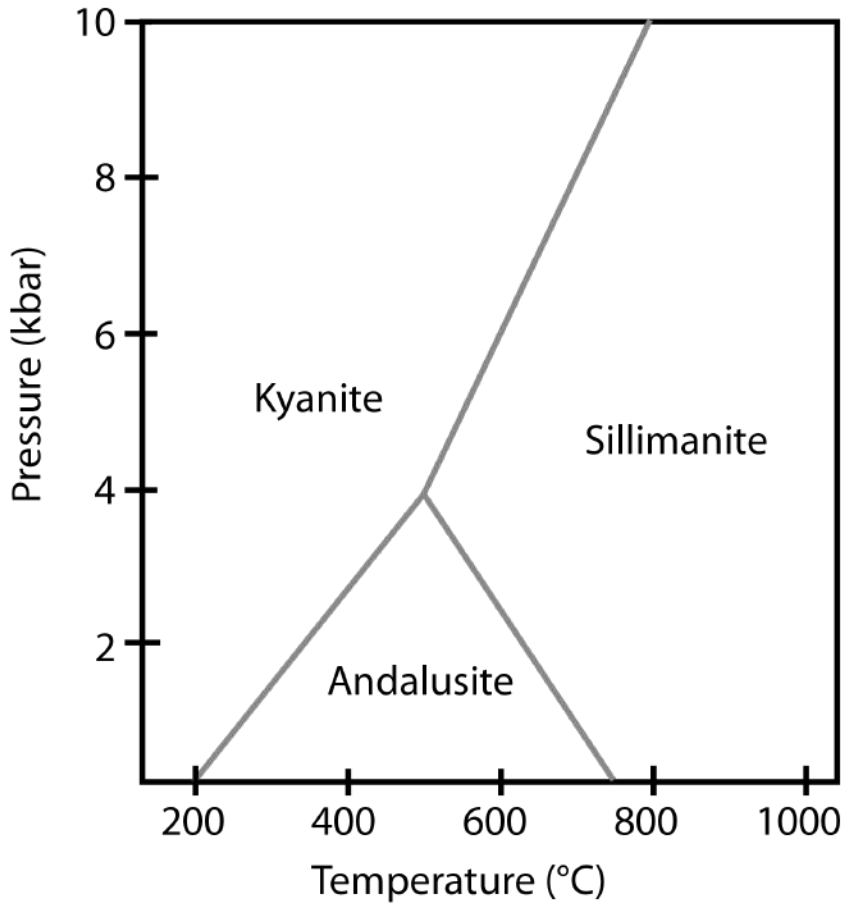


Figure 6.4.1: 

Andalusite

Figure 6.4.1: Kyanite

Figure 6.4.1: Sillimanite

Some metamorphic rocks are named based on the highest grade of index mineral present. Chlorite schist includes the low-grade index mineral chlorite. Muscovite schist contains a slightly higher grade muscovite, indicating a greater degree of metamorphism. Garnet schist includes the high-grade index mineral garnet and indicating it has experienced much higher pressures and temperatures than chlorite.

Contributions and Attributions

This page is licensed under a CC BY-NC-SA license and was authored, remixed, and/or curated by Chris Johnson, Matthew D. Affolter, Paul Inkenbrandt, & Cam Mosher. A detailed versioning history of the edits to source content is available upon request.

6.5: Metamorphic Environments

As with igneous processes, metamorphic rocks form at different zones of pressure (depth) and temperature as shown on the pressure-temperature (P-T) diagram. The term **facies** is an objective description of a rock. In metamorphic rocks, facies are groups of minerals called mineral assemblages. The names of **metamorphic facies** on the pressure-temperature diagram reflect minerals and mineral assemblages that are stable at these pressures and temperatures and provide information about the metamorphic processes that have affected the rocks. This is useful when interpreting the history of metamorphic rock.

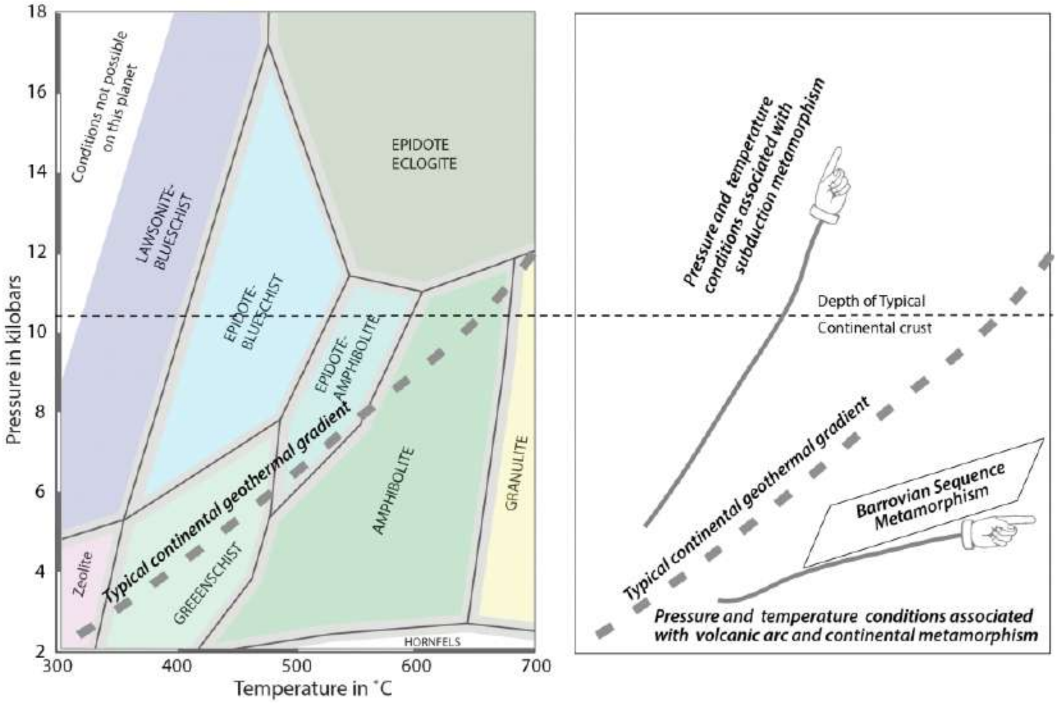


Figure 6.5.1: Pressure-temperature graph diagram showing metamorphic zones or facies.

In the late 1800s, British geologist George Barrow mapped zones of index minerals in different metamorphic zones of an area that underwent regional metamorphism. Barrow outlined a progression of index minerals, named the Barrovian Sequence, that represents increasing metamorphic grade: **chlorite (slates and phyllites) -> biotite (phyllites and schists) -> garnet (schists) -> staurolite (schists) -> kyanite (schists) -> sillimanite (schists and gneisses).**

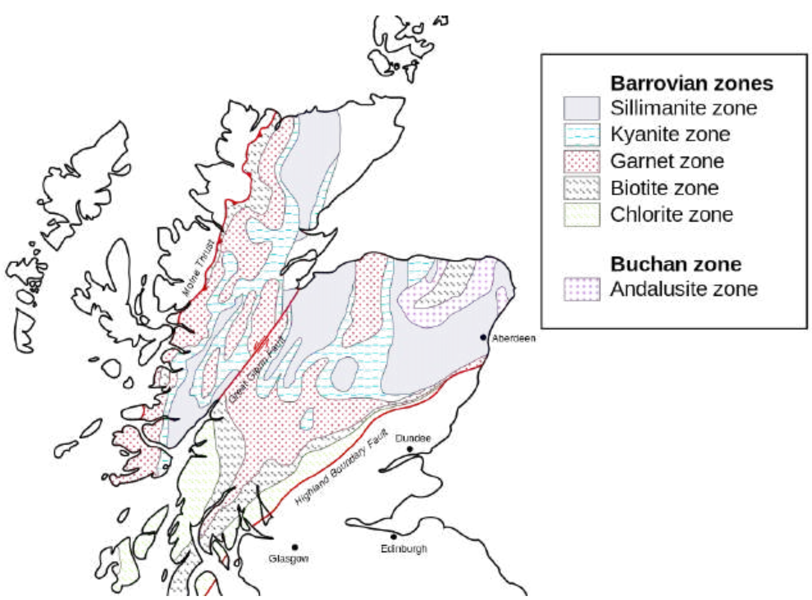


Figure 6.5.1: Barrovian sequence in Scotland.

The first of the Barrovian sequence has a mineral group that is commonly found in the metamorphic greenschist facies. Greenschist rocks form under relatively low pressure and temperatures and represent the fringes of regional metamorphism. The “green” part of the name is derived from green minerals like chlorite, serpentine, and epidote, and the “schist” part is applied due to the presence of platy minerals such as muscovite.

Many different styles of metamorphic facies are recognized, tied to different geologic and tectonic processes. Recognizing these facies is the most direct way to interpret the metamorphic history of rock. A simplified list of major metamorphic facies is given below.

Burial Metamorphism

Burial metamorphism occurs when rocks are deeply buried, at depths of more than 2000 meters (1.24 miles) [10]. Burial metamorphism commonly occurs in sedimentary basins, where rocks are buried deeply by overlying sediments. As an extension of diagenesis, a process that occurs during lithification (Chapter 5), burial metamorphism can cause clay minerals, such as smectite, in shales to change to another clay mineral illite. Or it can cause quartz sandstone to metamorphose into the quartzite such the Big Cottonwood Formation in the Wasatch Range of Utah. This formation was deposited as ancient near-shore sands in the late Proterozoic (see Chapter 7), deeply buried and metamorphosed to quartzite, folded, and later exposed at the surface in the Wasatch Range today. Increase of temperature with depth in combination and an increase of confining pressure produces low-grade metamorphic rocks with mineral assemblages are indicative of a zeolite facies [11; 12].

Contact Metamorphism

Contact metamorphism occurs in rock exposed to high temperature and low pressure, as might happen when hot magma intrudes into or lava flows over pre-existing protolith. This combination of high temperature and low pressure produces numerous metamorphic facies. The lowest pressure conditions produce hornfels facies, while higher pressure creates greenschist, amphibolite, or granulite facies.

As with all metamorphic rock, the parent rock texture and chemistry are major factors in determining the final outcome of the metamorphic process, including what index minerals are present. Fine-grained shale and basalt, which happen to be chemically similar, characteristically recrystallize to produce hornfels. Sandstone (silica) surrounding an igneous intrusion becomes quartzite via contact metamorphism, and limestone (carbonate) becomes marble.

Figure 6.5.1: Contact metamorphism in outcrop.

When contact metamorphism occurs deeper in the Earth, metamorphism can be seen as rings of facies around the intrusion, resulting in **aureoles**. These differences in metamorphism appear as distinct bands surrounding the intrusion, as can be seen around the Alta Stock in Little Cottonwood Canyon, Utah. The Alta Stock is a granite intrusion surrounded first by rings of the index minerals amphibole (tremolite) and olivine (forsterite), with a ring of talc (dolostone) located further away [13; 14].

Regional Metamorphism

Regional metamorphism occurs when parent rock is subjected to increased temperature and pressure over a large area and is often located in mountain ranges created by converging continental crustal plates. This is the setting for the Barrovian sequence of rock facies, with the lowest grade of metamorphism occurring on the flanks of the mountains and highest grade near the core of the mountain range, closest to the convergent boundary.

An example of an old regional metamorphic environment is visible in the northern Appalachian Mountains while driving east from New York state through Vermont and into New Hampshire. Along this route, the degree of metamorphism gradually increases from sedimentary parent rock to low-grade metamorphic rock, then higher-grade metamorphic rock, and eventually the igneous core.

The rock sequence is sedimentary rock, slate, phyllite, schist, gneiss, migmatite, and granite. In fact, New Hampshire is nicknamed the Granite State. The reverse sequence can be seen heading east, from eastern New Hampshire to the coast [15].

Subduction Zone Metamorphism

Figure 6.5.1: Blueschist

Subduction zone metamorphism is a type of regional metamorphism that occurs when a slab of oceanic crust is subducted under continental crust (see Chapter 2). Because rock is a good insulator, the temperature of the descending oceanic slab increases slowly relative to the more rapidly increasing pressure, creating a metamorphic environment of high pressure and low temperature. Glaucophane, which has a distinctive blue color, is an index mineral found in blueschist facies (see metamorphic facies diagram). The California Coast Range near San Francisco has blueschist-facies rocks created by subduction-zone metamorphism, which include rocks made of blueschist, greenstone, and red chert. Greenstone, which is metamorphized basalt, gets its color from the index mineral chlorite [16].

Fault Metamorphism

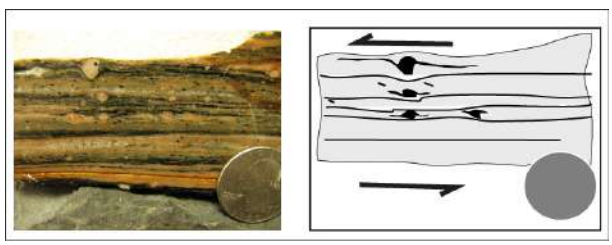


Figure 6.5.1: Mylonite

There are a range of metamorphic rocks made along faults. Near the surface, rocks are involved in repeated brittle faulting produce a material called *rock flour,* which is rock ground up to the particle size of flour used for food. At lower depths, faulting create **cataclastites** [17], chaotically-crushed mixes of rock material with little internal texture. At depths below cataclasites, where strain becomes ductile, mylonites are formed. **Mylonites** are metamorphic rocks created by dynamic recrystallization through directed shear forces, generally resulting in a reduction of grain size [18]. When larger, stronger crystals (like feldspar, quartz, garnet) embedded in a metamorphic matrix are sheared into an asymmetrical eye-shaped crystal, an **augen** is formed [19; 20].

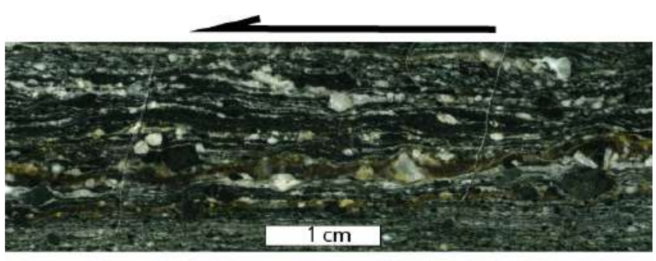


Figure 6.5.1: Examples of augens.

Shock Metamorphism

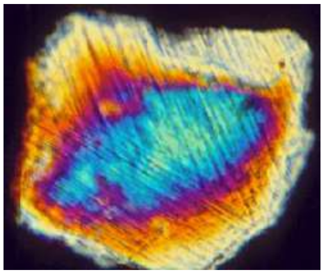


Figure 6.5.1: Shock lamellae in a quartz grain.

Shock (also known as impact) **metamorphism** is metamorphism resulting from meteor or other bolide impacts, or from a similar high-pressure shock event. Shock metamorphism is the result of very high pressures (and higher, but less extreme temperatures) delivered relatively rapidly. Shock metamorphism produces planar deformation features, tektites, shatter cones, and quartz polymorphs. Shock metamorphism produces planar deformation features (shock laminae), which are narrow planes of glassy material with distinct orientations found in silicate mineral grains. Shocked quartz has planar deformation features [21].

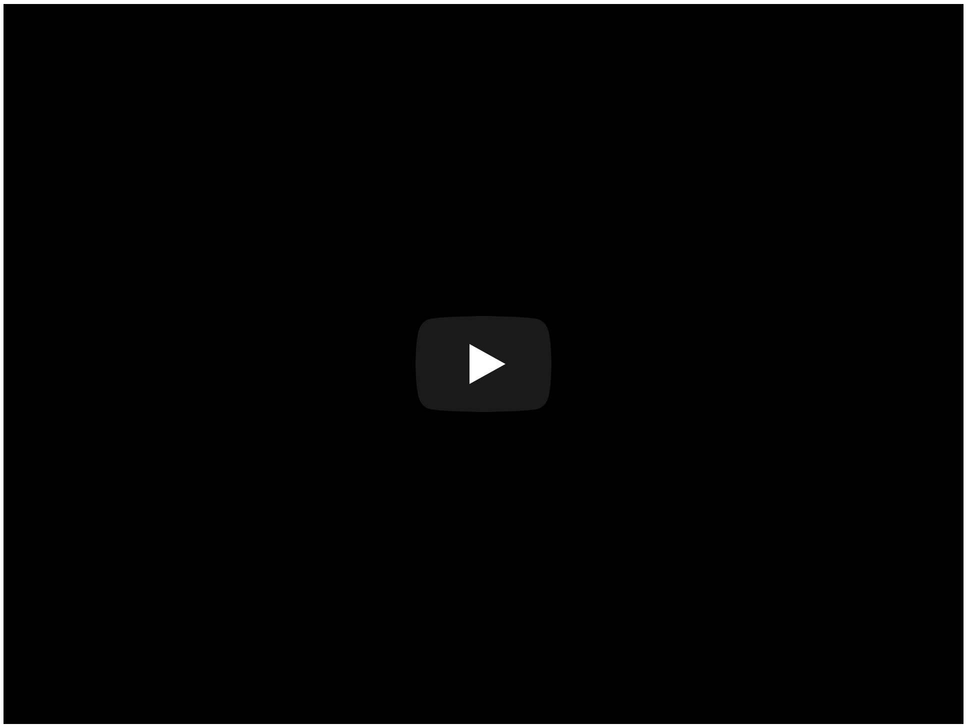
Figure 6.5.1: Shatter cone.

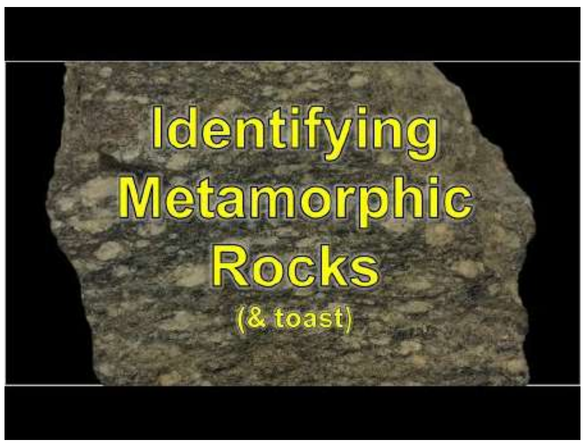
Shatter cones are cone-shaped pieces of rock created by dynamic branching fractures caused by impacts [22]. While not strictly a metamorphic structure, they are common around shock metamorphism. Their diameter can range from microscopic to several meters. Fine-grained rocks with shatter cones show a distinctive horsetail pattern.

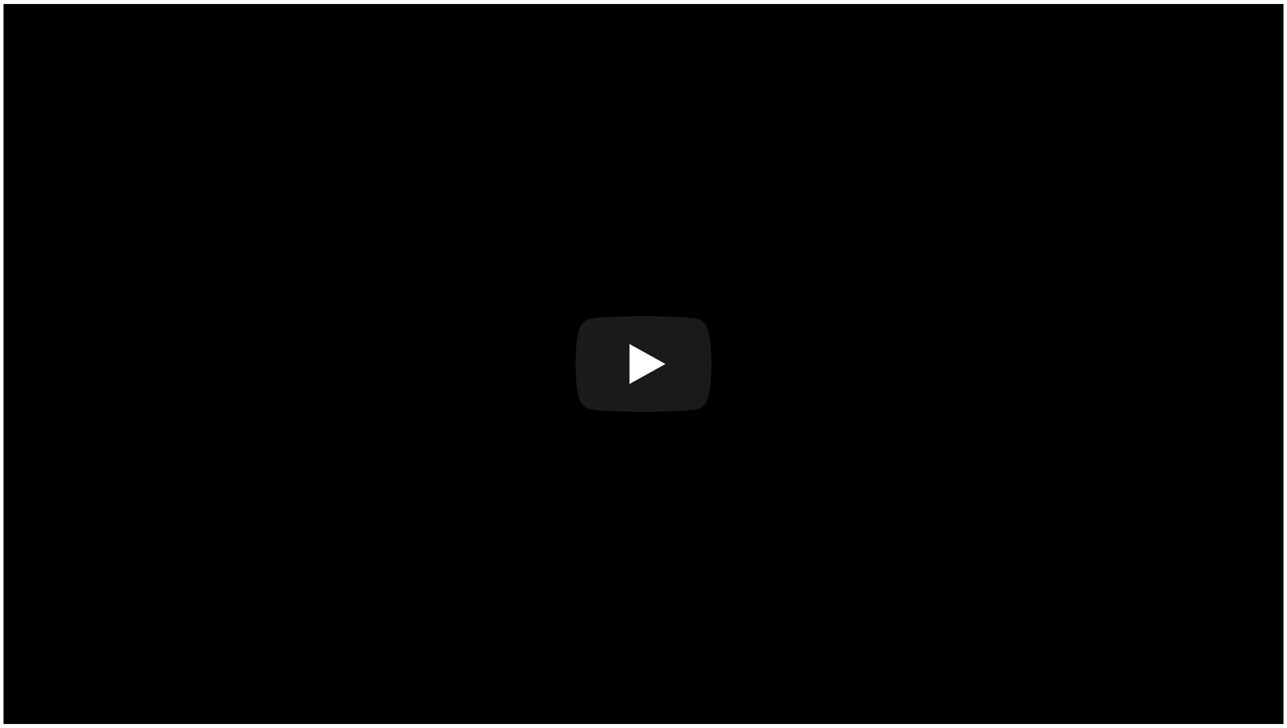
Shock metamorphism can also produce index minerals, though they are typically only found via microscopic analysis. The quartz polymorphs coesite and stishovite are indicative of impact metamorphism [21]. As discussed in chapter 3, polymorphs are minerals with the same composition but different crystal structures. Intense pressure (> 10 GPa) and moderate to high temperatures (700-1200 °C) are required to form these minerals.

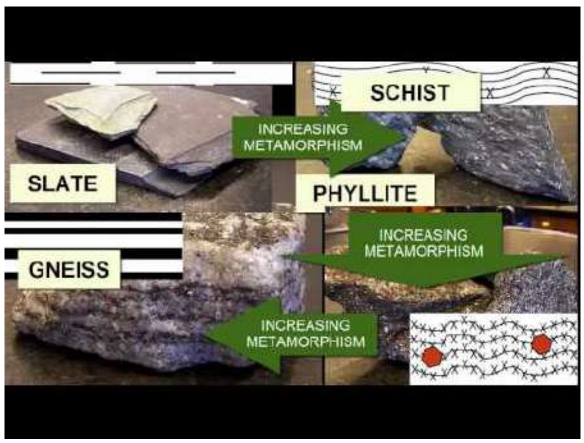
Figure 6.5.1: Tektites

Shock metamorphism can also produce glass. **Tektites** are gravel-size glass grains ejected during an impact event. They resemble volcanic glass but, unlike volcanic glass, tektites contain no water or phenocrysts, and have a different bulk and isotopic chemistry. Tektites contain partially melted inclusions of shocked mineral grains [23]. Although all are melt glasses, tektites are also chemically distinct from trinitite, which is produced from thermonuclear detonations [24], and fulgurites, which are produced by lightning strikes [25]. All geologic glasses not derived from volcanoes can be called with the general term pseudotachylyte [26], a name that can also be applied to glasses created by faulting. The term pseudo in this context means ‘false’ or ‘in the appearance of’, a volcanic rock called tachylite because the material observed looks like volcanic rock, but is produced by significant shear heating.




Identifying
Metamorphic
Rocks
(& toast)





References

10. Hower J, Eslinger EV, Hower ME, Perry EA (1976) Mechanism of burial metamorphism of argillaceous sediment: 1. Mineralogical and chemical evidence. Geol Soc Am Bull 87:725–737

11. ChO M, Liou JG, Maruyama S (1986) Transition from the zeolite to prehnite-pumpellyite facies in the Karmutsen metabasites, Vancouver Island, British Columbia. J Petrol

12. Coombs DS, Horodyski RJ, Naylor RS (1970) Occurrence of prehnite-pumpellyite facies metamorphism in northern Maine. Am J Sci 268:142–156

13. Cook SJ, Bowman JR (2000) Mineralogical evidence for fluid–rock interaction accompanying prograde contact metamorphism of siliceous dolomites: Alta Stock Aureole, Utah, USA. J Petrol 41:739–757

14. Moore JN, Kerrick DM (1976) Equilibria in siliceous dolomites of the Alta aureole, Utah. Am J Sci 276:502–524

15. Proctor BP, McAleer R, Kunk MJ, Wintsch RP (2013) Post-Taconic tilting and Acadian structural overprint of the classic Barrovian metamorphic gradient in Dutchess County, New York. Am J Sci 313:649–682

16. Wahrhaftig C (1984) A streetcar to subduction and other plate tectonic trips by public transport in San Francisco. American Geophysical Union

17. Brodie K, Fettes D, Harte B, Schmid R (2007) Recommendations by the IUGS Subcommission on the Systematics of Metamorphic Rocks. https://www.bgs.ac.uk/ scmr/docs/papers/paper_3.pdf

18. Trouw RAJ, Passchier CW, Wiersma DJ (2009) Atlas of mylonites- and related microstructures: Springer Berlin Heidelberg

19. Davis RA Jr, Fitzgerald DM (2009) Beaches and coasts. John Wiley & Sons

20. Simpson C, Schmid SM (1983) An evaluation of criteria to deduce the sense of movement in sheared rocks. Geol Soc Am Bull 94:1281–1288

21. Goltrant O, Leroux H, Doukhan J-C, Cordier P (1992) Formation mechanisms of planar deformation features in naturally shocked quartz. Physics of the Earth and Planetary Interiors 74:219–240. https://doi.org/10.1016/0031-9201(92)90012-K

22. Sagy A, Fineberg J, Reches Z (2004) Shatter cones: Branched, rapid fractures formed by shock impact. J Geophys Res 109:B10209

23. French BM (1998) Traces of Catastrophe: A Handbook of Shock-Metamorphic Effects in Terrestrial Meteorite Impact Structures

24. Eby N, Hermes R, Charnley N, Smoliga JA (2010) Trinitite—the atomic rock. Geology Today 26:180–185

25. Joseph ML (2012) A Geochemical Analysis of Fulgurites: from the inner glass to the outer crust. University of South Florida

26. Shand SJ (1916) The Pseudotachylyte of Parijs (Orange Free State), and its Relation to ‘Trap-Shotten Gneiss’ and ‘Flinty Crush-Rock.’ Q J Geol Soc London 72:198–221

Contributions and Attributions

CHAPTER OVERVIEW

7: GEOLOGIC TIME

Learning Objectives

Explain the difference between relative time and numeric time
Describe the five principles of stratigraphy
Apply relative dating principles to a block diagram and interpret the sequence of geologic events
Define an isotope, and explain alpha decay, beta decay, and electron capture as mechanisms of radioactive decay
Describe how radioisotopic dating is accomplished and list the four key isotopes used
Explain how carbon-14 forms in the atmosphere and how it is used in dating recent events
Explain how scientists know the numeric age of the Earth and other events in Earth history
Explain how sedimentary sequences can be dated using radioisotopes and other techniques
Define a fossil and describe types of fossils preservation
Outline how natural selection takes place as a mechanism of evolution
Describe stratigraphic correlation
List the eons, eras, and periods of the geologic time scale and explain the purpose behind the divisions
Explain the relationship between time units and corresponding rock units—chronostratigraphy versus lithostratigraphy

The geologic time scale and basic outline of Earth's history were worked out long before we had any scientific means of assigning numerical age units, like years, to events of Earth history. Working out Earth's history depended on realizing some key principles of relative time. Nicolaus Steno (1638-1686) introduced basic principles of stratigraphy, the study of layered rocks, in 1669 [1]. William Smith (1769-1839), working with the strata of English coal mines, noticed that strata and their sequence were consistent throughout the region. Eventually, he produced the first national geologic map of Britain [2], becoming known as "the Father of English Geology." Nineteenth-century scientists developed a relative time scale using Steno's principles, with names derived from the characteristics of the rocks in those areas. The figure of this geologic time scale shows the names of the units and subunits. Using this time scale, geologists can place all events of Earth history in order without ever knowing their numerical ages. The specific events within Earth history are discussed in Chapter 8.

Figure \PageIndex1: Nicolas Steno, c. 1670

7.1: RELATIVE DATING

Relative dating is the process of determining if one rock or geologic event is older or younger than another, without knowing their specific ages—i.e., how many years ago the object was formed. The principles of relative time are simple, even obvious now, but were not generally accepted by scholars until the scientific revolution of the 17th and 18th centuries.

7.2: ABSOLUTE DATING

Relative time allows scientists to tell the story of Earth events, but does not provide specific numeric ages, and thus, the rate at which geologic processes operate. Based on Hutton's principle of uniformitarianism, early geologists surmised geological processes work slowly and the Earth is very old. Relative dating principles was how scientists interpreted Earth history until the end of the 19th Century.

7.3: FOSSILS AND EVOLUTION

Fossils are any evidence of past life preserved in rocks. They may be actual remains of body parts (rare), impressions of soft body parts, casts and molds of body parts (more common), body parts replaced by mineral (common) or evidence of animal behavior such as footprints and burrows. The body parts of living organisms range from the hard bones and shells of animals, soft cellulose of plants, soft bodies of jellyfish, down to single cells of bacteria and algae.

7.4: CORRELATION

Correlation is the process of establishing which sedimentary strata are of the same age but geographically separated. Correlation can be determined by using magnetic polarity reversals (Chapter 2), rock types, unique rock sequences, or index fossils. There are four main types of correlation: stratigraphic, lithostratigraphic, chronostratigraphic, and biostratigraphic.

7.1: Relative Dating

Relative dating is the process of determining if one rock or geologic event is older or younger than another, without knowing their specific ages—i.e., how many years ago the object was formed. The principles of relative time are simple, even obvious now, but were not generally accepted by scholars until the scientific revolution of the 17th and 18th centuries [3]. James Hutton (see Chapter 1) realized geologic processes are slow and his ideas on uniformitarianism (i.e., "the present is the key to the past") provided a basis for interpreting rocks of the Earth using scientific principles.

EON	ERA	PERIOD	MILLIONS OF YEARS AGO
Phanerozoic	Cenozoic	Quaternary	1.6
		Tertiary	66
	Mesozoic	Cretaceous	138
		Jurassic	205
		Triassic	240
	Paleozoic	Permian	290
		Pennsylvanian	330
		Mississippian	360
		Devonian	410
		Silurian	435
		Ordovician	500
		Cambrian	570
Proterozoic	Late Proterozoic Middle Proterozoic Early Proterozoic		2500
Archean	Late Archean Middle Archean Early Archean		3800?
Pre-Archean			

Figure 7.1.1: A Geologic Time Scale

Relative Dating Principles

Stratigraphy is the study of layered sedimentary rocks. This section discusses principles of relative time used in all of geology, but are especially useful in stratigraphy.

Figure 7.1.1: Lower strata are older than those lying on top of them.

Principle of Superposition: In an otherwise undisturbed sequence of sedimentary strata, or rock layers, the layers on the bottom are the oldest and layers above them are younger.

Principle of Original Horizontality: Layers of rocks deposited from above, such as sediments and lava flows, are originally laid down horizontally. The exception to this principle is at the margins of basins, where the strata can slope slightly downward into the basin.

Figure 7.1.1: Lateral continuity

Principle of Lateral Continuity: Within the depositional basin, strata are continuous in all directions until they thin out at the edge of that basin. Of course, all strata eventually end, either by hitting a geographic barrier, such as a ridge, or when the depositional process extends too far from its source, either a sediment source or a volcano. Strata that are cut by a canyon later remain continuous on either side of the canyon.

Figure 7.1.1: Dark dike cutting across older rocks, the lighter of which is younger than the grey rock.

Principle of Cross-Cutting Relationships: Deformation events like folds, faults and igneous intrusions that cut across rocks are younger than the rocks they cut across.

Principle of Inclusions: When one rock formation contains pieces or inclusions of another rock, the included rock is older than the host rock.

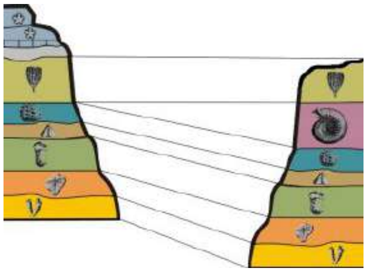


Figure 7.1.1: Fossil succession showing the correlation among strata.

Principle of Fossil Succession: Evolution has produced a succession of unique fossils that correlate to the units of the geologic time scale. Assemblages of fossils contained in strata are unique to the time they lived and can be used to correlate rocks of the same age across a wide geographic distribution. Assemblages of fossils refer to groups of several unique fossils occurring together.

Grand Canyon Example

Figure 7.1.1: The Grand Canyon of Arizona

The Grand Canyon of Arizona illustrates the stratigraphic principles. The photo shows layers of rock on top of one another in order, from the oldest at the bottom to the youngest at the top, based on the principle of superposition. The predominant white layer just below the canyon rim is the Coconino Sandstone. This layer is laterally continuous, even though the intervening canyon separates its outcrops. The rock layers exhibit the principle of lateral continuity, as they are found on both sides of the Grand Canyon which has been carved by the Colorado River.

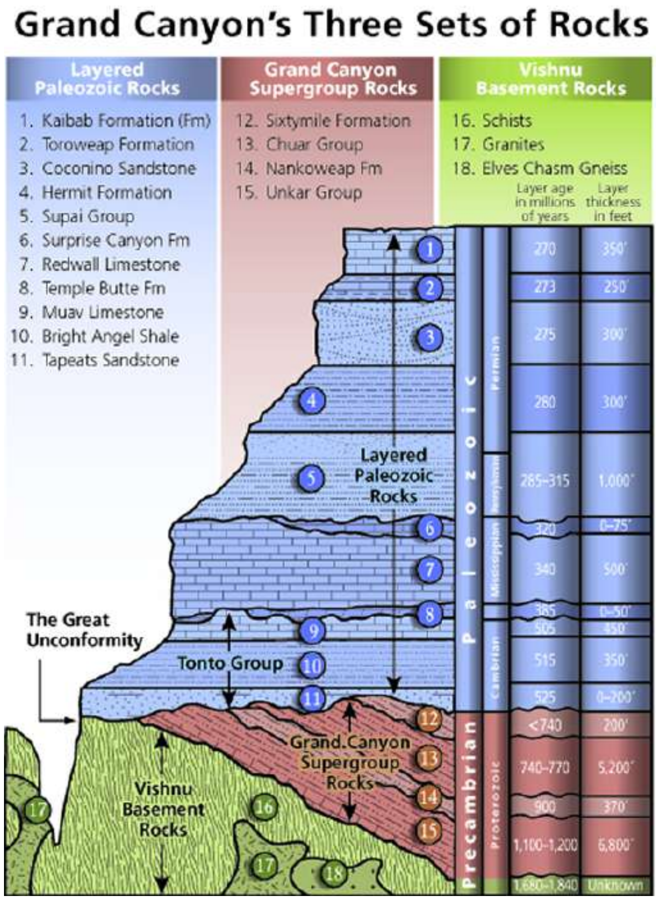


Figure 7.1.1: The rocks of the Grand Canyon

The diagram called "Grand Canyon's Three Sets of Rocks" shows a cross-section of the rocks exposed on the walls of the Grand Canyon, illustrating the principle of cross-cutting relationships, superposition, and original horizontality. In the lowest parts of the Grand Canyon are the oldest sedimentary formations, with igneous and metamorphic rocks at the bottom. The principle of cross-cutting relationships shows the sequence of these events. The metamorphic schist (#16) is the oldest rock formation and the cross-cutting granite intrusion (#17) is younger. As seen in the figure, the other layers on the walls of the Grand Canyon are numbered in reverse order with #15 being the oldest and #1 the youngest [4]. This illustrates the principle of superposition. The Grand Canyon region lies in Colorado Plateau, which is characterized by horizontal or nearly horizontal strata, which follows the principle of original horizontality. These rock strata have been barely disturbed from their original deposition, except by a broad regional uplift.

Figure 7.1.1: The red, layered rocks of the Grand Canyon Supergroup overlying the dark-colored rocks of the Vishnu schist represents a type of unconformity called a nonconformity.

The photo of the Grand Canyon here show strata that were originally deposited in a flat layer on top of older igneous and metamorphic "basement" rocks, per the original horizontality principle. Because the formation of the basement rocks and the deposition of the overlying strata is not continuous but broken by events of metamorphism, intrusion, and erosion, the contact between the strata and the older basement is termed an **unconformity**. An unconformity represents a period during which deposition did not occur or erosion removed rock that had been deposited, so there are no rocks that represent events of Earth history during that span of time at that place. Unconformities appear in cross-sections and stratigraphic columns as wavy lines between formations. Unconformities are discussed in the next section.

Unconformities

Figure 7.1.1: All three of these formations have a disconformity at the two contacts between them. The pinching Temple Butte is the easiest to see the erosion, but even between the Muav and Redwall, there is an unconformity.

There are three types of unconformities, nonconformity, disconformity, and angular unconformity. A nonconformity occurs when sedimentary rock is deposited on top of igneous and metamorphic rocks as is the case with the contact between the strata and basement rocks at the bottom of the Grand Canyon.

The strata in the Grand Canyon represent alternating marine transgressions and regressions where sea level rose and fell over millions of years. When the sea level was high marine strata formed. When sea-level fell, the land was exposed to erosion creating an unconformity. In the Grand Canyon cross-section, this erosion is shown as heavy wavy lines between the various numbered strata. This is a type of unconformity called a **disconformity**, where either non-deposition or erosion took place. In other words, layers of rock that could have been present, are absent. The time that could have been represented by such layers is instead represented by the disconformity. Disconformities are unconformities that occur between parallel layers of strata indicating either a period of no deposition or erosion.

Figure 7.1.1: In the lower part of the picture is an angular unconformity in the Grand Canyon known as the Great Unconformity. Notice the flat-lying strata over dipping strata (Source: Doug Dolde).

The Phanerozoic strata in most of the Grand Canyon are horizontal. However, near the bottom horizontal strata overlie tilted strata. This is known as the Great Unconformity and is an example of an **angular unconformity**. The lower strata were tilted by tectonic processes that disturbed their original horizontality and caused the strata to be eroded. Later, horizontal strata were deposited on top of the tilted strata creating an angular unconformity.

Here are three graphical illustrations of the three types of unconformity.

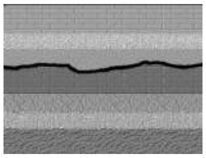


Figure 7.1.1: Disconformity

Disconformity, where is a break or stratigraphic absence between strata in an otherwise parallel sequence of strata.

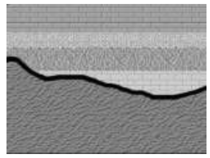


Figure 7.1.1: Nonconformity (the lower rocks are igneous or metamorphic)

Nonconformity, where sedimentary strata are deposited on crystalline (igneous or metamorphic) rocks.

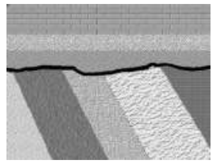


Figure 7.1.1: Angular unconformity

Angular unconformity, where sedimentary strata are deposited on a terrain developed on sedimentary strata that have been deformed by tilting, folding, and/or faulting. so that they are no longer horizontal.

Applying Relative Dating Principles

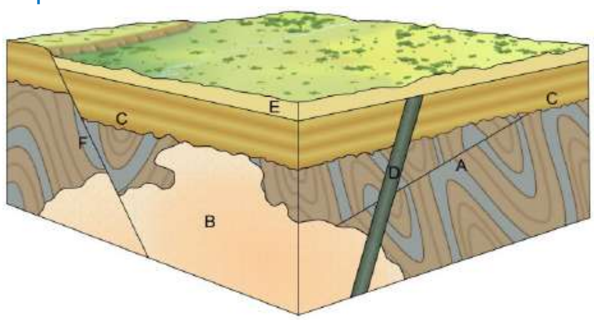


Figure 7.1.1: Block diagram to apply relative dating principles. The wavy rock is an old metamorphic gneiss, A and F are faults, B is an igneous granite, D is a basaltic dike, and C and E are sedimentary strata.

In the block diagram, the sequence of geological events can be determined by using the relative-dating principles and known properties of igneous, sedimentary, metamorphic rock (see Chapter 4, Chapter 5, and Chapter 6). The sequence begins with the folded metamorphic gneiss on the bottom. Next, the gneiss is cut and displaced by the fault labeled A. Both the gneiss and fault A are cut by the igneous granitic intrusion called batholith B; its irregular outline suggests it is an igneous granitic intrusion emplaced as magma into the gneiss. Since batholith B cuts both the gneiss and fault A, batholith B is younger than the other two rock formations. Next, the gneiss, fault A, and batholith B were eroded forming a nonconformity as shown with the wavy line. This unconformity was actually an ancient landscape surface on which sedimentary rock C was subsequently deposited perhaps by a marine transgression. Next, igneous basaltic dike D cut through all rocks except sedimentary rock E. This shows that there is a disconformity between sedimentary rocks C and E. The top of dike D is level with the top of layer C, which establishes that erosion flattened the landscape prior to the deposition of layer E, creating a disconformity between rocks D and E. Fault F cuts across all of the older rocks B, C and E, producing a fault scarp, which is the low ridge on the upper-left side of the diagram. The final events affecting this area are current erosion processes working on the land surface, rounding off the edge of the fault scarp, and producing the modern landscape at the top of the diagram.

References

3. Whewell, W. *History of the Inductive Sciences: From the Earliest to the Present Times*. (J.W. Parker, 1837).

4. Elston, D. P., Billingsley, G. H. & Young, R. A. *Geology of Grand Canyon, Northern Arizona (with Colorado River Guides): Lees Ferry to Pierce Ferry, Arizona*. **115**, (Amer Geophysical Union, 1989).

Contributions and Attributions

7.2: Absolute Dating

Figure 7.2.1: Canada's Nuvvuagittuq Greenstone Belt may have the oldest rocks and oldest evidence life on Earth, according to recent studies.

Relative time allows scientists to tell the story of Earth events, but does not provide specific numeric ages, and thus, the rate at which geologic processes operate. Based on Hutton's principle of uniformitarianism (see Chapter 1), early geologists surmised geological processes work slowly and the Earth is very old. Relative dating principles was how scientists interpreted Earth history until the end of the 19th Century. Because science advances as technology advances, the discovery of radioactivity in the late 1800s provided scientists with a new scientific tool called **radioisotopic dating**. Using this new technology, they could assign specific time units, in this case years, to mineral grains within a rock. These numerical values are not dependent on comparisons with other rocks such as with relative dating, so this dating method is called **absolute dating** [5]. There are several types of absolute dating discussed in this section but radioisotopic dating is the most common and therefore is the focus on this section.

Radioactive Decay

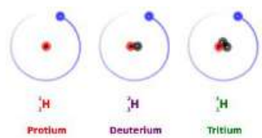


Figure 7.2.1: Three isotopes of hydrogen

All elements on the Periodic Table of Elements (see Chapter 3) contain isotopes. An isotope is an atom of an element with a different number of neutrons. For example, hydrogen (H) always has 1 proton in its nucleus (the atomic number), but the number of neutrons can vary among the isotopes (0, 1, 2). Recall that the number of neutrons added to the atomic number gives the atomic mass. When hydrogen has 1 proton and 0 neutrons it is sometimes called protium (^{1}H), when hydrogen has 1 proton and 1 neutron it is called deuterium (^{2}H), and when hydrogen has 1 proton and 2 neutrons it is called tritium (^{3}H).

Many elements have both stable and unstable isotopes. For the hydrogen example, ^{1}H and ^{2}H are stable, but ^{3}H is unstable. Unstable isotopes, called **radioactive isotopes**, spontaneously decay over time releasing subatomic particles or energy in a process called **radioactive decay**. When this occurs, an unstable isotope becomes a more stable isotope of another element. For example, carbon-14 (^{14}C) decays to nitrogen-14 (^{14}N).

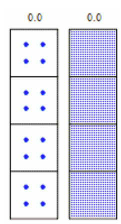


Figure 7.2.1: Simulation of half-life. On the left, 4 simulations with only a few atoms. On the right, 4 simulations with many atoms.

The radioactive decay of any individual atom is a completely unpredictable and random event. However, some rock specimens have an enormous number of radioactive isotopes, perhaps trillions of atoms and this large group of radioactive isotopes do have a predictable pattern of radioactive decay. The radioactive decay of *half* of the radioactive isotopes in this group takes a specific amount of time. The time it takes for half of the atoms in a substance to decay is called the **half-life**. In other words, the half-life of an isotope is the amount of time it takes for half of a group of unstable isotopes to decay to a stable isotope. The half-life is constant and measurable for a given radioactive isotope, so it can be used to calculate the age of a rock. For example, the half-life uranium-238 (^{238}U) is 4.5 billion years and the half-life of ^{14}C is 5,730 years.

The principles behind this dating method require two key assumptions. First, the mineral grains containing the isotope formed at the same time as the rock, such as minerals in an igneous rock that crystallized from magma. Second, the mineral crystals remain a closed system, meaning they are not subsequently altered by elements moving in or out of them.

Figure 7.2.1: Granite (left) and gneiss (right). Dating a mineral within the granite would give the crystallization age of the rock while dating the gneiss might reflect the timing of metamorphism.

These requirements place some constraints on the kinds of rock suitable for dating, with the igneous rock being the best. Metamorphic rocks are crystalline, but the processes of metamorphism may reset the clock and derived ages may represent a smear of different metamorphic events rather than the age of original crystallization. Detrital sedimentary rocks contain clasts from separate parent rocks from unknown locations and derived ages are thus meaningless. However, sedimentary rocks with precipitated minerals, such as evaporites, may contain elements suitable for radioisotopic dating. Igneous pyroclastic layers and lava flows within a sedimentary sequence can be used to date the sequence. Cross-cutting igneous rocks and sills can be used to bracket the ages of affected, older sedimentary rocks. The resistant mineral zircon, found as clasts in many ancient sedimentary rocks, has been successfully used for establishing very old dates, including the age of Earth's oldest known rocks [6] Knowing that zircon minerals in metamorphosed sediments came from older rocks that are no longer available for study, scientists can date zircon to establish the age of the pre-metamorphic source rocks.

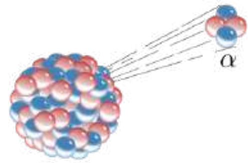
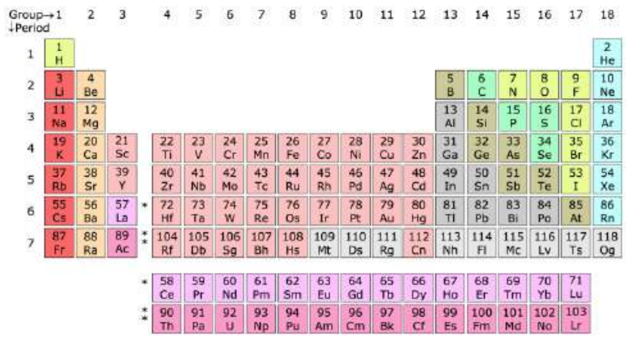


Figure 7.2.1: An alpha decay: Two protons and two neutrons leave the nucleus.

There are several ways radioactive atoms decay. We will consider three of them here—**alpha decay, beta decay,** and **electron capture**. **Alpha decay** is when an alpha particle, which consists of two protons and two neutrons, is emitted from the nucleus of an atom. This also happens to be the nucleus of a helium atom; helium gas may get trapped in the crystal lattice of a mineral in which alpha decay has taken place. When an atom loses two protons from its nucleus, lowering its atomic number, it is transformed into an element that is two atomic numbers lower on the Periodic Table of the Elements.



Periodic Table of the Elements

The loss of four particles, in this case, two neutrons and two protons, also lowers the mass of the atom by four. For example, alpha decay takes place in the unstable isotope ^{238}U, which has an atomic number of 92 (92 protons) and a mass number of 238 (total of all protons and neutrons). When ^{238}U spontaneously emits an alpha particle, it becomes thorium-234 (^{234}Th). The radioactive decay product of an element is called its **daughter isotope** and the original element is called the **parent isotope**. In this case, ^{238}U is the parent isotope and ^{234}Th is the daughter isotope. The half-life of ^{238}U is 4.5 billion years, i.e., the time it takes for half of the parent isotope atoms to decay into the daughter isotope. This isotope of uranium, ^{238}U, can be used for absolute dating the oldest materials found on Earth, and even meteorites and materials from the earliest events in our solar system.

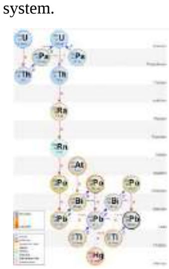


Figure 7.2.1: Decay chain of U-238 to stable Pb-206 through a series of alpha and beta decays.

Beta Decay is when a neutron in its nucleus splits into an electron and a proton. The electron is emitted from the nucleus as a beta ray. The new proton increases the element's atomic number by one, forming a new element with the same atomic mass as the parent isotope. For example, ^{234}Th is unstable and undergoes beta decay to form protactinium-234 (^{234}Pa), which also undergoes beta decay to form uranium-234 (^{234}U). Notice these are all isotopes of different elements but they have the same atomic mass of 234. The decay process of radioactive elements like uranium keeps producing radioactive parents and daughters until a stable, or non-radioactive, daughter is formed. Such a series is called a **decay chain**. The decay chain of the radioactive parent isotope ^{238}U progresses through a series of alpha (red arrows on the adjacent figure) and beta decays (blue arrows) until it forms the stable daughter isotope, lead-206 (^{206}Pb).

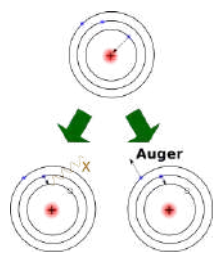


Figure 7.2.1: The two paths of electron capture

Electron capture is when a proton in the nucleus captures an electron from one of the electron shells and becomes a neutron. This produces one of two different effects: 1) an electron jumps in to fill the missing spot of the departed electron and emits an X-ray, or 2) in what is called the Auger process, another electron is released and changes the atom into an ion. The atomic number is reduced by one and the mass number remains the same. An example of an element that decays by electron capture is potassium-40 (^{40}K). Radioactive ^{40}K makes up a tiny percentage (0.012%) of naturally occurring potassium, most of which not radioactive. ^{40}K decays to argon-40 (^{40}Ar) with a half-life of 1.25 billion years, so it is very useful for dating geological events [7]. Below is a table of some of the more commonly-used radioactive dating isotopes and their half-lives.

Elements	Parent symbol	Daughter symbol	Half-life
Uranium-238/Lead-206	^{238}U	^{206}Pb	4.5 billion years
Uranium-235/Lead-207	^{235}U	^{207}Pb	704 million years
Potassium-40/Argon-40	^{40}K	^{40}Ar	1.25 billion years
Rubidium-87/Strontium-87	^{87}Rb	^{87}Sr	48.8 billion years
Carbon-14/Nitrogen-14	^{14}C	^{14}N	5,730 years

Some common isotopes used for radioisotopic dating [7; 8].

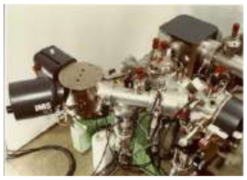


Figure 7.2.1: Mass spectrometer instrument

For a given sample of rock, how is the dating procedure carried out? The parent and daughter isotopes are separated out of the mineral using chemical extraction. In the case of uranium, ^{238}U and ^{235}U isotopes are separated out together, as are the ^{206}Pb and ^{207}Pb with an instrument called a mass spectrometer [9].

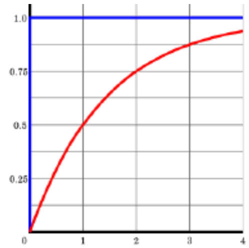


Figure 7.2.1: Graph of the number of half-lives vs. amount of daughter isotope in the sample up to 4 half-lives

Here is a simple example of age calculation using the daughter-to-parent ratio of isotopes. When the mineral initially forms, it consists of 0% daughter and 100% parent isotope, so the daughter-to-parent ratio (D/P) is 0. After one half-life, half the parent has decayed so there is 50% daughter and 50% parent, a 50/50 ratio, with D/P = 1. After two half-lives, there is 75% daughter and 25% parent (75/25 ratio) and D/P = 3. This can be further calculated for a series of half-lives as shown in the table. The table does not show more than 10 half-lives because, after about 10 half-lives, the amount of remaining parent is so small it becomes too difficult to accurately measure via chemical analysis. Modern applications of this method have achieved remarkable accuracies of plus or minus two million years in 2.5 billion years (that's ±0.055%) [10]. Applying the uranium/lead technique in any given sample analysis provides two separate clocks running at the same time, ^{238}U and ^{235}U. The existence of these two clocks in the same sample gives a cross-check between the two. Many geological samples contain multiple parent/daughter pairs, so cross-checking the clocks confirms that radioisotopic dating is highly reliable.

Half-lives (#)	Parent present (%)	Daughter present (%)	Daughter/ Parent ratio	Parent/ Daughter
Start the clock	100	0	0	infinite
1	50	50	1	1
2	25	75	3	0.33
3	12.5	87.5	7	0.143
4	6.25	93.75	15	0.0667
5	3.125	96.875	31	0.0325
10	0.098	99.9	1023	0.00098

The ratio of parent to a daughter in terms of half-life.

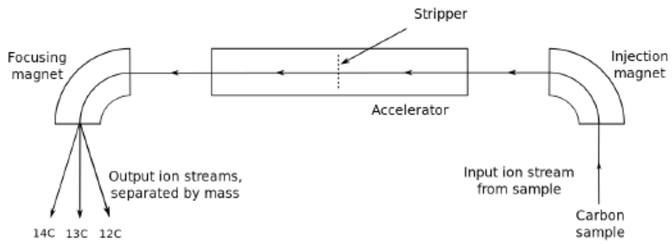


Figure 7.2.1: Schematic of carbon going through a mass spectrometer.

Another radioisotopic dating method involves carbon and is useful for dating archaeologically important samples containing organic substances like wood or bone. **Radiocarbon dating**, also called carbon dating, uses the unstable isotope carbon-14 (^{14}C) and the stable isotope carbon-12 (^{12}C). Carbon-14 is constantly being created in the atmosphere by the interaction of cosmic particles with atmospheric nitrogen-14 (^{14}N) [11]. Cosmic particles such as neutrons strike the nitrogen nucleus, kicking out a proton but leaving the neutron in the nucleus. The collision reduces the atomic number by one, changing it from seven to six, changing the nitrogen into carbon with the same mass number of 14. The ^{14}C quickly bonds with oxygen (O) in the atmosphere to form carbon dioxide ($^{14}CO_2$) that mixes with another atmospheric carbon dioxide ($^{12}CO_2$) while this mix of gases is incorporated into living matter. While an organism is alive, the ratio of $^{14}C/^{12}C$ in its body doesn't really change since CO_2 is constantly exchanged with the atmosphere. However, when it dies, the radiocarbon clock starts ticking as the ^{14}C decays back to ^{14}N by beta decay, which has a half-life of 5,730 years. The radiocarbon dating technique is thus useful for 57,300 years or so, about 10 half-lives back.

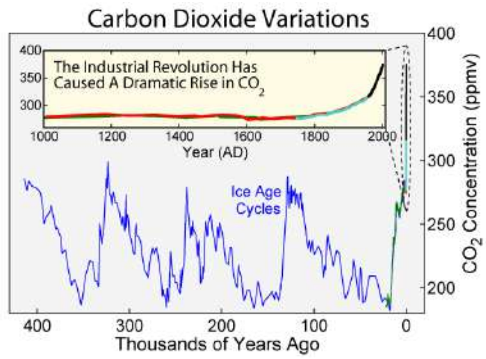


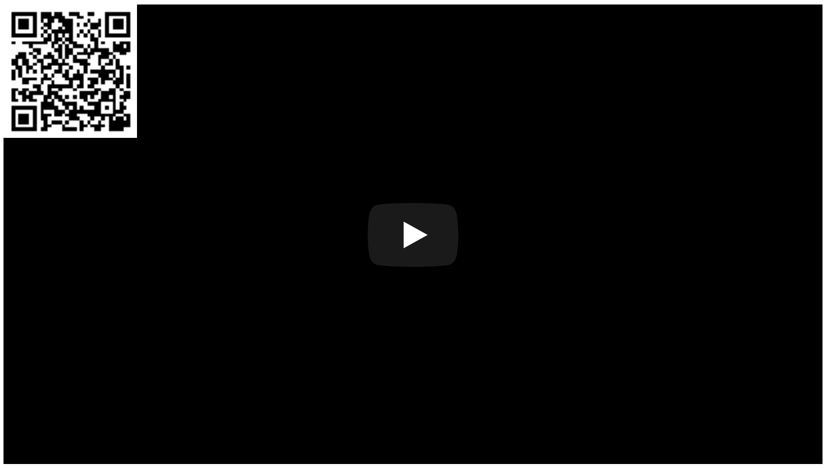
Figure 7.2.1: Carbon dioxide concentrations over the last 400,000 years.

Radiocarbon dating relies on daughter-to-parent ratios derived from a known quantity of parent ^{14}C. Early applications of carbon dating assumed the production and concentration of ^{14}C in the atmosphere remained fairly constant for the last 50,000 years. However, it is now known that the amount of parent ^{14}C levels in the atmosphere. Comparisons of carbon ages with tree-ring data and other data for known events have allowed reliable calibration of the radiocarbon dating method. Taking into account carbon-14 baseline levels must be calibrated against other reliable dating methods, carbon dating has been shown to be a reliable method for dating archaeological specimens and very recent geologic events.

Age of the Earth

Figure 7.2.1: Artist's impression of Earth in the Hadean Eon, early in Earth's history.

The work of Hutton and other scientists gained attention after the Renaissance (see Chapter 1), spurring exploration into the idea of an ancient Earth. In the late 19th century William Thompson, a.k.a. Lord Kelvin, applied his knowledge of physics to develop the assumption that the Earth started as a hot molten sphere. He estimated the Earth is 98 million years old, but because of uncertainties in his calculations stated the age as a range of between 20 and 400 million years [12; 13]. This animation (also shown below) illustrates how Kelvin calculated this range and why his numbers were so far off, which has to do with unequal heat transfer within the Earth. It has also been pointed out that Kelvin failed to consider pliability and convection in the Earth's mantle as a heat transfer mechanism. Kelvin's estimate for Earth's age was considered plausible but not without challenge, and the discovery of radioactivity provided a more accurate method for determining ancient ages [14].



In the 1950s, Clair Patterson (1922–1995) thought he could determine the age of the Earth using radioactive isotopes from meteorites, which he considered to be early solar system remnants that were present at the time Earth was forming. Patterson analyzed meteorite samples for uranium and lead using a mass spectrometer. He used the uranium/lead dating technique in determining the age of the Earth to be 4.55 billion years, give or take about 70 million (± 1.5%) [15]. The current estimate for the age of the Earth is 4.54 billion years, give or take 50 million (± 1.1%) [13]. It is remarkable that Patterson, who was still a graduate student at the University of Chicago, came up with a result that has been little altered in over 60 years, even as technology has improved dating methods.

Dating Geological Events

Figure 7.2.1: Photomicrograph of zircon crystal

Radioactive isotopes of elements that are common in mineral crystals are useful for radioisotopic dating. The uranium/lead method, with its two cross-checking clocks, is most often used with crystals of the mineral zircon ($ZrSiO_4$) where uranium can substitute for zirconium in the crystal lattice. Zircon is resistant to weathering which makes it useful for dating geological events in ancient rocks. During metamorphic events, zircon crystals may form multiple crystal layers, with each layer recording the isotopic age of an event, thus tracing the progress of the several metamorphic events [16].

Geologists have used zircon grains to do some amazing studies that illustrate how scientific conclusions can change with technological advancements. Zircon crystals from Western Australia that formed when the crust first differentiated from the mantle 4.4 billion years ago have been determined to be the oldest known rocks [6]. The zircon grains were incorporated into metasedimentary host rocks, sedimentary rocks showing signs of having undergone partial metamorphism. The host rocks were not very old but the embedded zircon grains were created 4.4 billion years ago and survived the subsequent processes of weathering, erosion, deposition, and metamorphism. From other properties of the zircon crystals, researchers concluded that not only were continental rocks exposed above sea level but also that conditions on the early Earth were cool enough for liquid water to exist on the surface. The presence of liquid water allowed the processes of weathering and erosion to take place [17]. Researchers at UCLA studied 4.1 billion-year-old zircon crystals and found carbon in the zircon crystals that may be biogenic in origin, meaning that life may have existed on Earth much earlier than previously thought [18].

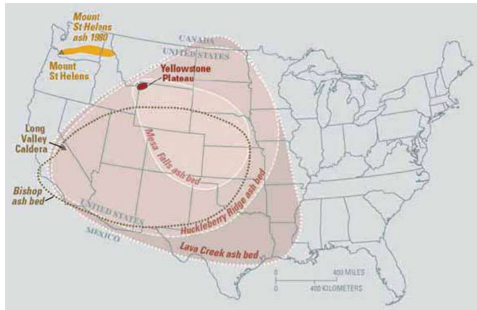


Figure 7.2.1: Several prominent ash beds found in North America, including three Yellowstone eruptions, shaded pink (Mesa Falls Tuff, Huckleberry Ridge Tuff, and Lava Creek Tuff), the Bishop Tuff ash bed (brown dashed line), and the May 18th, 1980 ash fall (yellow).

Igneous rocks best suited for radioisotopic dating because their primary minerals provide dates of crystallization from magma. Metamorphic processes tend to reset the clocks and smear the igneous rock's original date. Detrital sedimentary rocks are less useful because they are made of minerals derived from multiple parent sources with potentially many dates. However, scientists can use igneous events to date sedimentary sequences. For example, if sedimentary strata are between a lava flow and volcanic ash bed with radioisotopic dates of 54 million years and 50 million years, then geologists know the sedimentary strata and its fossils formed between 54 and 50 million years ago. Another example would be a 65 million-year-old volcanic dike that cut across sedimentary strata. This provides an upper limit age on the sedimentary strata, so these strata would be older than 65 million years. Potassium is common in evaporite sediments and has been used for potassium/argon dating [19]. Primary sedimentary minerals containing radioactive isotopes like ^{40}K has provided dates for important geologic events.

Other Absolute Dating Techniques

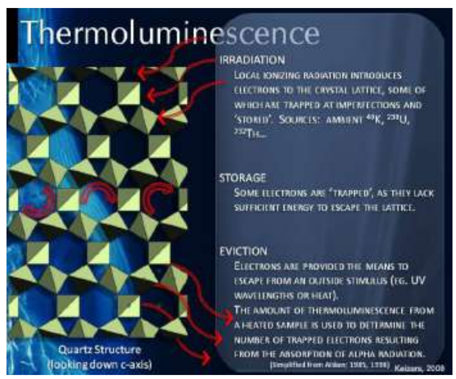


Figure 7.2.1: Thermoluminescence, a type of luminescence dating

Luminescence (aka Thermoluminescence): Radioisotopic dating is not the only way scientists determine numeric ages. Luminescence dating measures the time elapsed since some silicate minerals, such as coarse-sediments of silicate minerals, were last exposed to light or heat at the surface of Earth. All buried sediments are exposed to radiation from normal background radiation from the decay process described above. Some of these electrons get trapped in the crystal lattice of silicate minerals like quartz. When exposed at the surface, ultraviolet radiation and heat from the Sun release these electrons, but when the minerals are buried just a few inches below the surface, the electrons get trapped again. Samples of coarse sediments collected just a few feet below the surface are analyzed by stimulating them with light in a lab. This stimulation releases the trapped electrons as a photon of light which is called luminescence. The amount luminescence released indicates how long the sediment has been buried. Luminescence dating is only useful for dating sediments young sediment that is less than 1 million years old [20; 21]. In Utah, luminescence dating is used to determine when coarse-grained sediment layers were buried near a fault. This is one technique used to determine the recurrence interval of large earthquakes on faults like the Wasatch Fault that primarily cut coarse-grained material and lack buried organic soils for radiocarbon dating [22].

Figure 7.2.1: Apatite from Mexico.

Fission Track: Fission track dating relies on damage to the crystal lattice produced when unstable ^{238}U decays to the daughter product ^{234}Th and releases an alpha particle. These two decay products move in opposite directions from each other through the crystal lattice leaving a visible track of damage. This is common in uranium-bearing mineral grains such as apatite. The tracks are large and can be visually counted under an optical microscope. The number of tracks corresponds to the age of the grains. Fission track dating works from about 100,000 to 2 billion (1×10^5 to 2×10^9) years ago. Fission track dating has also been used as a second clock to confirm dates obtained by other methods [23; 7].

References

5. Geyh, M. A. & Schleicher, H. Absolute Age Determination: Physical and Chemical Dating Methods and Their Application, 503 pp. *Spring-er-Verlag, New York* (1990).

6. Wilde, S. A., Valley, J. W., Peck, W. H. & Graham, C. M. Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. *Nature* **409**, 175–178 (2001).

7. Dickin, A. P. *Radiogenic isotope geology*. (Cambridge University Press, 2005).

8. Jaffey, A. H., Flynn, K. F., Glendenin, L. E., Bentley, W. C. & others. Precision measurement of half-lives and specific activities of U 235 and U 238. *Phys. Rev. C Nucl. Phys.* (1971).

9. Dass, C. Basics of mass spectrometry. in *Fundamentals of Contemporary Mass Spectrometry* 1–14 (John Wiley & Sons, Inc., 2007).

10. Oberthür, T., Davis, D. W., Blenkinsop, T. G. & Höhndorf, A. Precise U–Pb mineral ages, Rb–Sr and Sm–Nd systematics for the Great Dyke, Zimbabwe—constraints on late Archean events in the Zimbabwe craton and Limpopo belt. *Precambrian Res.* **113**, 293–305 (2002).

11. Burleigh, R. W. F. Libby and the development of radiocarbon dating. *Antiquity* **55**, 96–98 (1981).

12. MacDougall, Doug. *Nature's clocks: how scientists measure the age of almost everything*. (2008).

13. Brent Dalrymple, G. *The Age of the Earth*. (Stanford University Press, 1994).

14. Stacey, F. D. Kelvin's age of the Earth paradox revisited. *J. Geophys. Res.* **105**, 13155–13158 (2000).

15. Patterson, C. Age of meteorites and the earth. *Geochim. Cosmochim. Acta* **10**, 230–237 (1956).

16. Ireland, T. New tools for isotopic analysis. *Science* **286**, 2289–2290 (1999).

17. Valley, J. W., Peck, W. H., King, E. M. & Wilde, S. A. A cool early Earth. *Geology* **30**, 351–354 (2002).

18. Bell, E. A., Boehnke, P., Harrison, T. M. & Mao, W. L. Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 14518–14521 (2015).

19. Léost, I., Féraud, G., Blanc-Valleron, M. M. & Rouchy, J. M. First absolute dating of Miocene Langbeinite evaporites by 40Ar/39Ar laser step-heating:[K2Mg2 (SO4) 3] Stebnyk Mine (Carpathian Foredeep Basin). *Geophys. Res. Lett.* **28**, 4347–4350 (2001).

20. Murray, A. S. & Olley, J. M. Precision and accuracy in the optically stimulated luminescence dating of sedimentary quartz: a status review. *Geochronometria* **21**, 1–16 (2002).

21. Murray, A. S. & Olley, J. M. Optically stimulated luminescence dating of sediments over the past 200,000 years. *Annu. Rev. Earth Planet. Sci.* **39**, 461–488 (2011).

22. Christopher B. DuRoss, Stephen F. Personius, Anthony J. Crone, Susan S. Olig & William R. Lund. Integration of Paleoseismic Data from Multiple Sites to Develop an Objective Earthquake Chronology: Application to the Weber Segment of the Wasatch Fault Zone, Utah. *Bulletin of the Seismological Society of America* **101**, 2765–2781 (2011).

23. P van den Haute & Frans De Corte. Advances in Fission-Track Geochronology. *Springer Science & Business Media* (2013).

Contributions and Attributions

7.3: Fossils and Evolution

Figure 7.3.1: Archaeopteryx lithographica, specimen displayed at the Museum für Naturkunde in Berlin.

Fossils are any evidence of past life preserved in rocks. They may be actual remains of body parts (rare), impressions of soft body parts, casts and molds of body parts (more common), body parts replaced by mineral (common) or evidence of animal behavior such as footprints and burrows. The body parts of living organisms range from the hard bones and shells of animals, soft cellulose of plants, soft bodies of jellyfish, down to single cells of bacteria and algae. Which body parts can be preserved? The vast majority of life today consists of soft-bodied and/or single-celled organisms, and will not likely be preserved in the geologic record except under unusual conditions. The best environment for preservation is the ocean, yet marine processes can dissolve hard parts and scavenging can reduce or eliminate remains. Thus, even under ideal conditions in the ocean, the likelihood of preservation is quite limited. For terrestrial life, the possibility of remains being buried and preserved is even more limited. In other words, the fossil record is incomplete and records only a small percentage of life that existed. Although incomplete, fossil records are used for stratigraphic correlation, using the Principle of Faunal Succession, and provide a method used for establishing the age of a formation on the Geologic Time Scale.

Types of Preservation

Figure 7.3.1: Trilobites had a hard exoskeleton and are often preserved by permineralization.

Remnants or impressions of hard parts, such as a marine clamshell or dinosaur bone, are the most common types of fossils [24]. The original material has almost always been replaced with new minerals that preserve much of the shape of the original shell, bone, or cell. The common types of fossil preservation are actual preservation, permineralization, molds and casts, carbonization, and trace fossils.

Actual preservation is a rare form of fossilization where the original materials or hard parts of the organism are preserved. Preservation of soft-tissue is very rare since these organic materials easily disappear because of bacterial decay [24]. Examples of actual preservation are unaltered biological materials like insects in amber or original minerals like mother-of-pearl on the interior of a shell. Another example is mammoth skin and hair preserved in post-glacial deposits in the Arctic regions [25]. Rare mummification [26] have left fragments of soft tissue, skin, and sometimes even blood vessels of dinosaurs, from which proteins have been isolated and evidence for DNA fragments have been discovered [27].

Figure 7.3.1: Mosquito preserved in amber

Figure 7.3.1: Permineralization in petrified wood

Permineralization occurs when an organism is buried, and then elements in groundwater completely impregnate all spaces within the body, even cells. Soft body structures can be preserved in great detail, but stronger materials like bone and teeth are the most likely to be preserved. Petrified wood is an example of detailed cellulose structures in the wood being preserved. The University of California Berkeley website has more information on permineralization.

Molds and **casts** form when the original material of the organism dissolves and leaves a cavity in the surrounding rock. The shape of this cavity is an external mold. If the mold is subsequently filled with sediments or a mineral precipitate, the organism’s external shape is preserved as a cast. Sometimes internal cavities of organisms, such internal casts of clams, snails, and even skulls are preserved as internal casts showing details of soft structures. If the chemistry is right, and burial is rapid, mineral nodules form around soft structures preserving the three-dimensional detail. This is called **authigenic mineralization**.

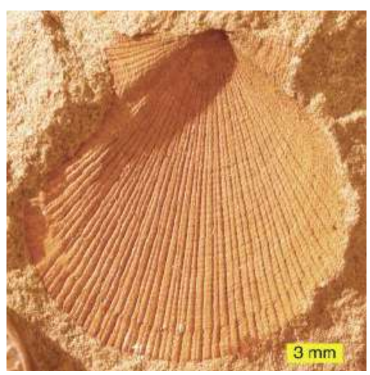


Figure 7.3.1: External mold of a clam

Figure 7.3.1: Carbonized leaf

Carbonization occurs when the organic tissues of an organism are compressed, the volatiles are driven out, and everything but the carbon disappears leaving a carbon silhouette of the original organism. Leaf and fern fossils are examples of carbonization [28].

Trace fossils are indirect evidence left behind by an organism, such as burrows and footprints, as it lived its life. **Ichnology** is specifically the study of prehistoric animal tracks. Dinosaur tracks testify to their presence and movement over an area and even provide information about their size, gait, speed, and behavior [29; 30]. Burrows dug by tunneling organisms tell of their presence and mode of life [31; 32; 33]. Other trace fossils include fossilized feces called **coprolites** [34] and stomach stones called **gastroliths** [35] that provide information about diet and habitat.

Figure 7.3.1: Footprints of the early crocodile Chirotherium

Figure 7.3.1: Fossil animal droppings (coprolite)

Evolution

Evolution has created a variety of ancient fossils that are important to stratigraphic correlation. (see chapter 7 and Chapter 5) This section is a brief discussion of the process of evolution. The British naturalist Charles Darwin (1809-1882) recognized that life forms evolve into progeny life forms. He proposed **natural selection**—which operated on organisms living under environmental conditions that posed challenges to survival—was the mechanism driving the process of evolution forward.

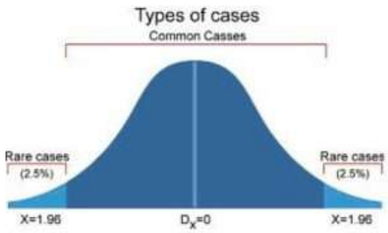


Figure 7.3.1: Variation within a population

The basic classification unit of life is the **species**: a population of organisms that exhibit shared characteristics and are capable of reproducing fertile offspring. For a species to survive, each individual within a particular population is faced with challenges posed by the environment and must survive them long enough to reproduce. Within the natural variations present in the population, there may be individuals possessing characteristics that give them some advantage in facing environmental challenges. These individuals are more likely to reproduce and pass these favored characteristics on to successive generations. If sufficient individuals in a population fail to surmount the challenges of the environment and the population cannot produce enough viable offspring, the species becomes extinct. The average lifespan of a species in the fossil record is around a million years. That life still exists on Earth shows the role and importance of evolution as a natural process in meeting the continual challenges posed by our dynamic

Earth. If the inheritance of certain distinctive characteristics is sufficiently favored over time, populations may become genetically isolated from one another, eventually resulting in the evolution of separate species. This genetic isolation may also be caused by a geographic barrier, such as an island surrounded by ocean. This theory of evolution by natural selection was elaborated by Darwin in his book *On the Origin of Species* (see Chapter 1) [36]. Since Darwin's original ideas, technology has provided many tools and mechanisms to study how evolution and speciation take place and this arsenal of tools is growing. Evolution is well beyond the hypothesis stage and is a well-established theory of modern science.

Variation within populations occurs by the natural mixing of genes through sexual reproduction or from naturally occurring mutations. Some of this genetic variation can introduce advantageous characteristics that increase the individual's chances of survival. While some species in the fossil record show little morphological change over time, others show gradual or punctuated changes, within which intermediate forms can be seen.

References

24. Allison, P. A. & Briggs, D. E. G. Exceptional fossil record: Distribution of soft-tissue preservation through the Phanerozoic. *Geology* **21**, 527–530 (1993).

25. Dale Guthrie, R. *Frozen Fauna of the Mammoth Steppe*. (pu3430623_3430810, 1989).

26. Manning, P. *Grave Secrets of Dinosaurs: Soft Tissues and Hard Science*. (National Geographic, 2009).

27. Schweitzer, M. H., Wittmeyer, J. L., Horner, J. R. & Toporski, J. K. Soft-tissue vessels and cellular preservation in Tyrannosaurus rex. *Science* **307**, 1952–1955 (2005).

28. Erickson, J., Coates, D. R. & Erickson, H. P. *An introduction to fossils and minerals: seeking clues to the Earth's past*. (Facts On File, Incorporated, 2014).

29. James O Farlow, R.E. Chapman, B Breithaupt & N Matthews. The Scientific Study of Dinosaur Footprints. in *The Complete Dinosaur* (2012).

30. Martin Lockley. *Tracking dinosaurs: a new look at an ancient world*. (1991).

31. Pemberton, S. G. & Frey, R. W. Trace Fossil Nomenclature and the Planolites-Palaeophycus Dilemma. *J. Paleontol.* **56**, 843–881 (1982).

32. *Palaeobiology II*. (Wiley-Blackwell, 2001).

33. *Biogenic structures: Their use in interpreting depositional environments*. (Sepm Society for Sedimentary, 1985).

34. Karen Chin. What did Dinosaurs Eat? Coprolites and other direct evidence of Dinosaur diets. in *The Complete Dinosaur* (2012).

35. Wm. Lee Stokes. Dinosaur gastroliths revisited. *J. Paleontol.* **61**, 1242–1246 (1987).

36. Darwin, C. *On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life*. (J. Murray, 1859).

Contributions and Attributions

7.4: Correlation

Correlation is the process of establishing which sedimentary strata are of the same age but geographically separated. Correlation can be determined by using magnetic polarity reversals (Chapter 2), rock types, unique rock sequences, or index fossils. There are four main types of correlation: stratigraphic, lithostratigraphic, chronostratigraphic, and biostratigraphic.

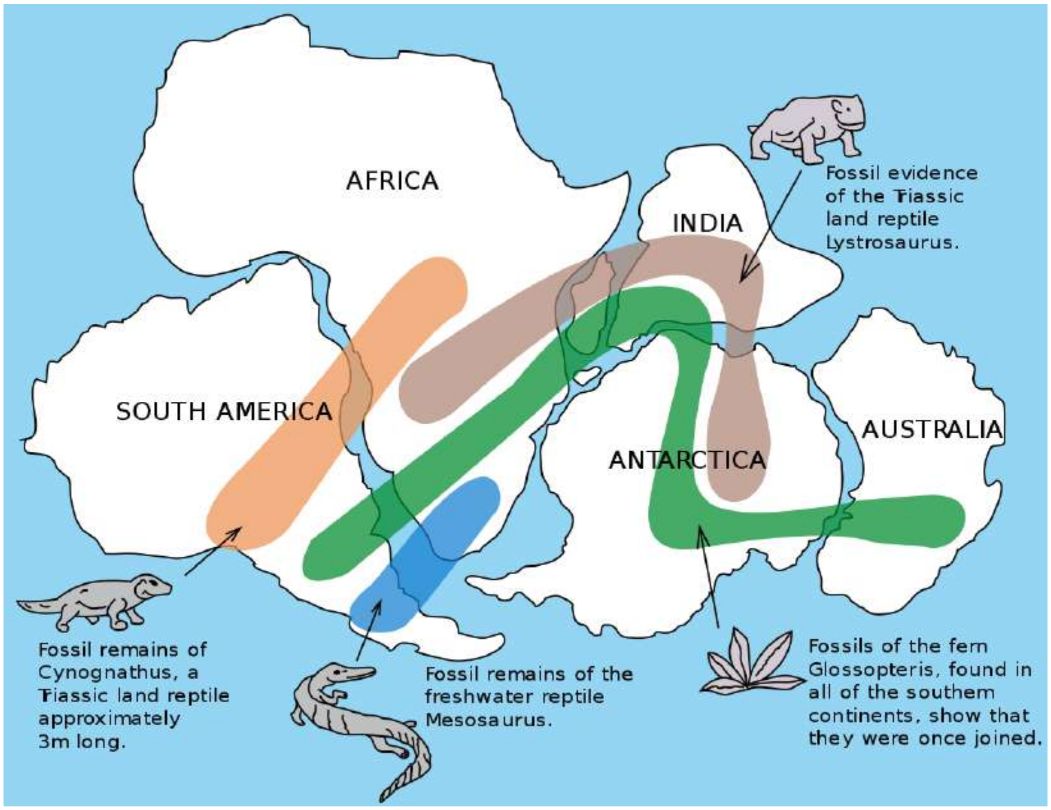


Figure 7.4.1: Image showing fossils that connect the continents of Gondwana (the southern continents of Pangea). Wegener used correlation to help develop the idea of continental drift.

Stratigraphic Correlation

Stratigraphic correlation is the process of establishing which sedimentary strata are the same age at distant geographical areas by means of their stratigraphic relationship. Geologists construct geologic histories of areas by mapping and making stratigraphic columns-a detailed description of the strata from bottom to top. An example of stratigraphic relationships and correlation between Canyonlands National Park and Zion National Park in Utah. At Canyonlands, the Navajo Sandstone overlies the Kayenta Formation which overlies the cliff-forming Wingate Formation. In Zion, the Navajo Sandstone overlies the Kayenta formation which overlies the cliff-forming Moenave Formation. Based on the stratigraphic relationship, the Wingate and Moenave Formations correlate. These two formations have unique names because their composition and outcrop pattern is slightly different. Other strata in the Colorado Plateau and their sequence can be recognized and correlated over thousands of square miles.

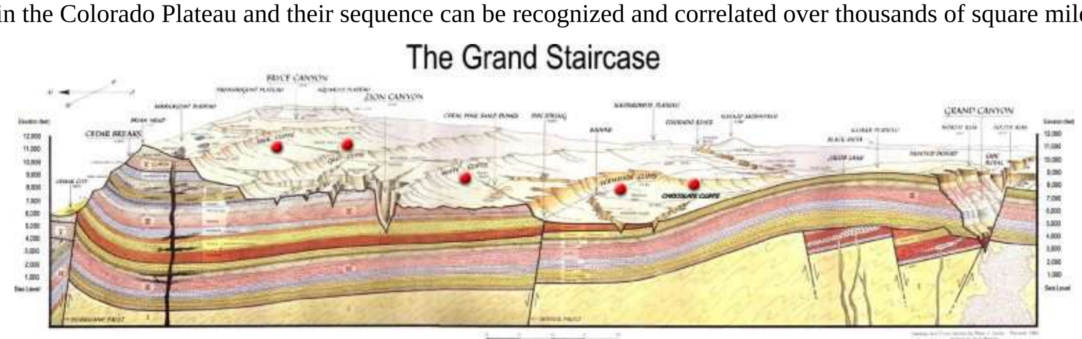


Figure 7.4.1: Correlation of Paleozoic and Mesozoic strata along the Grand Staircase from the Grand Canyon to Zion Canyon, Bryce Canyon, and Cedar Breaks. (Source: National Park Service)

Lithostratigraphic Correlation

Figure 7.4.1: View of Navajo Sandstone from Angel's Landing in Zion National Park

Lithostratigraphic correlation establishes a similar age of strata based on **lithology**, which is the composition and physical properties of that strata. *Lithos* is Greek for stone and -logy comes from the Greek word for doctrine or science. Lithostratigraphic correlation can be used to correlate whole formations long distances or can be used to correlate smaller strata within formations to trace their extent and regional depositional environments.

Figure 7.4.1: Stevens Arch in the Navajo Sandstone at Coyote Gulch some 125 miles away from Zion National Park

For example, the Navajo Sandstone, which makes up the prominent walls of Zion National Park, is the same Navajo Sandstone in Canyonlands because the lithology of the two are identical even though they are hundreds of miles apart. Extensions of the same Navajo Sandstone formation are found miles away in other parts of southern Utah, including Capitol Reef and Arches National Parks. Further, this same formation is called the Aztec Sandstone in Nevada and Nugget Sandstone near Salt Lake City because they are lithologically distinct enough to warrant new names.

Chronostratigraphic Correlation

Chronostratigraphic correlation matches rocks of the same age, even though they are made of different lithologies. Different lithologies of sedimentary rocks can form at the same time at different geographic locations because depositional environments vary geographically. For example, at any one time in a marine setting, there could be this sequence of depositional environments from the beach to deep marine: beach, nearshore area, shallow marine lagoon, reef, slope, and deep marine. Each depositional environment will have a unique sedimentary rock formation. On the figure of the Permian Capitan Reef at Guadalupe National Monument in West Texas, the red line shows a chronostratigraphic timeline that represents a snapshot in time. Shallow-water marine lagoon/back reef area is light blue, the main Capitan reef is dark blue, and deep-water marine siltstone is yellow. All three of these unique lithologies were forming at the same time in Permian along this red timeline.

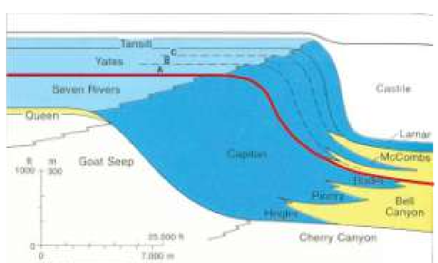


Figure 7.4.1: Cross-section of the Permian El Capitan Reef at Guadalupe National Monument, Texas. The red line shows a chronostratigraphic timeline that represents a snapshot in time in which the shallow marine lagoon/back reef area (light blue), main Capitan reef (dark blue), and deep marine siltstones (yellow) were all being deposited at the same time.

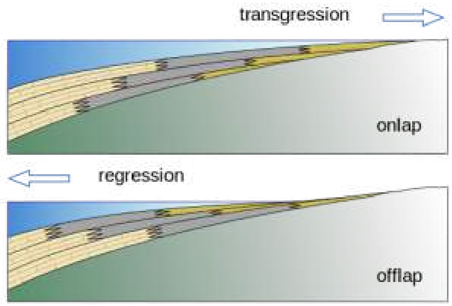


Figure 7.4.1: The rising sea levels of transgressions create onlapping sediments, regressions create off-lapping. Ocean water is shown in blue so the timeline is on the surface below the water. At the same time sandstone (buff color), limestone (gray), and shale (mustard color) are all forming at different depths of water.

Biostratigraphic Correlation

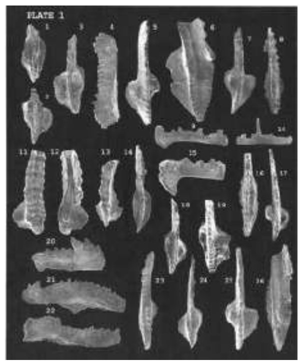


Figure 7.4.1: Conodonts

Biostratigraphic correlation uses index fossils to determine strata ages. Index fossils represent assemblages or groups of organisms that were uniquely present during specific intervals of geologic time. Assemblages refer a group of fossils. Fossils allow geologists to assign a formation to an absolute date range, such as the Jurassic Period (199 to 145 million years ago), rather than a relative time scale. In fact, most of the geologic time ranges are mapped to fossil assemblages. The most useful index fossils come from lifeforms that were geographically widespread and had a species lifespan that was limited to a narrow time interval. In other words, index fossils can be found in many places around the world, but only during a narrow time frame. Some of the best fossils for biostratigraphic correlation are microfossils**,** most of which came from single-celled organisms. As with microscopic organisms today, they were widely distributed across many environments throughout the world. Some of these microscopic organisms had hard parts, such as exoskeletons or outer shells, making them better candidates for preservation. Foraminifera, single-celled organisms with calcareous shells, are an example of an especially useful index fossil for the Cretaceous Period and Cenozoic Era [37].

Conodonts are another example of microfossils useful for biostratigraphic correlation of the Cambrian through Triassic Periods. Conodonts are tooth-like phosphatic structures of an eel-like multi-celled organism that had no other preservable hard parts. The conodont-bearing creatures lived in shallow marine environments all over the world. Upon death, the phosphatic hard parts were scattered into the rest of the marine sediments. These distinctive tooth-like structures are easily collected and separated from limestone in the laboratory.

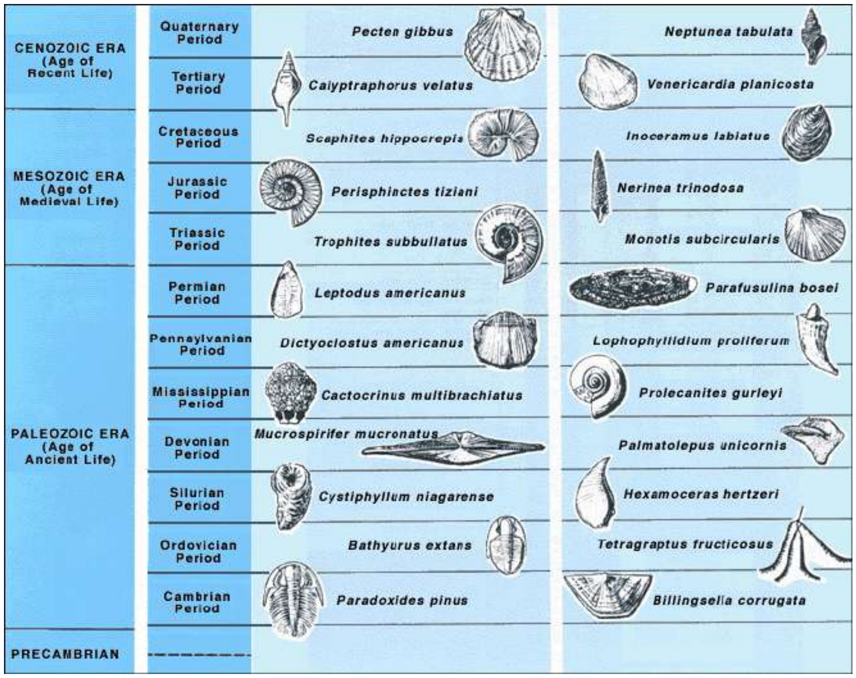


Figure 7.4.1: Index fossils used for biostratigraphic correlation

Figure 7.4.1: Foraminifera, microscopic creatures with hard shells

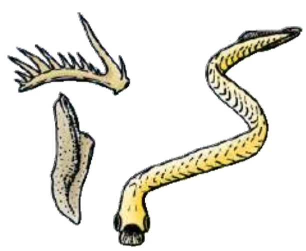


Figure 7.4.1: Artist reconstruction of the conodont animal (right) alongside its teeth

Because the conodont creatures were so widely abundant, rapidly evolving, and readily preserved in sediments, their fossils are especially useful for correlating strata, even though knowledge of the actual animal possessing them is sparse. Scientists in the 1960s carried out a fundamental biostratigraphic correlation that tied Triassic conodont zonation into ammonoids, which are extinct ancient cousins of the pearly nautilus. Up to that point, ammonoids were the only standard for Triassic correlation, so cross-referencing micro- and macro-index fossils enhanced the reliability of biostratigraphic correlation for either type [38]. That conodont study went on to establish the use of conodonts to internationally correlate Triassic strata located in Europe, Western North America, and the Arctic Islands of Canada [39].

Geologic Time Scale

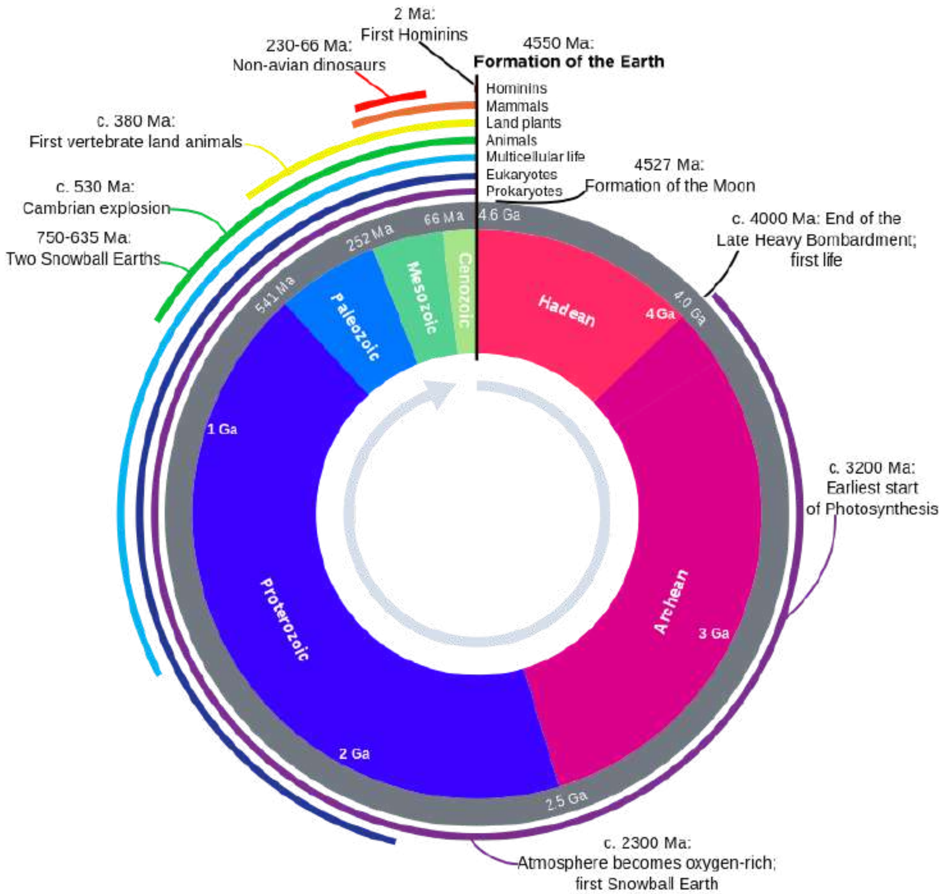


Figure 7.4.1: Geologic time on Earth, represented circularly, to show the individual time divisions and important events. Ga=billion years ago, Ma=million years ago.

Geologic time has been subdivided into a series of divisions by geologists. Eon is the largest division of time, followed by era, period, epoch, and age. The partitions of the geologic time scale are the same everywhere on Earth; however, rocks may or may not be present at a given location depending on the geologic activity going on during a particular period of time. Thus, we have the concept of time vs. rock, in which time is an unbroken continuum but rocks may be missing and/or unavailable for study. The figure of the geologic time scale represents time flowing continuously from the beginning of the Earth, with the time units presented in an unbroken sequence. But that does not mean there are rocks available for study for all of these time units.

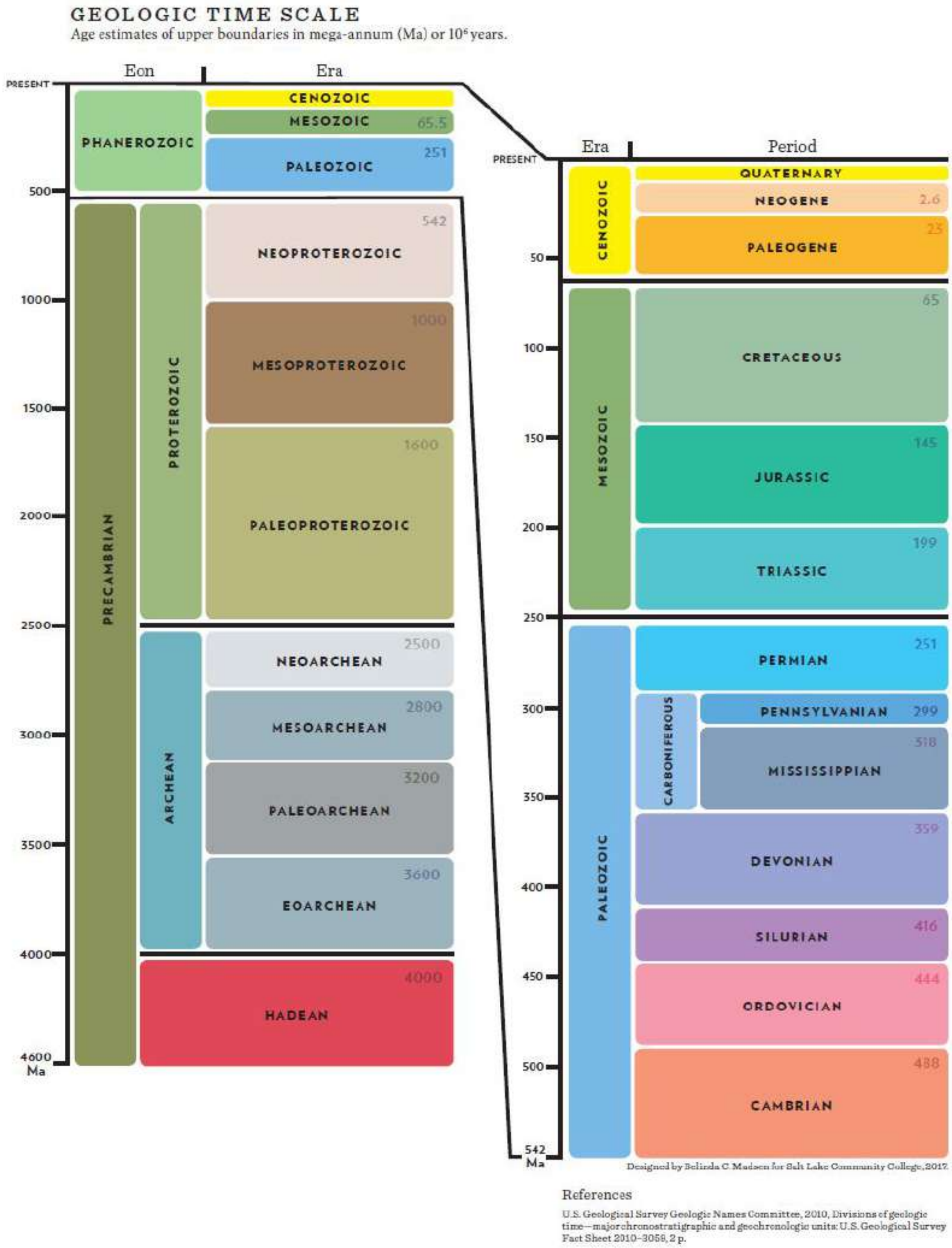


Figure 7.4.1: Geologic Time Scale with ages shown in millions of years ago (Ma).

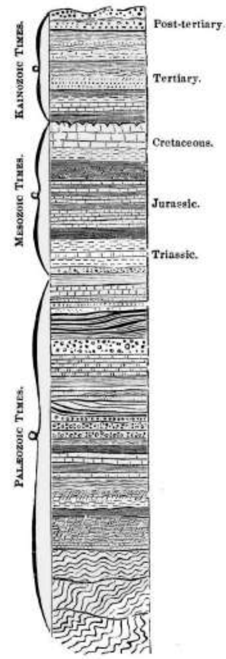


Figure 7.4.1: Names from the Geologic Time Scale applied to strata in a region

The geologic time scale was developed during the 19th century using the principles of stratigraphy. The relative order of the time units was determined before geologists had the tools to assign numerical ages to periods and events. Biostratigraphic correlation using fossils to assign era and period names to sedimentary rocks on a worldwide scale [40]. With the expansion of science and technology, some geologists think the influence of humanity on natural processes has become so great they are suggesting a new geologic time period, known as the **Anthropocene**. [39; 41].

References

37. Wade, B. S., Pearson, P. N., Berggren, W. A. & Pälike, H. Review and revision of Cenozoic tropical planktonic foraminiferal biostratigraphy and calibration to the geomagnetic polarity and astronomical time scale. *Earth-Sci. Rev.* **104**, 111–142 (2011).

38. Mosher, L. C. Triassic conodonts from Western North America and Europe and Their Correlation. *J. Paleontol.* **42**, 895–946 (1968).

39. Mosher, L. C. Triassic conodonts from British Columbia and the northern Arctic Islands. *Bull. Geol. Surv. Can.* **222**, 141–193 (1973).

40. Berry, W. B. N. *Growth of a Prehistoric Time Scale: Based on Organic Evolution (San Francisco, USA, and Folkestone, Kent, England*. (WH Freeman and Co, 1968).

41. Zalasiewicz, J. *et al.* Are we now living in the Anthropocene? *GSA Today* **18**, 4 (2008).

Contributions and Attributions

CHAPTER OVERVIEW

8: EARTH HISTORY

Learning Objectives

Explain the big-bang theory and origin of the elements
Explain the solar system's origin and the consequences for Earth.
Describe the turbulent beginning of Earth during the Hadean and Archean Eons
Identify the transition to the modern atmosphere, plate tectonics, and evolution that occurred in the Proterozoic Eon
Describe the Paleozoic evolution and extinction of invertebrates with hard parts, fish, amphibians, reptiles, tetrapods, and land plants; and tectonics and sedimentation associated with the supercontinent Pangea
Describe the Mesozoic evolution and extinction of birds, dinosaurs, and mammals; and tectonics and sedimentation associated with the breakup of Pangea
Describe the Cenozoic evolution of mammals and birds, paleoclimate, and tectonics that shaped the modern world

Entire courses and careers have been based on the wide-ranging topics covering Earth's history. Throughout the long history of Earth, change has been the norm. Looking back in time, an untrained eye would see many unfamiliar life forms and terrains. The main topics studied in Earth history are paleogeography, paleontology, and paleoecology and paleoclimatology—respectively, past landscapes, past organisms, past ecosystems, and past environments. The changes that have occurred since the inception of Earth are vast and significant. From the oxygenation of the atmosphere, the progression of life forms, the assembly and deconstruction of several supercontinents, to the extinction of more life forms than exist today, having a general understanding of these changes can put present change into a more rounded perspective. This chapter will cover briefly the origin of the universe and the 4.6 billion year history of Earth. This Earth history will focus on the major physical and biological events in each eon and Era.

8.6: PALEOZOIC

The Phanerozoic eon is the most recent eon and represents time in which fossils are common, 541 million years ago to today. The word Phanerozoic means “visible life.” Older rocks, collectively known as the Precambrian (sometimes referred to as the Cryptozoic, meaning “invisible life”), are less common and the fossils that exist represent soft-bodied life forms. The invention of hard parts like claws, scales, shells, and bones made fossils more easily preserved, and thus, easier to find.

8.7: MESOZOIC

Pangea started breaking up around 210 million years ago in the Late Triassic. Clear evidence for this includes the age of the sediments in the Newark Supergroup rift basins and the Palisades sill of the eastern part of North America and the age of the Atlantic ocean floor. Due to sea-floor spreading, the oldest rocks on the Atlantic’s floor are along the coast of northern Africa and the east coast of North America, while the youngest are along the mid-ocean ridge.

8.8: CENOZOIC

The Cenozoic, meaning “new life,” is known as the age of mammals because it is in this era that mammals came to be a dominant and large life form, including human ancestors. Birds, as well, flourished in the open niches left by the dinosaur’s demise. Most of the Cenozoic has been relatively warm, with the main exception being the ice age that started about 2.558 million years ago and (despite recent warming) continues today.

8.1: Origin of the Universe

The universe appears to have an infinite number of galaxies and solar systems and our solar system occupies a small section of this vast entirety. The origins of the universe and solar system set the context for conceptualizing the Earth's origin and early history.

Figure 8.1.1: The Hubble Deep Field. This image, released in 1996, is a composite long-exposure picture of one of the darkest parts of the night sky. Every light on this image that does not have diffraction spikes is believed to be an entire galaxy, with hundreds of billions of stars, demonstrating the immense size and scope of the universe.

Big-Bang Theory

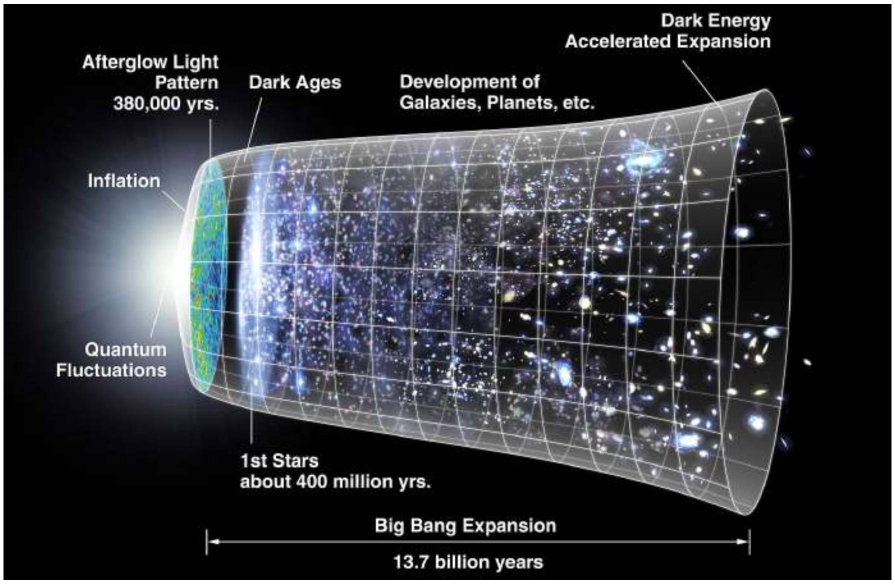


Figure 8.1.1: Timeline of expansion of the universe

The mysterious details of events prior to and during the origin of the universe are subject to great scientific debate. The prevailing idea about how the universe was created is called the **big-bang theory**. Although the ideas behind the big-bang theory feel almost mystical, they are supported by Einstein's theory of general relativity [1]. Other scientific evidence, grounded in empirical observations, supports the big-bang theory.

The big-bang theory proposes the universe was formed from an infinitely dense and hot core of the material. The bang in the title suggests there was an explosive, outward expansion of all matter and space that created atoms. Spectroscopy confirms that hydrogen makes up about 74% of all matter in the universe. Since its creation, the universe has been expanding for 13.8 billion years and recent observations suggest the rate of this expansion is increasing [2].

Spectroscopy

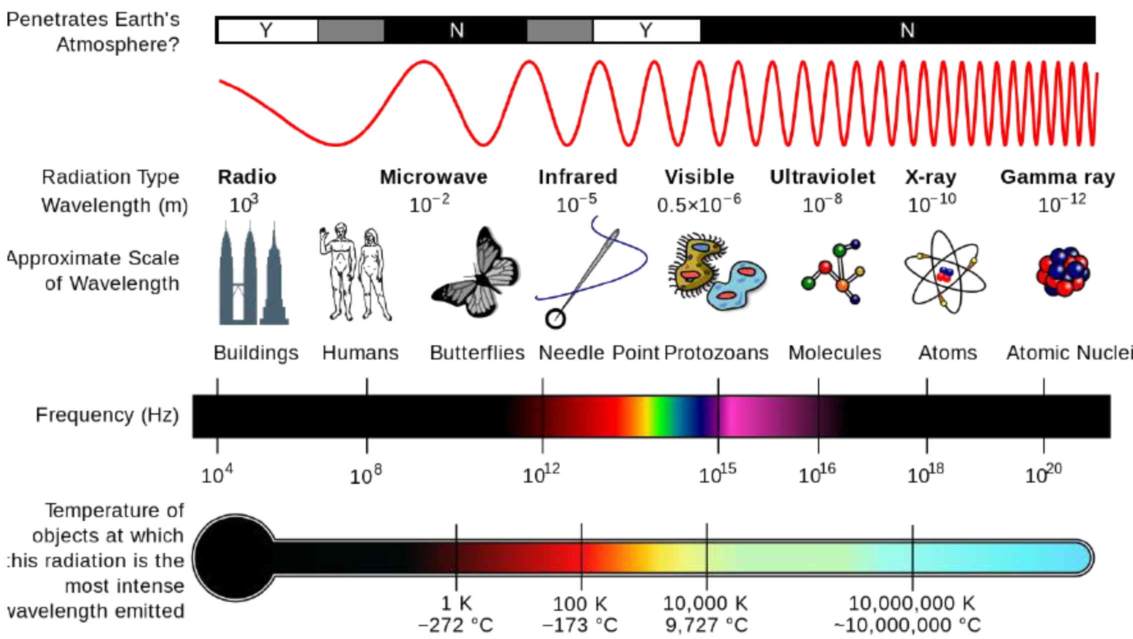


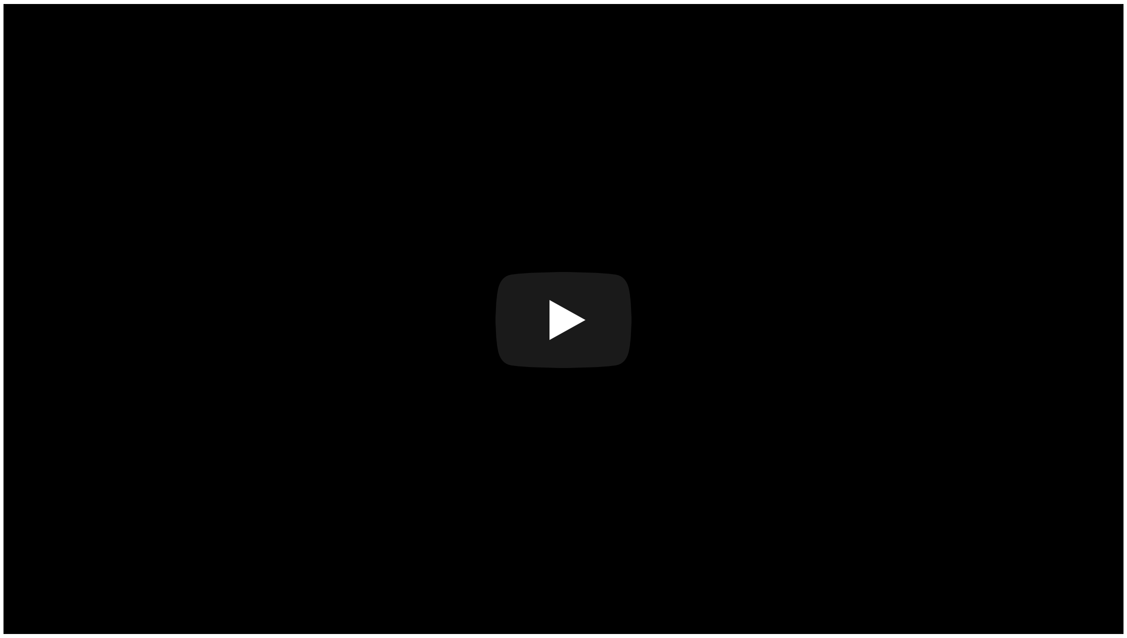
Figure 8.1.1: The electromagnetic spectrum and properties of light across the spectrum.

Spectroscopy is the investigation and measurement of spectra produced when materials interact with or emit electromagnetic radiation. *Spectra* is the plural for *spectrum* which is a particular wavelength from the **electromagnetic spectrum**. Common spectra include the different colors of visible light, X-rays, ultraviolet waves, microwaves, and radio waves. Each beam of light is a unique mixture of wavelengths that combine across the spectrum to make the color we see. The light wavelengths are created or absorbed inside atoms, and each wavelength signature matches a specific element. Even white light from the Sun, which seems like an uninterrupted continuum of wavelengths, has gaps in some wavelengths. The gaps correspond to elements present in the Earth's atmosphere that act as filters for specific wavelengths. These missing wavelengths were famously observed by Joseph von Fraunhofer (1787–1826) in the early 1800s [3], but it took decades before scientists were able to relate the missing wavelengths to atmospheric filtering. Spectroscopy shows that the Sun is mostly made of hydrogen and helium. Applying this process to light from distant stars, scientists can calculate the abundance of elements in a specific star and visible universe as a whole. Also, this spectroscopic information can be used as an interstellar speedometer.

Redshift

Figure 8.1.1: Click to animate. This animation demonstrates how the Doppler effect is heard as a car moves. The waves in front of the car are compressed together, making the pitch higher. The waves in the back of the car are stretched, and the pitch gets lower.

The **Doppler effect** is the same process that changes the pitch of the sound of an approaching car or ambulance from high to low as it passes. When an object emits waves, such as light or sound, while moving toward an observer, the wavelengths get compressed. In sound, this results in a shift to a higher pitch. When an object moves away from an observer, the wavelengths are extended, producing a lower-pitched sound. The Doppler effect is used on light emitted from stars and galaxies to determine their speed and direction of travel. Scientists, including Vesto Slipher (1875–1696) [6] and Edwin Hubble (1889–1953) [7], examined galaxies both near and far and found that almost all galaxies outside of our galaxy are moving away from each other, and us. Because the light wavelengths of receding objects are extended, visible light is shifted toward the red end of the spectrum, called a **redshift**. In addition, Hubble noticed that galaxies that were farther away from Earth also had a greater amount of redshift, and thus, the faster they are traveling away from us. The only way to reconcile this information is to deduce the universe is still expanding. Hubble's observation forms the basis of the big-bang theory.



Cosmic Microwave Background Radiation

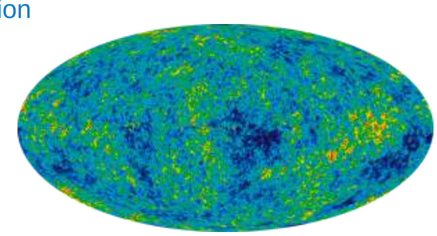


Figure 8.1.1: Heat map, showing slight variations in background heat, which is related to cosmic background radiation.

Another strong indication of the big-bang is **cosmic microwave background radiation**. Cosmic radiation was accidentally discovered by Arno Penzias (1933–) and Robert Woodrow Wilson (1936–) [8] when they were trying to eliminate background noise from a communication satellite. They discovered very faint traces of energy or heat that are omnipresent across the universe. This energy was left behind from the big bang, like an echo.

Stellar Evolution

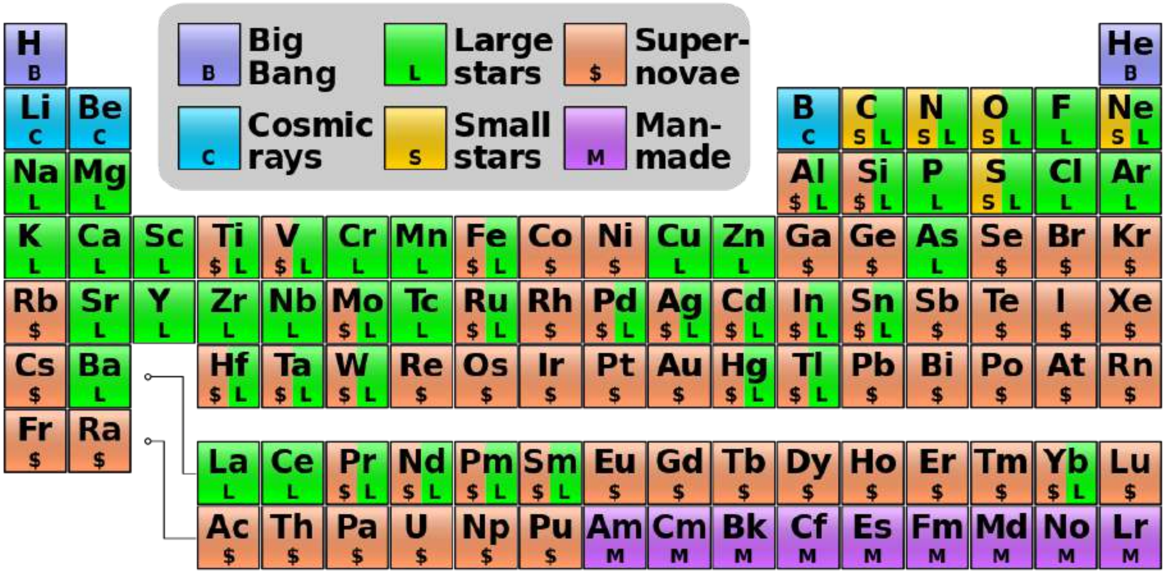


Figure 8.1.1: Origin of the elements on the periodic table, showing the important role the star life cycle plays.

Astronomers think the big bang created lighter elements, mostly hydrogen and smaller amounts of elements helium, lithium, and beryllium. Another process must be responsible for creating the other 90 heavier elements. The current model of stellar evolution explains the origins of these heavier elements.

Birth of a Star

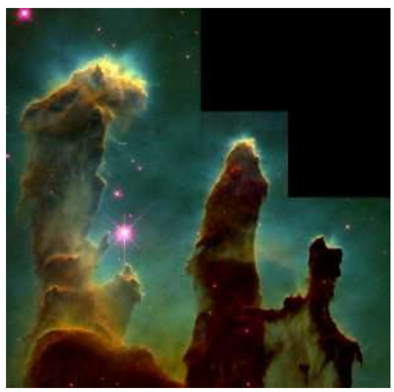


Figure 8.1.1: Section of the Eagle Nebula known as “The Pillars of Creation.”

Stars start their lives as elements floating in cold, spinning clouds of gas and dust known as **nebulas**. Gravitational attraction or perhaps a nearby stellar explosion causes the elements to condense and spin into a disk shape. In the center of this disk shape, a new star is born under the force of gravity. The spinning whirlpool concentrates material in the center, and the increasing gravitational forces collect even more mass. Eventually, the immensely concentrated mass of material reaches a critical point of such intense heat and pressure it initiates fusion.

Fusion

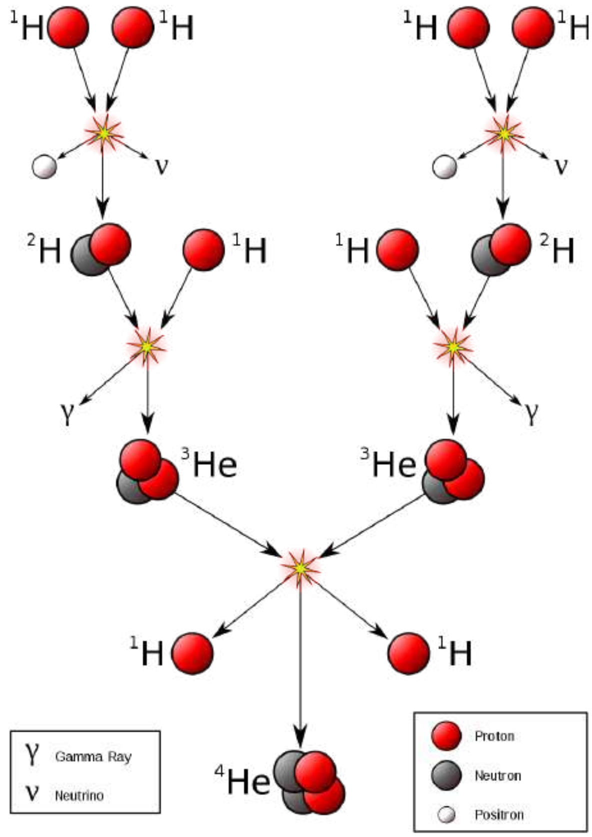


Figure 8.1.1: General diagram showing the series of fusion steps that occur in the sun.

Fusion is not a chemical reaction. **Fusion** is a nuclear reaction in which two or more nuclei, the centers of atoms, are forced together and combine creating a new larger atom. This reaction gives off a tremendous amount of energy, usually as light and solar radiation. An element such as hydrogen combines or fuses with other hydrogen atoms in the core of a star to become a new element, in this case, helium. Another product of this process is energy, such as solar radiation that leaves the Sun and comes to the Earth as light and heat. Fusion is a steady and predictable process, which is why we call this the main phase of a star's life. During its main phase, a star turns hydrogen into helium. Since most stars contain plentiful amounts of hydrogen, the main phase may last billions of years, during which their size and energy output remains relatively steady.

Figure 8.1.1: Two main paths of a star’s life cycle, depending on mass.

The giant phase in a star’s life occurs when the star runs out of hydrogen for fusion. If a star is large enough, it has sufficient heat and pressure to start fusing helium into heavier elements. This style of fusion is more energetic and the higher energy and temperature expand the star to a larger size and brightness. This giant phase is predicted to happen to our Sun in another few billion years, growing the radius of the Sun to Earth’s orbit, which will render life impossible. The mass of a star during its main phase is the primary factor in determining how it will evolve. If the star has enough mass and reaches a point at which the primary fusion element, such as helium, is exhausted, fusion continues using new, heavier elements. This occurs over and over in very large stars, forming progressively heavier elements like carbon and oxygen. Eventually, fusion reaches its limit as it forms iron and nickel. This progression explains the abundance of iron and nickel in rocky objects, like Earth, within the solar system. At this point, any further fusion absorbs energy instead of giving it off, which is the beginning of the end of the star’s life [9].

Death of a Star

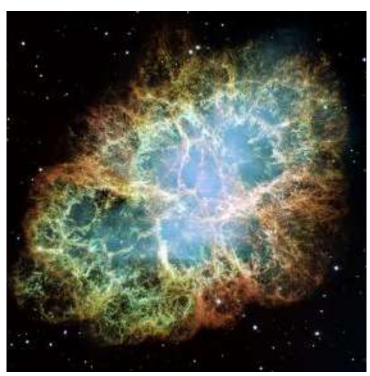


Figure 8.1.1: Hubble space telescope image of the Crab Nebula, the remnants of a supernova that occurred in 1054 C.E.

The death of a star can range from spectacular to other-worldly (see figure). Stars like the Sun form a planetary nebula, which comes from the collapse of the star’s outer layers in an event like the implosion of a building. In the tug-of-war between gravity’s inward pull and fusion’s outward push, gravity instantly takes over when fusion ends, with the outer gasses puffing away to form a nebula. More massive stars do this as well but with a more energetic collapse, which starts another type of energy release mixed with element creation known as a supernova. In a **supernova**, the collapse of the core suddenly halts, creating a massive outward-propagating shock wave. A supernova is the most energetic explosion in the universe short of the big bang. The energy release is so significant the ensuing fusion can make every element up through uranium [10].

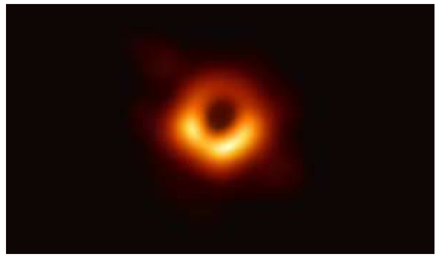


Figure 8.1.1: A black hole and its shadow have been captured in an image for the first time in 2019, a historic feat by an international network of radio telescopes called the Event Horizon Telescope (Source: NASA)

The death of the star can result in the creation of white dwarfs, neutron stars, or black holes. Following their deaths, stars like the Sun turn into white dwarfs.

White dwarfs are hot star embers, formed by packing most of a dying star's mass into a small and dense object about the size of Earth. Larger stars may explode in a supernova that packs their mass even tighter to become neutron stars. Neutron stars are so dense that protons combine with electrons to form neutrons. The largest stars collapse their mass even further, becoming objects so dense that light cannot escape their gravitational grasp. These are the infamous black holes and the details of the physics of what occurs in them are still up for debate.

References

1. Einstein A (1917) Cosmological Reflections on the General Relativity Theory. Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 142-152 142–152

2. Perlmutter S, Aldering G, Goldhaber G (1999) Measurements of Omega and Lambda from 42 high-redshift supernovae. Astrophys J 517:565–586

3. Fraunhofer J (1817) Bestimmung des Brechungs-und des Farbenzerstreungs-Vermögens verschiedener Glasarten, in Bezug auf die Vervollkommnung achromatischer Fernröhre. Ann Phys 56:264–313. https://doi.org/10.1002/andp.18170560706

6. Slipher VM (1913) The radial velocity of the Andromeda Nebula. Lowell Observatory Bulletin 2:56–57

7. Hubble E (1929) A relation between distance and radial velocity among extra-galactic nebulae. Proc Natl Acad Sci U S A 15:168–173

8. Penzias AA, Wilson RW (1965) A Measurement of Excess Antenna Temperature at 4080 Mc/s. Astrophys J 142:419–421

9. Salaris M, Cassisi S (2005) Evolution of stars and stellar populations. John Wiley & Sons

10. Timmes FX, Woosley SE, Weaver TA (1995) Galactic chemical evolution: Hydrogen through zinc. The Astrophysical journal Supplement series 98:617–658

Contributions and Attributions

8.2: Origin of the Solar System—The Nebular Hypothesis

Our solar system formed at the same time as our Sun as described in the nebular hypothesis. The **nebular hypothesis** is the idea that a spinning cloud of dust made of mostly light elements, called a nebula, flattened into a protoplanetary disk, and became a solar system consisting of a star with orbiting planets [12]. The spinning nebula collected the vast majority of material in its center, which is why the sun Accounts for over 99% of the mass in our solar system.

Figure 8.2.1: Small protoplanetary discs in the Orion *Nebula*

Planet Arrangement and Segregation

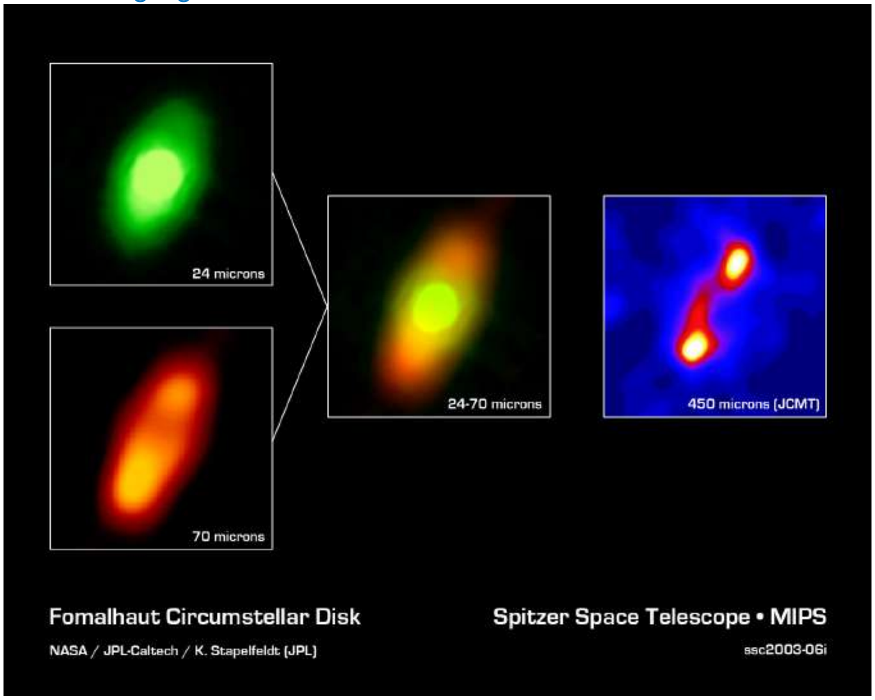


Figure 8.2.1: This disk is asymmetric, possibly because of a large gas giant planet orbiting relatively far from the star.

As our solar system formed, the nebular cloud of dispersed particles developed distinct temperature zones. Temperatures were very high close to the center, only allowing condensation of metals and silicate minerals with high melting points. Farther from the Sun, the temperatures were lower, allowing the condensation of lighter gaseous molecules such as methane, ammonia, carbon dioxide, and water [13]. This temperature differentiation resulted in the inner four planets of the solar system becoming rocky, and the outer four planets becoming gas giants.

Figure 8.2.1: Image by the ALMA telescope of HL Tauri and its protoplanetary disk, showing grooves formed as planets absorb the material in the disk.

Both rocky and gaseous planets have a similar growth model. Particles of dust, floating in the disc were attracted to each other by static charges and eventually, gravity. As the clumps of dust became bigger, they interacted with each other—colliding, sticking, and forming proto-planets. The planets continued to grow over the course of many thousands or millions of years, as material from the protoplanetary disc was added. Both rocky and gaseous planets started with a solid core. Rocky planets built more rock on that core, while gas planets added gas and ice. Ice giants formed later and on the furthest edges of the disc, accumulating less gas and more ice. That is why the gas-giant planets Jupiter and Saturn are composed of mostly hydrogen and helium gas, more than 90%. The ice giants Uranus and Neptune are composed of mostly methane ices and only about 20% hydrogen and helium gases.

Figure 8.2.1: This artist’s impression of the water snowline around the young star V883 Orionis, as detected with ALMA.

The planetary composition of the gas giants is clearly different from the rocky planets. Their size is also dramatically different for two reasons: First, the original planetary nebula contained more gases and ices than metals and rocks. There was abundant hydrogen, carbon, oxygen, nitrogen, and less silicon and iron, giving the outer planets more building material. Second, the stronger gravitational pull of these giant planets allowed them to collect large quantities of hydrogen and helium, which could not be collected by the weaker gravity of the smaller planets.

Figure 8.2.1: A polished fragment of the iron-rich Toluca Meteorite, with octahedral Widmanstätten Pattern.

Jupiter’s massive gravity further shaped the solar system and growth of the inner rocky planets. As the nebula started to coalesce into planets, Jupiter’s gravity accelerated the movement of nearby materials, generating destructive collisions rather than constructively gluing material together [14]. These collisions created the asteroid belt, an unfinished planet, located between Mars and Jupiter. This asteroid belt is the source of most **meteorites** that currently impact the Earth. Study of asteroids and meteorites help geologist to determine the age of Earth and the composition of its core, mantle, and crust. Jupiter's gravity may also explain Mars' smaller mass, with the larger planet consuming material as it migrated from the inner to the outer edge of the solar system [15].

Pluto and Planet Definition

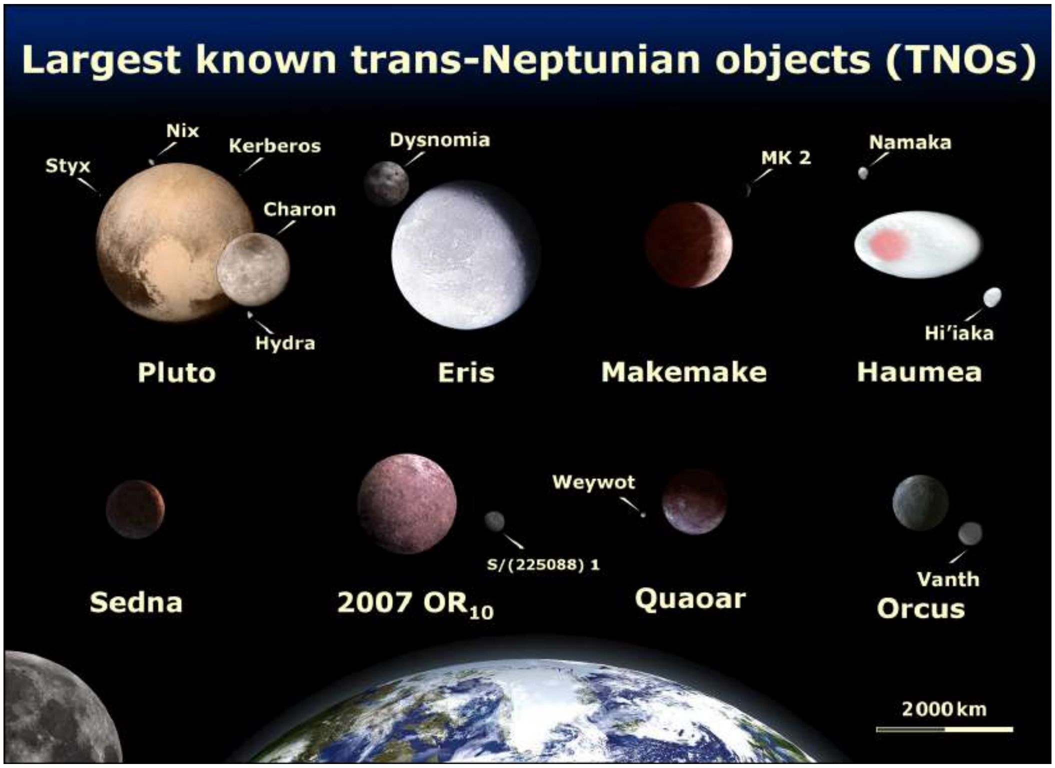


Figure 8.2.1: Eight largest objects discovered past Neptune.

The outermost part of the solar system is known as the Kuiper belt, which is a scattering of rocky and icy bodies. Beyond that is the Oort cloud, a zone filled with small and dispersed ice traces. These two locations are where most comets form and continue to orbit, and objects found here have relatively irregular orbits compared to the rest of the solar system. Pluto, formerly the ninth planet, is located in this region of space. The XXVIth General Assembly of the International Astronomical Union (IAU) stripped Pluto of planetary status in 2006 because scientists discovered an object more massive than Pluto, which they named Eris. The IAU decided against including Eris as a planet, and therefore, excluded Pluto as well. The IAU narrowed the definition of a planet to three criteria:

1. Enough mass to have gravitational forces that force it to be rounded
2. Not massive enough to create a fusion
3. Large enough to be in a cleared orbit, free of other planetesimals that should have been incorporated at the time the planet formed. Pluto passed the first two parts of the definition, but not the third. Pluto and Eris are currently classified as dwarf planets

References

12. Montmerle T, Augereau J-C, Chaussidon M, et al (2006) Solar System Formation and Early Evolution: the First 100 Million Years. In: From Suns to Life: A Chronological Approach to the History of Life on Earth. Springer New York, pp 39–95

13. Martin RG, Livio M (2012) On the evolution of the snow line in protoplanetary discs. Mon Not R Aston Soc Lett 425:L6–L9

14. Petit J-M, Morbidelli A, Chambers J (2001) The Primordial Excitation and Clearing of the Asteroid Belt. Icarus 153:338–347. https://doi.org/10.1006/icar.2001.6702

15. Walsh KJ, Morbidelli A, Raymond SN, et al (2011) A low mass for Mars from Jupiter's early gas-driven migration. Nature 475:206–209

Contributions and Attributions

This page is licensed under a CC BY-NC-SA license and was authored, remixed, and/or curated by Chris Johnson, Matthew D. Affolter, Paul Inkenbrandt, & Cam Mosher. A detailed versioning history of the edits to source content is available upon request.

8.3: Hadean Eon

Geoscientists use the geological time scale to assign relative age names to events and rocks, separating major events in Earth's history based on significant changes as recorded in rocks and fossils. This section summarizes the most notable events of each major time interval. For a breakdown of how these time intervals are chosen and organized, see chapter 7.

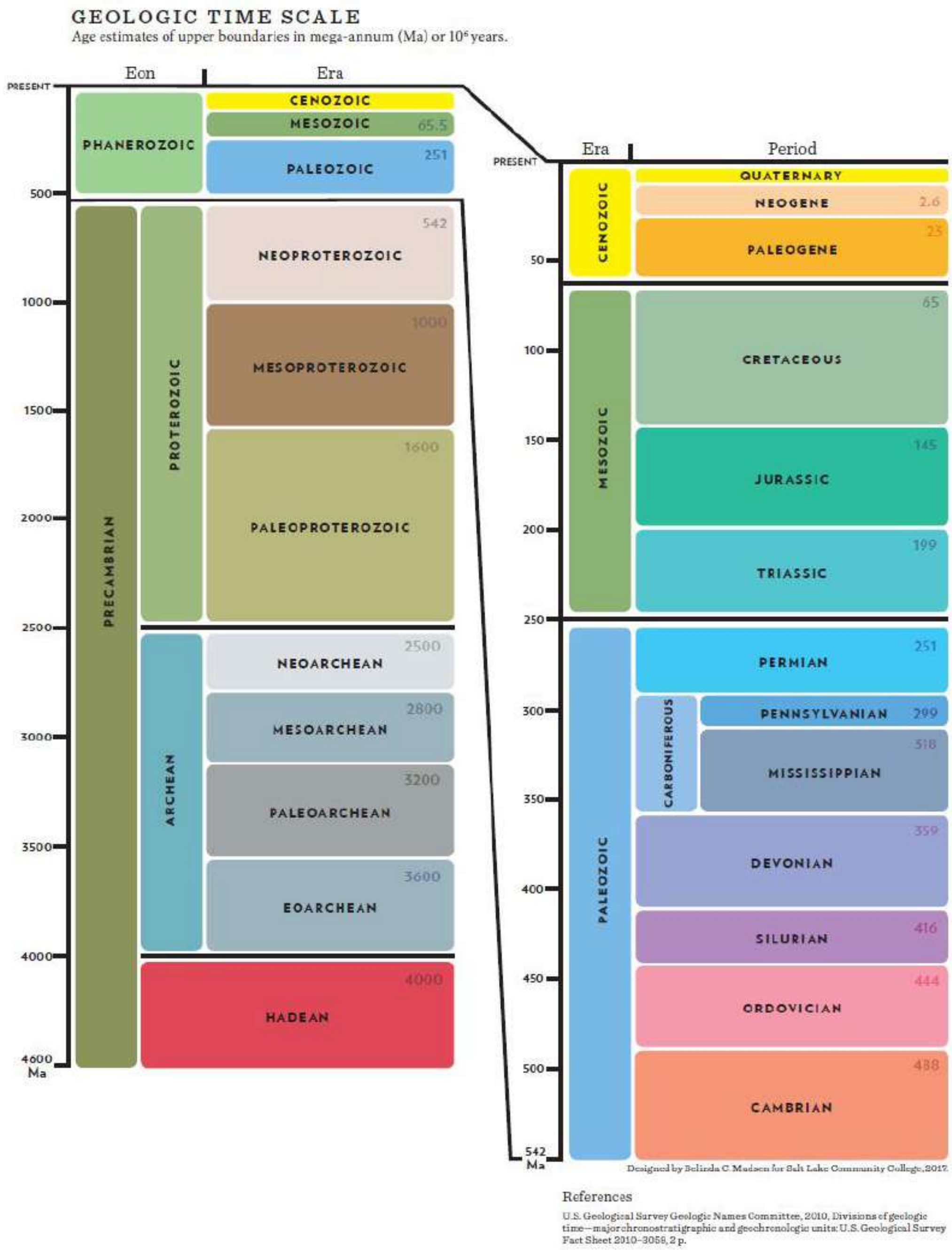


Figure 8.3.1: Geologic Time Scale with ages shown

The Hadean Eon, named after the Greek god and ruler of the underworld Hades, is the oldest eon and dates from 4.5–4.0 billion years ago.

This time represents Earth's earliest history, during which the planet was characterized by a partially molten surface, volcanism, and asteroid impacts. Several mechanisms made the newly forming Earth incredibly hot: gravitational compression, radioactive decay, and asteroid impacts. Most of this initial heat still exists inside the Earth. The Hadean was originally defined as the birth of the planet occurring 4.0 billion years ago and preceding the existence of many rocks and life forms. However, geologists have dated minerals at 4.4 billion years, with evidence that liquid water was present [18]. There is possibly even evidence of life existing over 4.0 billion years ago. However, the most reliable record for early life, the microfossil record, starts at 3.5 billion years ago.

Figure 8.3.1: Artist's impression of the Earth in the Hadean.

Origin of Earth's Crust

As Earth cooled from its molten state, minerals started to crystallize and settle resulting in a separation of minerals based on density and the creation of the crust, mantle, and core. The earliest Earth was chiefly molten material and would have been rounded by gravitational forces so it resembled a ball of lava floating in space. As the outer part of the Earth slowly cooled, the high melting-point minerals (see Bowen's Reaction Series in Chapter 4) formed solid slabs of early crust. These slabs were probably unstable and easily reabsorbed into the liquid magma until the Earth cooled enough to allow numerous larger fragments to form a thin primitive crust. Scientists generally assume this crust was oceanic and mafic in composition and littered with impacts, much like the Moon's current crust. There is still some debate over when plate tectonics started, which would have led to the formation of the continental and the felsic crust [23]. Regardless of this, as Earth cooled and solidified, less dense felsic minerals floated to the surface of the Earth to form the crust, while the denser mafic and ultramafic materials sank to form the mantle and the highest-density iron and nickel sank into the core. This differentiated the Earth from a homogenous planet into a heterogeneous one with layers of felsic crust, mafic crust, ultramafic mantle, and iron and nickel core.

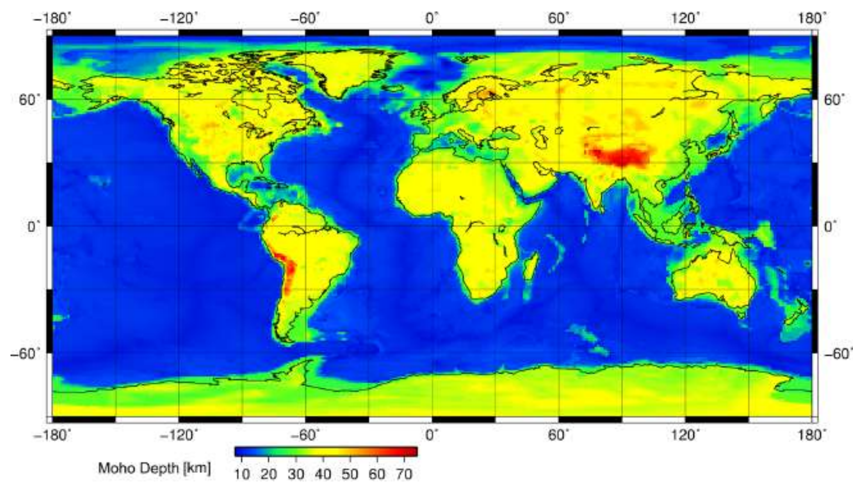


Figure 8.3.1: The global map of the depth of the moho, or thickness of the crust.

Origin of the Moon

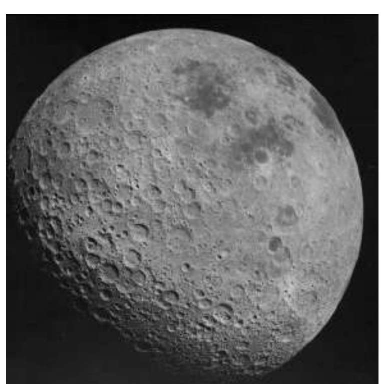
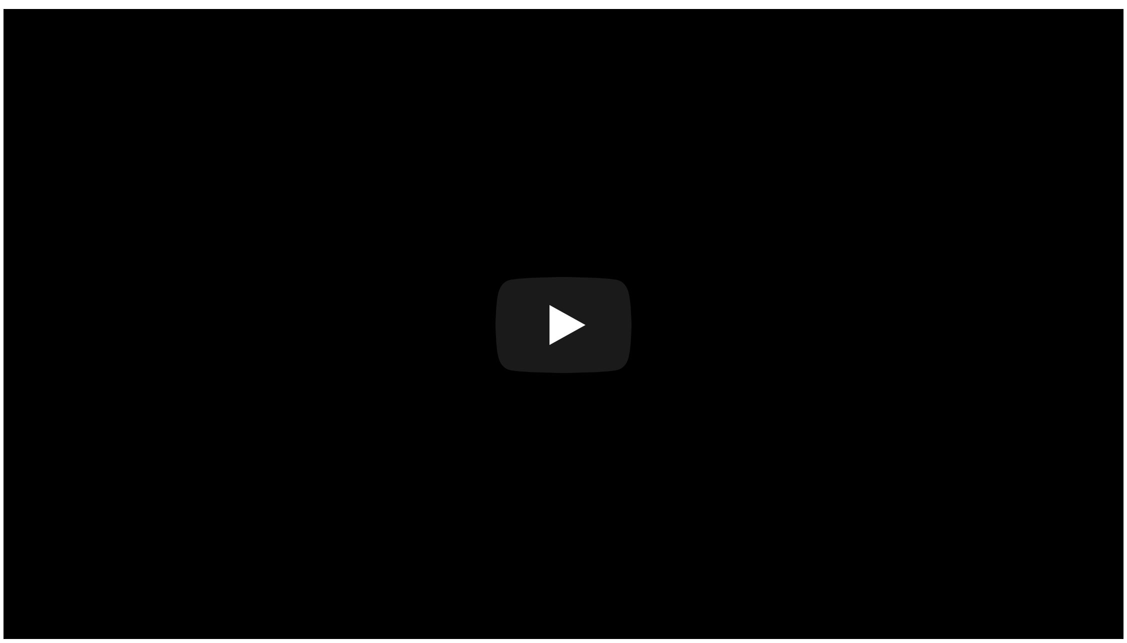


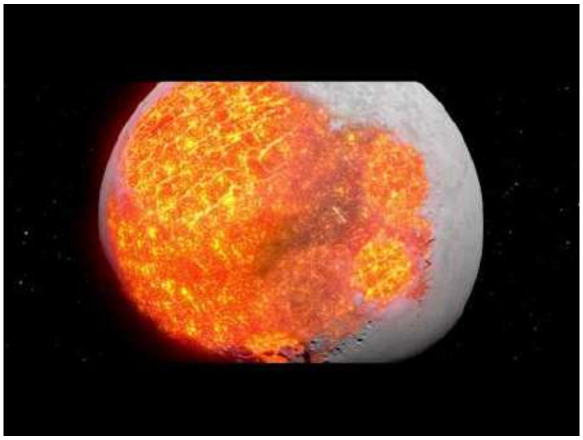
Figure 8.3.1: Dark side of the Moon

Several unique features of Earth's Moon have prompted scientists to develop the current hypothesis about its formation. The Earth and Moon are tidally locked, meaning that as the Moon orbits, one side always faces the Earth and the opposite side is not visible to us. Also and most importantly, the chemical compositions of the Earth and Moon show nearly identical isotope ratios [24] and volatile content. Apollo missions returned from the Moon with rocks that allowed scientists to conduct very precise comparisons between Moon and Earth rocks. Other bodies in the solar system and meteorites do not share the same degree of similarity and show much higher variability. If the Moon and Earth formed together, this would explain why they are so chemically similar.

Figure 8.3.1: Artist's concept of the giant impact from a Mars-sized object that could have formed the moon.

Many ideas have been proposed for the origin of the Moon: The Moon could have been captured from another part of the solar system and formed in place together with the Earth, or the Moon could have been ripped out of the early Earth. None of the proposed explanations can account for all the evidence. The currently prevailing hypothesis is the **giant-impact hypothesis**. It proposes a body about half of Earth's size must have shared at least parts of Earth's orbit and collided with it, resulting in a violent mixing and scattering of material from both objects. Both bodies would be composed of a combination of materials, with more of the lower density splatter coalescing into the Moon. This may explain why the Earth has a higher density and thicker core than the Moon.





Computer simulation of the evolution of the Moon (2 minutes).

Origin of Earth's Water

Explanations for the origin of Earth's water include volcanic outgassing, comets, and meteorites. The volcanic outgassing hypothesis for the origin of Earth's water is that it originated from inside the planet, and emerged via tectonic processes as vapor associated with volcanic eruptions [27]. Since all volcanic eruptions contain some water vapor, at times more than 1% of the volume, these alone could have created Earth's surface water. Another likely source of water was from space. Comets are a mixture of dust and ice, with some or most of that ice being frozen water. Seemingly dry meteors can contain small but measurable amounts of water, usually trapped in their mineral structures [28; 29]. During heavy bombardment periods later in Earth's history, its cooled surface was pummeled by comets and meteorites, which could be why so much water exists above ground. There isn't a definitive answer to what process is the source of ocean water. Earth's water isotopically matches water found in meteorites much better than that of comets [30]. However, it is hard to know if Earth processes could have changed the water's isotopic signature over the last 4-plus billion years. It is possible that all three sources contributed to the origin of Earth's water.

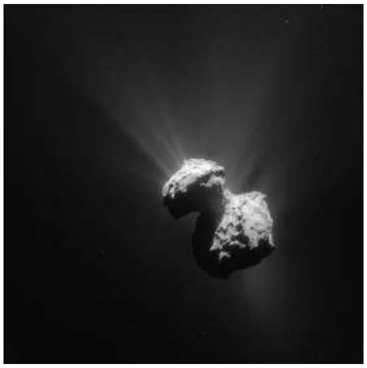


Figure 8.3.1: Water vapor leaves comet 67P/Churyumov–Gerasimenko.

References

18. Wilde SA, Valley JW, Peck WH, Graham CM (2001) Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. Nature 409:175–178

23. Johnson TE, Brown M, Gardiner NJ, et al (2017) Earth's first stable continents did not form by subduction. Nature 543:239–242

24. Wiechert U, Halliday AN, Lee DC, et al (2001) Oxygen isotopes and the moon-forming giant impact. Science 294:345–348

27. Pearson DG, Brenker FE, Nestola F, et al (2014) Hydrous mantle transition zone indicated by ringwoodite included within diamond. Nature 507:221–224

28. Morbidelli A, Chambers J, Lunine JI, et al (2000) Source regions and timescales for the delivery of water to the Earth. Meteorit Planet Sci 35:1309–1320

29. Hosono N, Karato S, Makino J, Saitoh TR (2019) Terrestrial magma ocean origin of the Moon. Nature Geoscience 1. https://doi.org/10.1038/s41561-019-0354-2

30. Altwegg K, Balsiger H, Bar-Nun A, et al (2015) Cometary science. 67P Churyumov-Gerasimenko, a Jupiter family comet with a high D-H ratio. Science 347:1261952

Contributions and Attributions

8.4: Archean Eon

The **Archean Eon**, which lasted from 4.0–2.5 billion years ago, is named after the Greek word for beginning. This eon represents the beginning of the rock record. Although there is current evidence that rocks and minerals existed during the Hadean Eon, the Archean has a much more robust rock and fossil record.

Figure 8.4.1: Artist's impression of the Archean.

Late Heavy Bombardment

Objects were chaotically flying around at the start of the solar system, building the planets and moons. There is evidence that after the planets formed, about 4.1–3.8 billion years ago, a second large spike of asteroid and comet impacted the Earth and Moon in an event called **late heavy bombardment**. Meteorites and comets in stable or semi-stable orbits became unstable and started impacting objects throughout the solar system. In addition, this event is called the lunar cataclysm because most of the Moons craters are from this event. During the late heavy bombardment, the Earth, Moon, and all planets in the solar system were pummeled by material from the asteroid and Kuiper belts. Evidence of this bombardment was found within samples collected from the Moon.

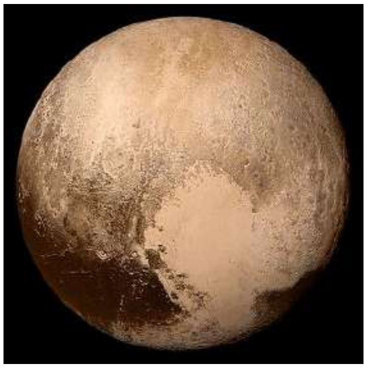


Figure 8.4.2: 2015 image from NASA's New Horizons probe of Pluto. The lack of impacts found on the Tombaugh Regio (the heart-shaped plain, lower right) has been inferred as being younger than the Late Heavy Bombardment and the surrounding surface due to its lack of impacts.

It is universally accepted that the solar system experienced extensive asteroid and comet bombardment at its start; however, some other process must have caused the second increase in impacts hundreds of millions of years later. A leading theory blames gravitational resonance between Jupiter and Saturn for disturbing orbits within the asteroid and Kuiper belts [33] based on a similar process observed in the Eta Corvi star system.

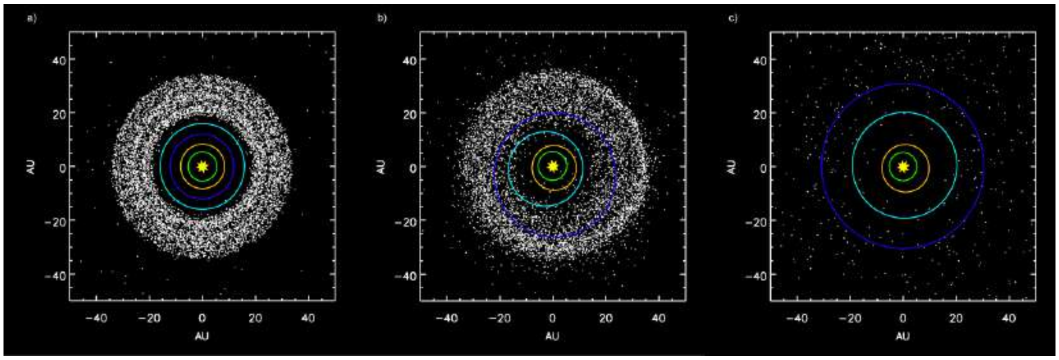


Figure 8.4.3: Simulation of before, during, and after the late heavy bombardment.

Origin of the Continents

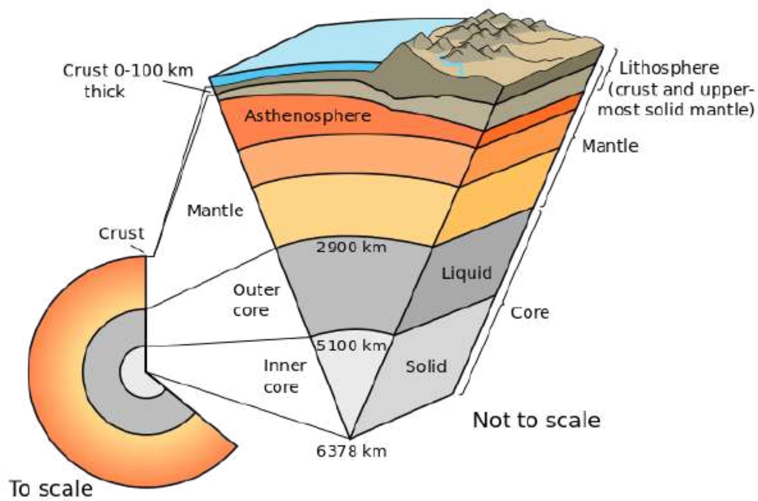


Figure 8.4.4: The layers of the Earth. Physical layers include the lithosphere and asthenosphere; chemical layers are crust, mantle, and core.

In order for plate tectonics to work as it does currently, it necessarily must have continents. However, the easiest way to create continental material is via assimilation and differentiation of existing continents (see Chapter 4). This chicken-and-egg quandary over how continents were made in the first place is not easily answered because of the great age of continental material and how much evidence has been lost during tectonics and erosion. While the timing and specific processes are still debated, volcanic action must have brought the first continental material to the Earth's surface during the Hadean, 4.4 billion years ago [18]. This model does not solve the problem of continent formation since magmatic differentiation seems to need a thicker crust. Nevertheless, the continents formed by some incremental process during the early history of Earth. The best idea is that density differences allowed lighter felsic materials to float upward and heavier ultramafic materials and metallic iron to sink. These density differences led to the layering of the Earth, the layers that are now detected by seismic studies. Early protocontinents accumulated felsic materials as developing plate-tectonic processes brought lighter material from the mantle to the surface [36].

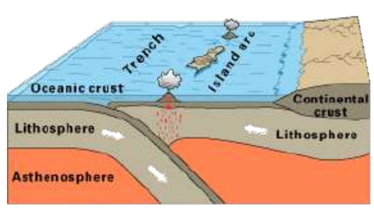


Figure 8.4.5: Subduction of an oceanic plate beneath another oceanic plate, forming a trench and an island arc. Several island arcs might combine and eventually evolve into a continent.

The first solid evidence of modern plate tectonics is found at the end of the Archean, indicating at least some continental lithosphere must have been in place. This evidence does not necessarily mark the starting point of plate tectonics; remnants of earlier tectonic activity could have been erased by the rock cycle.

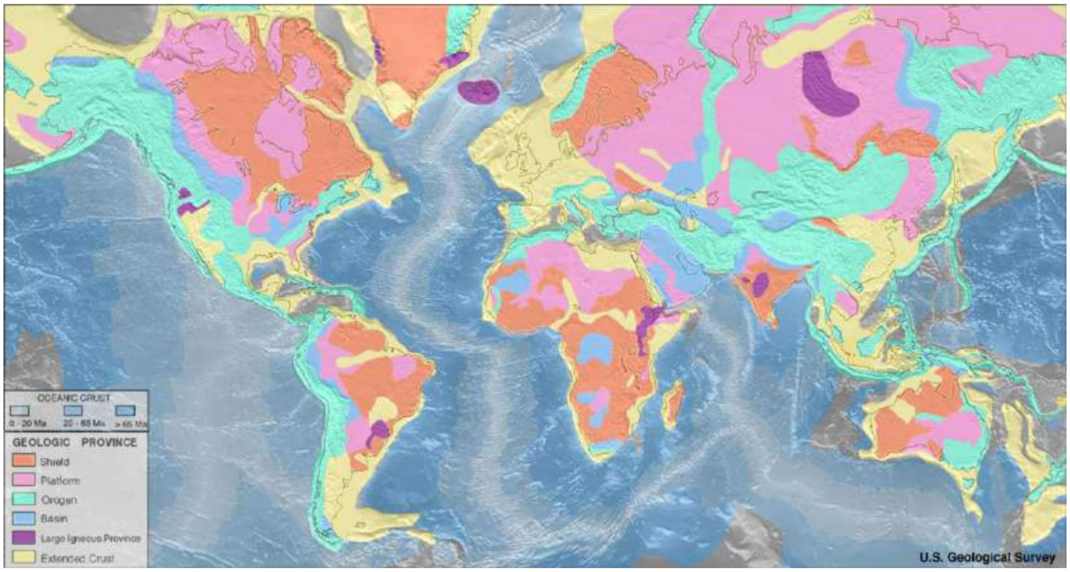


Figure 8.4.6: Geologic provinces of Earth. Cratons are pink and orange.

The stable interiors of the current continents are called **cartons** and were mostly formed in the Archean Eon. A craton has two main parts: the **shield**, which is crystalline basement rock near the surface, and the **platform** made of sedimentary rocks covering the shield. Most cratons have remained relatively unchanged with most tectonic activity having occurred around cratons instead of within them. Whether they were created by plate tectonics or another process, Archean continents gave rise to the Proterozoic continents that now dominate our planet.

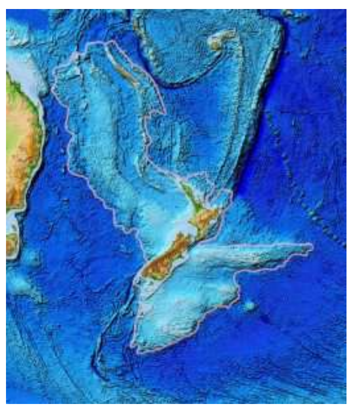


Figure 8.4.7: The continent of Zealandia

The general guideline as to what constitutes a continent and differentiates oceanic from the continental crust is under some debate. At passive margins, continental crust grades into the oceanic crust at passive margins, making a distinction difficult. Even island-arc and hot-spot material can seem more closely related to continental crust than oceanic. Continents usually have a craton in the middle with felsic igneous rocks. There is evidence that submerged masses like Zealandia, that includes present-day New Zealand, would be considered a continent [39]. Continental crust that does not contain a craton is called a continental fragment, such as the island of Madagascar off the east coast of Africa.

First Life on Earth

Life most likely started during the late Hadean or early Archean Eons. The earliest evidence of life is chemical signatures, microscopic filaments, and microbial mats. Carbon found in 4.1 billion-year-old zircon grains have a chemical signature suggesting an organic origin. Other evidence of early life is the 3.8–4.3 billion-year-old microscopic filaments from a hydrothermal vent deposit in Quebec, Canada. While the chemical and microscopic filaments evidence is not as robust as fossils, there is significant fossil evidence for life at 3.5 billion years ago. These first well-preserved fossils are photosynthetic microbial mats, called **stromatolites**, found in Australia [41].

Figure 8.4.8: Fossils of microbial mats from Sweden

Although the origin of life on Earth is unknown, hypotheses include a chemical origin in the early atmosphere and ocean, deep-sea hydrothermal vents, and delivery to Earth by comets or other objects. One hypothesis is that life arose from the chemical environment of the Earth's early atmosphere and oceans, which was very different than today. The oxygen-free atmosphere produced a reducing environment with abundant methane, carbon dioxide, sulfur, and nitrogen compounds. This is what the atmosphere is like on other bodies in the solar system.

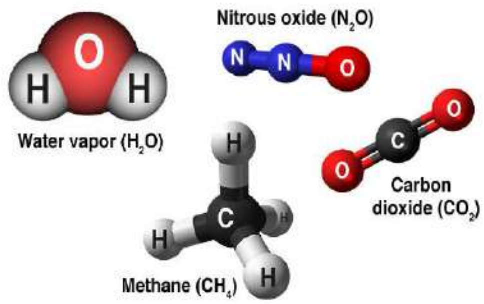


Figure 8.4.9: Greenhouse gases were more common in Earth's early atmosphere.

In the famous Miller-Urey experiment, researchers simulated early Earth's atmosphere and lightning within a sealed vessel. After igniting sparks within the vessel, they discovered the formation of amino acids, the fundamental building blocks of proteins [42]. In 1977, when scientists discovered an isolated ecosystem around hydrothermal vents on a deep-sea mid-ocean ridge (Chapter 4), it opened the door for another explanation of the origin of life. The hydrothermal vents have a unique ecosystem of critters with chemosynthesis as the foundation of the food chain instead of photosynthesis. The ecosystem is deriving its energy from hot chemical-rich waters pouring out of underground towers. This suggests that life could have started on the deep ocean floor and derived energy from the heat from the Earth's interior via chemosynthesis. Scientists have since expanded the search for life to more unconventional places, like Jupiter's icy moon Europa.

Animation of the original Miller-Urey 1959 experiment that simulated the early atmosphere and created amino acids from simple elements and compounds.

Another possibility is that life or its building blocks came to Earth from space, carried aboard comets or other objects. Amino acids, for example, have been found within comets and meteorites. This intriguing possibility also implies a high likelihood of life existing elsewhere in the cosmos.

References

- 18. Wilde SA, Valley JW, Peck WH, Graham CM (2001) Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. Nature 409:175–178
- 33. Desch SJ (2007) Mass distribution and planet formation in the solar nebula. ApJ 671:878. https://doi.org/10.1086/522825
- 36. Smithies RH, Champion DC, Cassidy KF (2003) Formation of Earth's early Archaean continental crust. Precambrian Res 127:89–101
- 39. Mortimer N, Campbell HJ, Tulloch AJ, et al (2017) Zealandia: Earth's Hidden Continent. GSA Today 27
- 41. Noffke N, Christian D, Wacey D, Hazen RM (2013) Microbially induced sedimentary structures recording an ancient ecosystem in the ca. 3.48 billion-year-old Dresser Formation, Pilbara, Western Australia. Astrobiology 13:1103–1124
- 42. Miller SL, Urey HC (1959) Organic compound synthesis on the primitive earth. Science 130:251

Contributions and Attributions

8.5: Proterozoic Eon

The **Proterozoic Eon**, meaning “earlier life,” is the eon of time after the Archean eon and ranges from 2.5 billion years old to 541 million years old. During this time, most of the central parts of the continents had formed and the plate tectonic process had started. Photosynthesis (in organisms like stromatolites) had already been adding oxygen slowly to the atmosphere, but it was quickly absorbed in minerals. Evolutionary advancements in multicellular cyanobacteria completely transformed the atmosphere by adding free oxygen gas (O_2) and causing the decimation of the anaerobic (non-oxygen) bacteria that existed at the time [47]. This is known as the **Great Oxygenation Event.** In an oxygenated world, organisms could thrive in ways they could not earlier. Oxygen also changed the chemistry of the planet in significant ways. For example, iron can be carried in solution in a non-oxygenated environment. However, when iron combines with free oxygen, it creates a solid precipitate to make minerals like hematite (iron oxide). This is the reason large deposits of iron known as banded iron formations are common during this time, ending around 2 billion years ago [48].

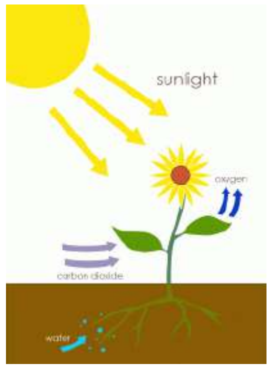


Figure 8.5.1: Diagram showing the main products and reactants in photosynthesis. The one product that is not shown is sugar, which is the chemical energy that goes into constructing the plant, and the energy that is stored in the plant which is used later by the plant or by animals that consume the plant.

The formation of the banded iron lasted a long time and prevented the oxygen level from increasing significantly in the oceans since the rocks literally took the oxygen out of the water and formed alternating layers of iron-oxide minerals and red chert. Eventually, as oxygen continued to be made, absorption of oxygen in mineral precipitation leveled off, and dissolved oxygen gas started filling the oceans and eventually bubbling out into the atmosphere. Oxygenation of the atmosphere is the single biggest event that distinguishes the Archean Earth and the Proterozoic Earth [49]. In addition to changing mineral and ocean chemistry, this event is also tabbed as the likely cause of Earth’s first glaciation, the Huron Glaciation that occurred around 2.1 billion years ago [50]. Free oxygen reacted with methane in the atmosphere, turning it into carbon dioxide. Methane is a more effective greenhouse gas than carbon dioxide, and as CO_2 increased in the atmosphere, the greenhouse effect actually decreased, thus cooling the planet.

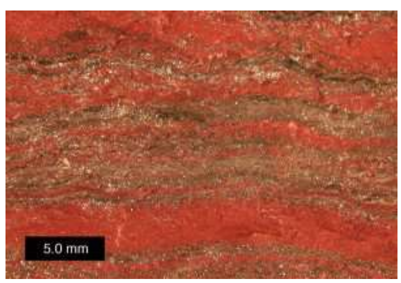


Figure 8.5.1: Alternating bands of iron-rich and silica-rich mud, formed as oxygen combined with dissolved iron.

Rodinia

By the Proterozoic eon, lithospheric plates had formed and started moving according to plate tectonic motions similar to today. As these plates formed and started moving, eventually a **supercontinent** formed from collisions as the ocean basins closed. The exact number of supercontinents during the Proterozoic (or earlier) is unknown, but **Rodinia** is the best understood. It formed about 1 billion years ago and broke up at the end of the Proterozoic, about 750-600 million years ago. **Laurentia**, the name for the continental mass that became North America, most likely was in the center of Rodinia. The reconstruction of Rodinia has been accomplished matching and aligning ancient mountain chains to assemble the pieces like a jigsaw puzzle, and paleomagnetic to orient them with respect to magnetic north.

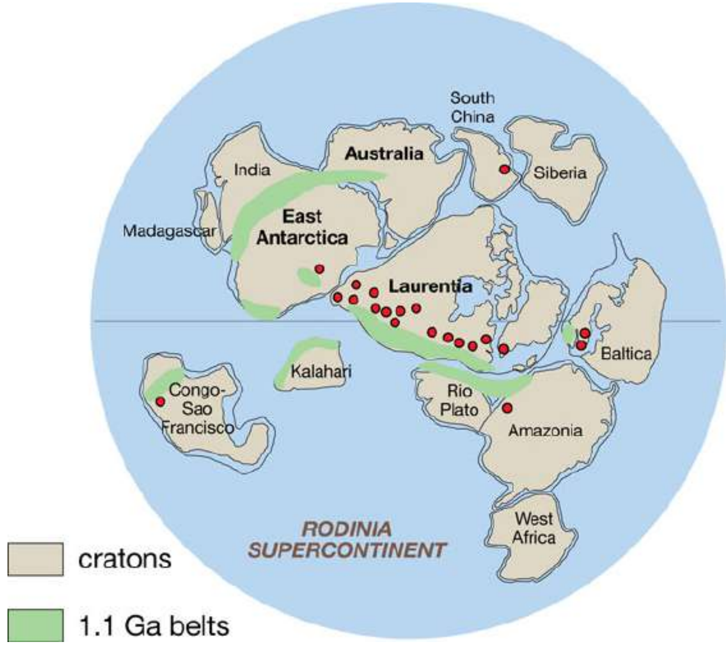


Figure 8.5.1: One possible reconstruction of Rodinia 1.1 billion years ago. Source: John Goodge, modified from Dalziel (1997).

As examples of the complexity of the issue and disagreements among geologists over the reconstructions, there are at least six different models of what broke away from Laurentia in the Panthallasa Ocean (early Pacific), including Australia, Antarctica [53], parts of China, the Tarim craton north of the Himalaya [54], Siberia [55], or the Kalahari craton of eastern Africa [56]. Regardless of the exact details, it was this breakup that created lots of biologically-favorable shallow water environments that fostered the evolutionary breakthroughs that mark the start of the next eon, the Phanerozoic.

Life Evolves

Early life in the Archean and earlier is poorly documented in the fossil record, but chemical evidence and evolutionary theory state that this life would have been single-celled photosynthetic organisms such as cyanobacteria in **stromatolites**. Fossil cyanobacteria in these stromatolites produced free oxygen in the atmosphere through photosynthesis. Cyanobacteria are **prokaryotes**, i.e. single-celled organisms (archaea and bacteria) with simple cells that lack a cell nucleus and other organelles.

Figure 8.5.1: Modern cyanobacteria (as stromatolites) in Shark Bay, Australia.

However, during the Proterozoic, a large evolutionary step occurred with the appearance of eukaryotes. Evolving around 2.1-1.6 billion years ago, **eukaryotic** cells are more complex with cell organelles and a nucleus with more complex DNA replication and regulation, mitochondria for additional energy, and chloroplasts to perform photosynthesis and produce energy. Certain organelles even have their own DNA, like mitochondria. Eukaryotes are the branch of the tree of life that gave rise to fungi, plants, and animals. About 1.2 billion years ago, another important event in Earth's biological history occurred when some eukaryotes invented sex [59]. By sharing genetic material between reproducing individuals (male and female), evolutionary change was enhanced by increasing genetic variability. This allowed more complexity among individual organisms, and eventually, ecosystems.

Figure 8.5.1: Fossil stromatolites in Saratoga Springs, New York.

It is important to realize that the Proterozoic land surfaces were barren, at least of plants like grasses, trees, and animals. Geologic processes were active just like today, but the application of the Uniformity Principle requires the realization of differences in the environments in which the processes operate. For example, rain and rivers were present but erosion on barren land surfaces would have operated at different rates than on modern land surfaces protected by plants.

Figure 8.5.1: Dickinsonia, a typical Ediacaran fossil.

The **Ediacaran fauna** (635.5-541 million years ago, [60]) offers the first glimpse at these evolving ecosystems toward the end of the Proterozoic. These organisms were among the first multicellular life forms and may have been similar to soft jellyfish or worm-like organisms [61; 62]. Since the Ediacaran fauna did not have hard parts like shells, they are not well preserved in Proterozoic rocks. However, studies suggest that they were widespread around the earth [63]. Scientists still debate how many of these are extinct evolutionary dead-ends or the ancestors to modern biological groups [61]. The transition of life from the soft-bodied Ediacaran forms to the explosion of forms with hard parts at the end of the Proterozoic and beginning of the Phanerozoic made a dramatic difference in our ability to understand earth history and the history of life.

References

47. Schirrmeister BE, de Vos JM, Antonelli A, Bagheri HC (2013) Evolution of multicellularity coincided with increased diversification of cyanobacteria and the Great Oxidation Event. Proc Natl Acad Sci U S A 110:1791–1796

48. Cloud P (1973) Paleoecological significance of the banded iron-formation. Econ Geol 68:1135–1143

49. Planavsky NJ, Asael D, Hofman A, et al (2014) Evidence for oxygenic photosynthesis half a billion years before the Great Oxidation Event. Nat Geosci 7:283–286

50. Rasmussen B, Bekker A, Fletcher IR (2013) Correlation of Paleoproterozoic glaciations based on U–Pb zircon ages for tuff beds in the Transvaal and Huronian Supergroups. Earth Planet Sci Lett 382:173–180

53. Moores EM (1991) Southwest U.S.-East Antarctic (SWEAT) connection: A hypothesis. Geology 19:425–428

54. Wen B, Evans DAD, Li Y-X (2017) Neoproterozoic paleogeography of the Tarim Block: An extended or alternative "missing-link" model for Rodinia? Earth Planet Sci Lett 458:92–106

55. Sears JW, Price RA (2000) New look at the Siberian connection: No SWEAT. Geology 28:423–42

56. Scotese CR (2009) Late Proterozoic plate tectonics and palaeogeography: a tale of two supercontinents, Rodinia and Pannotia. Geological Society, London, Special Publications 326:67–83

59. Sampson SD (2009) Dinosaur odyssey: Fossil threads in the web of life. University of California Press

60. Knoll AH, Walter MR, Narbonne GM, Christie-Blick N (2004) A new period for the geologic time scale. Science 305:621–622

61. McMenamin MAS (1986) The Garden of Ediacara. Palaios 1:178–182

62. McMenamin MA, Schulte McMenamin DL (1990) The Emergence of Animals: The Cambrian Breakthrough. Columbia University Press

63. Smith EF, Nelson LL, Strange MA, et al (2016) The end of the Ediacaran: Two new exceptionally preserved body fossil assemblages from Mount Dunfee, Nevada, USA. Geology 44:911–914

Contributions and Attributions

8.6: Paleozoic

Figure 8.6.1: The trilobites had a hard exoskeleton and were an early arthropod, the same group that includes modern insects, crustaceans, and arachnids.

The **Phanerozoic** eon is the most recent eon and represents time in which fossils are common, 541 million years ago to today. The word Phanerozoic means “visible life.” Older rocks, collectively known as the **Precambrian** (sometimes referred to as the Cryptozoic, meaning “invisible life”), are less common and have only rare fossils, and the fossils that exist represent soft-bodied life forms. The invention of hard parts like claws, scales, shells, and bones made fossils more easily preserved, and thus, easier to find. Since the younger rocks of the Phanerozoic are more common and contain the majority of fossils, the study of this eon yields much greater detail. It is further subdivided into three eras: Paleozoic (“ancient life”), Mesozoic (“middle life”), and Cenozoic (“recent life”).

Figure 8.6.1: Trilobites, by Heinrich Harder, 1916.

The **Paleozoic** era was dominated by marine organisms, but by the middle of the era, plants and animals had evolved to live and reproduce on land, including amphibians and reptiles. Fish evolved jaws and fins evolved into limbs. Lungs evolved and life emerged from the sea onto land to become the first four-legged tetrapods, amphibians. Amphibians eventually evolved into reptiles once they developed hard-shelled eggs. From reptiles evolved an early ancestor to mammals. The Carboniferous Period near the end of the Paleozoic had some of the most productive forests in the history of Earth and produced the coal that powered the industrial revolution in Europe and the United States. Tectonically, during the early Paleozoic, North America was separated from the other continents until the supercontinent Pangea formed towards the end of the era.

Paleozoic Tectonics and Paleogeography

Figure 8.6.1: Laurentia, which makes up the North American craton.

After the breakup of Rodinia toward the end of the Proterozoic, sea level remained high relative to land in the early Paleozoic. This resulted in much of **Laurentia** (considered mainly synonymous with North America) being inundated with water over the stable platforms surrounding the craton. While sea level fluctuated during transgressions and regressions after the Ordovician, many of the Paleozoic rocks found in the interior of the United States are marine in origin, due to overall relative high sea level throughout the Paleozoic.

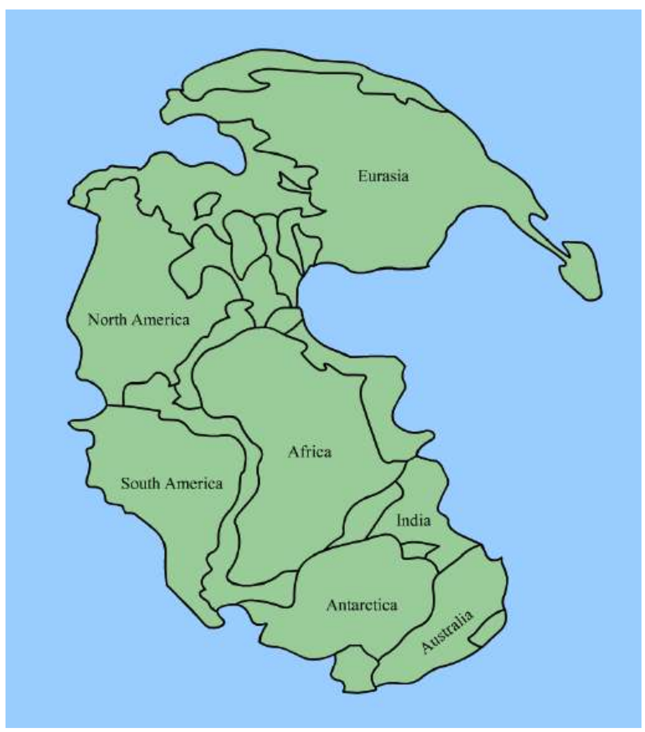
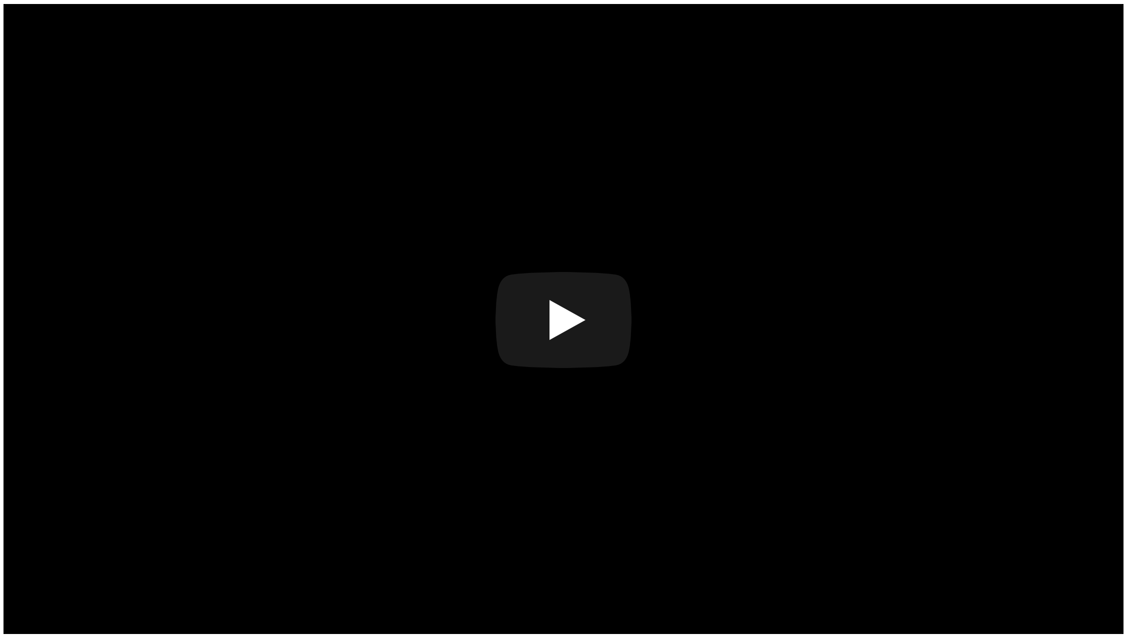
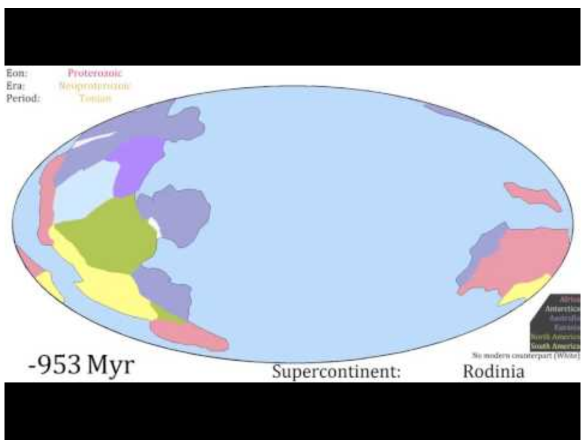


Figure 8.6.1: A reconstruction of Pangaea, showing approximate positions of modern continents.

In eastern North America, the assembly of **Pangea** (sometimes spelled Pangaea) started as early as the Cambrian with a series of events including subduction with island arcs and continental collisions and eventually ocean-basin closures known as the Taconic, Acadian, Caledonian, and Alleghanian (also known as Appalachian) orogenies [66; 68]. The name Pangea, originally coined by Alfred Wegener, means "all land." Colliding lithospheric plates formed the supercontinent, creating a series of mountain ranges and a broad fold-thrust belt, leaving a large global ocean basin known as the Panthalassa Ocean, with the Tethys Sea being the name of the large “bay” that formed between **Laurasia** (the northern continents of Laurentia and Eurasia) and **Gondwana** (the southern continents of India, Australia, Antarctica, and Africa). The eroded remains of the collisional mountains formed on Pangea are still in existence today as the Appalachian, Alleghanian, Scandinavian, Marathon, and Ouachita Mountain ranges. Stress from the Alleghanian orogeny reactivated faulting, produced uplifts, and deformation/folding as far west as the Pennsylvanian-aged Ancestral Rocky Mountains of Colorado.





Animation of plate movement in the last 3.3 billion years. Pangea occurs at the 4:40 mark.

Tectonics in western North America during the early part of the Paleozoic was mostly mild, as a long-lived passive margin developed. After the start of the Devonian, the Antler orogeny finally caused faulting and basin development, mostly seen across Nevada today. The Antler belt is most likely a result of an island arc crashing into western North America [71].

Paleozoic Evolution

Figure 8.6.1: Anomalocaris reconstruction by the MUSE science museum in Italy.

The earliest Paleozoic had a significant biological explosion and contains evidence of a wide variety of evolutionary paths, including the evolutionary invention of hard parts like shells, spikes, teeth, and scales. Paleontologists refer to this event as the **Cambrian Explosion**, named after the first period in the Paleozoic. Scientists debate whether this was a manifestation of a true evolutionary pattern of diversification, better preservation from easier to fossilize creatures, or simply an artifact of a more complete recent rock record. Ediacaran fauna, which lacked easily-fossilized hard parts, may have already been diverse and set the stage for the Cambrian Explosion [72]. Regardless, during the Cambrian period, 541-485 million years ago, a large majority of the phyla of modern marine animals appeared [73]. These new organisms had simple cone- or tube-shaped shells that quickly became more complex. Some of these life forms have survived to today, and some were “experimental” whose lineage did not continue past the Cambrian period. Fossil evidence of this time was first discovered by Charles Walcott in a rock layer called the Burgess Shale in western Canada in 1909.

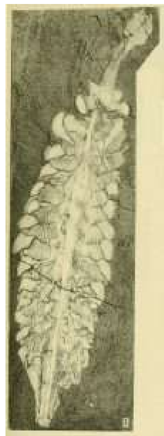


Figure 8.6.1: Original plate from Walcott’s 1912 description of Opabinia, with labels: fp = frontal appendage, e = eye, ths = thoracic somites, i = intestine, ab = abdominal segment.

The Burgess Shale is a **lagerstätte**, or fossil site of exceptional preservation, including impressions of soft body parts. This allowed scientists to learn immense details of the animals that existed at the time, in addition to their tough shells, spikes, and claws. Other lagerstätte sites of similar age in China and Utah have allowed the forming of a fairly detailed picture of what the biodiversity was like in the Cambrian. The biggest mystery is the animals that do not fit existing lineages and are unique to that time. This includes famous fossil creatures like the first compound-eyed trilobites, and many other strange ones, including *Wiwaxia*, a spiked shell creature; *Hallucigenia*, a walking worm with spikes; *Opabinia*, a 5-eyed lobed arthropod with a trunk and a grappling claw at the end; and the related *Anomalocaris*, the alpha predator of the time, complete with grabbing arms and a deadly circular mouth full of teeth. Most notably at this time, an important ancestor to humans evolved. *Pikaia*, a segmented worm, is thought to be the earliest ancestor of the **Chordata** phylum (including vertebrates; animals with backbones [76]). These astonishing creatures offer a glimpse at evolutionary creativity. At the end of the Cambrian, mollusks, brachiopods, nautiloids, gastropods, graptolites, echinoderms and trilobites had evolved and shared the seafloor.

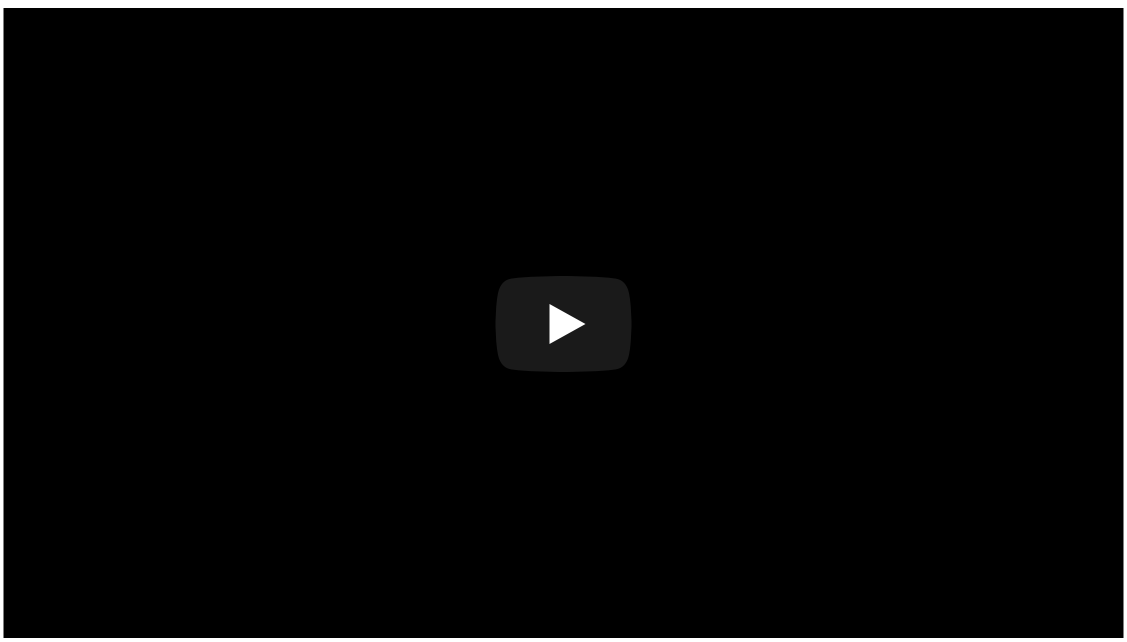
Figure 8.6.1: A modern coral reef.

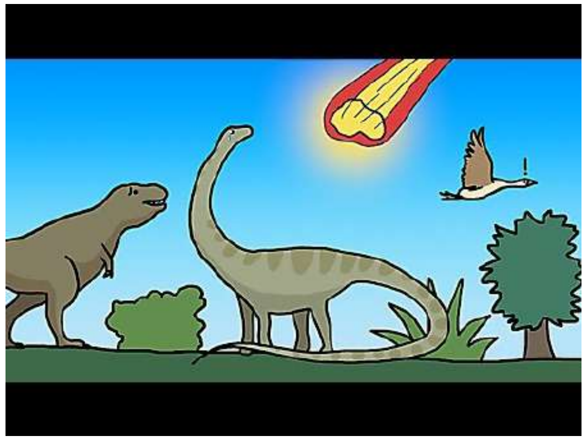
After the Cambrian Explosion, a similar event occurred which abandoned some of the Cambrian evolutionary animal lines and proliferated others. Known as the Ordovician Radiation or Great Ordovician Biodiversification Event, many common forms and ecosystems recognizable today became common. This includes invertebrates such as mollusks (clams and their relatives), corals, arthropods (insects and their relatives), and vertebrates became more diverse and complex and dominated the oceans [77].

Figure 8.6.1: Guadalupe National Park is made of a giant fossil reef complex.

The most important of these advancements may have been reef-building organisms. Mostly colonial coral, they took advantage of better ocean chemistry for calcite and built large structures [78], resembling modern reefs like the Great Barrier Reef off of Australia. Many of the organisms of this time swam around, hid in, or crawled over the reefs. Reefs are so important because of their preservation potential, size (some reef fossils are the size of mountains), and the ability to create an in-place ecosystem in and around them. Few other fossil assemblages in the geologic record can offer more diversity and complexity than reefs. Warm temperatures and high sea levels in the Ordovician most likely helped spur this diversification.

A small ice age based on evidence of glacial deposits and associated falling sea levels led to the dramatic mass extinction by the end of the Ordovician, the first one documented in the fossil record. **Mass extinction** is when an unusually large number of species abruptly vanish and go extinct, and this can be observed in the fossil record (see video below). Life bounced back in the Silurian [78]. The major evolutionary event was the development of the forward pair of gill arches into jaws, allowing fish new feeding strategies and opening up new ecological niches.





3-minute video describing mass extinctions and how they are defined.

Figure 8.6.1: The armor-plated fish (placoderm) Bothriolepis panderi from the Devonian of Russia.

The Silurian provides the first evidence of land plants and animals [80; 81]. This includes the first-ever vascular plant, *Cooksonia*, with woody tissues, veins for transporting water and food, seeds, and roots. The first bony fish and shark are also Silurian, which includes the first primitive jaws. This also saw the start of armored fish, known as the placoderms. Insects, spiders, scorpions, and crustaceans began to inhabit dry-land and freshwater habitats.

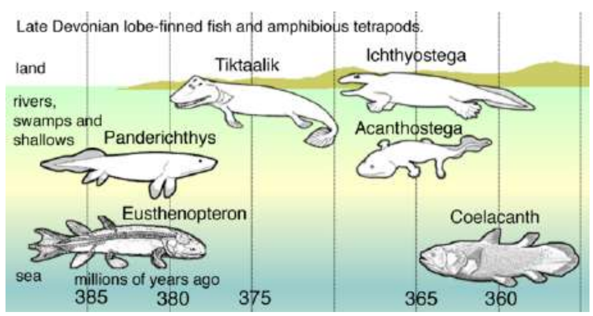


Figure 8.6.1: Several different types of fish and amphibians that led to walking on land.

The Devonian, called the age of fishes, saw a rise in plated fish and jawed fish [85], along with the lobe-finned fish. The lobe-finned fish (relatives of the modern lungfish and coelacanth) are important for their eventual evolution into tetrapods, the four-limbed creatures that went on to dominate the land. The first evidence of walking fish, named *Tiktaalik* (about 375 million years ago), gave rise to amphibians. [87]. Most amphibians live on land but lay soft eggs in water. They would later evolve into reptiles that lay hard-shelled eggs on land. Land plants had also evolved into the first trees and forests [88]. Toward the end of the Devonian, another mass extinction event occurred. This extinction, while severe, is the least temporally defined, with wide variations in the timing of the event or events. Reef-building organisms were the hardest hit, leading to dramatic changes in marine ecosystems [89].

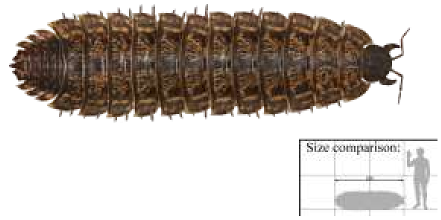


Figure 8.6.1: A reconstruction of the giant arthropod (insects and their relatives) Arthropleura.

The next time period called the Carboniferous (North American geologists have subdivided this into the Mississippian and Pennsylvanian periods), saw the highest levels of oxygen ever known, with forests (e.g., ferns, clubmosses) and swamps dominating the landscape [90]. This helped cause the largest arthropods ever, like the millipede *Arthropleura*, at 2.5 meters (6.4 feet) long! It also saw the rise of a new group of animals, the reptiles. The evolutionary advantage that reptiles have over amphibians is the amniote egg (egg with a protective shell), which allows them to rely on non-aquatic environments for reproduction. This widened the terrestrial reach of reptiles compared to amphibians. This booming life, especially plants life, created cooling temperatures as carbon dioxide was removed from the atmosphere [92]. By the middle Carboniferous, these cooler temperatures led to an ice age (called the Karoo Glaciation) and less-productive forests. The reptiles fared much better than the amphibians, leading to their diversification [93]. This glacial event lasted into the early Permian [94].

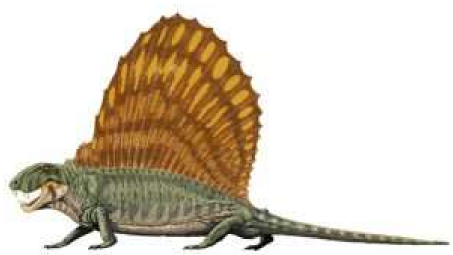


Figure 8.6.1: Reconstruction of Dimetrodon.

By the Permian, with Pangea assembled, the supercontinent led to a dryer climate and even more diversification and domination by the reptiles [95]. The groups that developed in this warm-climate eventually radiated into dinosaurs. Another group, known as the synapsids, eventually evolved into mammals. Synapsids, including the famous sail-backed *Dimetrodon* are commonly confused with dinosaurs. Pelycosaurs (of the Pennsylvanian to early Permian like *Dimetrodon)* are the first group of synapsids that exhibit the beginnings of mammalian characteristics such as well-differentiated dentition: incisors, highly developed canines in lower and upper jaws and cheek teeth, premolars and molars. Starting in the late Permian, the second group of synapsids, called the therapsids (or mammal-like reptiles) evolve, and become the ancestors to mammals.

Permian Mass Extinction

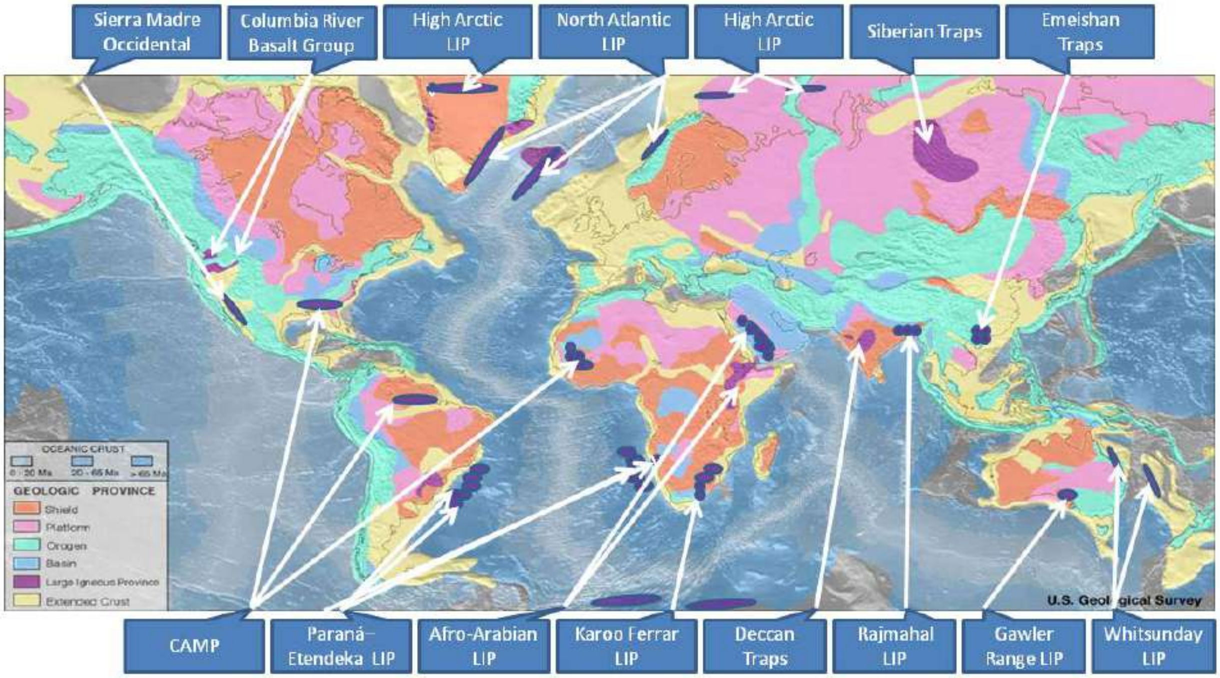


Figure 8.6.1: Map of global flood basalts. Note the largest is the Siberian Traps.

The end of the Paleozoic era is marked by the largest mass extinction in earth history. The Paleozoic era had two smaller mass extinctions, but these were not as large as the **Permian Mass Extinction**, also known as the Permian-Triassic Extinction Event. It is estimated that up to 96% of marine species and 70% of land-dwelling (terrestrial) vertebrates went extinct. Many famous organisms, like sea scorpions and trilobites, were never seen again in the fossil record. What caused such a widespread extinction event? The exact cause is still debated, though the leading idea relates to extensive volcanism associated with the **Siberian Traps**, which are one of the largest deposits of flood basalts known on Earth, dating to the time of the extinction event [100]. The eruption size is estimated at over 3 million cubic kilometers [101] that is approximately 4,000,000 times larger than the famous 1980 Mt. St. Helens eruption in Washington. The unusually large volcanic eruption would have contributed a large amount of toxic gases, aerosols, and greenhouse gasses into the atmosphere. Further, some evidence suggests that volcanism burned vast coal deposits releasing methane (a greenhouse gas) into the atmosphere. As discussed in Chapter 15, greenhouse gases cause the climate to warm. This extensive addition of greenhouse gases from the Siberian Traps may have caused a runaway greenhouse effect that rapidly changed the climate, acidified the oceans, disrupted food chains, disrupted carbon cycling, and caused the largest mass extinction].

References

66. Rodgers J (1971) The Taconic Orogeny. Geol Soc Am Bull 82:1141–1178

68. McKerrow WS, Mac Niocaill C, Dewey JF (2000) The Caledonian orogeny redefined. J Geol Soc London 157:1149–1154

71. Speed RC, Sleep NH (1982) Antler orogeny and foreland basin: A model. Geol Soc Am Bull 93:815–828

72. Schiffbauer JD, Huntley JW, O'Neil GR, et al (2016) The latest Ediacaran Wormworld fauna: Setting the ecological stage for the Cambrian Explosion. GSA Today 26

73. Morris SC (1998) The crucible of creation: the Burgess Shale and the rise of animals. Peterson's

76. Morris SC, Caron J-B (2012) Pikaia gracilens, a stem-group chordate from the Middle Cambrian of British Columbia. Biol Rev Camb Philos Soc 87:480–512

77. Webby BD, Paris F, Droser ML, Percival IG (2012) The Great Ordovician Biodiversification Event. Columbia University Press

78. Munnecke A, Calner M, Harper DAT, Servais T (2010) Ordovician and Silurian sea–water chemistry, sea level, and climate: A synopsis. Palaeogeogr Palaeoclimatol Palaeoecol 296:389–413

80. Niklas KJ (1976) Chemical examinations of some non-vascular Paleozoic plants. Brittonia 28:113–137

81. Selden P, Read H (2007) The oldest land animals: Silurian millipedes from Scotland. Editors 36

85. Rücklin M, Donoghue PCJ, Johanson Z, et al (2012) Development of teeth and jaws in the earliest jawed vertebrates. Nature 491:748–751

87. Niedźwiedzki G, Szrek P, Narkiewicz K, et al (2010) Tetrapod trackways from the early Middle Devonian period of Poland. Nature 463:43–48

88. Stein WE, Mannolini F, Hernick LV, et al (2007) Giant cladoxylopsid trees resolve the enigma of the Earth's earliest forest stumps at Gilboa. Nature 446:904–907

89. McGhee GR (1996) The Late Devonian Mass Extinction: The Frasnian/Famennian Crisis. Columbia University Press

90. Beerling D (2008) The emerald planet: how plants changed Earth's history. OUP Oxford

92. Montañez IP, Tabor NJ, Niemeier D, et al (2007) CO2-forced climate and vegetation instability during Late Paleozoic deglaciation. Science 315:87–91

93. Sahney S, Benton MJ, Falcon-Lang HJ (2010) Rainforest collapse triggered Carboniferous tetrapod diversification in Euramerica. Geology

94. López-Gamundí OR, Buatois LA (2010) Introduction: Late Paleozoic glacial events and postglacial transgressions in Gondwana. Geological Society of America Special Papers 468:v–viii

95. Parrish JT (1993) Climate of the supercontinent Pangea. J Geol 101:215–233

100. Burgess, S. D.; Muirhead, J. D.; Bowring, S. A. (2017) Initial pulse of Siberian Traps sills as the trigger of the end-Permian mass extinction. Nature Communications 8

101. Reichow MK, Pringle MS, Al'Mukhamedov AI, et al (2009) The timing and extent of the eruption of the Siberian Traps large igneous province: Implications for the end-Permian environmental crisis. Earth Planet Sci Lett 277:9–20

Contributions and Attributions

8.7: Mesozoic

Figure 8.7.1: Perhaps the greatest fossil ever found, a velociraptor attacked a protoceratops, and both were fossilized mid sequence.

Following the Permian Mass Extinction, the **Mesozoic** ("middle life") was from 252 million years ago to 66 million years ago. As Pangea started to break apart, mammals, birds, and flowering plants developed. The Mesozoic is probably best known as the age of reptiles, most notably, the dinosaurs.

Mesozoic Tectonics and Paleogeography

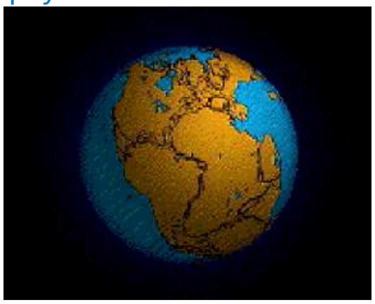


Figure 8.7.1: Animation showing Pangea breaking up.

Pangea started breaking up (in a region that would become eastern Canada and United States) around 210 million years ago in the Late Triassic. Clear evidence for this includes the age of the sediments in the Newark Supergroup rift basins and the Palisades sill of the eastern part of North America and the age of the Atlantic ocean floor. Due to sea-floor spreading (Chapter 3), the oldest rocks on the Atlantic's floor are along the coast of northern Africa and the east coast of North America, while the youngest are along the mid-ocean ridge.

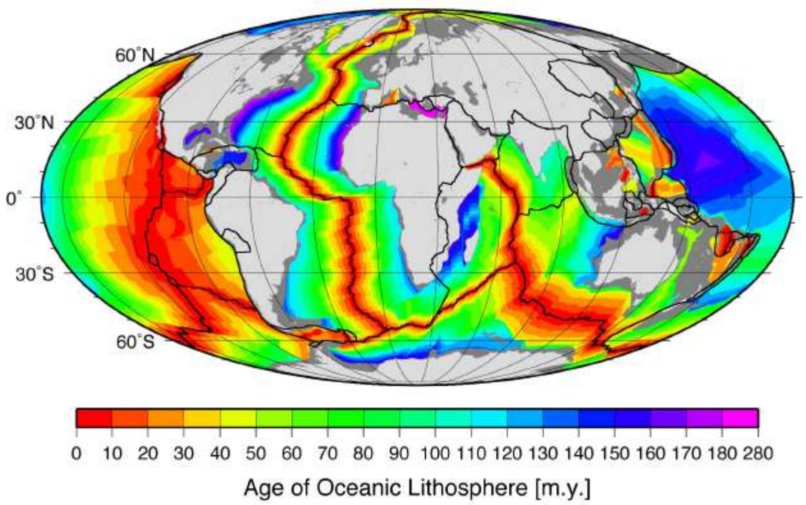


Figure 8.7.1: Age of oceanic lithosphere, in millions of years. Notice the differences in the Atlantic Ocean along the coasts of the continents.

This age pattern shows how the Atlantic Ocean opened as the young Mid-Atlantic Ridge began to create the seafloor. This means the Atlantic ocean started opening and was first formed here. The southern Atlantic opened next, with South America separating from central and southern Africa. Last (happening after the Mesozoic ended) was the northernmost Atlantic, with Greenland and

Scandinavia parting ways. The breaking points of each rifted plate margin eventually turned into the passive plate boundaries of the east coast of the Americas today

Video of Pangea breaking apart and plates moving to their present locations. By Tanya Atwater.

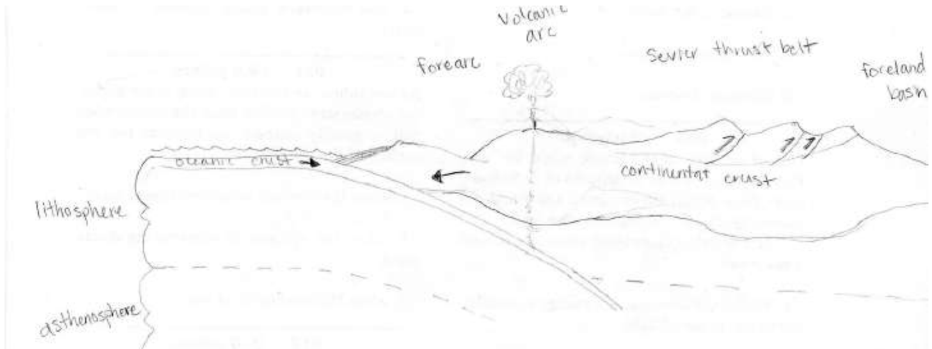


Figure 8.7.1: Sketch of the major features of the Sevier Orogeny.

In western North America, an active plate margin had started with subduction, controlling most of the tectonics of that region in the Mesozoic. Another possible island-arc collision created the Sonoman Orogeny in Nevada during the latest Paleozoic to the Triassic. In the Jurassic, another island-arc collision caused the Nevadan Orogeny, a large Andean-style volcanic arc and thrust belt [105]. The Sevier Orogeny followed in the Cretaceous, which was mainly a volcanic arc to the west and a thin-skinned fold and thrust belt to the east, meaning stacks of shallow faults and folds built up the topography. Many of the structures in the Rocky Mountains today date from this orogeny.

Figure 8.7.1: The Cretaceous Interior Seaway in the mid-Cretaceous.

Tectonics had an influence in one more important geographic feature in North America: the Cretaceous Western Interior Foreland Basin, which flooded during high sea levels forming the **Cretaceous Interior Seaway**. Subduction from the west was the Farallon Plate, an oceanic plate connected to the Pacific Plate (seen today as remnants such as the Juan de Fuca Plate, off the coast of the Pacific Northwest). Subduction was shallow at this time because a very young, hot and less dense portion of the Farallon plate was subducted. This shallow subduction caused a down warping in the central part of North America [107]. High sea levels due to shallow subduction, and increasing rates of seafloor spreading and subduction, high temperatures, and melted ice also contributed to the high sea levels [108]. These factors allowed a shallow epicontinental seaway that extended from the Gulf of Mexico to the Arctic Ocean to divide North America into two separate landmasses [109], Laramidia to the west and Appalachia to the east, for 25 million years [59]. Many of the coal deposits in Utah and Wyoming formed from swamps along the shores of this seaway [111]. By the end of the Cretaceous, cooling temperatures caused the seaway to regress [112].

Mesozoic Evolution

Figure 8.7.1: A Mesozoic scene from the late Jurassic.

The Mesozoic era is dominated by reptiles, and more specifically, the dinosaurs. The Triassic saw devastated ecosystems that took over 30 million years to fully re-emerge after the Permian Mass Extinction [113]. The first appearance of many modern groups of animals that would later flourish occurred at this time. This includes frogs (amphibians), turtles (reptiles), marine ichthyosaurs and plesiosaurs (marine reptiles), mammals, and the archosaurs. The archosaurs ("ruling reptiles") include ancestral groups that went extinct at the end of the Triassic, as well as the flying pterosaurs, crocodilians, and the dinosaurs. Archosaurs, like the placental mammals after them, occupied all major environments: terrestrial (dinosaurs), in the air (pterosaurs), aquatic (crocodilians) and even fully marine habitats (marine crocodiles). The pterosaurs, the first vertebrate group to take flight, like the dinosaurs and mammals, start small in the Triassic.

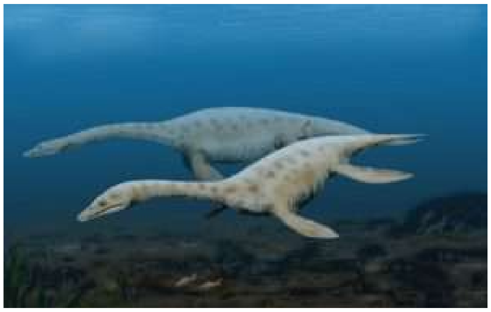


Figure 8.7.1: A drawing of the early plesiosaur Agustasaurus from the Triassic of Nevada.

At the end of the Triassic, another mass extinction event occurred [114], the fourth major mass extinction in the geologic record. This was perhaps caused by the Central Atlantic Magmatic Province flood basalt [115]. The end-Triassic extinction made certain lineages go extinct and helped spur the evolution of survivors like mammals, pterosaurs (flying reptiles), ichthyosaurs/plesiosaurs/mosasaurs (marine reptiles), and dinosaurs.

Figure 8.7.1: Reconstruction of the small (<5″) Megazostrodon, one of the first animals considered to be a true mammal.

Mammals, as previously mentioned, got their start from a reptilian synapsid ancestor possibly in the late Paleozoic. Mammals stayed small, in mainly nocturnal niches, with insects being their largest prey. The development of warm-blooded circulation and fur may have been a response to this lifestyle [118].

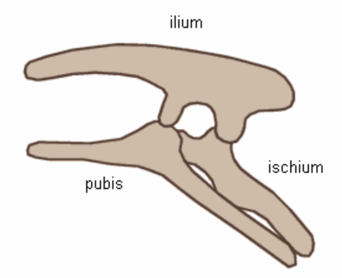


Figure 8.7.1: Closed structure of an ornithischian hip, which is similar to birds'.

In the Jurassic, species that were previously common flourished due to a warmer and more tropical climate. The dinosaurs were relatively small animals in the Triassic period of the Mesozoic but became truly massive in the Jurassic. Dinosaurs are split into two groups based on their hip structure [120], i.e. orientation of the pubis and ischium bones in relationship to each other. This is referred to as the "reptile hipped" saurischians and the "bird-hipped" ornithischians. This has recently been brought into question by a new idea for dinosaur lineage [121].

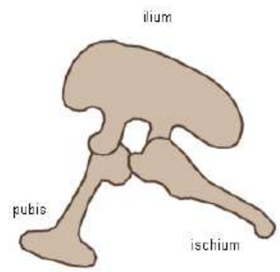


Figure 8.7.1: Open structure of a saurischian hip, which is similar to lizards'.

Most of the dinosaurs of the Triassic were saurischians, but all of them were bipedal. The major adaptive advantage dinosaurs had was changes in the hip and ankle bones, tucking the legs under the body for improved locomotion as opposed to the semi-erect gait of crocodiles or the sprawling posture of reptiles. In the Jurassic, limbs (or a lack thereof) were also important to another group of reptiles, leading to the evolution of *Eophis*, the oldest snake.

Figure 8.7.1: Therizinosaurs, like Beipiaosaurus (shown in this restoration), are known for their enormous hand claws.

There is a paucity of dinosaur fossils from the Early and Middle Jurassic but by the Late Jurassic, they were dominating the planet. The saurischians diversified into the giant herbivorous (plant-eating) long-necked sauropods weighing up to 100 tons and bipedal carnivorous theropods, with the possible exception of the *Therizinosaurs* [125]. All of the ornithischians (e.g *Stegosaurus, Iguanodon, Triceratops, Ankylosaurus, Pachycephalosaurus*) were herbivorous with a strong tendency to have a "turtle-like" beak at the tips of their mouths.

Figure 8.7.1: Iconic "Berlin specimen" Archaeopteryx lithographica fossil from Germany.

The pterosaurs grew and diversified in the Jurassic, and another notable arial organism developed and thrived in the Jurassic: birds. When *Archaeopteryx* was found in the Solnhofen Lagerstätte of Germany, a seeming dinosaur-bird hybrid, it started the conversation on the origin of birds. The idea that birds evolved from dinosaurs occurred very early in the history of research into evolution, only a few years after Darwin's *On the Origin of Species* [127]. This study used a remarkable fossil of *Archeopteryx* from a transitional animal between dinosaurs and birds. Small meat-eating theropod dinosaurs were likely the branch that became birds due to their similar features [128]. A significant debate still exists over how and when powered flight evolved. Some have stated a running-start model [129], while others have favored a tree-leaping gliding model or even a semi-combination: flapping to aid in climbing.

Figure 8.7.1: Reconstructed skeleton of Argentinosaurus, from Naturmuseum Senckenberg in Germany.

The Cretaceous saw a further diversification, specialization, and domination of the dinosaurs and other fauna. One of the biggest changes on land was the transition to angiosperm-dominated flora. Angiosperms, which are plants with flowers and seeds, had originated in the Cretaceous [132], switching many plains to grasslands by the end of the Mesozoic [133]. By the end of the period, they had replaced gymnosperms (evergreen trees) and ferns as the dominant plant in the world's forests. Haplodiploid eusocial insects (bees and ants) are descendants from Jurassic wasp-like ancestors that co-evolved with the flowering plants during this time period. The breakup of Pangea not only shaped our modern world's geography but biodiversity at the time as well. Throughout the Mesozoic, animals on the isolated, now separated island continents (formerly parts of Pangea), took strange evolutionary turns. This includes giant titanosaurian sauropods (*Argentinosaurus*) and theropods (*Giganotosaurus*) from South America.

K-T Extinction

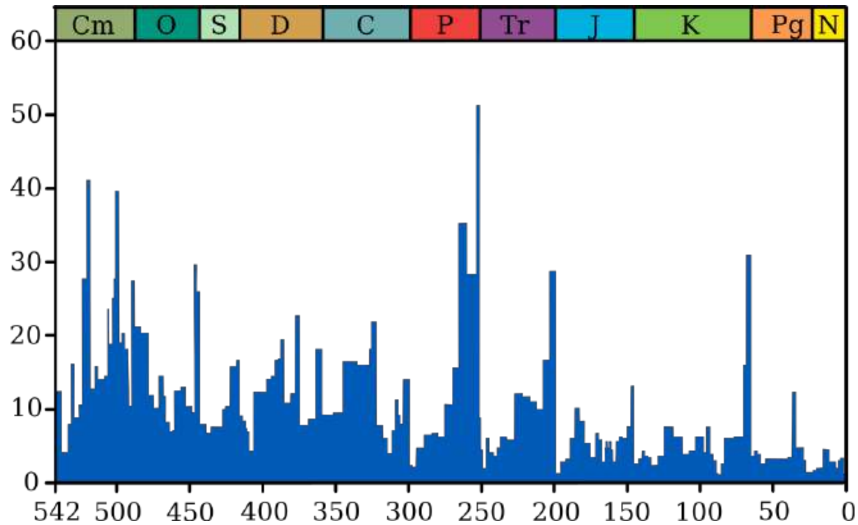


Figure 8.7.1: Graph of the rate of extinctions. Note the large spike at the end of the Cretaceous (labeled as K).

Similar to the end of the Paleozoic era, the Mesozoic Era ended with the **K-Pg Mass Extinction** (previously known as the **K-T Extinction**) 66 million years ago [136]. This extinction event was likely caused by a large **bolide** (an extraterrestrial impactor such as an asteroid, meteoroid, or comet) that collided with earth. Ninety percent of plankton species, 75% of plant species, and all the dinosaurs went extinct at this time.


Figure 8.7.1: Artist's depiction of the impact event

One of the strongest pieces of evidence comes from the element iridium. Quite rare on Earth, and more common in meteorites, it has been found all over the world in higher concentrations at a particular layer of rock that formed at the time of the K-T boundary. Soon other scientists started to find evidence to back up the claim. Melted rock spheres [138], a special type of "shocked" quartz called stishovite, that only is found at impact sites, was found in many places around the world. The huge impact created a strong thermal pulse that could be responsible for global forest fires [141], strong acid rains [142], a corresponding abundance of ferns, the first colonizing plants after a forest fire [143], enough debris thrown into the air to significantly cool temperatures afterward [144; 145], and a 2-km high tsunami inferred from deposits found from Texas to Alabama.

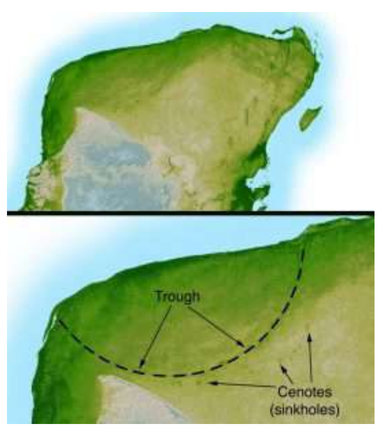


Figure 8.7.1: The land expression of the Chicxulub crater. The other side of the crater is within the Gulf of México.

Still, with all this evidence, one large piece remained missing: the crater where the bolide impact. It was not until 1991 that the crater was confirmed using petroleum company geophysical data. Even though it is the third-largest confirmed crater on Earth at roughly 180 km wide, the **Chicxulub Crater** was hard to find due to being partially underwater and partially obscured by the dense forest canopy of the Yucatan Peninsula. Coring of the center of the impact called the peak ring contained granite, indicating the impact was so powerful that it lifted basement sediment from the crust several miles toward the surface [149]. In 2010, an international team of scientists reviewed 20 years of research and blamed the impact for the extinction [150].

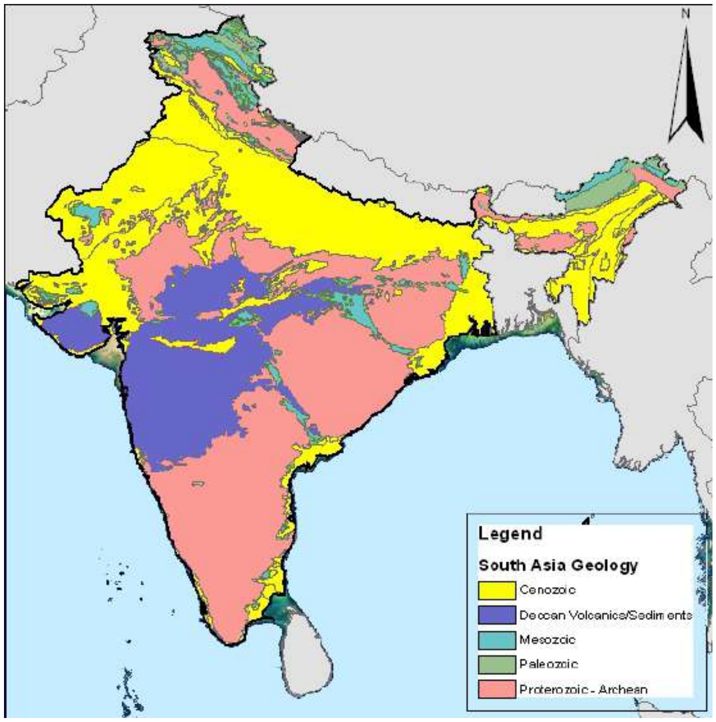


Figure 8.7.1: Geology of India, showing purple as Deccan Traps-related rocks.

With all of this information, it seems like the case would be closed. However, there are other events at this time which could have partially aided the demise of so many organisms. For example, sea levels are known to be slowly decreasing at the time of the K-T event, which is tied to marine extinctions [151], though any study on gradual vs. sudden changes in the fossil record is flawed due to the incomplete nature of the fossil record [152]. Another big event at this time was the **Deccan Traps** flood basalt volcanism in India. At over 1.3 million cubic kilometers of material, it was certainly a large source of material hazardous to ecosystems at the time, and it has been suggested as at least partially responsible for the extinction [153]. Some have found the impact and eruptions too much of a coincidence, and have even linked the two together [154].

References

59. Sampson SD (2009) Dinosaur odyssey: Fossil threads in the web of life. University of California Press

105. Schweickert RA, Bogen NL, Girty GH, et al (1984) Timing and structural expression of the Nevadan orogeny, Sierra Nevada, California. Geol Soc Am Bull 95:967–979

107. Miller KG, Kominz MA, Browning JV, et al (2005) The Phanerozoic record of global sea-level change. Science 310:1293–1298

108. Miall AD (2008) Chapter 5 The Paleozoic Western Craton Margin. In: Andrew D. Miall (ed) Sedimentary Basins of the World. Elsevier, pp 181–209

109. Mitrovica JX, Beaumont C, Jarvis GT (1989) Tilting of continental interiors by the dynamical effects of subduction. Tectonics

111. Ryer TA (1983) Transgressive-regressive cycles and the occurrence of coal in some Upper Cretaceous strata of Utah. Geology 11:207–210

112. Mörner N-A (1981) Revolution in Cretaceous sea-level analysis. Geology 9:344–346

113. Sahney S, Benton MJ (2008) Recovery from the most profound mass extinction of all time. Proc Biol Sci 275:759–765

114. Lucas SG, Tanner LH (2004) Late Triassic extinction events. Albertiana

115. Tanner LH, Lucas SG, Chapman MG (2004) Assessing the record and causes of Late Triassic extinctions. Earth-Sci Rev 65:103–139

118. Ruben JA, Jones TD (2000) Selective factors associated with the origin of fur and feathers. Am Zool 40:585–596

120. Seeley HG (1887) On the classification of the fossil animals commonly named Dinosauria. Proc R Soc Lond 43:165–171

121. Matthew G. Baron, David B. Norman, Paul M. Barrett (2017) A new hypothesis of dinosaur relationships and early dinosaur evolution. Nature 543:501–506

125. Maryanska T, Farlow JO, Brett-Surman MK (1997) Segnosaurs (Therizinosaurs). The Complete Dinosaur Indiana University Press, Bloomington, Indiana 234–241

127. Owen (1863) On the Archeopteryx of Von Meyer, with a Description of the Fossil Remains of a Long-Tailed Species, from the Lithographic Stone of Solenhofen. Philosophical Transactions of the Royal Society of London 153:33–47

128. Chiappe LM (2009) Downsized dinosaurs: The evolutionary transition to modern birds. Evo Edu Outreach 2:248–256

129. Ostrom JH (1974) Archaeopteryx and the Origin of Flight. Q Rev Biol 49:27–47

132. Sun G, Ji Q, Dilcher DL, et al (2002) Archaefructaceae, a new basal angiosperm family. Science 296:899–904

133. Piperno DR, Sues H-D (2005) Dinosaurs dined on grass. Science 310:1126–1128. https://doi.org/10.1126/science.1121020

136. Renne PR, Deino AL, Hilgen FJ, et al (2013) Time scales of critical events around the Cretaceous-Paleogene boundary

138. Smit J, Klaver G (1981) Sanidine spherules at the Cretaceous–Tertiary boundary indicate a large impact event. Nature

141. Melosh HJ, Schneider NM, Zahnle KJ, Latham D (1990) Ignition of global wildfires at the Cretaceous/Tertiary boundary. Nature 343:251–254

142. Ohno S, Kadono T, Kurosawa K, et al (2014) Production of sulphate-rich vapour during the Chicxulub impact and implications for ocean acidification. Nat Geosci 7:279–282

143. Vajda V, Raine JI, Hollis CJ (2001) Indication of global deforestation at the Cretaceous-Tertiary boundary by New Zealand fern spike. Science 294:1700–1702

144. Pope KO, Baines KH, Ocampo AC, Ivanov BA (1997) Energy, volatile production, and climatic effects of the Chicxulub Cretaceous/Tertiary impact. J Geophys Res 102:21645–21664

145. Pollack JB, Toon OB, Ackerman TP, et al (1983) Environmental effects of an impact-generated dust cloud: implications for the cretaceous-tertiary extinctions. Science 219:287–289

149. Morgan JV, Gulick S, Bralower T, et al (2016) The formation of peak rings in large impact craters

150. Schulte P, Alegret L, Arenillas I, et al (2010) The Chicxulub asteroid impact and mass extinction at the Cretaceous-Paleogene boundary. Science 327:1214–1218

151. Marshall CR, Ward PD (1996) Sudden and Gradual Molluscan Extinctions in the Latest Cretaceous of Western European Tethys. Science 274:1360–1363

152. MacLeod N, Rawson PF, Forey PL, et al (1997) The Cretaceous-tertiary biotic transition. J Geol Soc London 154:265–292

153. Schoene B, Samperton KM, Eddy MP, et al (2015) U-Pb geochronology of the Deccan Traps and relation to the end-Cretaceous mass extinction. Science 347:182–184

154. Richards MA, Alvarez W, Self S, et al (2015) Triggering of the largest Deccan eruptions by the Chicxulub impact. Geol Soc Am Bull 127:1507–1520

Contributions and Attributions

8.8: Cenozoic

Figure 8.8.1: Paraceratherium, seen in this reconstruction, was a massive (15-20 ton, 15 foot tall) ancestor of rhinos.

The **Cenozoic**, meaning “new life,” is known as the age of mammals because it is in this era that mammals came to be a dominant and large life form, including human ancestors. Birds, as well, flourished in the open niches left by the dinosaur’s demise. Most of the Cenozoic has been relatively warm, with the main exception being the ice age that started about 2.558 million years ago and (despite recent warming) continues today. Tectonic shifts in the west caused volcanism but eventually changed the long-standing subduction zone into a transform boundary.

Cenozoic Tectonics and Paleogeography

e last 38 million years of movement in western North America. Note, that after the ridge is subducted, convergent turns to transform (with divergent inland).

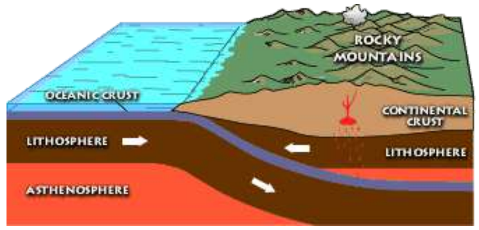


Figure 8.8.1: Shallow subduction during the Laramide Orogeny.

In the Cenozoic, the plates of the Earth moved into more familiar places, with the biggest change being the closing of the Tethys Sea with collisions such as the Alps, Zagros, and Himalaya, a collision that started about 57 million years ago and continues today. Maybe the most significant tectonic feature that occurred in the Cenozoic of North America was the conversion of the west coast of California from a convergent boundary subduction zone to a transform boundary. Subduction off the coast of the western United States, which had occurred throughout the Mesozoic, had continued in the Cenozoic. After the Sevier Orogeny in the late Mesozoic, a subsequent orogeny called the Laramide Orogeny occurred in the early Cenozoic. The Laramide was thick-skinned, different than the Sevier Orogeny. It involved deeper crustal rocks and produced bulges that would become mountain ranges like the Rockies, Black Hills, Wind River Range, Uinta Mountains, and the San Rafael Swell. Instead of descending directly into the mantle, the subducting plate shallowed out and moved eastward beneath the continental plate affecting the overlying continent hundreds of miles east of the continental margin and building high mountains. This occurred because the subducting plate was so young and near the spreading center and the density of the plate was therefore low and subduction was hindered [156].

Figure 8.8.1: Map of the San Andreas fault, showing relative motion.

As the mid-ocean ridge itself started to subduct, the relative motion had changed. Subduction caused a relative convergence between the subducting Farallon plate and the North American plate. On the other side of the mid-ocean ridge from the Farallon plate was the Pacific plate, which was moving away from the North American plate. Thus, as the subduction zone consumed the mid-ocean ridge, the relative movement became transform instead of convergent, which went on to become the San Andreas Fault System. As the San Andreas grew, it caused east-west directed extensional forces to spread over the western United States, creating the Basin and Range province. The transform fault switched position over the last 18 million years, twisting the mountains around Los Angeles, and new faults in the southeastern California deserts may become a future San Andreas-style fault [160]. During this switch from subduction to transform, the nearly horizontal Farallon slab began to sink into the mantle. This caused magmatism as the subducting slab sank, allowing asthenosphere material to rise around it. This event is called the Oligocene ignimbrite flare-up, which was one of the most significant periods of volcanism ever, including the largest single confirmed eruption, the 5000 cubic kilometer Fish Canyon Tuff [162].

Cenozoic Evolution

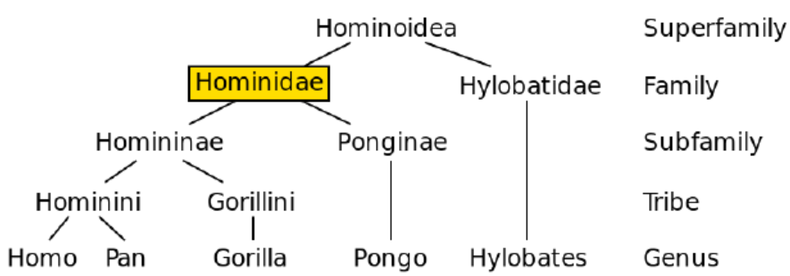


Figure 8.8.1: Family tree of Hominids (Hominidae).

There are five groups of early mammals in the fossil record, based primarily on fossil teeth, the hardest bone in vertebrate skeletons. For the purpose of this text, the most important group is the Eupantotheres, which diverges into the two main groups of mammals, the marsupials (*Sinodelphys*) and placentals or eutherians (*Eomaia*) in the Cretaceous and then diversified in the Cenozoic. The marsupials dominated on the isolated island continents of South America and Australia, and many went extinct in South America with the introduction of placental mammals. Some well-known mammal groups have been highly studied with interesting evolutionary stories in the Cenozoic. For example, horses started small with four toes, ended up larger and having just one toe [163]. Cetaceans (marine mammals like whales and dolphins) started on land from small bear-like (mesonychids) creatures in the early Cenozoic and gradually took to water [164]. However, no study of evolution has been more studied than human evolution. **Hominids**, the name for human-like primates, started in eastern Africa several million years ago.

Figure 8.8.1: Lucy skeleton from the Cleveland Natural History Museum, showing real fossil (brown) and reconstructed skeleton (white).

The first critical event in this story is an environmental change from jungle to more of a savanna, probably caused by changes in Indian Ocean circulation. While bipedalism is known to have evolved before this shift, it is generally believed that our bipedal ancestors (like *Australopithecus*) had an advantage by covering ground more easily in a more open environment compared to their non-bipedal evolutionary cousins. There is also a growing body of evidence, including the famous “Lucy” fossil of an Australopithecine, that our early ancestors lived in trees. Arboreal animals usually demand a high intelligence to navigate through a three-dimensional world. It is from this lineage that humans evolved, using endurance running as a means to acquire more resources and possibly even hunt. This can explain many uniquely human features, from our long legs, strong achilles, lack of lower gut protection, and our wide range of running efficiencies.

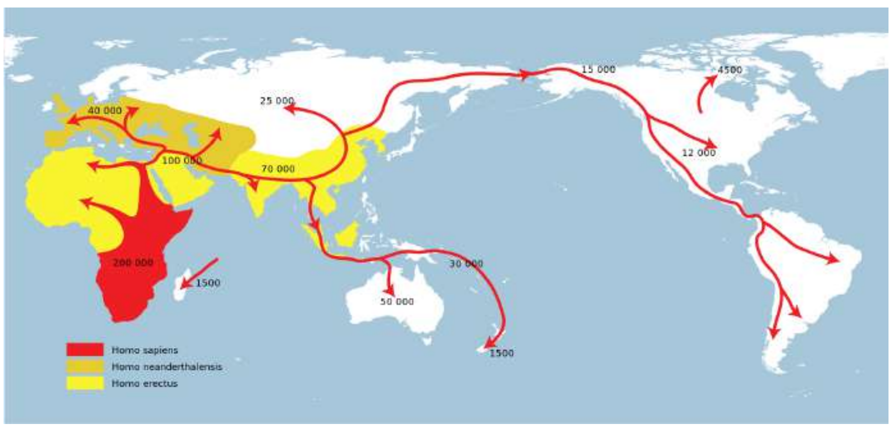


Figure 8.8.1: The hypothesized movement of the homo genus. Years are marked as to the best guess of the timing of movement.

Now that the hands are freed up, the next big step is a large brain. There have been arguments from a switch to more meat-eating, cooking with fire [170], tool usage, and even the construct of society itself to explain this increase in brain size. Regardless of how, it was this increased cognitive power that allowed humans to reign as their ancestors moved out of Africa and explored the world, ultimately entering the Americas through land bridges like the Bering Land Bridge. The details of this worldwide migration and the different branches of the hominid evolutionary tree are very complex, and best reserved for its own course.

Anthropocene and Extinction

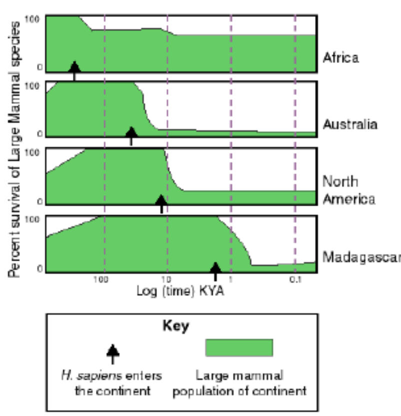


Figure 8.8.1: Graph showing the abundance of large mammals and the introduction of humans.

Humans have had an influence on the Earth, its ecosystems and climate. Yet, human activity can not explain all of the changes that have occurred in the recent past. The start of the Quaternary period, the last and current period of the Cenozoic, is marked by the start of our current ice age 2.58 million years ago. During this time period, ice sheets advanced and retreated, most likely due to Milankovitch cycles (see Chapter 15). Also at this time, various cold-adapted megafauna emerged (like giant sloths, saber-tooth cats, and woolly mammoths), and most of them went extinct as the Earth warmed from the most recent glacial maximum.

A long-standing debate is over the cause of these and other extinctions. Is climate warming to blame, or were they caused by humans [175]? Certainly, we know of recent human extinctions of animals like the dodo or passenger pigeon. Can we connect modern extinctions to extinctions in the recent past? If so, there are several ideas as to how this happened. Possibly the most widely accepted and oldest is the hunting/overkill hypothesis [176]. The idea behind this hypothesis is that humans hunted large herbivores for food, then carnivores could not find food, and human arrival times in locations have been shown to be tied to increased extinction rates in many cases.

Figure 8.8.1: Bingham Canyon Mine, Utah. This open-pit mine is the largest man-made removal of rock in the world.

Modern human impact on the environment and the Earth as a whole is unquestioned. In fact, many scientists are starting to suggest that the rise of human civilization ended and/or replaced the Holocene epoch and defines a new geologic time interval: the **Anthropocene** [177]. Evidence for this change includes extinctions, increased tritium (hydrogen with two neutrons) due to nuclear testing, rising pollutants like carbon dioxide, more than 200 never-before-seen mineral species that have occurred only in this epoch, materials such as plastic and metals which will be long-lasting "fossils" in the geologic record, and large amounts of earthen material moved. The biggest scientific debate with this topic is the starting point. Some say that humans' invention of agriculture would be recognized in geologic strata and that should be the starting point, around 12,000 years ago. Others link the start of the industrial revolution and the subsequent addition of vast amounts of carbon dioxide in the atmosphere [180]. Either way, the idea is that alien geologists visiting Earth in the distant future would easily recognize the impact of humans on the Earth as the beginning of a new geologic period.

References

156. Schellart WP, Stegman DR, Farrington RJ, et al (2010) Cenozoic tectonics of western North America controlled by evolving width of Farallon slab. Science 329:316–319. https://doi.org/10.1126/science.1190366

160. Powell RE, Weldon RJ (1992) Evolution of the San Andreas fault. Annu Rev Earth Planet Sci 20:431

162. Johnson MC, Rutherford MJ (1989) Experimentally determined conditions in the Fish Canyon Tuff, Colorado, magma chamber. J Petrol 30:711–737. https://doi.org/10.1093/petrology/30.3.711

163. MacFadden BJ (2005) Fossil Horses--Evidence for Evolution. Science 307:1728–1730. https://doi.org/10.1126/science.1105458

164. Uhen MD (2010) The origin(s) of whales. Annu Rev Earth Planet Sci 38:189–219

170. Organ C, Nunn CL, Machanda Z, Wrangham RW (2011) Phylogenetic rate shifts in feeding time during the evolution of Homo. Proc Natl Acad Sci U S A 108:14555–14559

175. Koch PL, Barnosky AD (2006) Late Quaternary extinctions: State of the debate. Annu Rev Ecol Evol Syst

176. Martin PS (1984) Prehistoric overkill: the global model. Quaternary extinctions: a prehistoric revolution 354–403

177. Waters CN, Zalasiewicz J, Summerhayes C, et al (2016) The Anthropocene is functionally and stratigraphically distinct from the Holocene. Science 351:aad2622

180. Zalasiewicz J, Williams M, Smith A, et al (2008) Are we now living in the Anthropocene? GSA Today 18:4

Contributions and Attributions

CHAPTER OVERVIEW

9: CRUSTAL DEFORMATION AND EARTHQUAKES

Learning Objectives

Differentiate between stress and strain
Identify the three major types of stress
Differentiate between brittle, ductile, and elastic deformation
Describe the geological map symbol used for strike and dip of strata
Name and describe different fold types
Differentiate the three major fault types and describe their associated movements
Explain how elastic rebound relates to earthquakes
Describe different seismic wave types and how they are measured
Explain how humans can induce seismicity
Describe how seismographs work to record earthquake waves
From seismograph records, locate the epicenter of an earthquake
Explain the difference between earthquake magnitude and intensity
List earthquake factors that determine ground shaking and destruction
Identify secondary earthquake hazards
Describe notable historical earthquakes

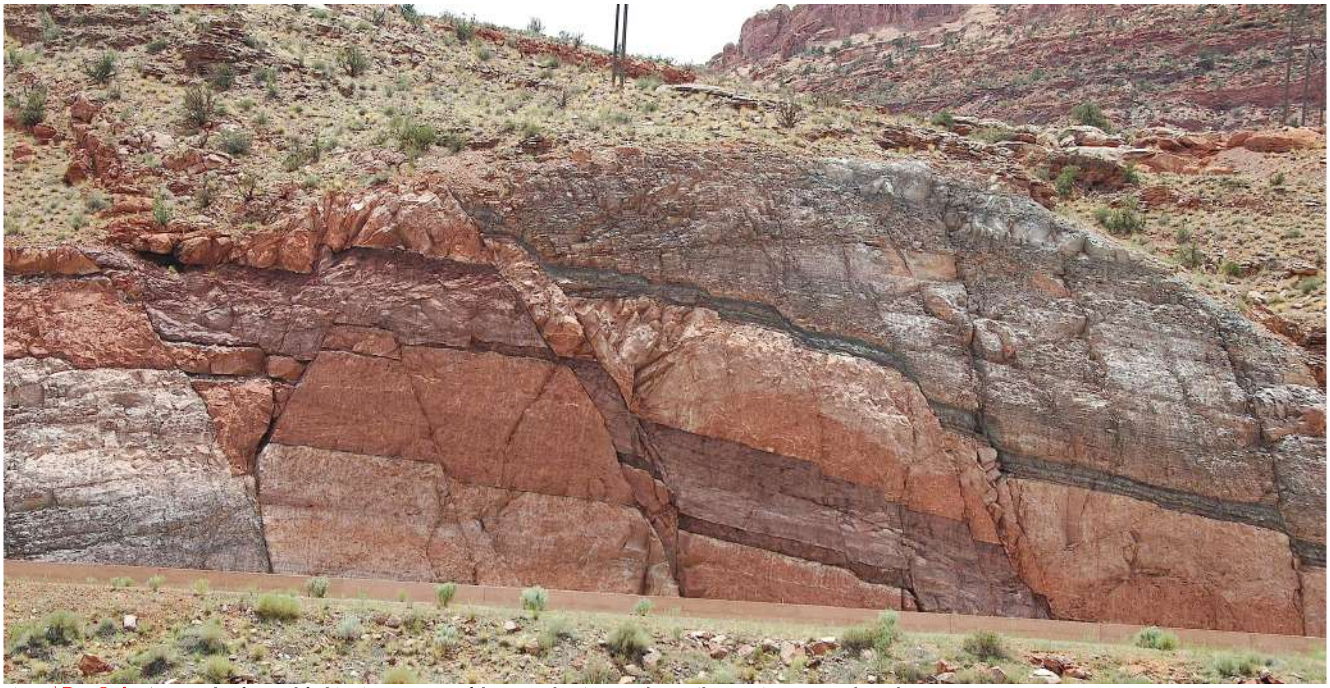


Figure \PageIndex1: Example of normal faulting in an outcrop of the Pennsylvanian Honaker Trail Formation near Moab, Utah.

Crustal deformation occurs when applied forces exceed the internal strength of rocks, physically changing their shapes. These forces are called stress, and the physical changes they create are called strain. Forces involved in tectonic processes as well as gravity and igneous pluton emplacement produce strains in rocks that include folds, fractures, and faults. When rock experiences large amounts of shear stress and breaks with rapid, brittle deformation, energy is released in the form of seismic waves, commonly known as an earthquake.

9.1: STRESS AND STRAIN

Stress is the force exerted per unit area and strain is the physical change that results in response to that force. When the applied stress is greater than the internal strength of rock, strain results in the form of deformation of the rock caused by the stress. Strain in rocks can be represented as a change in rock volume and/or rock shape, as well as fracturing the rock. There are three types of stress: tensional, compressional, and shear.

9.2: DEFORMATION

When rocks are stressed, the resulting strain can be elastic, ductile, or brittle. This change is generally called deformation. Elastic deformation is a strain that is reversible after the stress is released. For example, when you stretch a rubber band, it elastically returns to its original shape after you release it. The type of deformation a rock undergoes depends on pore pressure, strain rate, rock strength, temperature, stress intensity, time, and confining pressure.

9.3: GEOLOGICAL MAPS

Geologic maps are two dimensional (2D) representations of geologic formations and structures at the Earth’s surface, including formations, faults, folds, inclined strata, and rock types. Formations are recognizable rock units. Geologists use geologic maps to represent where geologic formations, faults, folds, and inclined rock units are. Geologic formations are recognizable, mappable rock units.

9.4: FOLDS

Geologic folds are layers of rock that are curved or bent by ductile deformation. Folds are most commonly formed by compressional forces at depth, where hotter temperatures and higher confining pressures allow ductile deformation to occur. Folds are described by the orientation of their axes, axial planes, and limbs. There are many types of folds, including symmetrical folds, asymmetrical folds, overturned folds, recumbent folds, and plunging folds.

9.5: FAULTS

Faults are the places in the crust where brittle deformation occurs as two blocks of rocks move relative to one another. Normal and reverse faults display vertical, also known as dip-slip, motion. Dip-slip motion consists of relative up-and-down movement along a dipping fault between two blocks, the hanging wall, and footwall. In a dip-slip system, the footwall is below the fault plane and the hanging wall is above the fault plane.

9.6: EARTHQUAKE ESSENTIALS

Earthquakes are felt at the surface of the Earth when energy is released by blocks of rock sliding past each other, i.e. faulting has occurred. Seismic energy thus released travels through the Earth in the form of seismic waves. Most earthquakes occur along active plate boundaries. Intraplate earthquakes (not along plate boundaries) occur and are still poorly understood. The USGS Earthquakes Hazards Program has a real-time map showing the most recent earthquakes.

9.7: MEASURING EARTHQUAKES

People feel approximately 1 million earthquakes a year, usually when they are close to the source and the earthquake registers at least moment magnitude 2.5. Major earthquakes of moment magnitude 7.0 and higher are extremely rare. The U. S. Geological Survey (USGS) Earthquakes Hazards Program real-time map shows the location and magnitude of recent earthquakes around the world.

9.8: EARTHQUAKE RISK

Earthquake magnitude is an absolute value that measures pure energy release. Intensity, however, i.e. how much the ground shakes, is determined by several factors. In general, the larger the magnitude, the stronger the shaking and the longer the shaking will last. Shaking is more severe closer to the epicenter. The severity of shaking is influenced by the location of the observer relative to the epicenter, the direction of rupture propagation, and the path of greatest rupture.

9.9: CASE STUDIES

This section contains multiple case studies of major earthquakes in North America as well as historically important earthquakes in the whole world.

9.1: Stress and Strain

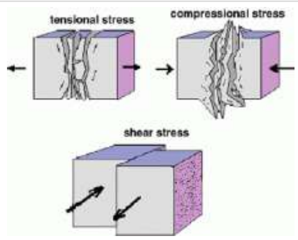


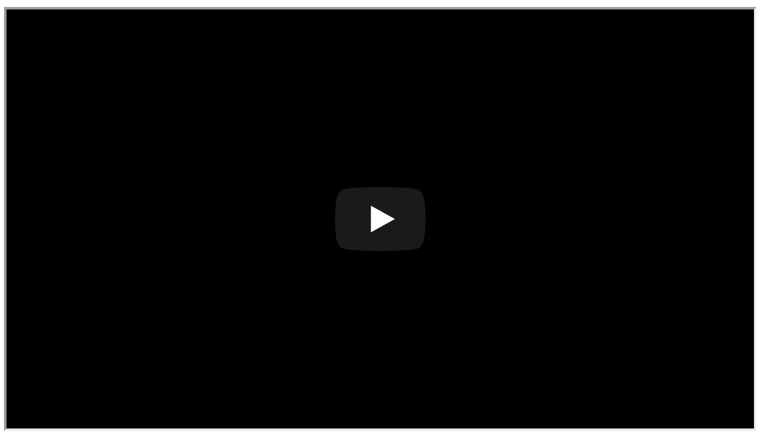
Figure 9.1.1: Types of stress. Clockwise from top left: tensional stress, compressional stress, and shear stress, and some examples of resulting strain.

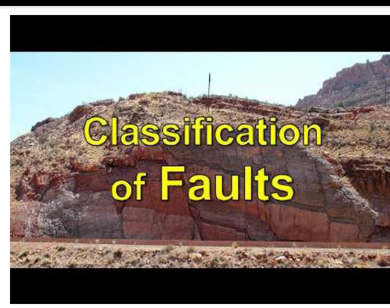
Stress is the force exerted per unit area and **strain** is the physical change that results in response to that force. When the applied stress is greater than the internal strength of rock, strain results in the form of deformation of the rock caused by the stress. Strain in rocks can be represented as a change in rock volume and/or rock shape, as well as fracturing the rock. There are three types of stress: **tensional**, **compressional**, and **shear** [1]. Tensional stress involves forces pulling in opposite directions, which results in strain that stretches and thins rock. Compressional stress involves forces pushing together, and the compressional strain shows up as rock folding and thickening. Shear stress involves transverse forces; the strain shows up as opposing blocks or regions of the material moving past each other.

Table showing types of stress and resulting strain:

Type of Stress	Associated Plate Boundary type (see Ch. 2)	Resulting Strain	Associated fault and offset types
Tensional	divergent	Stretching and thinning	Normal
Compressional	convergent	Shortening and thickening	Reverse
Shear	transform	Tearing	Strike-slip

Video showing types and classification of faults:





References

1. Earle S (2015) Physical geology OER textbook. BC Campus OpenEd

Contributions and Attributions

9.2: Deformation

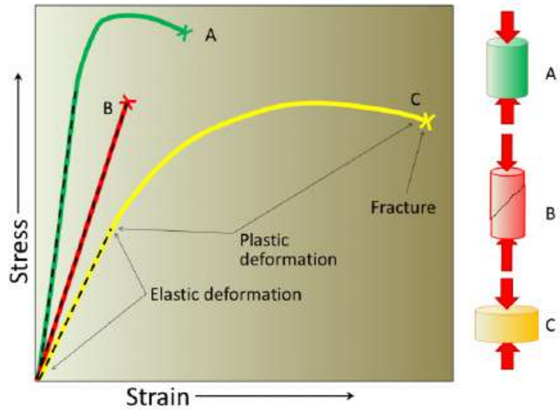


Figure 9.2.1: Different materials deform differently when stress is applied. Material A has relatively little deformation when undergoing large amounts of stress, before undergoing plastic deformation, and finally brittle failure. Material B only elastically deforms before brittle failure. Material C undergoes significant plastic deformation before finally brittle failure.

When rocks are stressed, the resulting strain can be elastic, ductile, or brittle. This change is generally called **deformation**. **Elastic deformation** is a strain that is reversible after the stress is released. For example, when you stretch a rubber band, it elastically returns to its original shape after you release it. **Ductile deformation** occurs when enough stress is applied to a material that the changes in its shape are permanent, and the material is no longer able to revert to its original shape. For example, if you bend a metal bar too far, it can be permanently bent out of shape. The point at which elastic deformation is surpassed and strain becomes permanent is called the **yield point**. In the figure, the yield point is where the line transitions from elastic deformation to ductile deformation (the end of the dashed line). Brittle deformation is another critical point of no return when rock integrity fails and the rock fractures under increasing stress.

The type of deformation a rock undergoes depends on pore pressure, strain rate, rock strength, temperature, stress intensity, time, and confining pressure. Pore pressure is exerted on the rock by fluids in the open spaces or pores embedded within rock or sediment. Strain rate measures how quickly material is deformed. For example, applying stress slowly makes it is easier to bend a piece of wood without breaking it. Rock strength measures how easily a rock deforms under stress. Shale has low strength and granite has high strength. Removing heat, or decreasing the temperature, makes materials more rigid and susceptible to brittle deformation. On the other hand, heating materials make them more ductile and less brittle. Heated glass can be bent and stretched.

Table showing the relationship between factors operating on rock and the resulting strains:

Factor	**Strain Response**
Increase Temperature	More Ductile
Increase Strain Rate	More Brittle
Increase Rock Strength	More Brittle

Contributions and Attributions

9.3: Geological Maps

Geologic maps are two dimensional (2D) representations of geologic formations and structures at the Earth's surface, including formations, faults, folds, inclined strata, and rock types. Formations are recognizable rock units. Geologists use geologic maps to represent where geologic formations, faults, folds, and inclined rock units are. Geologic formations are recognizable, mappable rock units. Each formation on the map is indicated by a color and a label. For examples of geologic maps, see the Utah Geological Survey (UGS) geologic map viewer.

Formation labels include symbols that follow a specific protocol. The first one or more letters are uppercase and represent the geologic time period of the formation. More than one uppercase letter indicates the formation is associated with multiple time periods. The following lowercase letters represent the formation name, abbreviated rock description, or both.

Cross-sections

Cross-sections are subsurface interpretations made from surface and subsurface measurements. Maps display geology in the horizontal plane while cross-sections show subsurface geology in the vertical plane. For more information on cross-sections, check out the AAPG wiki.

Strike and Dip

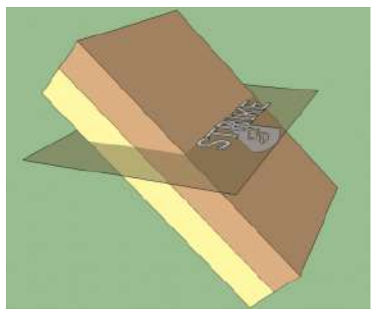


Figure 9.3.1: "Strike" and "dip" are words used to describe the orientation of rock layers with respect to North/South and Horizontal.

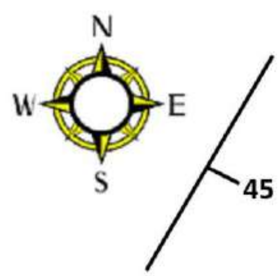
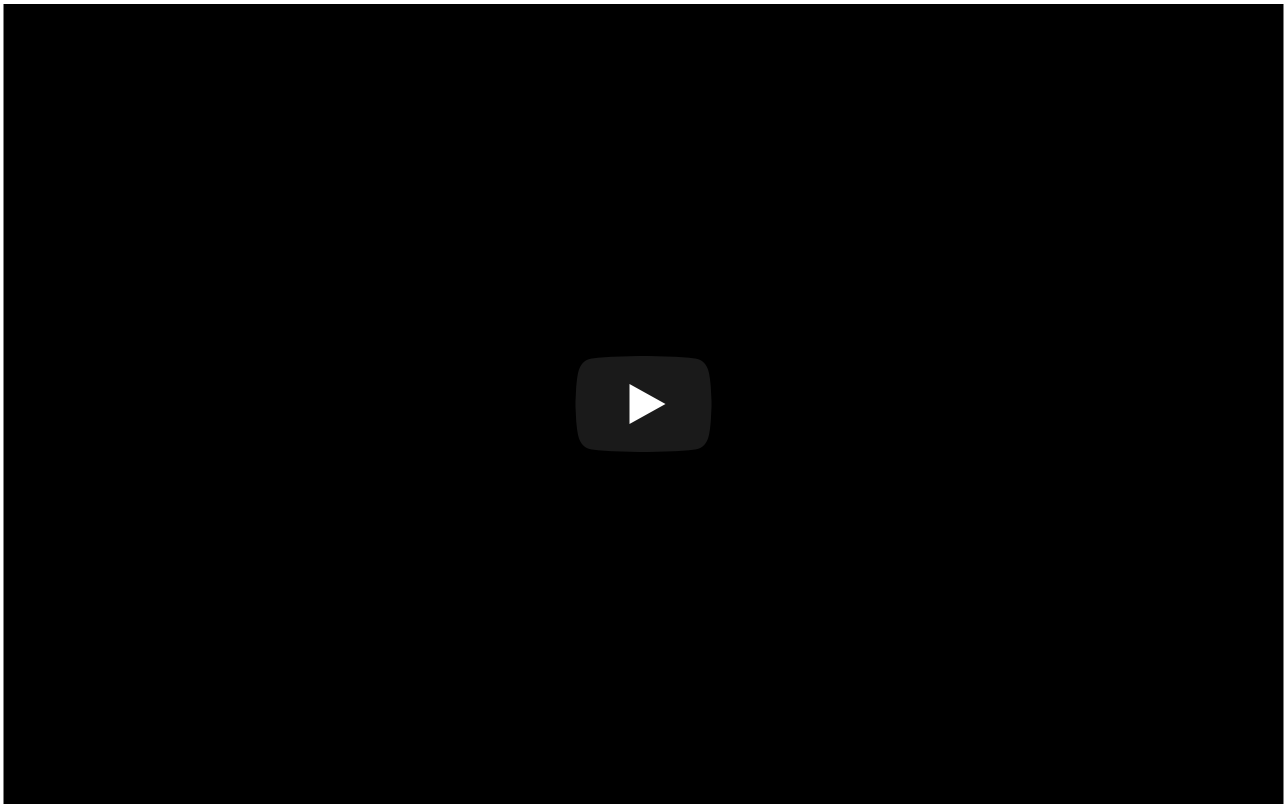
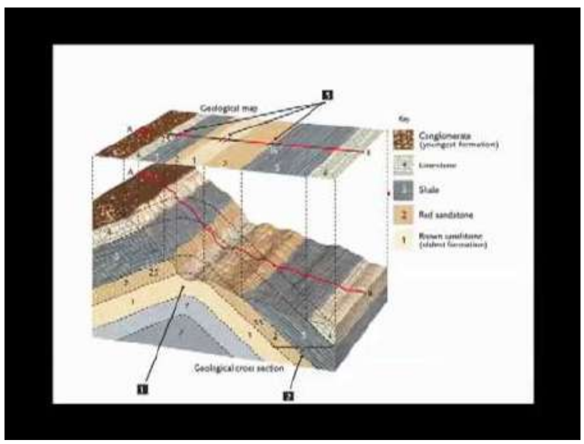


Figure 9.3.1: Attitude symbol on a geologic map (with compass directions for reference) showing strike of N30^{o}E and dip of 45 to the SE.

Geologists use a special symbol called strike and dip to represent inclined beds. Strike and dip map symbols look like the capital letter *T*, with a short trunk and extra-wide top line. The short trunk represents the dip and the top line represents the strike. A Dip is an angle that a bed plunges into the Earth from the horizontal. A number next to the symbol represents a dip angle. One way to visualize the strike is to think about a line made by standing water on the inclined layer. That line is horizontal and lies on a compass direction that has some angle with respect to true north or south (see figure). The strike angle is that angle measured by a special compass. E.g., N 30° E (read north 30 degrees east) means the horizontal line points northeast at an angle of 30° from true north. The strike and dip symbol is drawn on the map at the strike angle with respect to true north on the map. The dip of the inclined layer represents the angle down to the layer from horizontal, in the figure 45^{o} SE (read dipping 45 degrees to the SE). The direction of dip would be the direction a ball would roll if set on the layer and released. A horizontal rock bed has a dip of 0° and a vertical bed has a dip of 90°. Strike and dip considered together are called **rock attitude**.

This video illustrates geologic structures and associated map symbols.





Contributions and Attributions

9.4: Folds

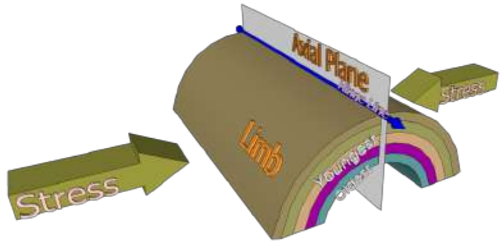


Figure 9.4.1: Model of anticline. Oldest beds are in the center and youngest on the outside. The axial plane intersects the center angle of bend. The hinge line follows the line of greatest bend, where the axial plane intersects the outside of the fold.

Geologic folds are layers of rock that are curved or bent by ductile deformation. Folds are most commonly formed by compressional forces at depth, where hotter temperatures and higher confining pressures allow ductile deformation to occur.

Folds are described by the orientation of their axes, axial planes, and limbs. The plane that splits the fold into two halves is known as the **axial plane**. The **fold axis** is the line along which the bending occurs and is where the axial plane intersects the folded strata. The **hinge line** follows the line of greatest bend in a fold. The two sides of the fold are the fold **limbs**.

Symmetrical folds have a vertical axial plane and limbs have equal but opposite dips. **Asymmetrical folds** have dipping, non-vertical axial planes, where the limbs dip at different angles. **Overturned folds** have steeply dipping axial planes and the limbs dip in the same direction but usually at different dip angles. **Recumbent folds** have horizontal or nearly horizontal axial planes. When the axis of the fold plunges into the ground, the fold is called a **plunging fold**. Folds are classified into five categories: anticline, syncline, monocline, dome, and basin.

Anticline

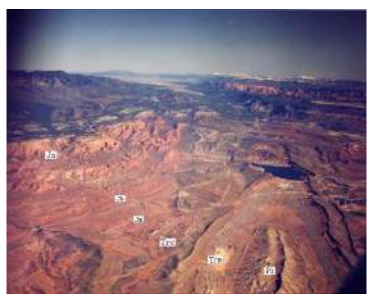


Figure 9.4.1: Oblique view of the Virgin Anticline (bottom right of photo) looking north. The anticline is plunging into the ground to the north. Units from youngest to oldest Jn = Jurassic Navajo Sandstone; Jk = Jurassic Kayenta Formation; Trc = Triassic Chinle Formation; Trm = Triassic Moenkopi; Pk = Permian Kaibab Formation.

Anticlines are arch-like, or A-shaped folds that are convex-upward in shape. They have downward curving limbs and beds that dip down and away from the central fold axis. In anticlines, the oldest rock strata are in the center of the fold, along the axis, and the younger beds are on the outside. Since geologic maps show the intersection of surface topography with underlying geologic structures, an anticline on a geologic map can be identified by both the attitude of the strata forming the fold and the older age of the rocks inside the fold. An **antiform** has the same shape as an anticline, but the relative ages of the beds in the fold cannot be determined. Oil geologists are interested in anticlines because they can form oil traps, where oil migrates up along the limbs of the fold and accumulates in the high point along the fold axis.

Syncline

Synclinal fold – Macigno Formation by **alanpitts** on **Sketchfab**

Synclines are trough-like, or U shaped, folds that are concave-upward in shape. They have beds that dip down and in toward the central fold axis. In synclines, older rock is on the outside of the fold and the youngest rock is inside of the fold axis. A **synform** has the shape of a syncline but like an antiform, does not have distinguishable age zones.

Monocline

Figure 9.4.1: Oblique aerial photograph of Capitol Reef National Park's Water Pocket fold. The perspective is looking southwest toward 50-Mile Mountain and Navajo Mountain.

Monoclines are step-like folds, in which flat rocks are upwarped or downwarped, then continue flat. Monoclines are relatively common on the Colorado Plateau where they form "reefs," which are ridges that act as topographic barriers and should not be confused with ocean reefs (see Chapter 5). Capitol Reef is an example of a monocline in Utah. Monoclines can be caused by bending of shallower sedimentary strata as faults grow below them. These faults are commonly called "blind faults" because they end before reaching the surface and can be either normal or reverse faults.

Dome

Figure 9.4.1: View of the San Rafael Swell from space. In this photograph, north is to the left. Dipping beds of rock will have lines of shadow around them. Note that the center part of the dome is eroded away.

A dome is symmetrical to semi-symmetrical upwarping of rock beds. Domes have a shape like an inverted bowl, similar to an architectural dome on a building. Examples of domes in Utah include the San Rafael Swell, Harrisburg Junction Dome, and Henry Mountains [2; 3]. Domes are formed from compressional forces, underlying igneous intrusions [2] (see Chapter 4), by salt diapirs, or even impacts, like upheaval dome in Canyonlands National Park.

Basin

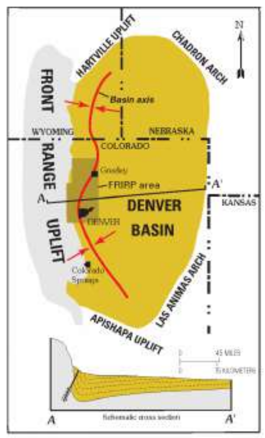


Figure 9.4.1: The Denver Basin is an active sedimentary basin at the eastern extent of the Rocky Mountains. As sediment accumulates, the basin subsides, creating a basin-shaped of beds that are all dipping towards the center of the basin.

A basin is the inverse of a dome, a bowl-shaped depression in a rock bed. The Uinta Basin in Utah is an example of a basin. Some structural basins are also sedimentary basins that collect large quantities of sediment over time. Sedimentary basins can form as a result of folding but are much more commonly produced in mountain building, forming between mountain blocks or via faulting. Regardless of the cause, as the basin sinks or subsides, it can accumulate more sediment because the weight of the sediment causes more subsidence in a positive-feedback loop. There are active sedimentary basins all over the world [4]. An example of a rapidly subsiding basin in Utah is the Oquirrh Basin, dated to the Pennsylvanian-Permian age, which has accumulated over 9,144 m (30,000 ft) of fossiliferous sandstones, shales, and limestones. These strata can be seen in the Wasatch Mountains along the east side of Utah Valley, especially on Mt. Timpanogos and in Provo Canyon.

References

2. Gilbert GK, Dutton CE (1877) Report on the geology of the Henry Mountains. U.S. Government Printing Office, Washington

3. Ressetar R (2013) The San Rafael Swell and Henry Mountains Basin: geologic centerpiece of Utah. Utah Geological Association, Utah Geological Association

4. Coleman JL, Cahan SM (2012) Preliminary catalog of the sedimentary basins of the United States. U.S. Geological Survey

Contributions and Attributions

9.5: Faults

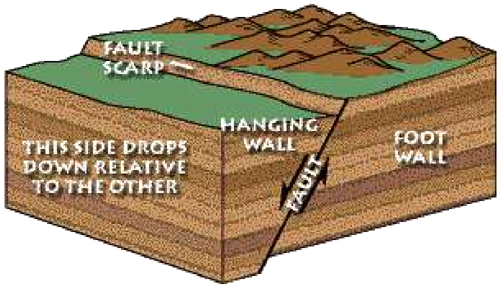


Figure 9.5.1: Common terms used for normal faults. Normal faults form when the hanging wall moves down relative to the footwall.

Faults are the places in the crust where brittle deformation occurs as two blocks of rocks move relative to one another. Normal and reverse faults display vertical, also known as dip-slip, motion. Dip-slip motion consists of relative up-and-down movement along a dipping fault between two blocks, the hanging wall, and footwall. In a dip-slip system, the footwall is below the fault plane and the hanging wall is above the fault plane. A good way to remember this is to imagine a mine tunnel running along a fault; the hanging wall would be where a miner would hang a lantern and the footwall would be at the miner's feet.

Faulting as a term refers to the rupture of rocks. Such ruptures occur at plate boundaries but can also occur in plate interiors as well. Faults slip along the fault plane. The fault scarp is the offset of the surface produced where the fault breaks through the surface. Slickensides are polished, often grooved surfaces along the fault plane created by friction during the movement.

A joint or fracture is a plane of brittle deformation in the rock created by the movement that is not offset or sheared. Joints can result from many processes, such as cooling, depressurizing, or folding. Joint systems may be regional affecting many square miles.

Normal Faults

Normal faults move by a vertical motion where the hanging-wall moves downward relative to the footwall along the dip of the fault. Normal faults are created by tensional forces in the crust. Normal faults and tensional forces commonly occur at divergent plate boundaries, where the crust is being stretched by tensional stresses (see Chapter 2). Examples of normal faults in Utah are the Wasatch Fault, the Hurricane Fault, and other faults bounding the valleys in the Basin and Range province.

Figure 9.5.1: Example of a normal fault in an outcrop of the Pennsylvanian Honaker Trail Formation near Moab, Utah.

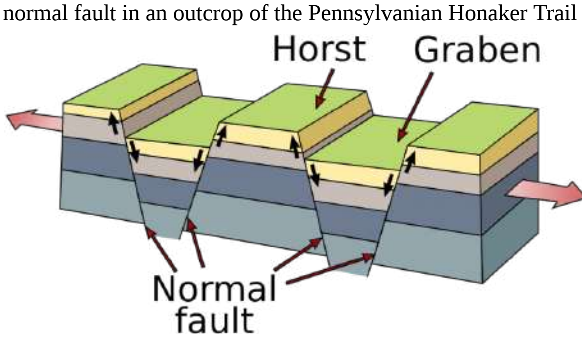


Figure 9.5.1: Faulting that occurs in the crust under tensional stress.

Grabens, horsts, and half-grabens are blocks of crust or rock bounded by normal faults (see Chapter 2). Grabens drop down relative to adjacent blocks and create valleys. Horsts rise up relative to adjacent down-dropped blocks and become areas of higher topography. Where occurring together, horsts and grabens create a symmetrical pattern of valleys surrounded by normal faults on both sides and mountains. Half-grabens are a one-sided version of a horst and graben, where blocks are tilted by a normal fault on one side, creating an asymmetrical valley-mountain arrangement. The mountain-valleys of the Basin and Range Province of Western Utah and Nevada consist of a series of full and half-grabens from the Salt Lake Valley to the Sierra Nevada Mountains.

Normal faults do not continue to clear into the mantle. In the Basin and Range Province, the dip of a normal fault tends to decrease with depth, i.e., the fault angle becomes shallower and more horizontal as it goes deeper. Such decreasing dips happen when large amounts of extension occur along very low-angle normal faults, known as **detachment faults**. The normal faults of the Basin and Range, produced by tension in the crust, appear to become detachment faults at greater depths.

Reverse Faults

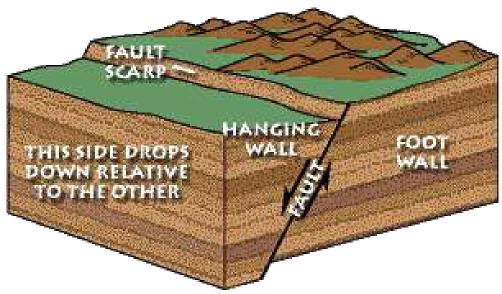


Figure 9.5.1: Simplified block diagram of a reverse fault.

In reverse faults, compressional forces cause the hanging wall to move up relative to the footwall. A thrust fault is a reverse fault where the fault plane has a low dip angle of less than 45°. Thrust faults carry older rocks on top of younger rocks and can even cause the repetition of rock units in the stratigraphic record.

Convergent plate boundaries with subduction zones create a special type of “reverse” fault called a megathrust fault where denser oceanic crust drives down beneath less dense overlying crust. Megathrust faults cause the largest magnitude earthquakes yet measured and commonly cause massive destruction and tsunamis.

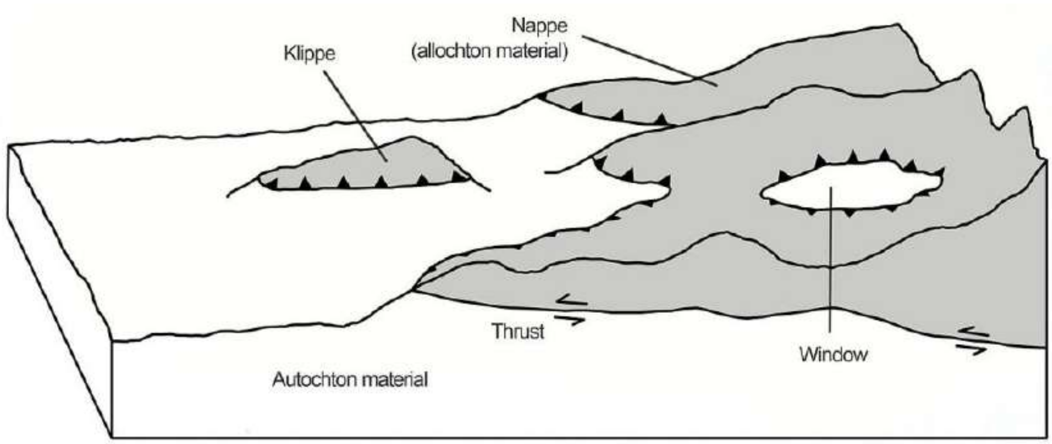


Figure 9.5.1: Terminology of thrust faults (low-angle reverse faults). A klippe is the remnant of the hangingwall (aka nappe), where the surrounding material has been eroded away. A window is where part of the hangingwall has been eroded away to expose the footwall (autochthon). Note the symbol shows flags on the overlying thrust plate.

Figure 9.5.1: Ketobe Knob in the San Rafael Swell of Utah displays an example of a thrust fault.

Strike-slip Faults

Strike-slip faults have side-to-side motion. Strike-slip faults are most commonly associated with transform plate boundaries and are prevalent in transform fracture zones along mid-ocean ridges. In pure strike-slip motion, fault blocks on either side of the fault do not move up or down relative to each other, rather move laterally, side to side. The direction of the strike-slip movement is determined by an observer standing on a block on one side of the fault. If the block on the opposing side of the fault moves left relative to the observer's block, this is called sinistral motion. If the opposing block moves right, it is dextral motion.

Video showing motion in a strike-slip fault.

Bends along strike-slip faults create areas of compression or tension between the sliding blocks (see Chapter 2). Tensional stresses create transtensional features with normal faults and basins, such as the Salton Sea in California. Compressional stresses create transpressional features with reverse faults and cause small-scale mountain building, such as the San Gabriel Mountains in California. The faults that splay off transpression or transtension features are known as flower structures.

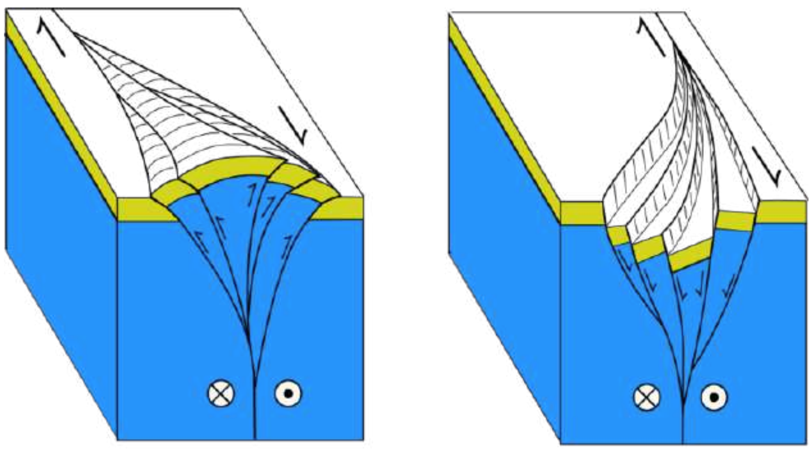
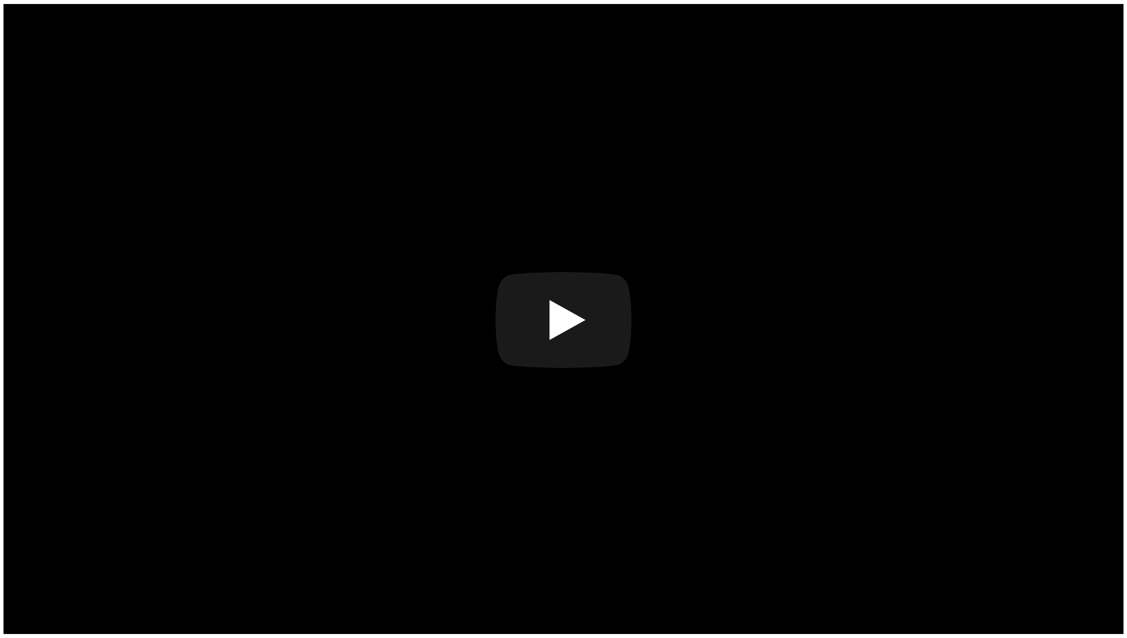


Figure 9.5.1: Flower structures created by strike-slip faults. Depending on the relative movement in relation to the bend in the fault, flower structures can create basins or mountains.

An example of a dextral, right-lateral strike-slip fault is the San Andreas Fault, which denotes a transform boundary between the North American and Pacific plates. An example of a sinistral, left-lateral strike-slip fault is the Dead Sea fault in Jordan and Israel.

Video showing how faults are classified:



Contributions and Attributions

9.6: Earthquake Essentials

Earthquakes are felt at the surface of the Earth when energy is released by blocks of rock sliding past each other, i.e. faulting has occurred. Seismic energy thus released travels through the Earth in the form of seismic waves. Most earthquakes occur along active plate boundaries. Intraplate earthquakes (not along plate boundaries) occur and are still poorly understood. The USGS Earthquakes Hazards Program has a real-time map showing the most recent earthquakes.

How Earthquakes Happen

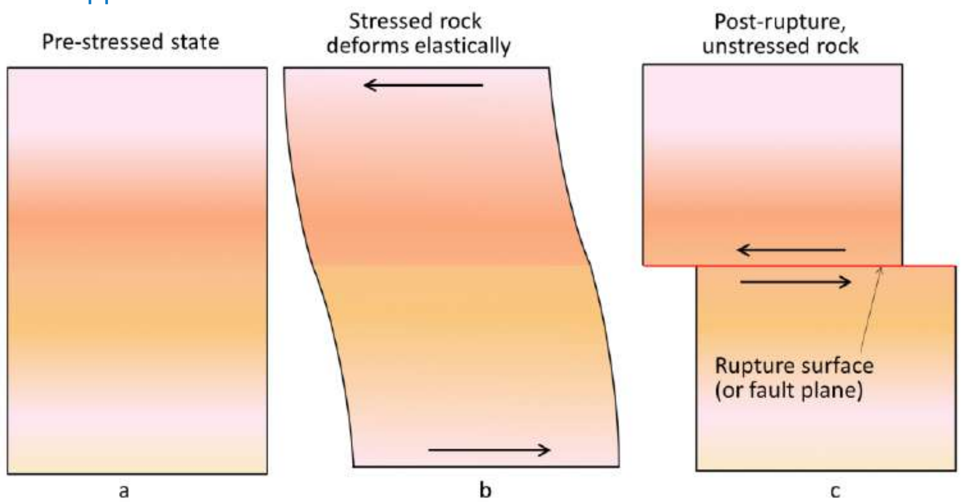


Figure 9.6.1: Process of elastic rebound: a) Undeformed state, b) accumulation of elastic strain, and c) brittle failure and release of elastic strain.

The release of seismic energy is explained by the elastic rebound theory. When the rock is strained to the point that it undergoes brittle deformation, The place where the initial offsetting rupture takes place between the fault blocks is called the focus. This offset propagates along the fault, which is known as the fault plane.

The fault blocks of persistent faults like the Wasatch Fault (Utah), that show recurring movements, are locked together by friction. Over hundreds to thousands of years, stress builds up along the fault until it overcomes frictional resistance, rupturing the rock and initiating fault movement. The deformed unbroken rocks snap back toward their original shape in a process called elastic rebound**.** Think of bending a stick until it breaks; stored energy is released, and the broken pieces return to near their original orientation.

Bending, the ductile deformation of the rocks near a fault reflects a build-up of stress. In earthquake-prone areas like California, strain gauges are used to measure this bending and help seismologists, scientists who study earthquakes, understand more about predicting them. In locations where the fault is not locked, seismic stress causes continuous, the gradual displacement between the fault blocks called fault creep. Fault creep occurs along some parts of the San Andreas Fault (California).

After an initial earthquake, continuous application of stress in the crust causes elastic energy to begin to build again during a period of inactivity along the fault. The accumulating elastic strain may be periodically released to produce small earthquakes on or near the main fault called foreshocks. Foreshocks can occur hours or days before a large earthquake, or may not occur at all. The main release of energy during the major earthquake is known as the mainshock. Aftershocks may follow the mainshock to adjust new strain produced during the fault movement and generally decrease over time.

Focus and Epicenter

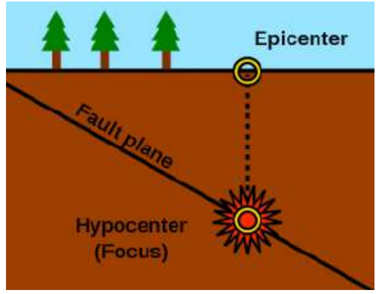


Figure 9.6.1: The hypocenter is the point along the fault plane in the subsurface from which seismic energy emanates. The epicenter is the point on a land surface vertically above the hypocenter.

The earthquake focus, also called the **hypocenter**, is the initial point of rupture and displacement of the rock moves from the hypocenter along the fault surface. The earthquake focus or hypocenter is the point along the fault plane from which initial seismic waves spread outward and is always at some depth below the ground surface. From the focus, rock displacement propagates up, down, and laterally along the fault plane. This displacement produces shock waves called seismic waves. The larger the displacement between the opposing fault blocks and the further the displacement propagates along the fault surface, the more seismic energy is released and the greater the amount and time of shaking is produced. The epicenter is the location on the Earth's surface vertically above the focus. This is the location that most news reports give because it is the center of the area where people are affected.

Seismic Waves

To understand earthquakes and how earthquake energy moves through the Earth, consider the basic properties of waves. Waves describe how energy moves through a medium, such as rock or unconsolidated sediments in the case of earthquakes. Wave amplitude indicates the magnitude or height of earthquake motion. Wavelength is the distance between two successive peaks of a wave. Wave frequency is the number of repetitions of the motion over a period of time, cycles per time unit. Period, which is the amount of time for a wave to travel one wavelength, is the inverse of frequency. When multiple waves combine, they can interfere with each other (see figure). When waves combine in sync, they produce constructive interference, where the influence of one wave adds to and magnifies the other. If waves are out of sync, they produce destructive interference, which diminishes the amplitudes of both waves. If two combined waves have the same amplitude and frequency but are one-half wavelength out of sync, the resulting destructive interference can eliminate each wave. These processes of wave amplitude, frequency, period, and constructive and destructive interference determine the magnitude and intensity of earthquakes.

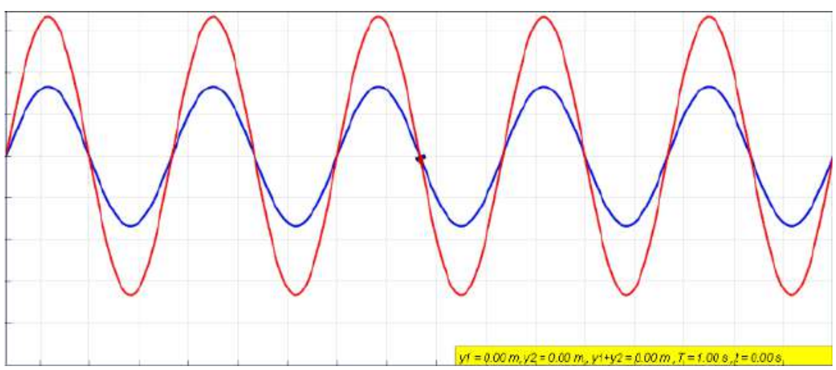
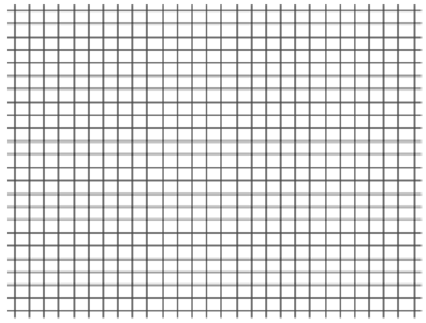


Figure 9.6.1: Example of constructive and destructive interference; note red line representing the results of interference.

Seismic waves are the physical expression of energy released by the elastic rebound of rock within displaced fault blocks and are felt as an earthquake. Seismic waves occur as body waves and surface waves. Body waves pass underground through the Earth's interior body and are the first seismic waves to propagate out from the focus. Body waves include primary (P) waves and secondary (S) waves. P waves are the fastest body waves and move through the rock via compression, very much like sound waves move through air. Rock particles move forward and back during the passage of the P waves, enabling them to travel through solids, liquids, plasma, and gases. S waves travel slower, following P waves, and propagate as shear waves that move rock particles from side to side. Because they are restricted to lateral movement, S waves can only travel through solids but not liquids, plasma, or gases.



P-waves are compressional.

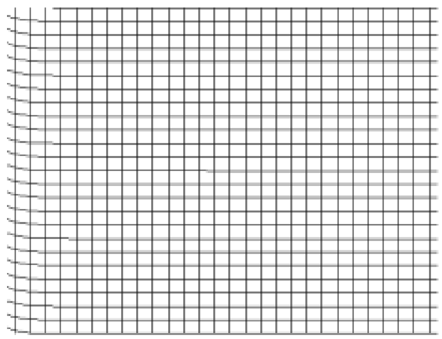
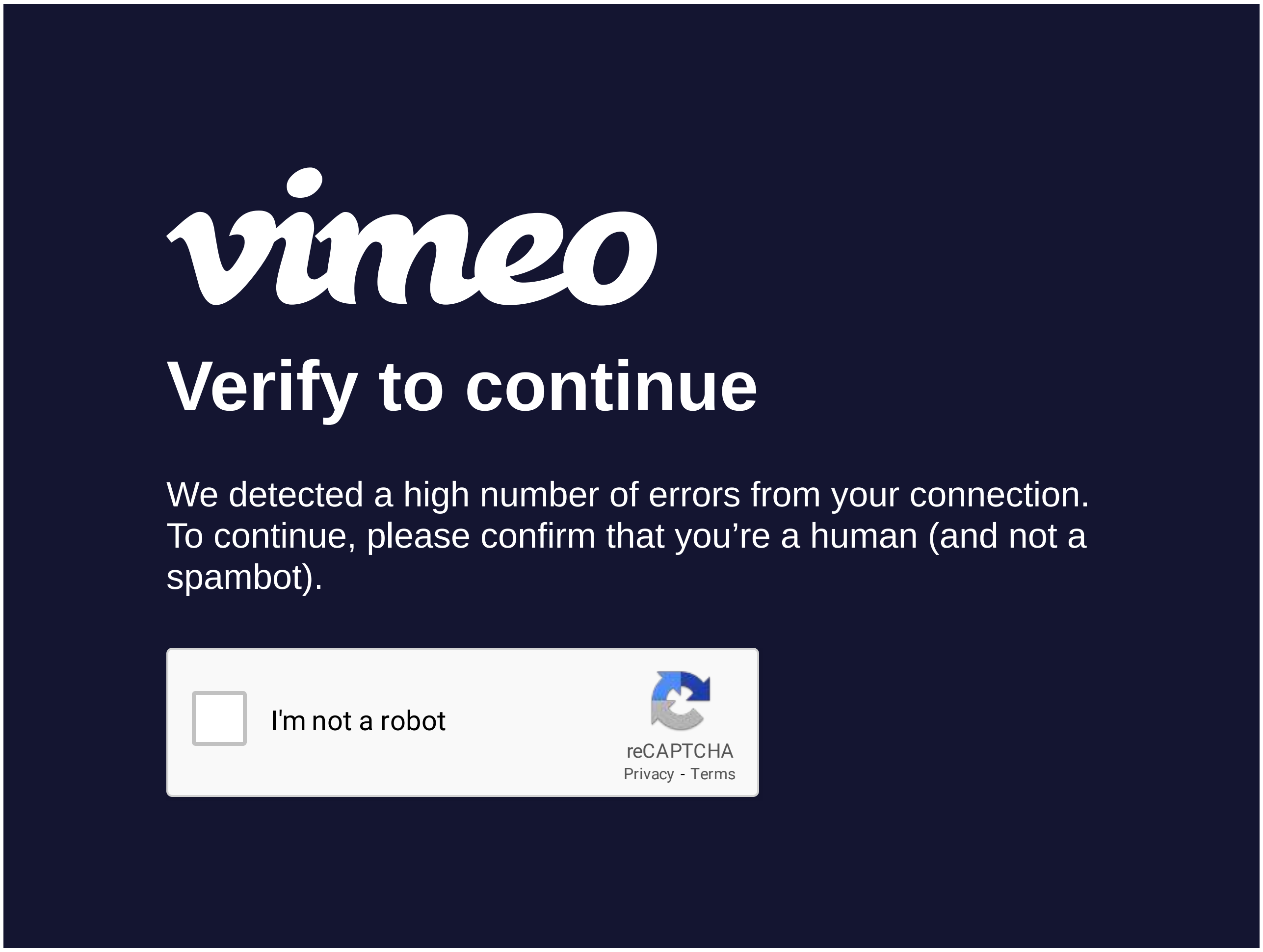


Figure 9.6.1: S waves are shear.

During an earthquake, body waves pass through the Earth and into the mantle as a sub-spherical wavefront. Considering a point on a wavefront, the path followed by a specific point on the spreading wavefront is called a seismic ray and a seismic ray reaches a specific seismograph located at one of thousands of seismic monitoring stations scattered over the Earth. Density increases with depth in the Earth, and since seismic velocity increases with density, a process called **refraction** causes earthquake rays to curve away from the vertical and bend back toward the surface, passing through different bodies of rock along the way.

Surface waves are produced when body waves from the focus strike the Earth's surface. Surface waves travel along the Earth's surface, radiating outward from the epicenter. Surface waves take the form of rolling waves called Raleigh Waves and side to side waves called Love Waves (watch videos for wave propagation animations). Surface waves are produced primarily as the more energetic S waves strike the surface from below with some surface wave energy contributed by P waves (videos courtesy blog.Wolfram.com). Surface waves travel more slowly than body waves and because of their complex horizontal and vertical movement, surface waves are responsible for most of the damage caused by an earthquake. Love waves produce predominantly horizontal ground shaking and, ironically from their name, are the most destructive. Rayleigh waves produce an elliptical motion with longitudinal dilation and compression, like ocean waves. However, Raleigh waves cause rock particles to move in a direction opposite to that of water particles in ocean waves.

The Earth has been described as ringing like a bell after an earthquake with earthquake energy reverberating inside it. Like other waves, seismic waves refract (bend) and bounce (reflect) when passing through rocks of differing densities. S waves, which cannot move through liquids, are blocked by the Earth's liquid outer core, creating an S wave shadow zone on the side of the planet opposite to the earthquake focus. P waves, on the other hand, pass through the core but are refracted into the core by the difference of density at the core-mantle boundary. This has the effect of creating a cone-shaped P wave shadow zone on parts of the other side of the Earth from the focus.



2011 Tohoku Earthquake, Mag. 9.0. Body and Surface Waves from seismicsoundlab on Vimeo.

Induced Seismicity

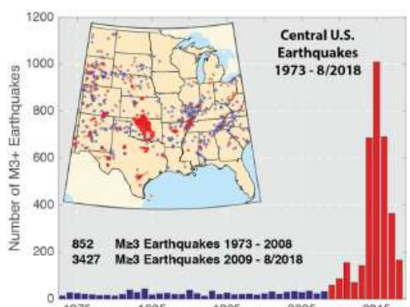


Figure 9.6.1: Frequency of earthquakes in the central United States. Note the sharp increase in the number of earthquakes from 2010 to 2015.

Earthquakes known as induced seismicity occur near natural gas extraction sites because of human activity. Injection of waste fluids in the ground, commonly a byproduct of an extraction process for natural gas known as **fracking**, can increase the outward pressure that liquid in the pores of a rock exerts, known as pore pressure [5; 6]. The increase in pore pressure decreases the frictional forces that keep rocks from sliding past each other, essentially lubricating fault planes. This effect is causing earthquakes to occur near injection sites, in a human-induced activity known as **induced seismicity** [5]. The significant increase in drilling activity in the central United States has spurred the requirement for the disposal of significant amounts of waste drilling fluid, resulting in a measurable change in the cumulative number of earthquakes experienced in the region.

Reference

5. Ellsworth WL (2013) Injection-Induced Earthquakes. Science 341:1225942. https://doi.org/10.1126/science.1225942

6. Rubinstein JL, Mahani AB (2015) Myths and facts on wastewater injection, hydraulic fracturing, enhanced oil recovery, and induced seismicity. Seismological research letters 86:1060–1067

Contributions and Attributions

9.7: Measuring Earthquakes

Seismographs

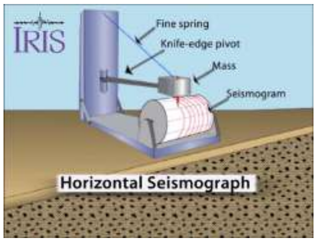


Figure 9.7.1: Animation of a horizontal seismograph.

People feel approximately 1 million earthquakes a year, usually when they are close to the source and the earthquake registers at least moment magnitude 2.5. Major earthquakes of moment magnitude 7.0 and higher are extremely rare. The U. S. Geological Survey (USGS) Earthquakes Hazards Program real-time map shows the location and magnitude of recent earthquakes around the world.

To accurately study seismic waves, geologists use **seismographs** that can measure even the slightest ground vibrations. Early 20th-century seismograms use a weighted pen (pendulum) suspended by a long spring above a recording device fixed solidly to the ground. The recording device is a rotating drum mounted with a continuous strip of paper. During an earthquake, the suspended pen stays motionless and records ground movement on the paper strip. The resulting graph a seismogram. Digital versions use magnets, wire coils, electrical sensors, and digital signals instead of mechanical pens, springs, drums, and paper. A seismograph array is multiple seismographs configured to measure vibrations in three directions: north-south (x-axis), east-west (y-axis), and up-down (z-axis).

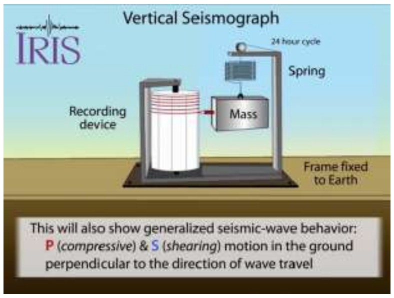


Figure 9.7.1: Animation of how a basic vertical seismograph records seismic waves.

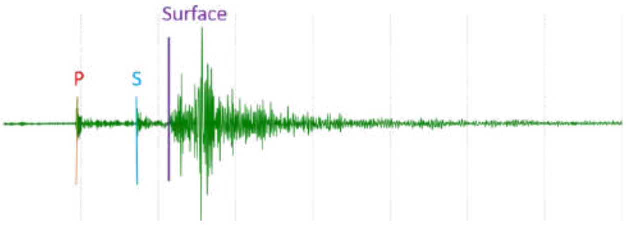
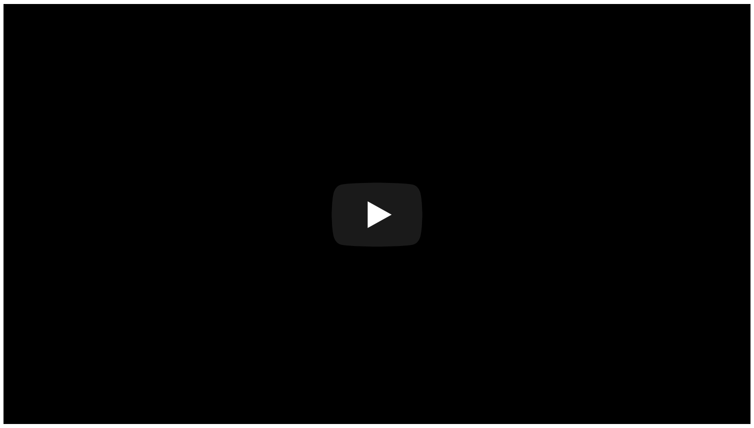
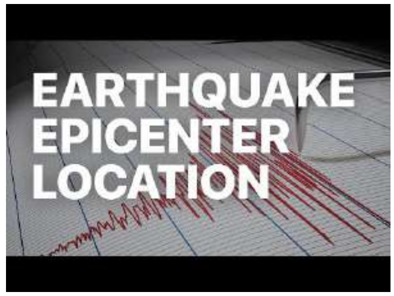


Figure 9.7.1: A seismogram showing the arrivals of the P, S, and surface waves

To pinpoint the location of an earthquake epicenter, seismologists use the differences in arrival times of the P, S, and surface waves. After an earthquake, P waves will appear first on a seismogram, followed by S waves, and finally surface waves, which have the largest amplitude. It is important to note that surface waves lose energy quickly, so they are not measurable at great distances from the epicenter. These time differences determine the distance but not the direction of the epicenter. By using wave arrival times recorded on seismographs at multiple stations, seismologists can apply triangulation to pinpoint the location of the epicenter of an earthquake. At least three seismograph stations are needed for triangulation. The distance from each station to the epicenter is plotted as the radius of a circle. The epicenter is demarked where the circles intersect. This method also works in 3D, using multi-axis seismographs and sphere radii to calculate the underground depth of the focus.

This video shows the method of triangulation to locate the epicenter of an earthquake.





Seismograph Network

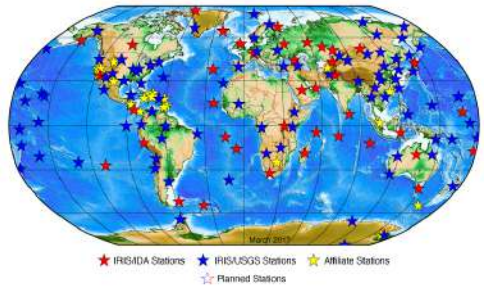
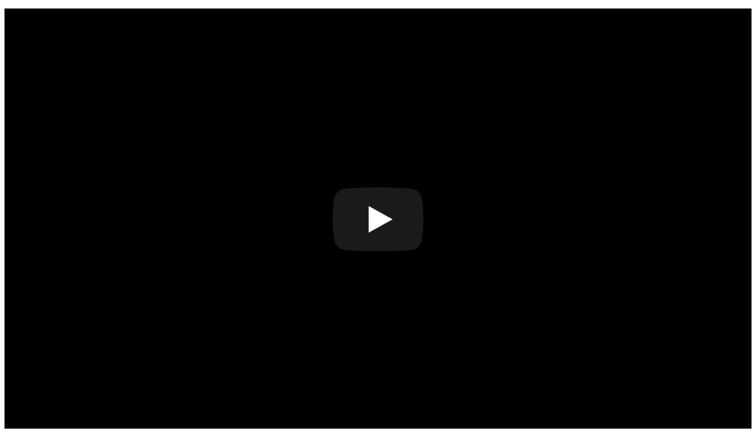
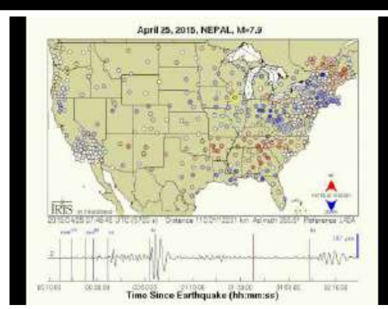


Figure 9.7.1: Global network of seismic stations. Note that this map does not show all of the world's seismic stations, just one of the networks of stations scientists use to measure seismic activity.

The International Registry of Seismograph Stations lists more than 20,000 seismographs on the planet. By comparing data from multiple seismographs, scientists can map the properties of the inside of the Earth, detect detonations of large explosive devices, and predict tsunamis. The Global Seismic Network, a worldwide set of linked seismographs that electronically distribute real-time data, includes more than 150 stations that meet specific design and precision standards. The USArray is a network of hundreds of permanent and transportable seismographs in the United States that are used to map the subsurface activity of earthquakes (see video).

Along with monitoring for earthquakes and related hazards, the Global Seismograph Network helps detect nuclear weapons testing, which is monitored by the Comprehensive Nuclear Test Ban Treaty Organization. Most recently, seismographs have been used to determine nuclear weapons testing by North Korea.





Nepal Earthquake M7.9 Ground Motion Visualization

Seismic Tomography

Very much like a CT (Computed Tomography) scan uses X-rays at different angles to image the inside of a body, **seismic tomography** uses seismic rays from thousands of earthquakes that occur each year, passing at all angles through masses of rock, to generate images of internal Earth structures.

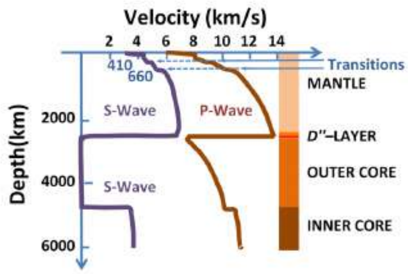


Figure 9.7.1: Speed of seismic waves with depth in the earth as predicted by the PREM. Two thousand kilometers is 1240 miles.

Using the assumption that the earth consists of homogenous layers, geologists developed a model of expected properties of earth materials at every depth within the earth called the PREM (Preliminary Reference Earth Model). These properties include seismic wave transmission velocity, which is dependent on rock density and elasticity. In the mantle, temperature differences affect rock density. Cooler rocks have a higher density and therefore transmit seismic waves faster. Warmer rocks have a lower density and transmit earthquake waves slower. When the arrival times of seismic rays at individual seismic stations are compared to arrival times predicted by PREM, differences are called **seismic anomalies** and can be measured for bodies of rock within the earth from seismic rays passing through them at stations of the seismic network. Because seismic rays travel at all angles from lots of earthquakes and arrive at lots of stations of the seismic network, like CT scans of the body, variations in the properties of the rock bodies allow 3D images to be constructed of the rock bodies through which the rays passed. Seismologists are thus able to construct 3D images of the interior of the Earth.

For example, seismologists have mapped the Farallon Plate, a tectonic plate that subducted beneath North America during the last several million years, and the Yellowstone magma chamber, which is a product of the Yellowstone hot spot under the North American continent. Peculiarities of the Farallon Plate subduction are thought to be responsible for many features of western North America including the Rocky Mountains (See chapter 8).

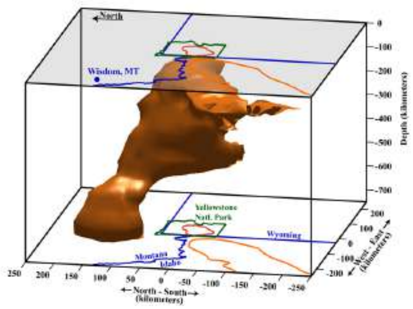


Figure 9.7.1: Seismic tomograph showing the magma chamber beneath Yellowstone National Park.

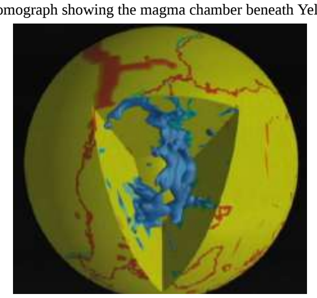


Figure 9.7.1: Tomographic image of the Farallon plate in the mantle below North America.

Earthquake Magnitude and Intensity

Richter Scale

Magnitude is the measure of the energy released by an earthquake. The **Richter scale** (M_L), the first and most well-known magnitude scale, was developed by Charles F. Richter (1900-1985) at the California Institute of Technology. This was the magnitude scale used historically by early seismologists. Used by early seismologists, Richter magnitude (M_L) is determined from the maximum amplitude of the pen tracing on the seismogram recording. Adjustments for epicenter distance from the seismograph are made using the arrival-time differences of S and P waves [7].

The Richter Scale is logarithmic, based on powers of 10. This means an increase of one Richter unit represents a 10-fold increase in seismic-wave amplitude or in other words, a magnitude 6 earthquake shakes the ground 10 times more than a magnitude 5. However, the *actual energy released* for each magnitude unit is 32 times greater, which means a magnitude 6 earthquake releases 32 times more energy than a magnitude 5.

The Richter Scale was developed for earthquakes in Southern California, using local seismographs. It has limited applications for larger distances and very large earthquakes. Therefore, most agencies no longer use the Richter Scale. The moment magnitude (M_W), which is measured using seismic arrays and generates values comparable to the Richter Scale, is more accurate for measuring earthquakes across the Earth, including large earthquakes, although they require more time to calculate. News media often report Richter magnitudes right after an earthquake occurs even though scientific calculations now use moment magnitudes.

Moment Magnitude Scale

The **Moment Magnitude scale** depicts the absolute size of earthquakes, comparing information from multiple locations and using a measurement of actual energy released calculated from the cross-sectional area of rupture, amount of slippage, and the rigidity of the rocks. Because each earthquake occurs in a unique geologic setting and the rupture area is often hard to measure, estimates of moment magnitude can take days or even months to calculate.

Like the Richter Scale, the moment magnitude scale is logarithmic. The magnitude values of the two scales are approximately equal, except for very large earthquakes. Both scales are used for reporting earthquake magnitude. The Richter Scale provides a quick magnitude estimate immediately following the quake and thus, is usually reported in news accounts. Moment magnitude calculations take much longer but are more accurate and thus, more useful for scientific analysis.

Intensity	Shaking	Description/Damage
I	Not felt	Not felt except by a very few under especially favorable conditions.
II	Weak	Felt only by a few persons at rest, especially on upper floors of buildings.
III	Weak	Felt quite noticeably by persons indoors, especially on upper floors of buildings. Many people do not recognize it as an earthquake. Standing motor cars may rock slightly. Vibrations similar to the passing of a truck. Duration estimated.
IV	Light	Felt indoors by many, outdoors by few during the day. At night, some awakened. Dishes, windows, doors disturbed; walls make cracking sound. Sensation like heavy truck striking building. Standing motor cars rocked noticeably.
V	Moderate	Felt by nearly everyone; many awakened. Some dishes, windows broken. Unstable objects overturned. Pendulum clocks may stop.
VI	Strong	Felt by all, many frightened. Some heavy furniture moved; a few instances of fallen plaster. Damage slight.
VII	Very strong	Damage negligible in buildings of good design and construction; slight to moderate in well-built ordinary structures; considerable damage in poorly built or badly designed structu
VIII	Severe	Damage slight in specially designed structures; considerable damage in ordinary substantial buildings with partial collapse. Damage great in poorly built structures. Fall of chimn walls. Heavy furniture overturned.
IX	Violent	Damage considerable in specially designed structures; well-designed frame structures thrown out of plumb. Damage great in substantial buildings, with partial collapse. Buildings
X	Extreme	Some well-built wooden structures destroyed; most masonry and frame structures destroyed with foundations. Rails bent.

Table. Abridged Mercalli Scale from USGS General Interest Publication 1989-288-913.

ShakeMaps

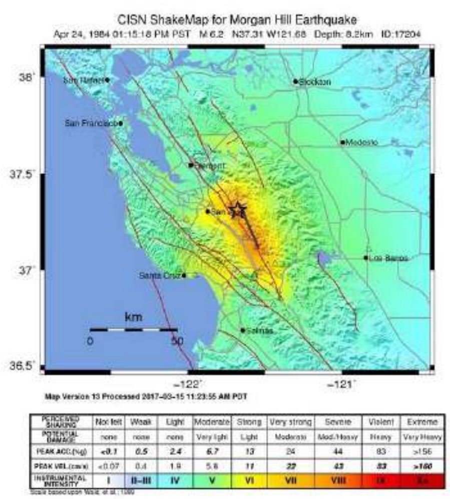


Figure 9.7.1: Example of a shake map.

Shake maps, written ShakeMaps by the USGS, use high-quality, computer-interpolated data from seismograph networks to show areas of intense shaking. Shake maps are useful in the crucial minutes after an earthquake, as they show emergency personnel where the greatest damage likely occurred and help them locate possibly damaged gas lines and other utility facilities.

Reference

7. Richter CF (1935) An instrumental earthquake magnitude scale. Bulletin of the Seismological Society of America 25:1–32

Contributions and Attributions

9.8: Earthquake Risk

Factors that Determine Shaking

Earthquake magnitude is an absolute value that measures pure energy release. Intensity, however, i.e. how much the ground shakes, is determined by several factors.

Earthquake Magnitude—In general, the larger the magnitude, the stronger the shaking and the longer the shaking will last.

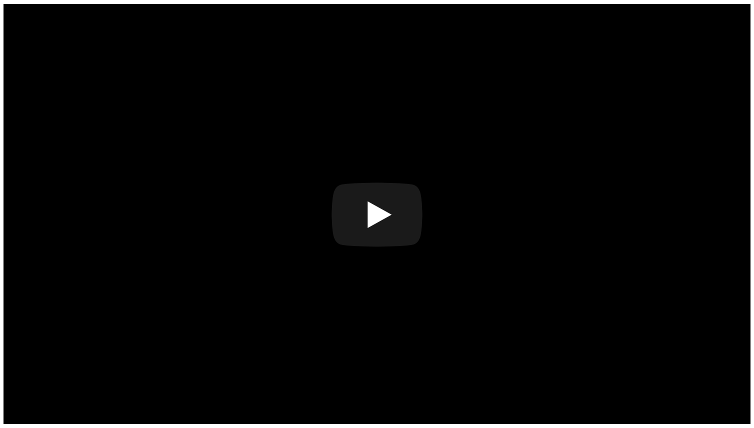
This table is taken from the USGS and shows scales of magnitude and Mercalli Intensity and descriptions of shaking and resulting damage.

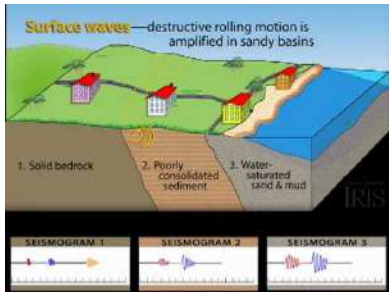
Magnitude	Modified Mercalli Intensity	Shaking/Damage Description
1.0–3.0	**I**	Only felt by a very few.
3.0–3.9	**II – III**	Noticeable indoors, especially on upper floors.
4.0–4.9	**IV – V**	Most to all feel it. Dishes, doors, cars shake and possibly break.
5.0–5.9	**VI – VII**	Everyone feels it. Some items knocked over or broken. Building damage possible.
6.0–6.9	**VII – IX**	Frightening amounts of shaking. Significant damage especially with poorly constructed buildings
≥ 7.0	**≥ VIII**	Significant destruction of buildings. Potential for objects to be thrown in the air from shaking.

Location and Direction—Shaking is more severe closer to the epicenter. The severity of shaking is influenced by the location of the observer relative to the epicenter, the direction of rupture propagation, and the path of greatest rupture. The severity of shaking is influenced by the location of the observer relative to the epicenter, the direction of rupture propagation, and the path of greatest rupture.

Local Geologic Conditions—Seismic waves are affected by the nature of the ground materials through which they pass. Different materials respond differently to an earthquake. Think of shaking a block of Jello versus a meatloaf, one will jiggle much more when hit by waves of the same amplitude. The ground's response to shaking depends on the degree of substrate consolidation. Solid sedimentary, igneous, or metamorphic bedrock shakes less than unconsolidated sediments.

This video shows how different substrates behave in response to different seismic waves and their potential for destruction.





Seismic waves move fastest through consolidated bedrock, slower through unconsolidated sediments, and slowest through unconsolidated sediments with a high water content. Seismic energy is transmitted by wave velocity and amplitude. When seismic waves slow down, energy is transferred to the amplitude, increasing the motion of surface waves, which in turn amplifies ground shaking.

Focus depth—Deeper earthquakes cause less surface shaking because much of their energy, transmitted as body waves, is lost before reaching the surface. Recall that surface waves are generated by P and S waves impacting the Earth’s surface.

Factors that Determine Destruction

Just as certain conditions will impact the intensity of ground-shaking, several factors affect how much destruction is caused.

Figure 9.8.1: Example of devastation on unreinforced masonry by seismic motion.

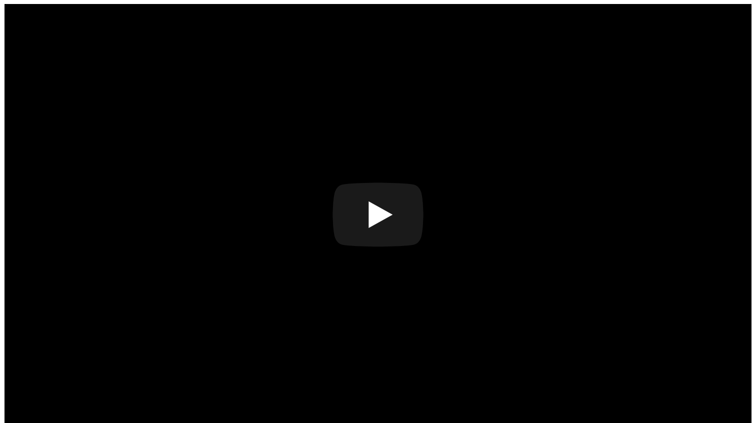
Building Materials—The flexibility of a building material determines its resistance to earthquake damage. Unreinforced masonry (URM) is the material most devastated by ground shaking. Wood framing fastened with nails bends and flexes during seismic wave passage and is more likely to survive intact. Steel also has the ability to deform elastically before brittle failure. The Fix the Bricks campaign in Salt Lake City, Utah has good information on URMs and earthquake safety.

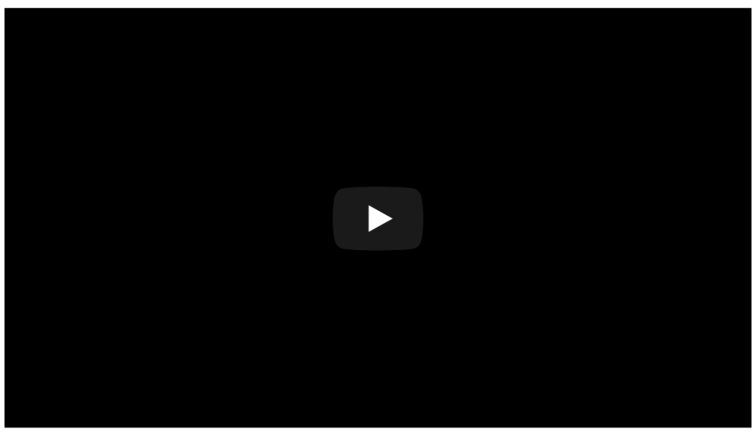
Intensity and Duration—Greater shaking and duration of shaking causes more destruction than lower and shorter shaking.

Resonance—Resonance occurs when seismic wave frequency matches a building’s natural shaking frequency and increases the shaking happened in the 1985 Mexico City Earthquake, where buildings of heights between 6 and 15 stories were especially vulnerable to earthquake damage. Skyscrapers designed with earthquake resilience have dampers and base isolation features to reduce resonance.

Resonance is influenced by the properties of the building materials. Changes in the structural integrity of a building can alter resonance [8]. Conversely, changes in measured resonance can indicate a potentially altered structural integrity.

These two videos discuss why buildings fall during earthquakes and a modern procedure to reduce potential earthquake destruction for larger buildings.





Earthquake Recurrence

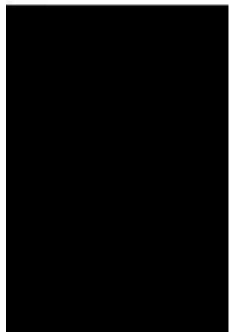
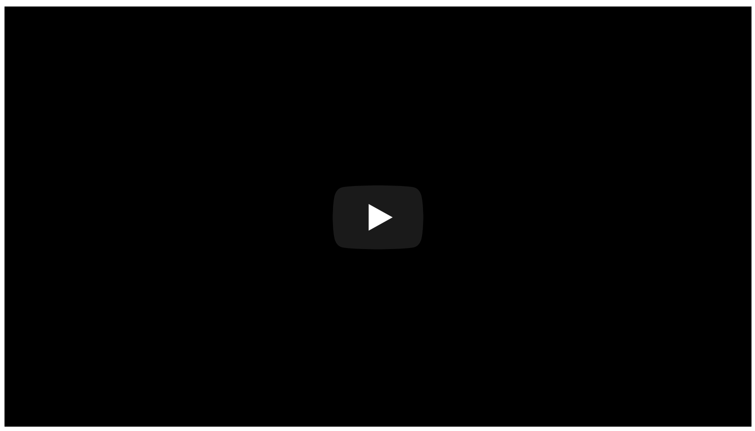


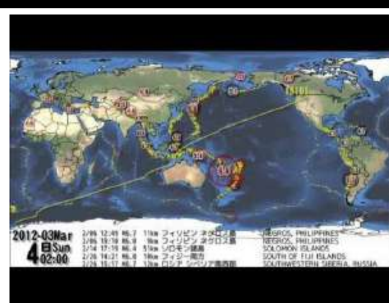
Figure 9.8.1: Fault trench near Draper Utah. Trenches allow geologists to see a cross-section of a fault and to use dating techniques to determine how frequently earthquakes occur.

A long hiatus in activity along a fault segment with a history of recurring earthquakes is known as a **seismic gap**. The lack of activity may indicate the fault segment is locked, which may produce a buildup of strain and a higher probability of an earthquake recurring. Geologists dig earthquake trenches across faults to estimate the frequency of past earthquake occurrences. Trenches are effective for faults with relatively long **recurrence intervals**, roughly 100s to 10,000s of years between significant earthquakes. Trenches are less useful in areas with more frequent earthquakes because they usually have more recorded data.

Earthquake Distribution

This video shows the distribution of significant earthquakes on Earth during the years 2010 through 2012. Like volcanoes, earthquakes tend to aggregate around active boundaries of tectonic plates. The exception is intraplate earthquakes, which are comparatively rare.





Subduction Zones—Subduction zones, found at convergent plate boundaries, are where the largest and deepest earthquakes, called megathrust earthquakes, occur. Examples of subduction-zone earthquake areas include the Sumatran Islands, Aleutian Islands, west coast of South America, and Cascadia Subduction Zone off the coast of Washington and Oregon. See Chapter 2 for more information about subduction zones.

Collision Zones—Collisions between converging continental plates create broad earthquake zones that may generate deep, large earthquakes from the remnants of past subduction events or other deep-crustal processes. The Himalayan Mountains (northern border of the Indian subcontinent) and Alps (southern Europe and Asia) are active regions of collision-zone earthquakes. See Chapter 2 for more information about collision zones.

Transform Boundaries—Strike-slip faults created at transform boundaries produce moderate-to-large earthquakes, usually having a maximum moment magnitude of about 8. Transform fault boundaries create moderate and large earthquakes, usually having a maximum magnitude of about 8. The San Andreas fault (California) is an example of a transform-boundary earthquake zone. Other examples are the Alpine Fault (New Zealand) and Anatolian Faults (Turkey). See Chapter 2 for more information about transform boundaries.

Divergent Boundaries—Continental rifts and mid-ocean ridges found at divergent boundaries generally produce moderate earthquakes. Examples of active earthquake zones include the East African Rift System (southwestern Asia through eastern Africa), Iceland, and Basin and Range province (Nevada, Utah, California, Arizona, and northwestern Mexico). See Chapter 2 for more information about divergent boundaries.

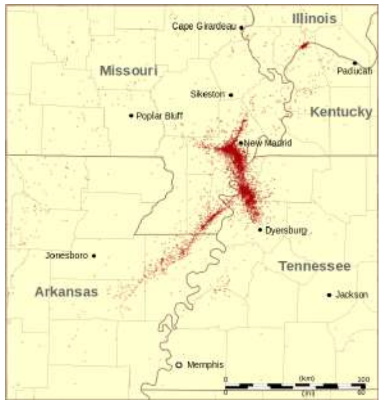


Figure 9.8.1: High density of earthquakes in the New Madrid seismic zone.

Intraplate Earthquakes—Intraplate earthquakes are not found near tectonic plate boundaries, but generally occur in areas of weakened crust or concentrated tectonic stress. The New Madrid seismic zone, which covers Missouri, Illinois, Tennessee, Arkansas, and Indiana, is thought to represent the failed Reelfoot rift [9]. The failed rift zone weakened the crust, making it more responsive to tectonic plate movement and interaction. Geologists theorize the infrequently occurring earthquakes are produced by low strain rates

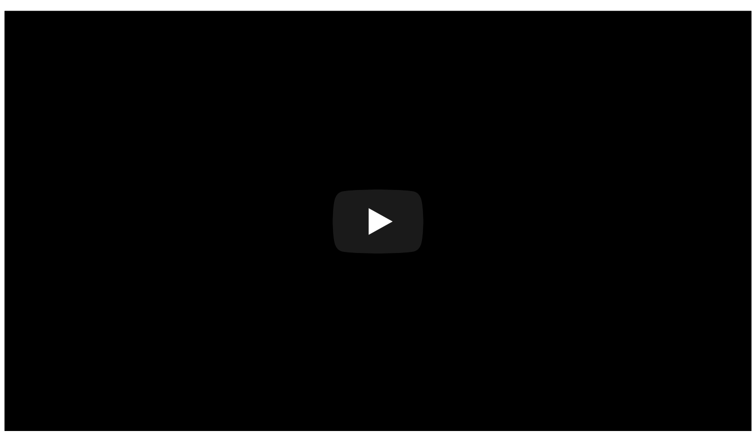
Secondary Hazards Caused by Earthquakes

Most earthquake damage is caused by ground shaking and fault-block displacement In addition, there are secondary hazards that endanger structures and people, in some cases after the shaking stops.

Figure 9.8.1: Buildings toppled from liquefaction during a 7.5 magnitude earthquake in Japan.

Liquefaction—Liquefaction occurs when water-saturated, unconsolidated sediments, usually silt or sand, become fluid-like from shaking. The shaking breaks the cohesion between grains of sediment, creating a slurry of particles suspended in water. Buildings settle or tilt in the liquified sediment, which looks very much like quicksand in the movies. Liquefaction also creates sand volcanoes, cone-shaped features created when liquefied sand is squirted through an overlying and usually finer-grained layer.

This video demonstrates how liquefaction takes place.



Tsunamis—Among the most devastating natural disasters are tsunamis, earthquake-induced ocean waves. When the seafloor is offset by fault movement or an underwater landslide, the ground displacement lifts a volume of ocean water and generates the tsunami wave. Ocean wave behavior, which includes tsunamis, is covered in Chapter 12. Tsunami waves are fast-moving with low amplitude in deep ocean water but grow significantly in amplitude in the shallower waters approaching the shore. When a tsunami is about to strike land, the drawback of the trough preceding the wave crest causes the water to recede dramatically from shore. Tragically, curious people wander out and follow the disappearing water, only to be overcome by an oncoming wall of water that can be upwards of a 30 m (100 ft) high. Early warning systems help mitigate the loss of life caused by tsunamis.

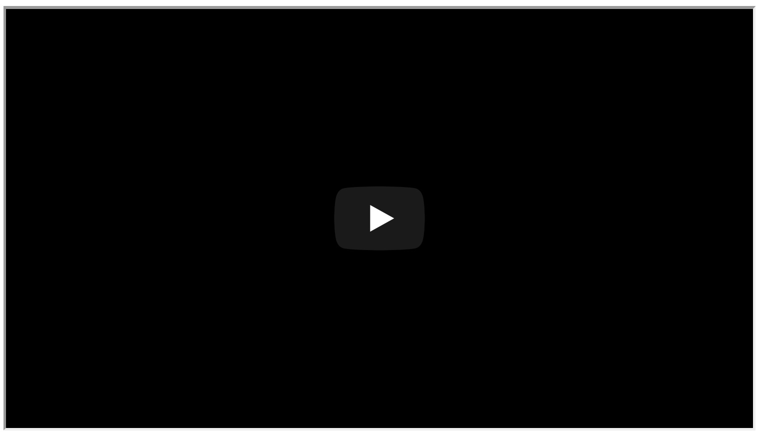
Figure 9.8.1: As the ocean depth becomes shallower, the wave slows down and pile up on top of itself, making large, high-amplitude waves.

Figure 9.8.1: House in Springdale, Utah destroyed by an earthquake-triggered landslide.

Landslides—Shaking can trigger landslides (see Chapter 10). In 1992 a moment magnitude 5.9 earthquake in St. George, Utah, caused a landslide that destroyed several structures in the Balanced Rock Hills subdivision in Springville, Utah [10].

Seiches—Seiches are waves generated in lakes by earthquakes. The shaking may cause water to slosh back-and-forth or sometimes change the lake depth. Seiches in Hebgen Lake during a 1959 earthquake caused major destruction to nearby structures and roads.

This video shows a seiche generated in a swimming pool by an earthquake in Nepal in 2015.



Land Elevation Changes—Elastic rebound and displacement along the fault plane can cause significant land elevation changes, such as subsidence or upheaval. The 1964 Alaska earthquake produced significant land elevation changes, with the differences in height between the hanging wall and footwall ranging from one to several meters (3–30 ft). The Wasatch Mountains in Utah represent an accumulation of fault scarps created a few meters at a time, over a few million years.

References

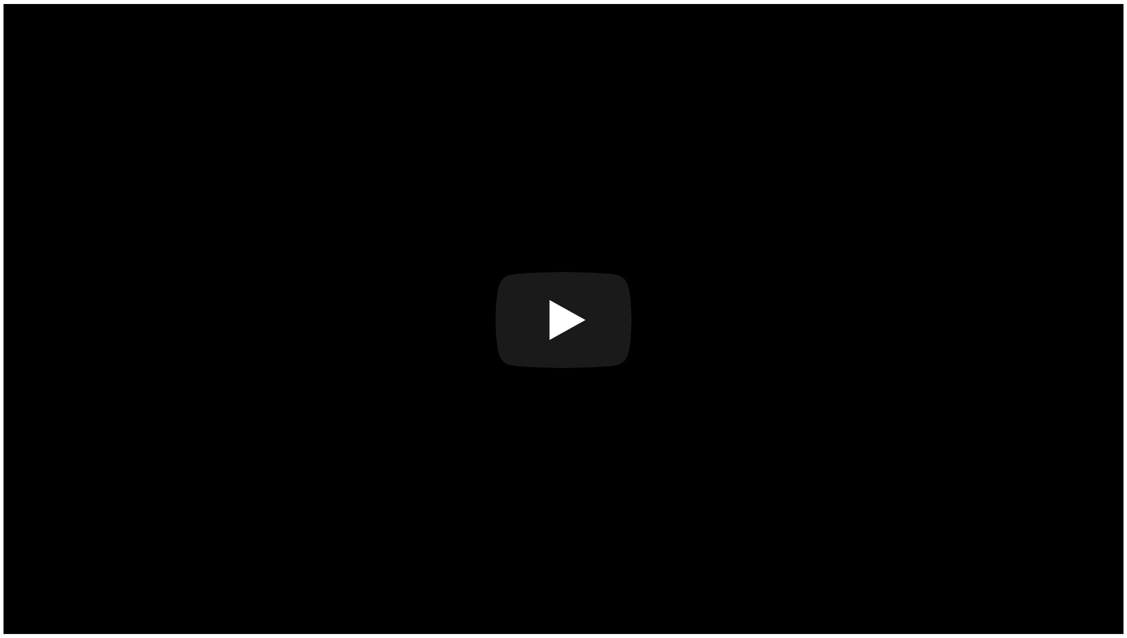
8. Starr AM (2015) Ambient resonance of rock arches. M.S. thesis, University of Utah

9. Hildenbrand TG, Hendricks JD (1995) Geophysical setting of the Reelfoot rift and relations between rift structures and the New Madrid seismic zone. U.S. Geological Survey, Washington; Denver, CO

10. Christenson GE (1995) The September 2, 1992 ML 5.8 St. George earthquake, Washington County, Utah. Utah Geological Survey, Salt Lake City, Utah

Contributions and Attributions

9.9: Case Studies



Video explaining the seismic activity and hazards of the Intermountain Seismic Belt and the Wasatch Fault, a large intraplate area of seismic activity.

North American Earthquakes

Basin and Range Earthquakes—Earthquakes in the Basin and Range Province, from the Wasatch Fault (Utah) to the Sierra Nevada (California), occur primarily in normal faults created by tensional forces. The Wasatch Fault, which defines the eastern extent of the Basin and Range province, has been studied as an earthquake hazard for more than 100 years.

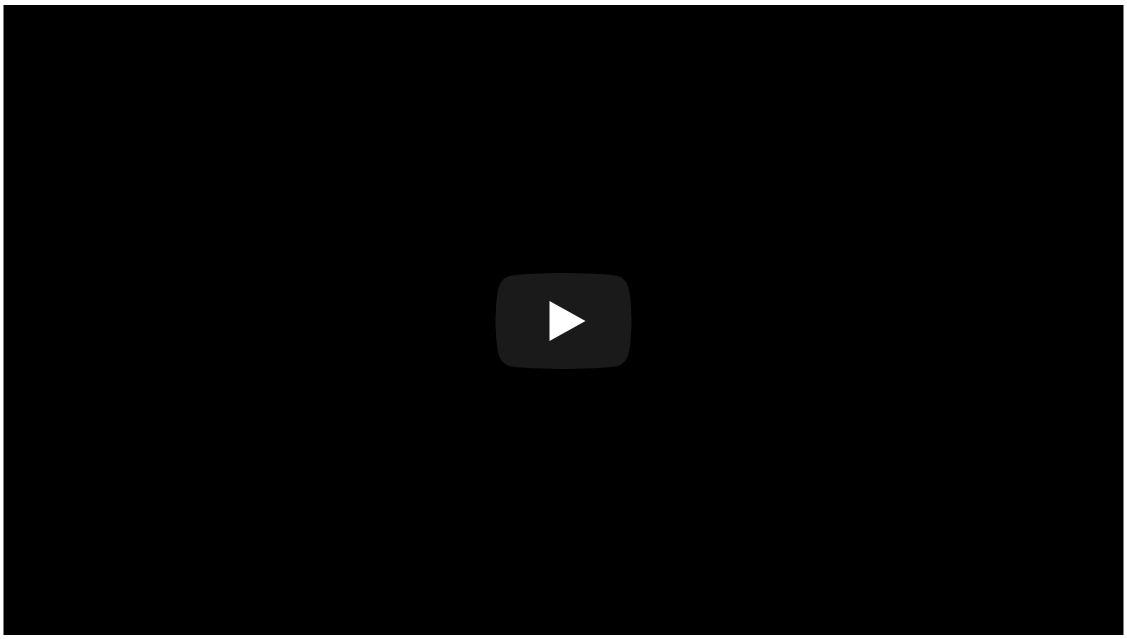
New Madrid Earthquakes (1811-1812)—Historical accounts of earthquakes in the New Madrid seismic zone date as far back as 1699 and earthquakes continue to be reported in modern times [11]. A sequence of large (M_w >7) occurred from December 1811 to February 1812 in the New Madrid area of Missouri [12]. The earthquakes damaged houses in St. Louis, affected the stream course of the Mississippi River, and leveled the town of New Madrid. These earthquakes were the result of intraplate seismic activity [9].

Charleston (1868)—The 1868 earthquake in Charleston South Carolina was a moment magnitude 7.0, with a Mercalli intensity of X, caused significant ground motion, and killed at least 60 people. This intraplate earthquake was likely associated with ancient faults created during the breakup of Pangea. The earthquake caused significant liquefaction [13]. Scientists estimate the recurrence of destructive earthquakes in this area with an interval of approximately 1500 to 1800 years.

Great San Francisco Earthquake and Fire (1906)—On April 18, 1906, a large earthquake, with an estimated moment magnitude of 7.8 and MMI of X, occurred along the San Andreas fault near San Francisco California. There were multiple aftershocks followed by devastating fires, resulting in about 80% of the city being destroyed. Geologists G.K. Gilbert and Richard L. Humphrey, working independently, arrived the day following the earthquake and took measurements and photographs [14].

Figure 9.9.1: Remains of San Francisco after the 1906 earthquake and fire.

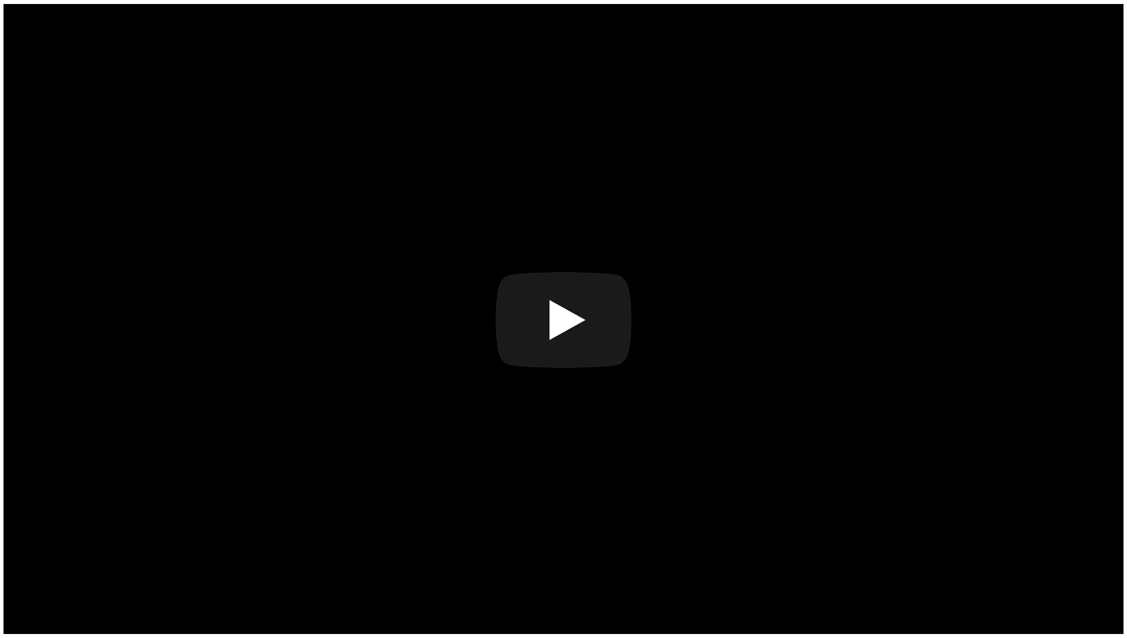
Alaska (1964)—The 1964 Alaska earthquake, moment magnitude 9.2, was one of the most powerful earthquakes ever recorded. The earthquake originated in a megathrust fault along the Aleutian subduction zone. The earthquake caused large areas of land subsidence and uplift, as well as significant mass wasting.



Video from the USGS about the 1964 Alaska earthquake.

Loma Prieta (1989)—The Loma Prieta, California, earthquake was created by movement along the San Andreas Fault. The moment magnitude 6.9 earthquake was followed by a magnitude of 5.2 aftershock. It caused 63 deaths, buckled portions of the several freeways, and collapsed part of the San Francisco-Oakland Bay Bridge.

This video shows how shaking propagated across the Bay Area during the 1989 Loma Prieta earthquake.



This video shows the destruction caused by the 1989 Loma Prieta earthquake.

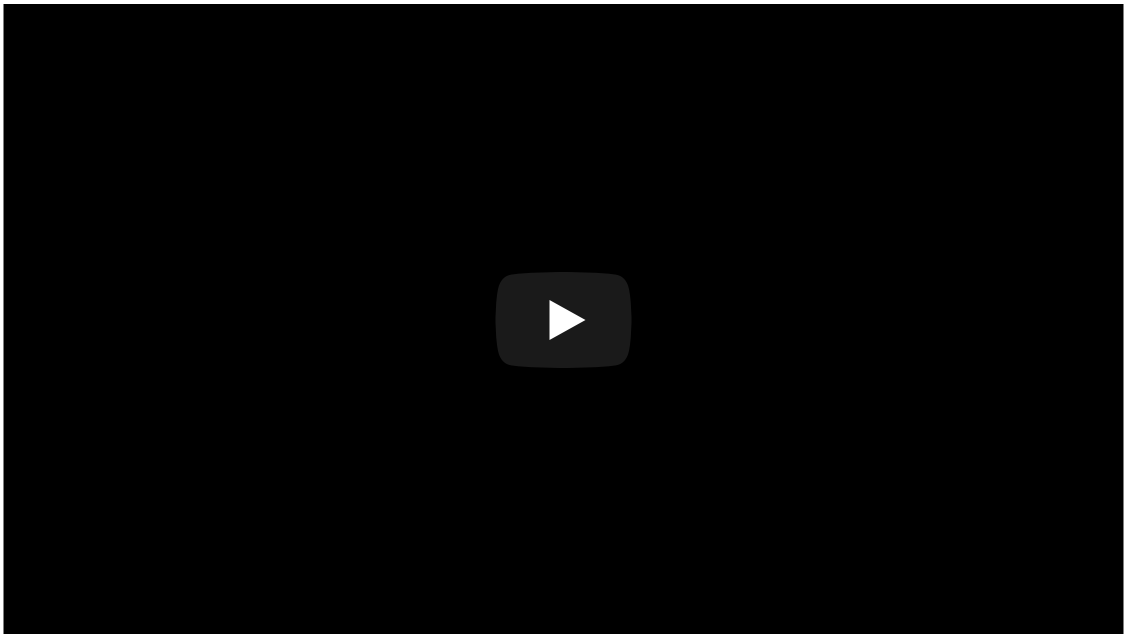
Global Earthquakes

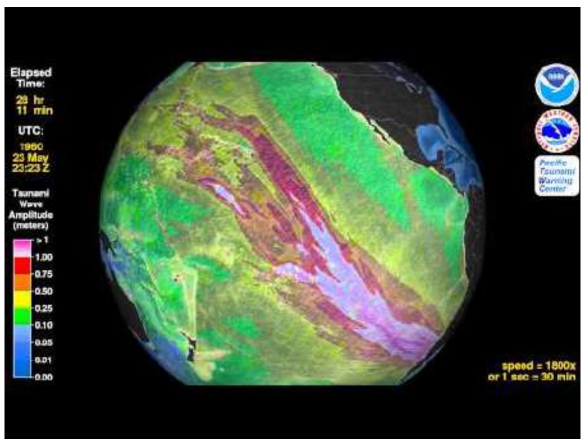
Many of history’s largest earthquakes occurred in megathrust zones, such as the Cascadia Subduction Zone (Washington and Oregon coasts) and Mt. Rainier (Washington).

Shaanxi, China (1556)—On January 23, 1556 an earthquake of an approximate moment magnitude 8 hit central China, killing approximately 830,000 people in what is considered the most deadly earthquake in history. The high death toll was attributed to the collapse of cave dwellings (*yaodong*) built in loess deposits, which are large banks of windblown, compacted sediment (see Chapter 5). Earthquakes in this are region are believed to have a recurrence interval of 1000 years. [15].

Lisbon, Portugal (1755)—On November 1, 1755 an earthquake with an estimated moment magnitude range of 8–9 struck Lisbon, Portugal [13], killing between 10,000 to 17,400 people [16]. The earthquake was followed by a tsunami.

Valdivia, Chile (1960)—The May 22, 1960 earthquake was the most powerful earthquake ever measured, with a moment magnitude of 9.4–9.6 and lasting an estimated 10 minutes. It triggered tsunamis that destroyed houses across the Pacific Ocean in Japan and Hawaii and caused vents to erupt on the Puyehue-Cordón Caulle (Chile).





Video describing the tsunami produced by the 1960 Chili earthquake.

Tangshan, China (1976)—Just before 4 a.m. (Beijing time) on July 28, 1976 a moment magnitude 7.8 earthquake struck Tangshan (Hebei Province), China, and killed more than 240,000 people. The high death toll is attributed to people still being asleep or at home and most buildings being made of URM.

Sumatra, Indonesia (2004)—On December 26, 2004, slippage of the Sunda megathrust fault generated a moment magnitude 9.0–9.3 earthquake off the coast of Sumatra, Indonesia [17]. This megathrust fault is created by the Australia plate subducting below the Sunda plate in the Indian Ocean [18]. The resultant tsunamis created massive waves as tall as 24 m (79 ft) when they reached the shore and killed more than an estimated 200,000 people along the Indian Ocean coastline.

Haiti (2010)—The moment magnitude 7 earthquake that occurred on January 12, 2010, was followed by many aftershocks of magnitude 4.5 or higher. More than 200,000 people are estimated to have died as a result of the earthquake. The widespread infrastructure damage and crowded conditions contributed to a cholera outbreak, which is estimated to have caused thousands more deaths.

Tōhoku, Japan (2011)—Because most Japanese buildings are designed to tolerate earthquakes, the moment magnitude 9.0 earthquake on March 11, 2011, was not as destructive as the tsunami it created. The tsunami caused more than 15,000 deaths and tens of billions of dollars in damage, including the destructive meltdown of the Fukushima nuclear power plant.

References

9. Hildenbrand TG, Hendricks JD (1995) Geophysical setting of the Reelfoot rift and relations between rift structures and the New Madrid seismic zone. U.S. Geological Survey, Washington; Denver, CO

11. Feldman J (2012) When the Mississippi Ran Backwards: Empire, Intrigue, Murder, and the New Madrid Earthquakes of 1811 and 1812. Free Press

12. Fuller ML (1912) The New Madrid earthquake. Central United States Earthquake Consortium, Washington, D.C.

13. Talwani P, Cox J (1985) Paleoseismic evidence for recurrence of Earthquakes near Charleston, South Carolina. Science 229:379–381

14. Gilbert GK, Holmes JA, Humphrey RL, et al (1907) The San Francisco earthquake and fire of April 18, 1906 and their effects on structures and structural materials. U.S. Geological Survey, Washington, D.C.

15. Boer JZ de, Sanders DT (2007) Earthquakes in human history: The far-reaching effects of seismic disruptions. Princeton University Press, Princeton

16. Aguirre B.E. (2012) Better disaster statistics: The Lisbon earthquake. J Interdiscip Hist 43:27–42

17. Rossetto T, Peiris N, Pomonis A, et al (2007) The Indian Ocean tsunami of December 26, 2004: observations in Sri Lanka and Thailand. Nat Hazards 42:105–124

18. Satake K, Atwater BF (2007) Long-Term Perspectives on Giant Earthquakes and Tsunamis at Subduction Zones. Annual Review of Earth and Planetary Sciences 35:349–374. https://doi.org/10.1146/annurev.earth.35.031306.140302

Contributions and Attributions

CHAPTER OVERVIEW

10: MASS WASTING

This chapter discusses the fundamental processes driving mass-wasting, types of mass wasting, examples and lessons learned from famous mass-wasting events, how mass wasting can be predicted, and how people can be protected from this potential hazard. Mass wasting is the downhill movement of rock and soil material due to gravity. The term landslide is often used as a synonym for mass wasting, but mass wasting is a much broader term referring to all movement downslope.

10.1: Slope Strength

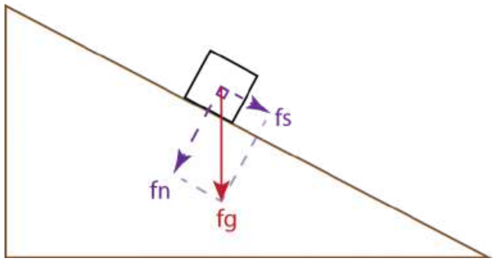


Figure 10.1.1: Forces on a block on an inclined plane (fg = force of gravity; fn = normal force; fs = shear force).

Mass wasting occurs when a slope fails. A slope fails when it is too steep and unstable for existing materials and conditions. Slope stability is ultimately determined by two principal factors: the slope angle and the strength of the underlying material. Force of gravity, which plays a part in mass wasting, is constant on the Earth's surface for the most part, although small variations exist depending on the elevation and density of the underlying rock. In the figure, a block of rock situated on a slope is pulled down toward the Earth's center by the force of gravity (fg). The gravitational force acting on a slope can be divided into two components: the shear or driving force (fs) pushing the block down the slope, and the normal or resisting force (fn) pushing into the slope, which produces friction. The relationship between shear force and normal force is called shear strength. When the normal force, i.e., friction, is greater than the shear force, then the block does *not* move downslope. However, if the slope angle becomes steeper or if the earth material is weakened, shear force exceeds normal force, compromising shear strength, and downslope movement occurs.

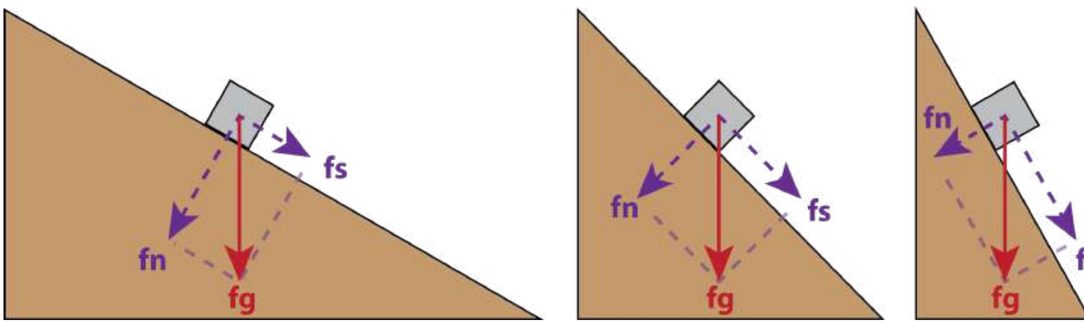


Figure 10.1.1: As slope increases, the force of gravity (fg) stays the same and the normal force decreases while the shear force proportionately increases.

In the figure, the force vectors change as the slope angle increases. The gravitational force doesn't change, but the shear force increases while the normal force decreases. The steepest angle at which rock and soil material is stable and will *not* move downslope is called the **angle of repose.** The angle of repose is measured relative from the horizontal. When a slope is at the angle of repose, the shear force is in equilibrium with the normal force. If the slope becomes just slightly steeper, the shear force exceeds the normal force, and the material starts to move downhill. The angle of repose varies for all material and slopes depending on many factors such as grain size, grain composition, and water content. The figure shows the angle of repose for sand that is poured into a pile on a flat surface. The sand grains cascade down the sides of the pile until coming to rest at the angle of repose. At that angle, the base and height of the pile continue to increase, but the angle of the sides remains the same.

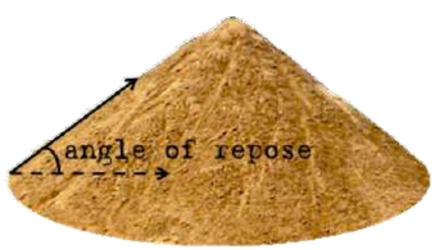


Figure 10.1.1: Angle of repose in a pile of sand.

Water is a common factor that can significantly change the shear strength of a particular slope. Water is located in **pore spaces**, which are empty air spaces in sediments or rocks between the grains. For example, assume a dry sand pile has an angle of repose of 30 degrees. If water is added to the sand, the angle of repose will increase, possibly to 60 degrees or even 90 degrees, such as a sandcastle being built at a beach. But if too much water is added to the pore spaces of the sandcastle, the water decreases the shear strength, lowers the angle of repose, and the sandcastle collapses.

Another factor influencing shear strength are planes of weakness in sedimentary rocks. Bedding planes (see Chapter 5) can act as significant planes of weakness when they are parallel to the slope but less so if they are perpendicular to the slope [1]. locations A and B, the bedding is nearly perpendicular to the slope and relatively stable. At location D, the bedding is nearly parallel to the slope and quite unstable. At location C, the bedding is nearly horizontal, and the stability is intermediate between the other two extremes [1]. Additionally, if clay minerals form along bedding planes, they can absorb water and become slick. When a bedding plane of shale (clay and silt) becomes saturated, it can lower the shear strength of the rock mass and cause a landslide, such as at the 1925 Gros Ventre, Wyoming rock slide. See the case studies section for details on this and other landslides.

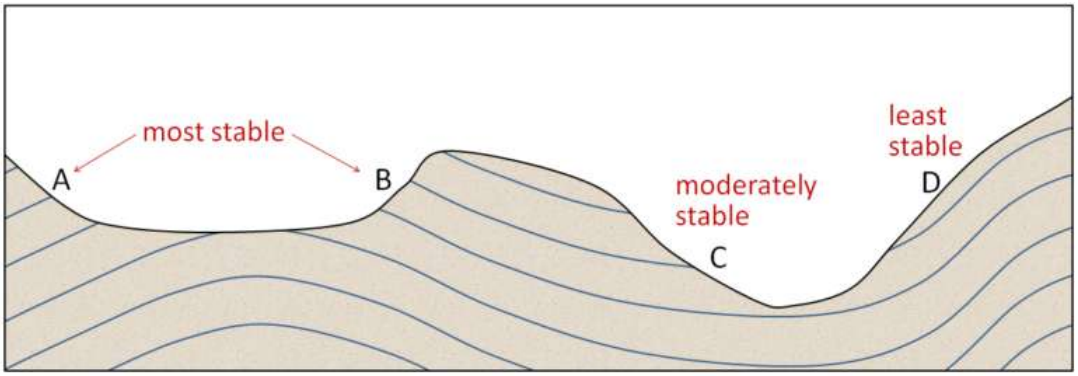


Figure 10.1.1: Locations A and B have bedding nearly perpendicular to the slope, making for a relatively stable slope. Location D has bedding nearly parallel to the slope, increasing the risk of slope failure. Location C has bedding nearly horizontal and the stability is relatively intermediate.

Reference

1. Earle, S. *Physical geology OER textbook*. (BC Campus OpenEd, 2015).

Contributions and Attributions

10.2: Mass-Wasting Triggers and Mitigation

Mass-wasting events often have a **trigger**: something changes that cause a landslide to occur at a specific time. It could be rapid snowmelt, intense rainfall, earthquake shaking, volcanic eruption, storm waves, rapid-stream erosion, or human activities, such as grading a new road. Increased water content within the slope is the most common mass-wasting trigger. Water content can increase due to rapidly melting snow or ice or an intense rain event. Intense rain events can occur more often during El Niño years. Then, the west coast of North America receives more precipitation than normal, and landslides become more common. Changes in surface-water conditions resulting from earthquakes, previous slope failures that dam up streams or human structures that interfere with runoff, such as buildings, roads, or parking lots [2] can provide additional water to a slope. In the case of the 1959 Hebgen Lake rock slide, Madison Canyon, Montana, the shear strength of the slope may have been weakened by earthquake shaking. Most landslide mitigation diverts and drains water away from slide areas. Tarps and plastic sheeting are often used to drain water off of slide bodies and prevent infiltration into the slide. Drains are used to dewater landslides and shallow wells are used to monitor the water content of some active landslides.

An **oversteepened** slope may also trigger landslides. Slopes can be made excessively steep by natural processes of erosion or when humans modify the landscape for building construction. An example of how a slope may be oversteepened during development occurs where the bottom of the slope is cut into, perhaps to build a road or level a building lot, and the top of the slope is modified by depositing excavated material from below. If done carefully, this practice can be very useful in land development, but in some cases, this can result in devastating consequences. For example, this might have been a contributing factor in the 2014 North Salt Lake City, Utah landslide. A former gravel pit was regraded to provide a road and several building lots. These activities may have oversteepened the slope, which resulted in a slow-moving landslide that destroyed one home at the bottom of the slope. Natural processes such as excessive stream erosion from a flood or coastal erosion during a storm can also oversteepened slopes. For example, natural undercutting of the riverbank was proposed as part of the trigger for the famous 1925 Gros Ventre, Wyoming rock slide.

Slope reinforcement can help prevent and mitigate landslides. For rockfall-prone areas, sometimes it is economical to use long steel bolts. Bolts, drilled a few meters into a rock face, can secure loose pieces of material that could pose a hazard. Shockcrete, a reinforced spray-on form of concrete, can strengthen a slope face when applied properly. Buttressing a slide by adding weight at the toe of the slide and removing weight from the head of the slide, can stabilize a landslide. Terracing, which creates a stairstep topography, can be applied to help with slope stabilization, but it must be applied at the proper scale to be effective.

A different approach in reducing landslide hazard is to shield, catch, and divert the runout material. Sometimes the most economical way to deal with a landslide hazard is to divert and slow the falling material. Special stretchable fencing can be applied in areas where rockfall is common to protect pedestrians and vehicles. Runout channels, diversion structures, and check dams can be used to slow debris flows and divert them around structures. Some highways have special tunnels that divert landslides over the highway. In all of these cases, the shielding has to be engineered to a scale that is greater than the slide, or catastrophic loss in property and life could result.

Reference

2. Highland, L. *Landslide types and processes*. (pubs.er.usgs.gov, 2004).

Contributions and Attributions

10.3: Landslide Classification and Identification

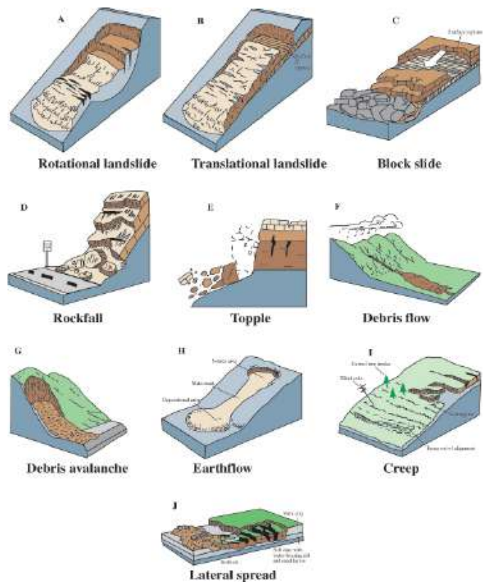
Mass-wasting events are classified by type of movement and type of material, and there are several ways to classify these events. The figure and table show the terms used. In addition, mass-wasting types often share common morphological features observed on the surface, such as the head scarp—commonly seen as crescent shapes on a cliff face; hummocky or uneven surfaces; accumulations of talus—loose rocky material falling from above; and toe of the slope, which covers existing surface material.

Types of Mass Wasting

The most common mass-wasting types are falls, rotational and translational slides, flows, and creep [2]. **Falls** are abrupt rock movements that detach from steep slopes or cliffs. Rocks separate along existing natural breaks such as fractures or bedding planes. Movement occurs as free-falling, bouncing, and rolling. Falls are strongly influenced by gravity, mechanical weathering, and water. **Rotational slides** commonly show slow movement along a curved rupture surface. **Translational slides** often are rapid movements along a plane of distinct weakness between the overlying slide material and the more stable underlying material. Slides can be further subdivided into rock slides, debris slides, or earth slides depending on the type of the material involved (see table).

Type of Movement	Primary Material Type and Common Name of Slide		
	Bedrock	Soil Types	
		Mostly Coarse-Grained	M
Falls	Rock Fall	—	
Rock Avalanche	Rock Avalanche	—	
Rotational Slide (Slump)	—	Rotational Debris Slide (Slump)	Rotat
Translational Slide	Translational Rock Slide	Translational Debris Slide	Tr
Flows	—	Debris Flow	
Soil Creep	—	Creep	

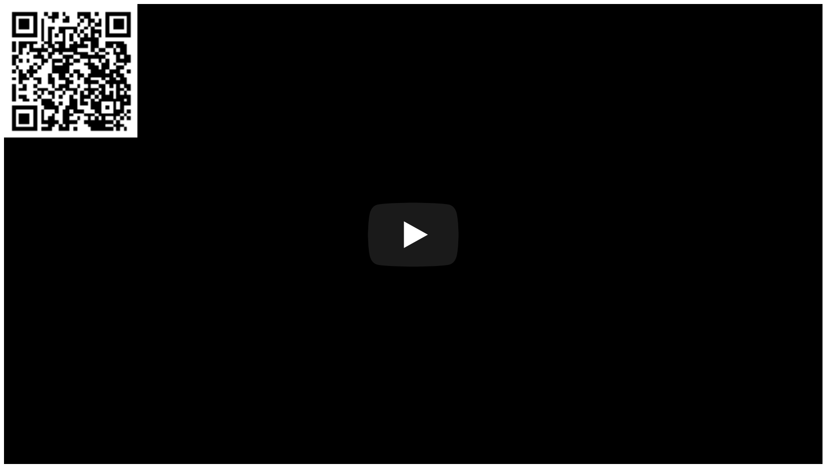
Table of Mass Wasting Types. Mass wasting movement type and primary earth material. Modified from [2; 3].



Examples of some of the types of landslides.

Flows are rapidly moving mass-wasting events in which the loose material is typically mixed with abundant water, creating long runouts at the slope base. Flows are commonly separated into **debris flow** (coarse material) and **earthflow** (fine material) depending on the type of material involved and the amount of water. Some of the largest and fastest flows on land are called **sturzstroms**, or long-runout landslides. They are still poorly understood but are known to travel for long distances, even in places without significant atmospheres like the Moon.

Creep is the imperceptibly slow downward movement of material caused by shear stress sufficient to produce permanent deformation in unconsolidated material [2]. Creep is indicated by curved tree trunks, bent fences or retaining walls, tilted poles or fences, and small soil ripples or ridges. A special type of soil creep is solifluction, which is the slow movement of soil lobes on low-angle slopes due to soil repeatedly freezing and thawing in high-latitude, typically sub-Arctic, Arctic, and Antarctic locations.



Landslide Hazards, David Applegate

Parts of a Landslide

Landslides have several identifying features that can be common across the different types of mass wasting. Note that there are many exceptions, and a landslide does not have to have these features. Displacement of material by landslides causes the absence of material uphill and the deposition of new material downhill, and careful observation can identify the evidence of that displacement. Other signs of landslides include tilted or offset structures or natural features that would normally be vertical or in place.

Many landslides have escarpments or scarps. Landslide scarps, like fault scarps, are steep terrain created when the movement of the adjacent land exposes a part of the subsurface. The most prominent scarp is the main scarp, which marks the uphill extent of the landslide. As the disturbed material moves out of place, a step slope forms and develops a new hillside escarpment for the undisturbed material. Main scarps are formed by the movement of the displaced material away from the undisturbed ground and are the visible part of the slide rupture surface.

Diagram of a rotational landslide

Parts	Description
Crown	The undisturbed material uphill of the scarp. (i.e. the brown house at the top of the hill)
Main Scarp	Steep slope at the upper edge of the landslide (at the head), caused by the movement of displaced material away from the undisturbed ground. The visible part of the slide surface.
Flank	Undisplaced materials adjacent to the sides of the landslide. Flank usually describes the left and right lateral extents of the mass-wasting material.
Rupture Surface/Slide Surface	The lower boundary of the movement below the original ground surface. This is the surface along which material slides. Also known as the surface of rupture.
Main Body	Part of the landslide that overlies the rupture surface.
Tension Cracks	Cracks formed as the result of the middle part of the slide being pulled apart. Usually found in the middle of the slide.
Separation Surface	Part of the original ground surface that is now covered by landslide material (i.e. the location of the landslide where the green and grey surfaces meet)
Foot	The part of the landslide that overlies the original ground surface (i.e. right below the separation surface).
Toe	The downhill end of the slide. The most distant part of the slide from the main scarp.

For more information, visit http://opengeology.org/textbook/10-mass-wasting/#1032_Parts_of_a_Landslide

The slide rupture surface is the boundary of the body of the movement of the landslide. The geologic material below the slide surface does not move, and is marked on the sides by the flanks of the landslide and at the end by the toe of the landslide.

The toe of the landslide marks the end of the moving material. The toe marks the runout, or maximum distance traveled, of the landslide. In rotational landslides, the toe is often a large, disturbed mound of geologic material, forming as the landslide moves past its original rupture surface.

Rotational and translational landslides often have extensional cracks, sag ponds, hummocky terrain, and pressure ridges. Extensional cracks form when a landslide’s toe moves forward faster than the rest of landslide, resulting in tensional forces. Sag ponds are small bodies of water filling depressions formed where landslide movement has impounded drainage. Hummocky terrain is undulating and uneven topography that results from the ground being disturbed. Pressure ridges develop on the margins of the landslide where the material is forced upward into a ridge structure [4].

References

2. Highland, L. *Landslide types and processes*. (pubs.er.usgs.gov, 2004).

3. Hungr, O., Leroueil, S. & Picarelli, L. The Varnes classification of landslide types, an update. *Landslides* **11**, 167–194 (2013).

4. Highland, L. M. & Bobrowsky, P. *The Landslide Handbook - A Guide to Understanding Landslides*. 147 (U.S. Geological Survey, 2008).

Contributions and Attributions

10.4: Examples of Landslides

Figure 10.4.1: Scar of the Gros Ventre landslide in the background with landslide deposits in the foreground.

1925, Gros Ventre, Wyoming: On June 23, 1925, a 38 million cubic meter (50 million cu yd) translational rock slide occurred next to the Gros Ventre River (pronounced “grow vont”) near Jackson Hole, Wyoming. Large boulders dammed the Gros Ventre River and ran up the opposite side of the valley several hundred vertical feet. The dammed river created Slide Lake, and two years later in 1927, lake levels rose high enough to destabilize the dam. The dam failed and caused a catastrophic flood that killed six people in the small downstream community of Kelly, Wyoming [5].

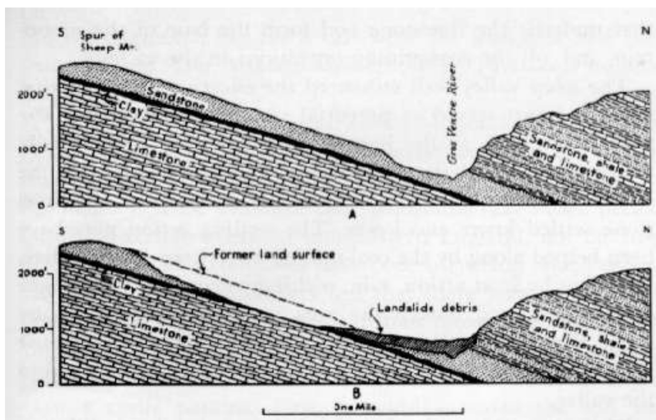


Figure 10.4.1: Cross-section of 1925 Gros Ventre slide showing sedimentary layers parallel with the surface and undercutting (oversteepening) of the slope by the river.

A combination of three factors caused the rock slide: 1) heavy rains and rapidly melting snow saturated the Tensleep Sandstone causing the underlying shale of the Amsden Formation to lose its shear strength, 2) the Gros Ventre River cut through the sandstone creating an oversteepened slope, and 3) soil on top of the mountain became saturated with water due to poor drainage [6]. The cross-section diagram shows how the parallel bedding planes between the Tensleep Sandstone and Amsden Formation offered little friction against the slope surface as the river undercut the sandstone. Lastly, the rockslide may have been triggered by an earthquake.

1959, Madison Canyon, Montana: In 1959, the largest earthquake in Rocky Mountain recorded history, magnitude 7.5, struck the Hebgen Lake, Montana area. The earthquake caused a rock avalanche that dammed the Madison River and ran up the other side of the valley hundreds of vertical feet. Today, there are still house-sized boulders visible on the slope opposite their starting point. The slide moved at a velocity of up to 160.9 kph (100 mph), creating an incredible air blast that swept through the Rock Creek Campground. The slide killed 28 people, most of whom were in the campground and remain buried there [5]. In a manner like the Gros Ventre slide, foliation planes of weakness in metamorphic rock outcrops were parallel with the surface, compromising shear strength.

Figure 10.4.1: 1959 Madison Canyon landslide scar. Photo taken from landslide material.

1980, Mount Saint Helens, Washington: On May 18, 1980, a 5.1-magnitude earthquake triggered the largest landslide observed in the historical record. This landslide was followed by the lateral eruption of Mount Saint Helens volcano, and the eruption was followed by volcanic debris flows known as lahars. The volume of material moved by the landslide was 2.8 cubic kilometers (0.67 mi^3) [7].

1995 and 2005, La Conchita, California: On March 4, 1995, a fast-moving earthflow damaged nine houses in the southern California coastal community of La Conchita. A week later, debris flow in the same location damaged five more houses. Surface-tension cracks at the top of the slide gave early warning signs in the summer of 1994. During the rainy winter season of 1994/1995, the cracks grew larger. The likely trigger of the 1995 event was unusually heavy rainfall during the winter of 1994/1995 and rising groundwater levels. Ten years later, in 2005, a rapid-debris flow occurred at the end of a 15-day period of near-record rainfall in southern California. Vegetation remained relatively intact as it was rafted on the surface of the rapid flow, indicating that much of the landslide mass simply was being carried on a presumably much more saturated and fluidized layer beneath. The 2005 slide damaged 36 houses and killed 10 people [8].

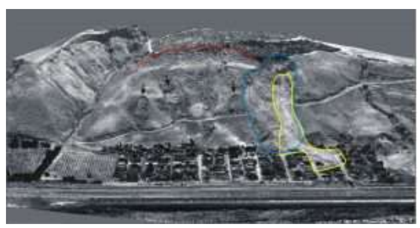
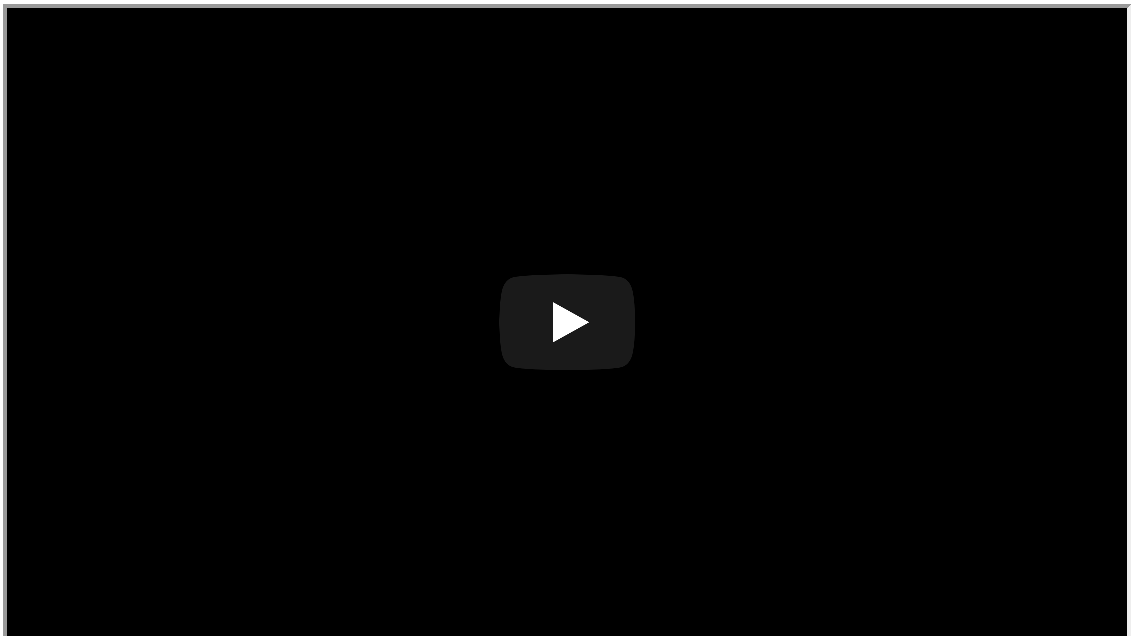


Figure 10.4.1: Oblique LIDAR image of La Conchita after the 2005 landslide. Outline of 1995 (blue) and 2005 (yellow) landslides shown; arrows show examples of other landslides in the area; red line outlines main scarp of an ancient landslide for the entire bluff. Source: Todd Stennett, Airborne 1 Corp., El Segundo. Public domain

Figure 10.4.1: 1995 La Conchita slide. Source: USGS.



La Conchita Landslide

Figure 10.4.1: 2014 Oso slide in Washington killed 43 people and buried many homes (source: USGS, public domain).

2014, Oso Landslide, Washington: On March 22, 2014, a landslide of approximately 18 million tons (10 million yd^3) traveled at 64 kph (40 mph), extended for nearly a 1.6 km (1 m), and dammed the North Fork of the Stillaguamish River. The landslide covered 40 homes and killed 43 people in the Steelhead Haven community near Oso, Washington. It produced a volume of material equivalent to 600 football fields covered in material 3 m (10 ft) deep. The winter of 2013-2014 was unusually wet with almost double the average amount of precipitation. The landslide occurred in an area of the Stillaguamish River Valley historically active with many landslides, but previous events had been small [11].

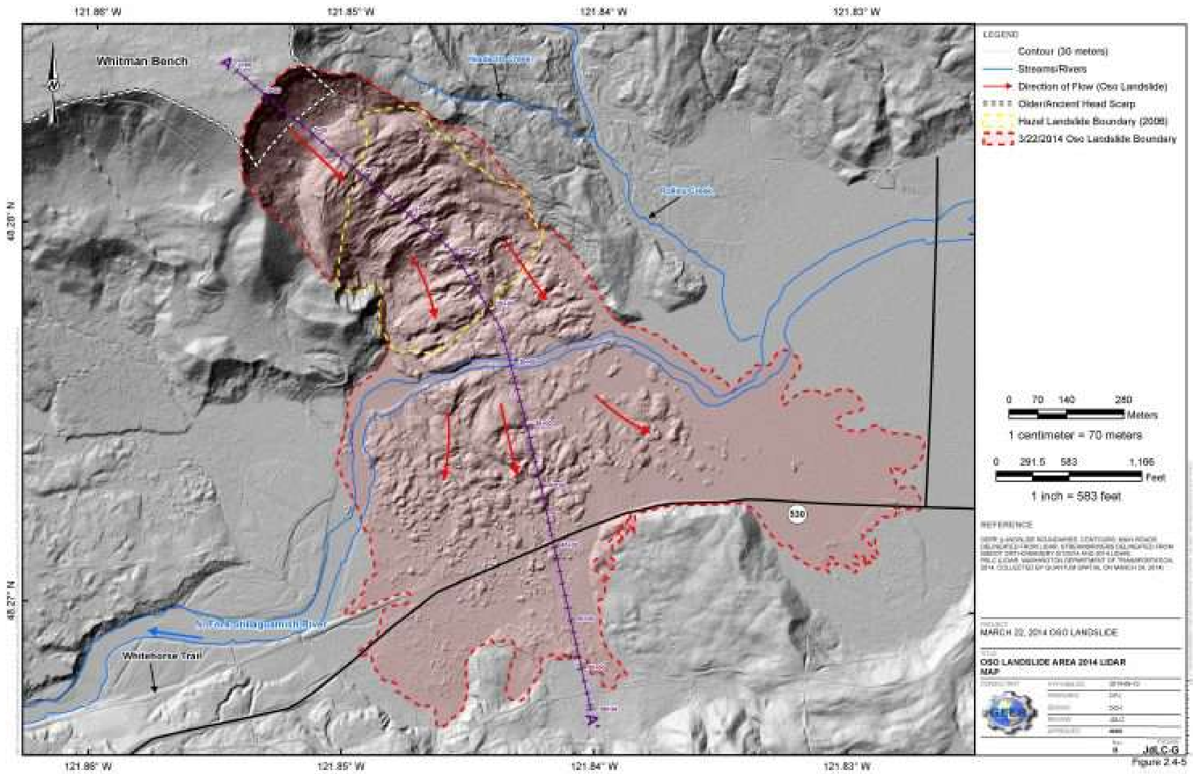
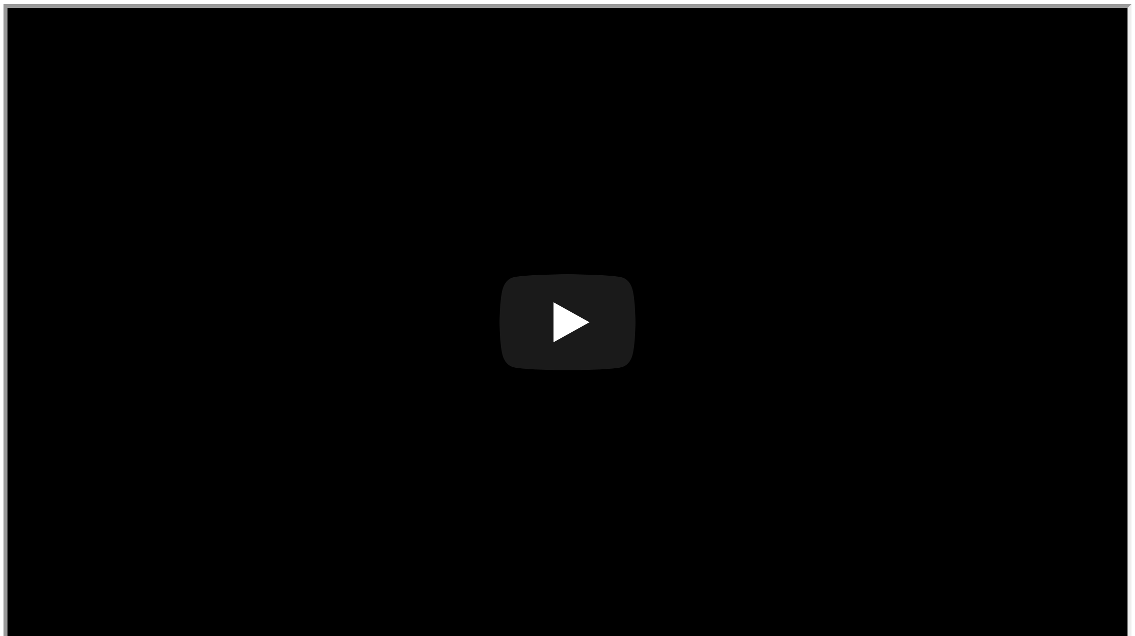


Figure 10.4.1: Annotated LiDAR map of 2014 Oso slide in Washington.

Yosemite National Park Rock Falls: The steep cliffs of Yosemite National Park cause frequent rockfalls. Fractures created to tectonic stresses and exfoliation and expanded by frost wedging can cause house-sized blocks of granite to detach from the cliff-faces of Yosemite National Park. The park models potential runout, the distance landslide material travels, to better assess the risk posed to the millions of park visitors.



Rockfalls in Yosemite.

Utah Landslides

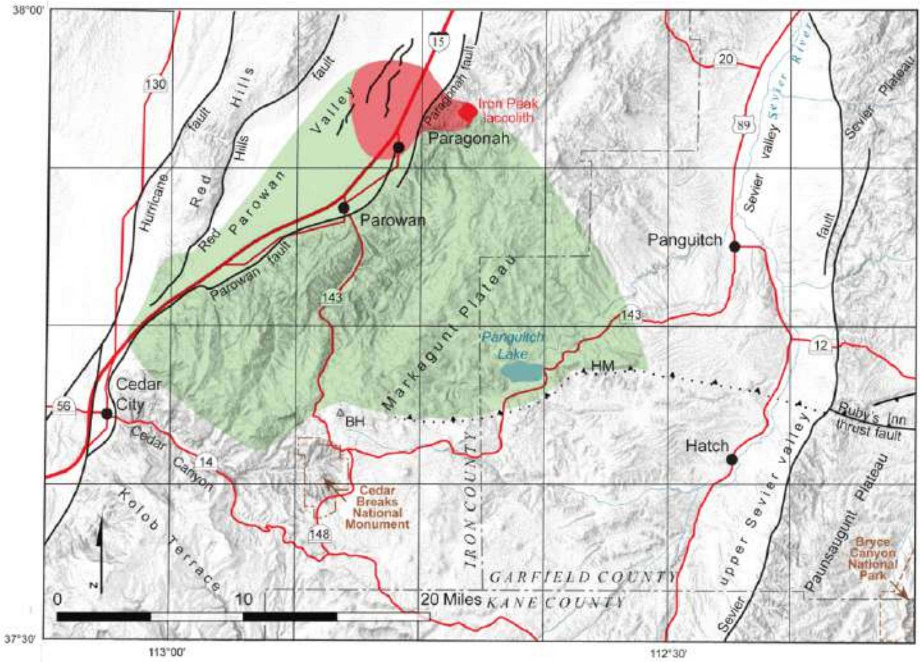


Figure 10.4.1: Approximate extent of Markagunt Gravity slide.

Markagunt Gravity Slide: About 21–22 million years ago, one of the biggest land-based landslides yet discovered in the geologic record displaced more than 1,700 cu km (408 cu mi) of material in one relatively fast event. Evidence for this slide includes breccia conglomerates (see Chapter 5), glassy pseudotachylyte, (see Chapter 6), slip surfaces (similar to faults) see Chapter 9), and dikes (see Chapter 7). The landslide is estimated to encompass an area the size of Rhode Island and to extend from near Cedar City, Utah to Panguitch, Utah. This landslide was likely the result of material released from the side of a growing laccolith (a type of igneous intrusion) see Chapter 4), after being triggered by an eruption-related earthquake.

Figure 10.4.1: The 1983 Thistle landslide (foreground) dammed the Spanish Fork river creating a lake.

1983, Thistle Slide: Starting in April of 1983 and continuing into May of that year, a slow-moving landslide traveled 305 m (1,000 ft) downhill and blocked Spanish Fork Canyon with an earthflow dam 61 m (200 ft) high. This caused disastrous flooding upstream in the Soldier Creek and Thistle Creek valleys, submerging the town of Thistle. As part of the emergency response, a spillway was constructed to prevent the newly formed lake from breaching the dam. Later, a tunnel was constructed to drain the lake, and currently, the river continues to flow through this tunnel. The rail line and US-6 highway had to be relocated at a cost of more than $200 million [13].

Figure 10.4.1: House before and after destruction from 2013 Rockville rockfall.

2013, Rockville Rock Fall: Rockville, Utah is a small community near the entrance to Zion National Park. In December of 2013, a 2,700 ton (1,400 yd^3) block of Shinarump Conglomerate fell from the Rockville Bench cliff, landed on the steep 35-degree slope below, and shattered into several large pieces that continued downslope at a high speed. These boulders completely destroyed a house located 375 feet below the cliff (see the before and after photographs) and killed two people inside the home. The topographic map shows other rockfalls in the area prior to this catastrophic event [14].

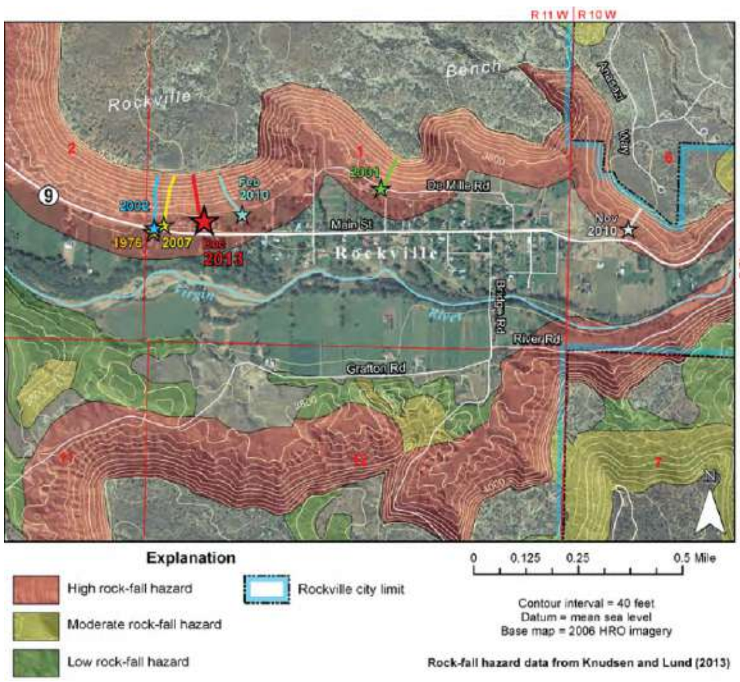
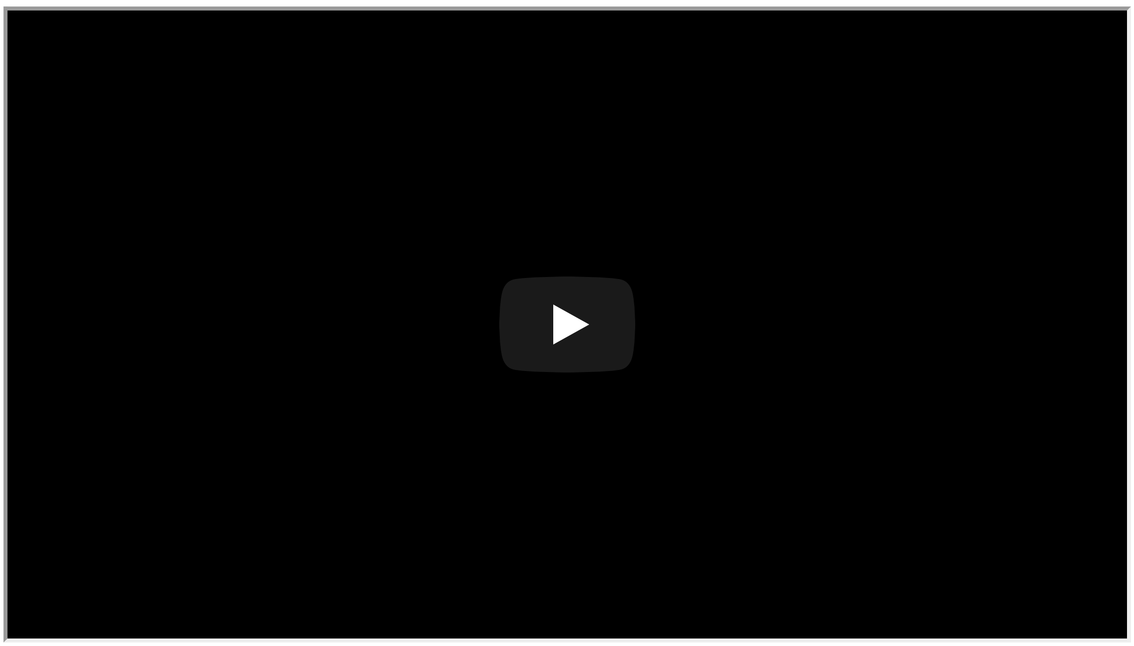


Figure 10.4.1: Tracks of deadly 2013 Rockville rockfall and earlier documented rockfall events.

2014, North Salt Lake Slide: In August 2014 after a particularly wet period, a slow-moving rotational landslide destroyed one home and damaged nearby tennis courts.

Figure 10.4.1: Scarp and displaced material from the North Salt Lake (Parkview) slide of 2014.

Reports from residents suggested that ground cracks had been seen near the top of the slope at least a year prior to the catastrophic movement. The presence of easily-drained sands and gravels overlying more impermeable clays weathered from volcanic ash, along with recent regrading of the slope, may have been contributing causes of this slide. Local heavy rains seem to have provided the trigger. In the two years after the landslide, the slope has been partially regraded to increase its stability. Unfortunately, in January 2017, parts of the slope have shown reactivation movement. Similarly, in 1996 residents in a nearby subdivision started reporting distress to their homes. This distress continued until 2012 when 18 homes became uninhabitable due to extensive damage and were removed. A geologic park was constructed in the now vacant area.



North Salt Lake Landslide

2013, Bingham Canyon Copper Mine Landslide, Utah: At 9:30 pm on April 10, 2013, more than 65 million cubic meters of steep terraced mine wall slid down into the engineered pit of Bingham Canyon mine, making it one of the largest historic landslides not associated with volcanoes. Radar systems maintained by the mine operator warned of movement of the wall, preventing the loss of life and limiting the loss of property.

Figure 10.4.1: Before (left) vs. After (right)

References

5. Highland, L. M. & Schuster, R. L. *Significant landslide events in the United States*. (United States Geological Survey, 2000).

6. United States Forest Service. *A Brief History of the Gros Ventre Slide Geological Site*. (United States Forest Service, 2016).

7. Lipman, P. W. & Mullineaux, D. R. *The 1980 eruptions of Mount St. Helens, Washington*. 844 (US Geological Survey, 1981). doi:10.3133/pp1250

8. Jibson, R. W. *Landslide hazards at La Conchita, California*. (United States Geological Survey, 2005).

11. Haugerud, R. A. *Preliminary interpretation of pre-2014 landslide deposits in the vicinity of Oso, Washington*. (US Geological Survey, 2014).

13. Duncan, J. M., Fleming, R. W. & Patton, F. D. *Report of the Thistle Slide Committee to State of Utah, Department of Natural Resources, Division of Water Rights*. (U.S. Geological Survey, 1986).

14. Lund, W. R., Knudsen, T. R. & Bowman, S. D. *Investigation of the December 12, 2013, Fatal Rock Fall at 368 West Main Street, Rockville, Utah*. 24 (Utah Geological Survey, 2014).

Contributions and Attributions

CHAPTER OVERVIEW

11: WATER

Learning Objectives

Components and definition of the hydrologic cycle
Water users and volume of water used
How water is shared among people
Distribution of water on the planet
Define aquifer and confining layer
Properties required for a good aquifer

All life requires water. The hydrosphere (Earth's water) is an important agent of geologic change. It shapes our planet through weathering and erosion, deposits minerals that aid in lithification, and alters rocks after they are lithified. Water carried by subducted oceanic plates causes melting in the upper mantle material. Communities rely on suitable water sources for consumption, power generation, crop production, and many other things.

11.8: GROUNDWATER

Most rocks are not entirely solid and contain a certain amount of open space between grains or crystals, known as pores. Porosity is a measure of the open space in rocks –expressed as the percentage of open space that makes up the total volume of the rock or sediment material. Porosity can occur as primary porosity, which represents the original pore spaces in the rock (e.g. space between sand grains), or secondary porosity which occurs after the rock forms (e.g. dissolved portions of rock).

11.9: WATER CONTAMINATION

Water can be contaminated by various human activities or by existing natural features, like mineral-rich geologic formations. Agricultural activities, industrial operations, landfills, animal operations, and small and large scale sewage treatment processes, among many other things, all can potentially contribute to contamination. As water runs over the land or infiltrates into the ground, it dissolves material left behind by these potential contaminant sources.

11.10: KARST

Karst refers to landscapes and hydrologic features created by the dissolution of limestone. Karst can be found anywhere where there are limestone and other soluble subterranean substances like salt deposits. The dissolution of limestone creates features like sinkholes, caverns, disappearing streams, and towers. Karst forms when natural water, in combination with carbon dioxide, creates carbonic acid and dissolves calcite (calcium carbonate) in limestone.

11.1: Prelude to Water

Figure 11.1.1: Entrenched meander of the Colorado River, downstream of Page, Arizona.

All life requires water. The hydrosphere (Earth's water) is an important agent of geologic change. It shapes our planet through weathering and erosion, deposits minerals that aid in lithification, and alters rocks after they are lithified. Water carried by subducted oceanic plates causes melting in the upper mantle material. Communities rely on suitable water sources for consumption, power generation, crop production, and many other things.

Figure 11.1.1: Example of a Roman aqueduct in Segovia, Spain.

In pre-industrial civilizations, control of water resources was a symbol of power [1; 2]. Two thousand-year-old Roman aqueducts still grace European, Middle Eastern, and North African skylines. Ancient Mayan kings used water imagery such as frogs, water-lilies, waterfowl to show their divine power over their societies' water resources [3]. Mask facades of the hooked-nosed rain god, *Chac*, are prominent on Mayan buildings such as the Kodz-Poop (Temple of the Masks) at Kabah in the drier northern lowlands of the Yucatan peninsula but much rarer in the tropical, wet regions to the south. Control over water continues to be an integral part of the governmental duties of most modern societies.


Figure 11.1.1: Chac mask in Mexico.

References

1. Fagan B (2012) Elixir: A history of water and humankind, Reprint edition. Bloomsbury Press

2. Solomon S (2011) Water: The Epic Struggle for Wealth, Power, and Civilization, Reprint edition. Harper Perennial

3. Schele L, Miller ME, Kerr J, et al (1992) The Blood of Kings: Dynasty and Ritual in Maya Art, Reprint edition. George Braziller Inc.

11.2: Properties of Water

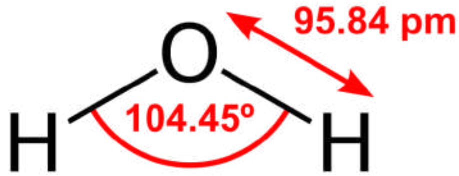


Figure 11.2.1: A model of a water molecule, showing the bonds between the hydrogen and oxygen.

The physical and chemical properties of water are what make it essential to life and useful to civilization. Water is a molecule made of one negatively charged (-2) oxygen ion and two positively-charged (+1) hydrogen ions, giving it the chemical formula H^2O, with strong covalent bonds between the oxygen and two hydrogen ions. The shape of the water molecule allows for an uneven distribution of charge, where one side is slightly positive and one side is slightly negative. Because of this **polarity**, water molecules form hydrogen bonds with each other. Hydrogen bonds are electrostatic intermolecular bonds that are weaker than ionic and covalent bonds (see discussion in the Minerals chapter). Water is amphoteric, that is it can self-ionize, breaking down into an acidic hydrogen ion (H^+) and a hydroxyl ion ($OH^–$), chemically a base. Because of its polarity and its ability to be **amphoteric**, water is a **universal solvent—**a chemical that can dissolve a wide range of other chemicals.

Figure 11.2.1: Surface tension keeps the paperclip on top of the water.

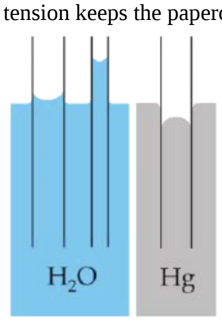


Figure 11.2.1: Capillary action by water in a narrow tube. Unlike water, mercury does not rise up in the tube.

Other side-effects of water’s polarity are **cohesion** (water likes to stick to itself) and **adhesion** (water likes to stick to other things). Water has the highest cohesion of all nonmetallic liquids. Cohesion gives water surface tension, allowing water glider insects to float on the water surface. Surface tension is what gives raindrops a spherical shape. Capillary action occurs when a combination of adhesive and cohesive forces causes water to move up narrow passages and tubes, rising higher than surrounding liquid. Capillary action happens when the adhesion of water to the tube is greater than the water’s internal cohesive forces. Paper towels have small pores that use capillary forces to clean up water spills. Plants use capillary forces to pump water into their tissues.

Water has a high specific heat capacity. Specific-heat is the amount of heat required to raise the temperature of a substance. Compared to many other substances, water requires a large amount of heat to raise its temperature. The high specific heat of water allows it to act as an energy buffer to extreme changes in air temperature. It also allows the oceans to soak up solar heat without changing temperature much and distribute that heat over the Earth by currents thus making the Earth habitable.

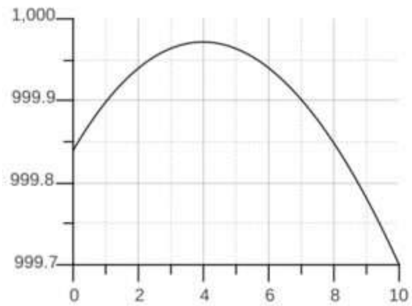


Figure 11.2.1: Density curve of water showing the greatest density at 3.98 degrees Celcius. Density is shown as kg/m3.

The density curve of water shows that as water is cooled, it becomes more dense, as do most other substances, but its greatest density occurs at about 4 degrees Celsius while most other substances continue to increase in density until they freeze. This unique density curve means that water is most dense just above its freezing point and sinks. Thus the oceans remain liquid. If water behaved like other substances, the oceans would be frozen.

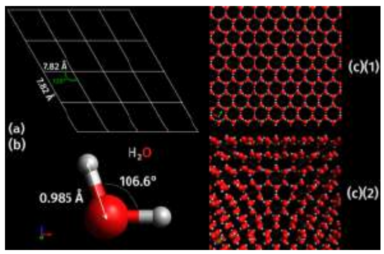


Figure 11.2.1: Molecular arrangement of hydrogen dioxide (water) molecules in ice.

When water freezes, the molecules arrange themselves in a well-ordered crystal structure, creating a spacing between molecules that is greater than if the water is in liquid form. The difference in molecular spacing causes ice to be less dense than water, making it more buoyant than liquid water, causing it to float on water. Because of its high specific heat capacity, ice floating on a lake's surface insulates the liquid water beneath and keeps it from freezing.

Because of its hydrogen bonds, water also has a high heat of vaporization. A significant amount of energy is required to evaporate water. As water evaporates, energy is absorbed by the breaking of hydrogen bonds and the air around the evaporating water is cooled. This energy is stored in the water vapor.

Contributions and Attributions

11.3: Water Cycle

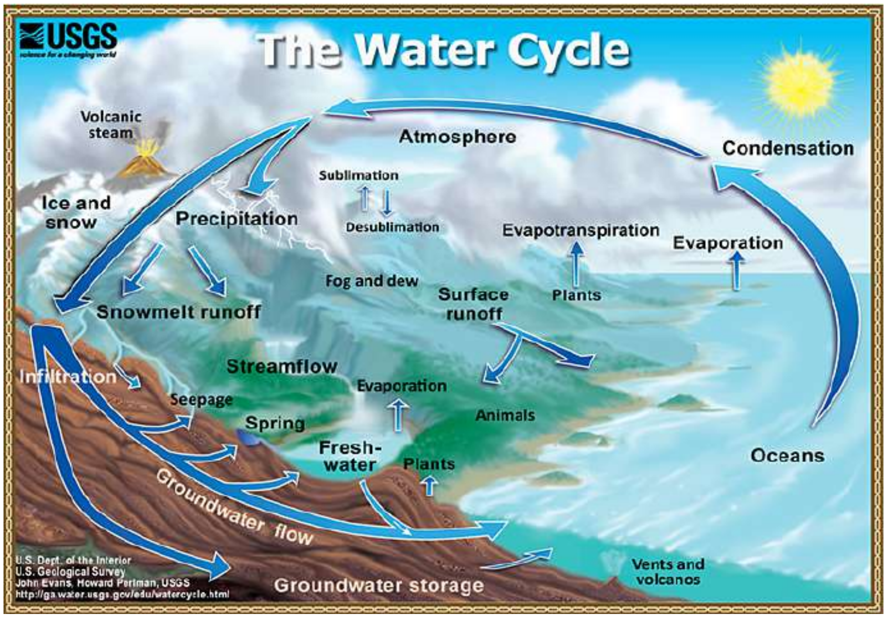
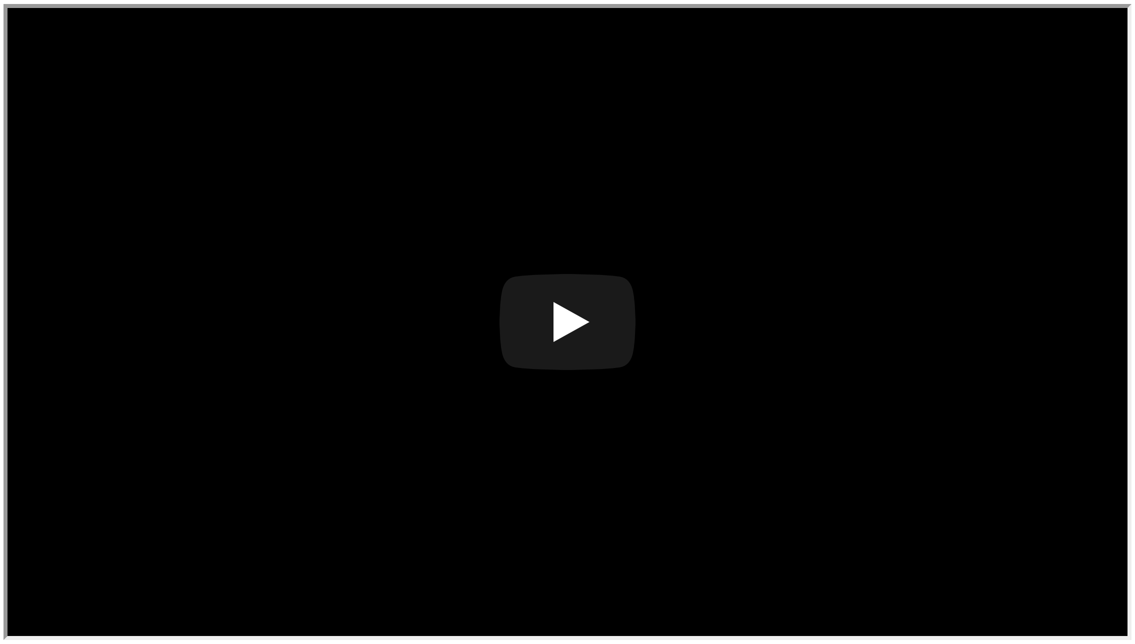


Figure 11.3.1: The water cycle.

The water cycle describes how water changes between solid, liquid, and gas (water vapor) phases and changes location. Water can be evaporated, which is the process where a liquid is converted to a gas. Solar energy warms the water sufficiently to excite the water molecules to the point of vaporization. Evaporation occurs from surface water bodies such as oceans, lakes, and streams and the land surface. Plants contribute significant amounts of water vapor as a byproduct of photosynthesis in a process called transpiration. Geologists commonly combine these two sources of water entering the atmosphere in a term called **evapotranspiration**.

Water vapor in the atmosphere can migrate long-distance from ocean to over land by way of prevailing winds. Over the ocean or land, the air can cool and cause the water to condense back into liquid water. This usually happens in the form of very small water droplets that form around a microscopic piece of dust or salt called condensation nuclei. These small water droplets are visible as a cloud. Clouds build and once the water droplets are big enough, they fall to earth as precipitation. Precipitation can take the form of rain, snow, hail, or sleet.

Once it has reached the surface it does two important things relevant to the geology of this chapter. At the surface, liquid water can flow as **runoff** into streams, lakes, and eventually back to the oceans (in most cases). Water in streams and lakes is called surface water. In addition, water can also **infiltrate** into the soil and finally fill the pore spaces in the rock or sediment deep underground to become **groundwater**, the name given to all subsurface water. Groundwater slowly moves through rock and unconsolidated materials and some of it eventually reaches the surface again, where it discharges as springs and into streams, lakes, and the ocean. Also, surface water in streams and lakes can infiltrate again to recharge groundwater. Therefore, the surface water and groundwater systems are connected.



Contributions and Attributions

11.4: Water Basin and Budgets

Drainage Basins

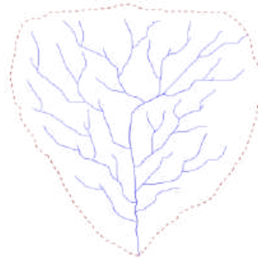


Figure 11.4.1: Map view of a drainage basin with main trunk streams and many tributaries with drainage divide in the dashed red line.

The basic unit of division of the landscape is the drainage basin. A **drainage basin**, also known as a **catchment** or **watershed**, is the area of land that captures precipitation and contributes runoff to a stream or stream segment (Figure). **Drainage divides** are local topographic high points that separate one drainage basin from another [5]. If water falls on one side of the divide, that water goes to one stream, and if it falls on the other side of the divide, then the water goes to a different stream. Each stream has its own drainage basin. Further, a drainage basin for each tributary can also be designated. In areas with flatter topography, drainage divides are not as easily identified but they still exist [6].

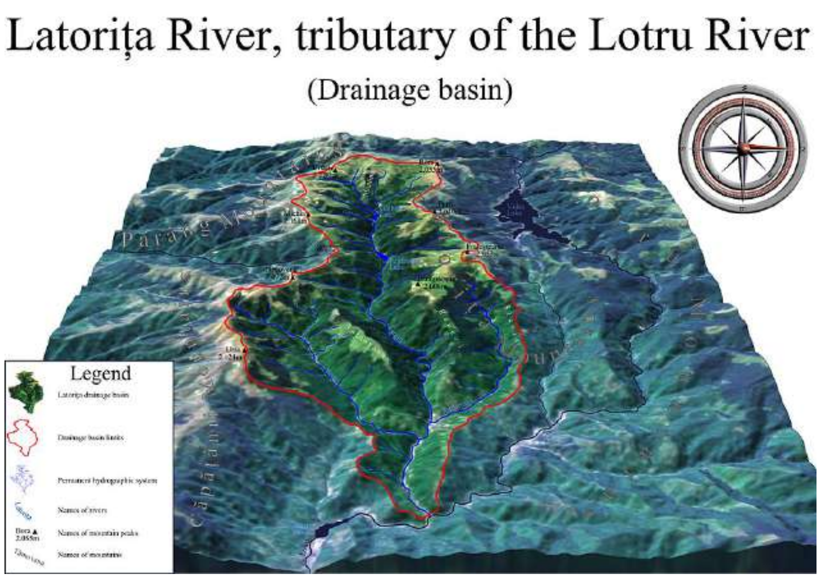


Figure 11.4.1: Oblique view of the drainage basin and divide of the Latorita River, Romania.

Streams only flow downhill and smaller **tributary** streams combine downhill to make the larger trunk of the stream. Where a stream begins is called the **headwaters** and where it finally reaches its end is called the **mouth**. Most streams have the mouth of the stream at the ocean. However, a rare number of streams do not flow to the ocean, but rather end in a **closed basin** (or endorheic basin) where the water evaporates from a stream or lake before it can reach the ocean. Most streams in the Great Basin are in closed basins. For example, Little Cottonwood Creek and the Jordan River flow into the Great Salt Lake where the water evaporates.

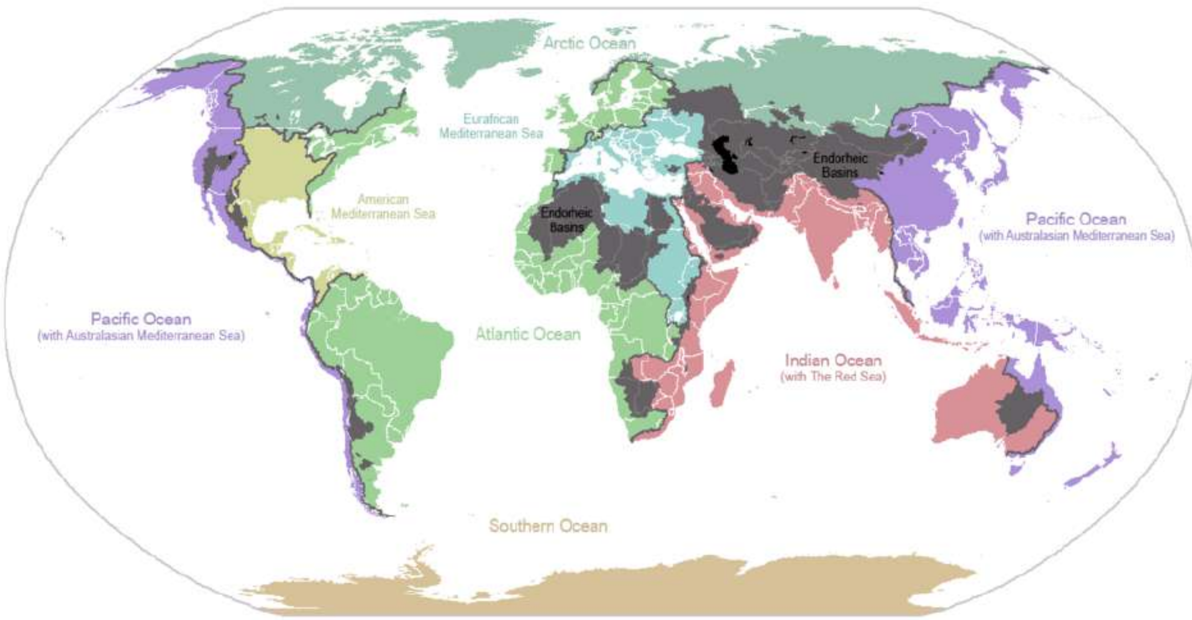


Figure 11.4.1: Major drainage basins color-coded to match the related ocean. Closed basins (or endorheic basins) are shown in gray.

In humid climates, many streams are perennial and in arid climates like Utah, many streams are ephemeral. **Perennial streams** flow all year round. They occur in humid or temperate climates where there are sufficient rainfall and low evaporation rates. Water levels rise and fall with the seasons, depending on the discharge. **Ephemeral streams** flow only during rain events or the wet season. These are often **dry washes** or **arroyos** for much of the year. They are above the water table and occur in dry climates with low amounts of rainfall and high evaporation rates. They flow mostly during flash floods.

Special Topic: Watershed Protection Areas

Along Utah's Wasatch Front, there are several watersheds that are designated as "Watershed Protection Areas" that limit the type of use allowed in those drainages. Dogs and swimming are limited in those watersheds because they could contribute harmful bacteria and substances to the drinking supply of Salt Lake City and surrounding municipalities.

Water in an area is very much like currency in a personal budget. There is income in the form of precipitation, stream inflow, and groundwater inflow, and there are expenses, in the form of groundwater withdrawal, evaporation, and stream and groundwater outflow. If the expenses outweigh the income, then the water budget is not balanced, and, if available, water will be removed from storage. Reservoirs, snow and ice, soil moisture, and aquifers can all act as storage in a water budget.

Water Budgets

Scientists can make groundwater budgets within any designated boundary, but they are generally made for watershed (basin) boundaries because groundwater and surface water are easier to account for within these boundaries. However, water budgets can be created for state, county, or aquifer extent boundaries as well. The groundwater budget is an essential component of the hydrologic model, where measured data are used with a conceptual workflow of the model to better understand the water system.

Like budgets for families and organizations, there is income from precipitation and expense from discharge and evaporation, with the ultimate expense for freshwater by returning to the salty ocean. For dry regions, the water budget is critical for sustaining human activities and understanding and managing it is an ongoing political challenge.

References

5. Seaber PR, Kapinos FP, Knapp GL (1987) Hydrologic unit maps

6. Charlton R (2007) Fundamentals of fluvial geomorphology. Taylor & Francis

Contributions and Attributions

11.5: Water Use and Distribution

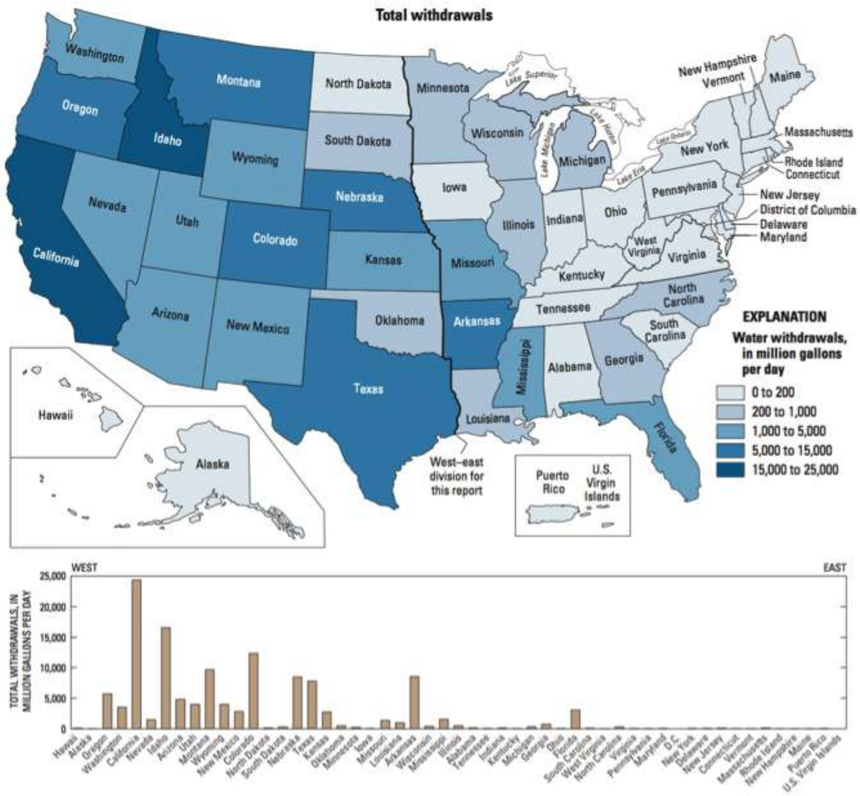


Figure 11.5.1: Agricultural water use in the United States by state.

In the United States, 355 billion gallons of ground and surface water are withdrawn for use each day, of which 76 billion gallons are fresh groundwater. The state of California accounts for 16% of national groundwater withdrawals [8].

Utah is the second driest state in the United States behind its neighbor Nevada, having a mean statewide precipitation of 12.2 inches per year. Utah also has the second-highest per capita rate of total domestic water use [8] of 167 gallons per day per person. With the combination of relatively high demand and limited quantity, Utah is at risk for water budget deficits.

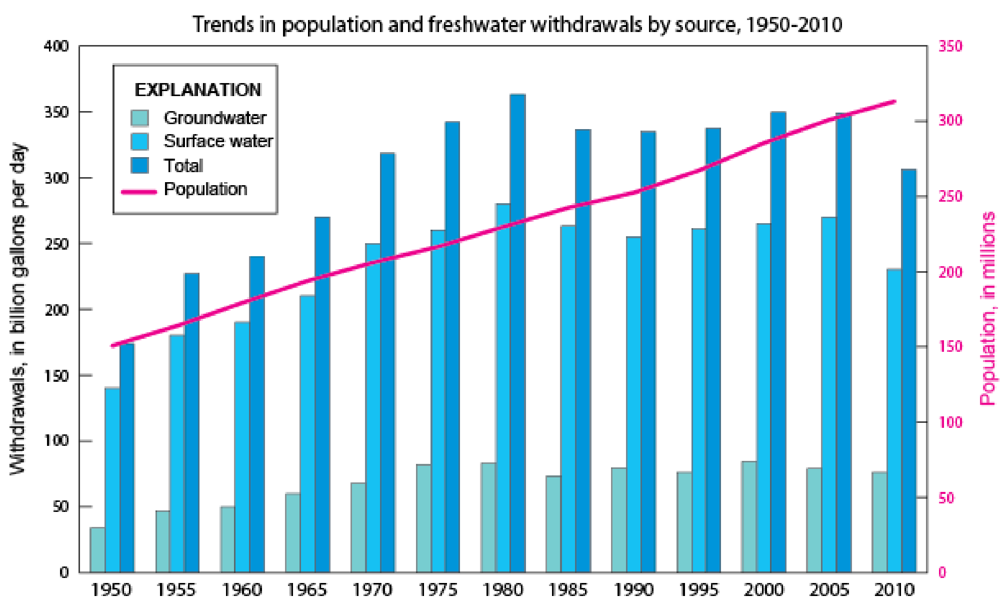


Figure 11.5.1: Trends in water use by source

Surface Water Distribution

Surface water makes up only 1.2% of the freshwater available on the planet, and 69% of that surface water is trapped in ground ice and permafrost. Water in rivers accounts for only 0.006% of all freshwater and lakes contain only 0.26% of the world's freshwater [9]. Freshwater is a precious resource and should not be taken for granted, especially in dry climate areas.

Global circulation patterns are the most important factor in precipitation, and thus, the distribution of surface water. In general, due to the Coriolis effect and the uneven heating of the Earth, air rises near the equator and near 60° north and south latitude and sinks at the poles and 30° north and south latitude (see Chapter 13 on Deserts). Landmasses near rising air are more prone to humid and wet climates while sinking air inhibits precipitation and creates dry conditions [10; 11]. Prevailing winds, ocean circulation patterns (e.g. the Gulf Stream's effects on eastern North America, rain shadows (dry leeward sides of mountains), and even the proximity of bodies of water can affect local climate patterns. For example, when cold winds blow across the relatively warm Great Salt Lake, the air warms, which causes it to pick up moisture. This local increase in the moisture content of the air may eventually fall as snow or rain on nearby mountains, a phenomenon known as "lake-effect precipitation" [12].

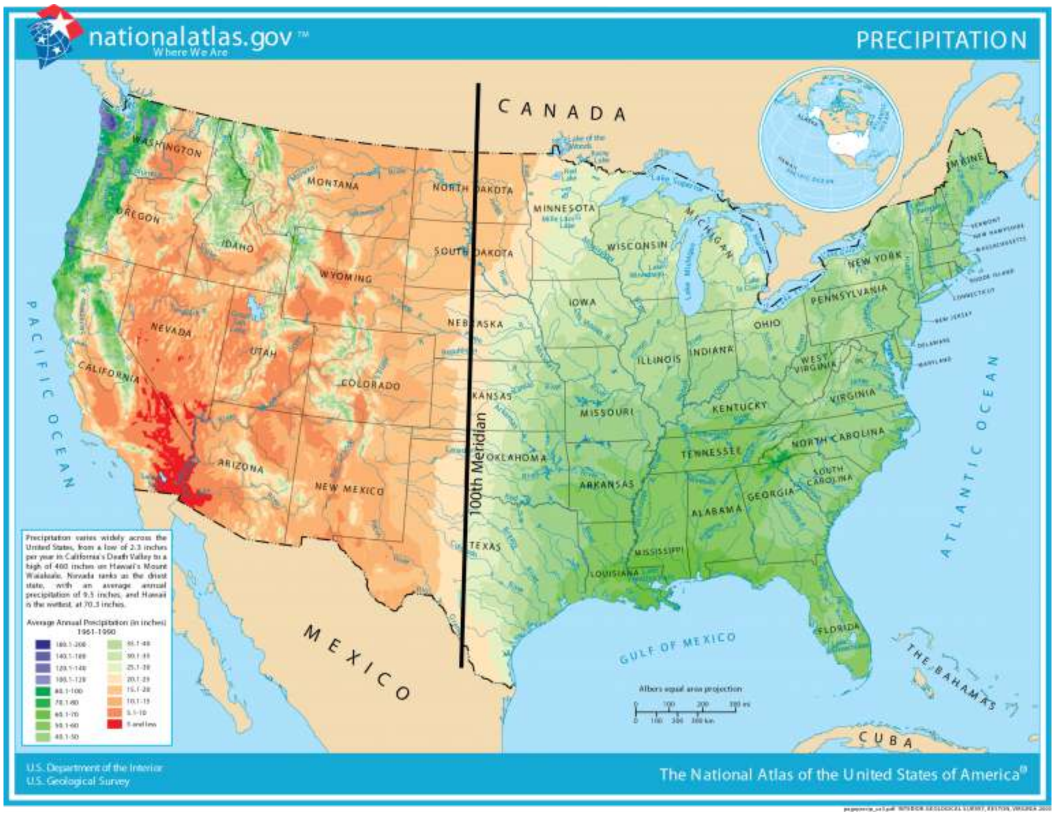


Figure 11.5.1: Distribution of precipitation in the United States. The 100th Meridian is approximately where the average precipitation transitions from relatively wet to dry. (Source: U.S. Geological Survey)

In the United States, the 100th Meridian roughly marks the boundary between the humid and arid parts of the country while west of the 100th Meridian, irrigation is required to grow crops [13]. In the West, surface water is stored in reservoirs and mountain snowpacks [14], then strategically released through a system of canals during times of high use.

Some of the driest parts of the western United States are in the Basin and Range. The Basin and Range have multiple mountain ranges that are oriented north to south. Most of the basin valleys in the Basin and Range are dry, receiving less than 12 inches of precipitation per year. However, some of the mountain ranges can receive more than 60 inches of water as snow (snow-water-equivalent). The snow-water equivalent is the amount of water that would result if the snow were melted, as the snowpack is generally much thicker than the equivalent amount of water that it would produce [12].

Groundwater Distribution

Water source	Water volume (cubic miles)	Freshwater (%)	Total water (%)

Water source	Water volume (cubic miles)	Freshwater (%)	Total water (%)
Oceans, Seas, & Bays	321,000,000	—	96.5
Ice caps, Glaciers, & Permanent Snow	5,773,000	68.7	1.74
Groundwater	5,614,000	—	1.69
— Fresh	2,526,000	30.1	0.76
— Saline	3,088,000	—	0.93
Soil Moisture	3,959	0.05	0.001
Ground Ice & Permafrost	71,970	0.86	0.022
Lakes	42,320	—	0.013
— Fresh	21,830	0.26	0.007
— Saline	20,490	—	0.006
Atmosphere	3,095	0.04	0.001
Swamp Water	2,752	0.03	0.0008
Rivers	509	0.006	0.0002
Biological Water	269	0.003	0.0001

Source: Igor Shiklomanov's chapter "World freshwater resources" in Peter H. Gleick (editor), 1993, Water in Crisis: A Guide to the World's Fresh Water Resources (Oxford University Press, New York)

Groundwater makes up 30.1% of the freshwater on the planet, making it the most abundant reservoir of freshwater accessible to most humans. The majority of freshwater, 68.7%, is stored in glaciers and ice caps as ice [9]. As the glaciers and ice caps melt due to global warming, this freshwater is lost as it flows into the oceans.

References

8. Maupin MA, Kenny JF, Hutson SS, et al (2014) Estimated use of water in the United States in 2010. US Geological Survey

9. Gleick PH (1993) Water in Crisis: A Guide to the World's Fresh Water Resources, 1 edition. Oxford University Press

10. Hadley G (1735) Concerning the cause of the general trade-winds: By Geo. Hadley, Esq; FRS. Philosophical Transactions 39:58–62

11. Lorenz EN (1955) Available potential energy and the maintenance of the general circulation. Tell'Us 7:157–167

12. Halvorson SF, James Steenburgh W (1999) Climatology of lake-effect snowstorms of the Great Salt Lake. University of Utah

13. Powell JW (1879) Report on the lands of the arid region of the United States with a more detailed account of the land of Utah with maps. Monograph

14. U.S. Geological Survey (2012) Snowmelt - The Water Cycle, from USGS Water-Science School

Contributions and Attributions

11.6: Water Law

Federal and state governments have put laws in place to ensure the fair and equitable use of water. Based on the distribution of precipitation in the United States, the states are in a position that requires them to create a fair and legal system for sharing water.

Water Rights

Because of the limited supply of water, especially in the western United States, some states have adopted a system of legally dispersing ownership of natural waters. A claim to a portion or all of a water source, such as a spring, stream, well, or lake is known as a **water right**. Federal law mandates that states control water rights, with the special exception of federally reserved water rights, such as those associated with national parks and Native American tribes, and navigation servitude, which maintains navigable water bodies. Each state in the United States has a different way to disperse and manage water rights.

A person or entity (company, organization, etc.) must have a water right to legally extract or use surface or groundwater in their state. Water rights in some western states are dictated by the concept of prior appropriation, or "first in time, first in right," where the person with the oldest water right gets priority water use during times when there is not enough water to fulfill every water right.

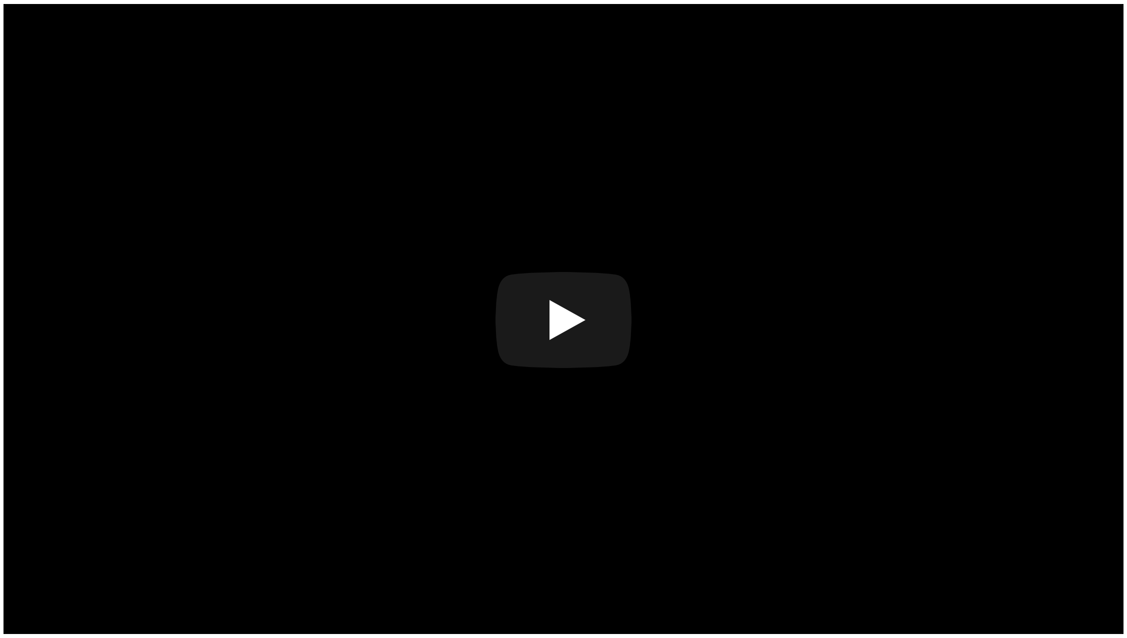
The Law of the River and the Colorado River Compact

The Colorado River and its tributaries pass through a desert region, including seven states (Wyoming, Colorado, Utah, New Mexico, Arizona, Nevada, California), Native American reservations, and Mexico. As the western United States became populated and while California was becoming a key agricultural producer, the states along the Colorado River realized that the river was important to sustaining population and agriculture in the West.

The states recognized a water budget was necessary for the Colorado River Basin to guarantee each state's certain perceived water rights. Thus was enacted the Colorado River Compact in 1922 to ensure that each state got a fair share of the river water. The Compact granted each state a specific volume of water based on the total measured flow at the time. However, in 1922, the flow of the river was higher than its long-term average flow, consequently, more water was allocated to each state than is typically available in the river.

Over the next several decades, many other agreements and modifications would follow the Colorado River Compact, including agreements that brought about the Hoover (formerly Boulder) and Glen Canyon Dams, and a treaty between the American and Mexican governments. Combined, the agreements became known as "The Law of the River." Despite adjustments to the Compact, many believe that over-allocation is still prevalent, as the Colorado River no longer reaches the Pacific Ocean, its original terminus (base level). Dams causing diversion and evaporation of Colorado River water have resulted in serious water budget concerns in the Colorado River basin. Predicted drought associated with global warming causes additional concerns about future over-allocation of Colorado River flow.

The Law of the River highlights the complex and prolonged nature of interstate water rights agreements, as well as the importance of water.



Snake Valley

In 1989, the Southern Nevada Water Authority (SNWA) submitted applications for water rights to pipe up to 155,000 acre-feet of water per year (an acre-foot of water is one acre covered with water one foot deep) from Spring, Snake, Delamar, Dry Lake, and Cave valleys to southern Nevada (mostly Las Vegas) [17]. Unlike the other valleys, Snake Valley straddles the border of Utah and Nevada, where more of the irrigable land area is on the Utah side of the border. Nevada and Utah attempted a comprehensive agreement, but negotiations have yet to be settled.

To print this story, please use the print button located in the share dialog.

A story map

Groundwater Issues of Utah

Most water-related issues are related to water quality and/or water quantity. Usually, groundwater problems are a combination of the two issues where one is a more dominant.

This story map was inspired by requirements for two Coursera courses:

Water in the Western United States

Maps and the Geospatial Revolution

15

NPR story on Snake Valley

npr

share

struggling over water

Las Vegas Water Battle: 'Crops vs. Craps'

listen

© 2007 npr

Howard Berkes, NPR

SNWA History

Dean Baker Story

Quality Protection

Two major federal laws that protect water quality in the United States are the Clean Water Act and the Safe Drinking Water Act. The Clean Water Act, an amendment of the Federal Water Pollution Control Act, protects navigable waters from dumpage and point-source pollution. The Safe Drinking Water Act ensures that water that is provided by public water suppliers, like cities and towns, is safe to drink [18].

The Superfund program ensures the cleanup of hazardous contamination and can be applied to situations of surface water and groundwater contamination. It is part of the Comprehensive Environmental Response, Compensation, and Liability Act of 1980. It allows state governments and/or the U.S. Environmental Protection Agency power to remediate polluted sites through either actions or funds provided by the polluter that caused the contamination.

References

17. (2013) Utah/Nevada Draft Snake Valley Agreement

18. EPA (2009) Water on Tap-What You Need to Know. U.S. Environmental Protection Agency

Contributions and Attributions

11.7: Surface Water

A **stream** or **river** is a body of flowing surface water confined to a channel. Terms such as creeks and brooks are social terms not used in geology. Streams are the most important agents of erosion and transportation of sediments on the earth's surface. They create much of the surface topography and are an important water resource. Most of this section will focus on stream location, processes, landforms, and flood hazards. Water resources and groundwater processes will be discussed in later sections.

Discharge

Several factors cause streams to erode and transport sediment, but the two main factors are stream channel gradient and velocity. The stream **gradient** is the slope of the river channel. A steeper gradient promotes downward stream erosion. When tectonic forces lift up a mountain, the increased stream gradient causes the stream to erode downward and make a valley. Stream velocity is the speed of the flowing water in the channel. Velocity can increase by increasing the gradient, decreasing cross-sectional area (narrowing) of the channel (reducing friction), or by increasing the discharge.

Stream size is measured in terms of **discharge,** i.e. the volume of water flowing past a point in the stream over a defined time interval. Smaller streams have a smaller discharge, therefore generally stream discharge increase downstream**.** Volume is commonly measured in cubic feet (length x width x depth), shown as feet^3 or ft^3. Therefore, the units of discharge are cubic feet per second (ft^3/sec or cfs). Smaller streams have less discharge than larger streams. For example, the Mississippi River is the largest river in North America, with an average flow of about 600,000 cfs [19]. For comparison, the average discharge for the Jordan River at Utah Lake is about 574 cfs [20] and for the Amazon River (the world's largest river), annual discharge is about 6,200,000 cfs [21].

Discharge can be expressed by the following equation:

Q = V A

- Q = discharge (ft^3/sec),
- A = cross-sectional area of the stream channel [width times average depth] (in ft^2),
- V = average velocity (ft/sec).

When the channel narrows but discharge remains constant, the same volume of water flows through a narrower space causing the velocity to increase, similar to putting a thumb over the end of a backyard water hose. In addition, during rainstorms or heavy snowmelt, runoffs will increase which increases stream discharge and thus velocity.

Velocity varies within the stream channel as well. Generally, when the channel is straight and uniform in-depth, the highest velocity is in the center of the channel along the top of the water where it is the farthest from frictional contact with the channel bottom and sides. When the channel curves, the highest velocity will be on the outside of the bend.

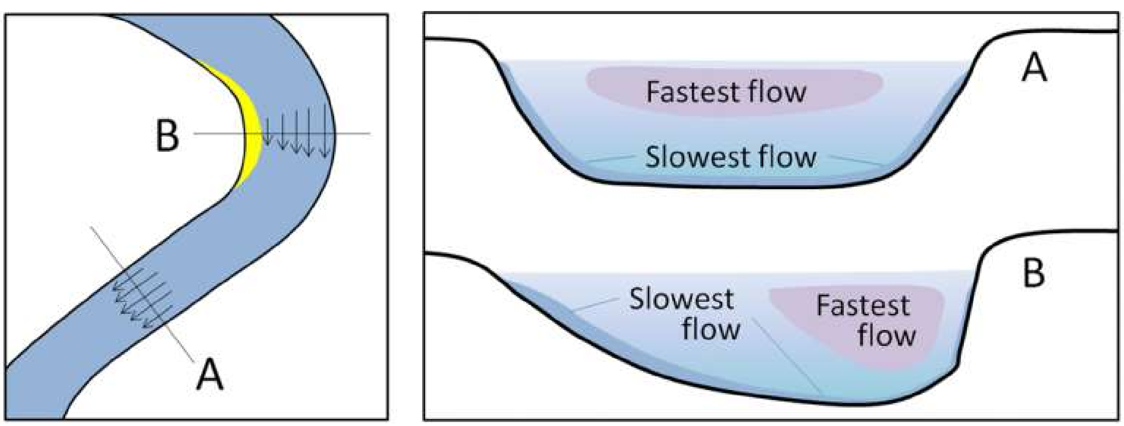


Figure 11.7.1: Thalweg of a river. In a river bend, the fastest moving particles are on the outside of the bend, near the cutbank. Stream velocity is higher on the outside bend and the surface which is farthest from the friction of the stream bed. Longer arrows indicate faster velocity (Earle 2015).

Runoff vs. Infiltration

There are many factors dictating whether water will infiltrate into the ground or runoff over the land after precipitation. These include but are not limited to the amount, type, and intensity of precipitation, the type and amount of vegetative cover, the slope of the land, the temperature and aspect of the land, preexisting conditions, and the type of soil in the area of infiltration. High-intensity precipitation as rain will cause more runoff than the same amount of rain spread out over a longer duration. If the rain falls faster than the properties of the soil allow it to infiltrate, then the water that cannot infiltrate becomes runoff. Dense vegetation can increase infiltration, as the vegetative cover slows the overland flow of water particles, giving them more time to infiltrate. If a parcel of land has more direct solar radiation and/or higher seasonal temperatures, there will likely be less infiltration and runoff, as evapotranspiration rates will be higher. As the slope of the land increases, so does runoff, as the water is more inclined to move downslope than infiltrate into the ground. Extreme examples are a basin and a cliff, where water infiltrates much quicker into a basin than a cliff having the same soil properties. Because saturated soil does not have the capacity to take more water, runoff is generally greater over-saturated soil. Clay rich soil cannot accept infiltration as quickly as gravel-rich soil.

Drainage Patterns

The pattern of tributaries within a region is called a **drainage pattern**. They depend largely on the type of rock beneath, and on structures within that rock (such as folds and faults). The main types of drainage patterns are dendritic, trellis, rectangular, radial, and deranged. **Dendritic patterns** are the most common and develop in areas where the underlying rock or sediments are uniform in character, mostly flat-lying, and can be eroded equally easily in all directions. Examples are alluvial sediments or flat-lying sedimentary rocks. **Trellis patterns** typically develop where sedimentary rocks have been folded or tilted and then eroded to varying degrees depending on their strength. The Appalachian Mountains in the eastern United States have many good examples of trellis drainage. **Rectangular patterns** develop in areas that have very little topography and a system of bedding planes, joints, or faults that form a rectangular network. A **radial pattern** forms when streams flow away from a central high point such as a mountain top or volcano, with the individual streams typically having dendritic drainage patterns. In places with extensive limestone deposits, streams can disappear into the groundwater via caves and subterranean drainage and this creates a **deranged pattern**.

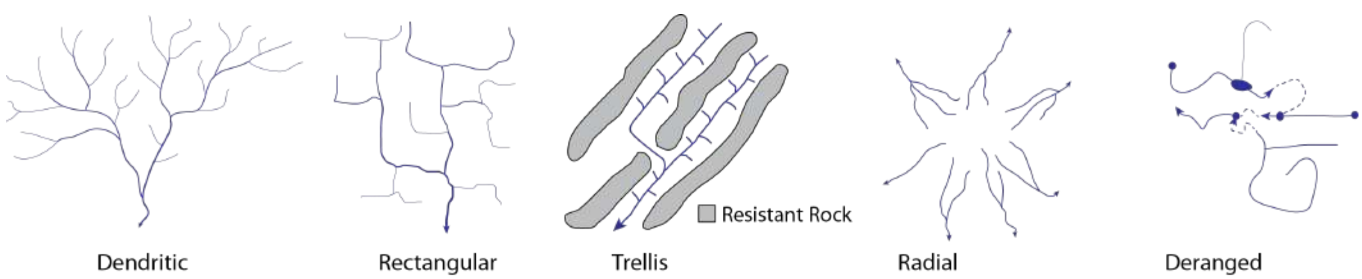


Figure 11.7.1: Various stream drainage patterns.

Fluvial Processes

Fluvial processes are the mechanisms that dictate how a stream functions and include factors controlling fluvial sediment production, transport, and deposition. Fluvial processes include velocity, slope and gradient, erosion, transportation, deposition, stream equilibrium, and base level.

Streams can be divided into three main sections: the many smaller tributaries in the source area, the main trunk stream in the floodplain and the distributaries at the mouth of the stream. These can be defined as zones of sediment production (erosion), transport, and deposition. The zone of sediment production is located in the headwaters of the stream. Downstream of the **headwaters**, the stream erodes less sediment but transports the sediment provided from the headwaters in the zone of sediment transfer. Lastly, most streams eventually flow into the ocean by a delta which is a zone of sediment deposition located at the **mouth** of a stream [6]. The **longitudinal profile** of a stream is a plot of the elevation of the stream channel at all points along its course and illustrates the location of the three zones [22].

Zone of Sediment Production (Erosion)

The zone of sediment production is located in the headwaters of a stream where rills and gullies erode sediment and contribute to larger tributary streams. These tributaries carry sediment and water further downstream to the main trunk of the stream. Tributaries at the headwaters have the steepest gradient and most sediment production and erosion, especially downward erosion, occur in the headwaters zone. Headwater streams tend to be narrow and straight with small or non-existent floodplains adjacent to the channel. Since the zone of sediment production is generally the steepest part of the stream, many headwaters are located in relatively high elevations. For example, the Rocky Mountains of Wyoming and Colorado contain much of the headwaters for the Colorado River which then flows from Colorado through Utah, Arizona, to Mexico.

Zone of Sediment Transfer (Transportation)

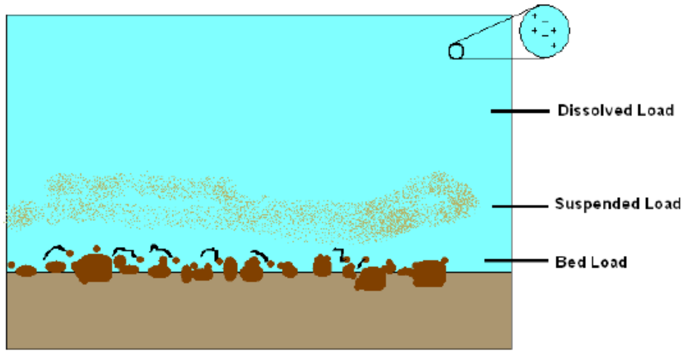
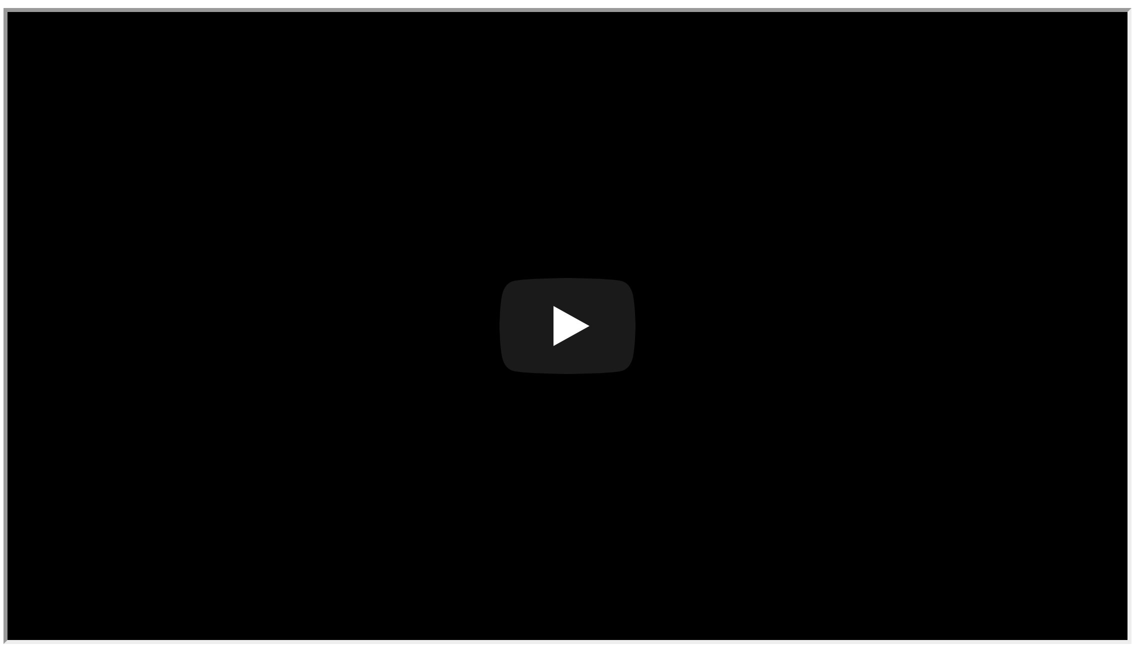


Figure 11.7.1: A stream carries dissolved load, suspended load, and bedload.

Streams transport sediment great distances from the headwaters to the ocean, the ultimate depositional basins. Sediment transportation is directly related to stream gradient and velocity. Faster and steeper streams can transport larger sediment grains. When velocity slows down, larger sediments settle to the channel bottom. When the velocity increases, those larger sediments are entrained and move again.

Transported sediments are grouped into bedload, suspended load, and dissolved load. Sediments moved along the channel bed are the **bedload** and typically are the largest and densest. Bedload is moved by saltation (bouncing) and traction (being pushed or rolled along by the force of the flow). When stream velocity increases, smaller bedload sediments can be picked up by flowing water and held in suspension as **suspended load**. The faster streams can carry larger grains as suspended load. **Dissolved load** in a stream is the sum of the ions in solution from chemical weathering. The dissolved load includes ions such as bicarbonate (HCO_3^-), calcium (Ca^{2+}), chloride (Cl^-), potassium (K^+), and sodium (Na^+). The solubility of these ions is not affected by flow velocity.



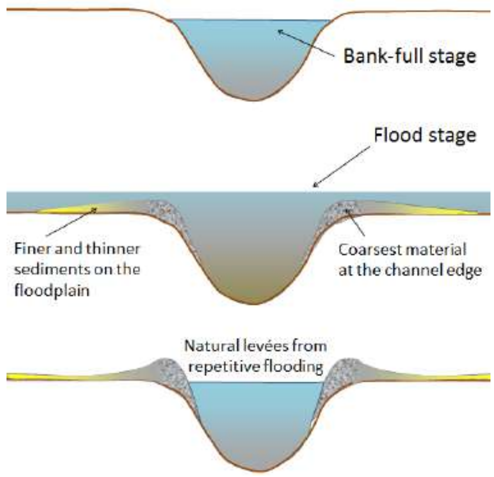


Figure 11.7.1: Profile of stream channel at bank-full stage, flood stage, and deposition of natural levee

Stream flooding is a natural process that adds sediment to floodplains. A **floodplain** is the generally flat area of land located adjacent to a stream channel that is inundated with floodwater on a regular basis. A stream typically reaches its greatest velocity when it is close to flooding, known as the **bank-full stage**. As soon as the flooding stream overtops its banks and occupies the wide area of its floodplain, the velocity decreases. At this point, sediment that was being carried by the swiftly moving water is deposited near the edge of the channel, forming a low ridge or **natural levée.** In addition, sediments are added to the floodplain during this flooding process.

Zone of Disposition

The process of **deposition** occurs when bedload and suspended load come to rest on the bottom of the water column in a stream channel, lake, or ocean. The two major factors causing deposition are the decrease in stream gradient and the reduction in velocity. These can be associated with a decrease in discharge or increased in cross-sectional area. Deposition occurs temporarily in the zone of transportation such as along meandering stream point bars, floodplains, and alluvial fans (discussed later), however, ultimate deposition occurs at the mouth of the stream where it reaches a lake or ocean. These deposits at the mouth of a stream form landforms called **deltas**. Deposition at the mouth of a stream is generally of the finest sediment such as fine sand, silt, and clay, because as the stream exits its channel, the energy of the water is completely dispersed, causing the deposition of all particles in the stream.

Equilibrium and Base Level

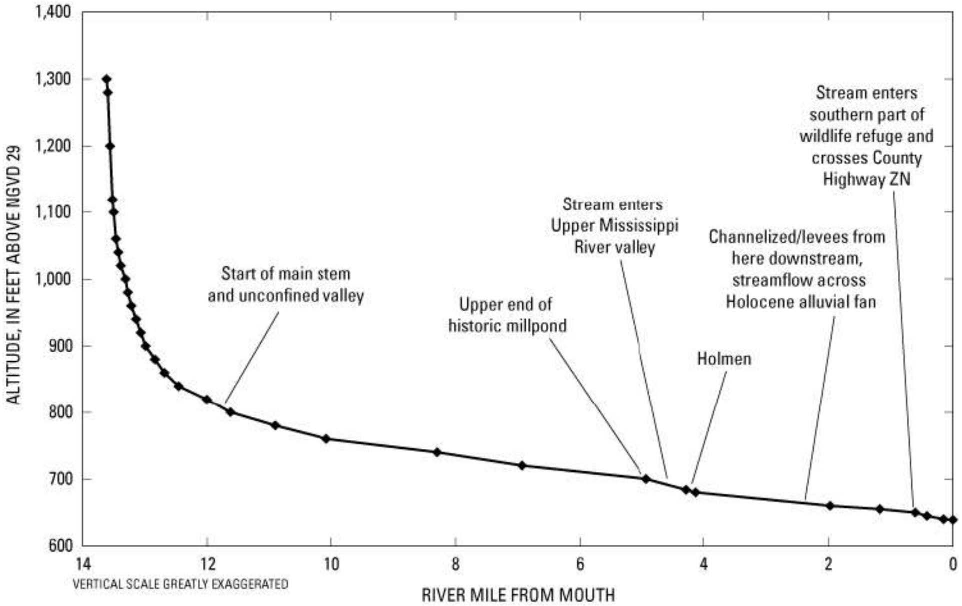


Figure 11.7.1: Example of a longitudinal profile of a stream; Halfway Creek, Indiana

All three stream zones are present in the typical **longitudinal profile** of a stream which plots the elevation of the channel at all points along its course (see figure). All streams have a long profile, some of which have been measured, plotted, and published. The long profile shows the stream gradient from headwater to mouth and represents the balance among erosion, transport, gradient, velocity, discharge, and channel characteristics at each point along the stream's course. This balance is called **equilibrium**. When mountains are uplifted, streams become steeper which erodes downward cutting a valley. This uplift is balanced against downward erosion of the stream. Eventually, streams erode enough downward that the gradient is reduced, downward erosion slows, and the river starts to erode from side to side. This point is generally characterized by a stream with a **floodplain** [6].

Another factor influencing equilibrium is the **base level**, the elevation of the stream's mouth. The base level represents the lowest level to which a stream can erode. The ultimate base level is, of course, sea-level. A lake or reservoir may also represent the base level for a stream entering it. The Great Basin of western Utah, Nevada, and parts of some surrounding states contains no outlets to the sea and provides internal base levels for streams within it. The base level for a stream entering the ocean can change if sea-level rises or falls or if a natural or human-made dam is added along its profile. When base level is lowered, a stream will downcut and deepen its channel, perhaps into a canyon. When base level is raised, deposition increases along the stream profile as the river adjusts to the change and establishes a new state of equilibrium. River equilibrium is dynamic as the river adjusts to changes in base level, tectonics, climate, precipitation, sea level, and human activities along its course.

Landforms

Stream landforms are the land features formed on the surface by either erosion or deposition. The primarily stream-related landforms described here are related to channel types.

Channel Types

Figure 11.7.1: Braided stream pattern on the Waimakariri River in New Zealand.

Figure 11.7.1: Air photo of the meandering river, Río Cauto, Cuba.

Stream channels can be straight, braided, meandering, or entrenched. The gradient, sediment load, discharge, and location of the base level all influence channel type. **Straight channels** are relatively straight, located near the headwaters, have steep gradients, low discharge, and narrow V-shaped valleys. Good examples of these are located in mountainous areas. Anastomosing streams, forming a network of branching and reconnecting channels are a variety of straight channels, formed in areas of high vegetation where the plant growth keeps the channel straight.

Braided channels have multiple smaller channels splitting and recombining downstream creating numerous mid-channel bars. These are found in broad terrain with low gradients near sediment source areas such as mountains or in front of glaciers, for example in Alaska.

Meandering channels are composed of a single channel that curves back and forth like a snake within its floodplain. Meandering channels tend to have a wide floodplain, high discharge, natural levees, and flood regularly. Meandering channels are usually located on low gradient slopes where the stream emerges from its headwaters into the zone of transportation and extends close to the zone of deposition at the stream's mouth. In areas of uplift, like has occurred on the Colorado Plateau, meanders that formed on the upland can become entrenched or incised as the stream cuts its meandering pattern down into bedrock.

Figure 11.7.1: Panoramic view of incised meanders of the San Juan River at Gooseneck State Park, Utah.

Alluvial Fans

Figure 11.7.1: Alluvial fan in Iraq seen by a NASA satellite. A stream emerges from the canyon and creates this cone-shaped deposit.

Alluvial fans are a depositional landform created where streams emerge from mountain canyons into a valley. The channel that had been confined by the canyon walls and suddenly is no longer confined slows down and spreads out, dropping its bedload of all sizes, forming a delta in the air of the valley. As distributary channels fill with sediment, the stream is diverted laterally, and the alluvial fan develops into a cone shape with distributaries radiating from the canyon mouth. Alluvial fans are common in the dry climates of the West where ephemeral streams emerge from canyons in the ranges of the Basin and Range.

Floodplains, Meandering Levels, and Natural Levees

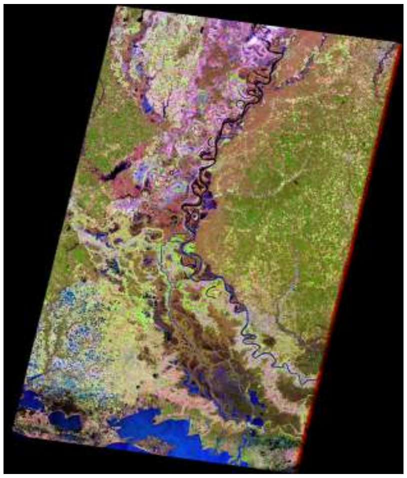
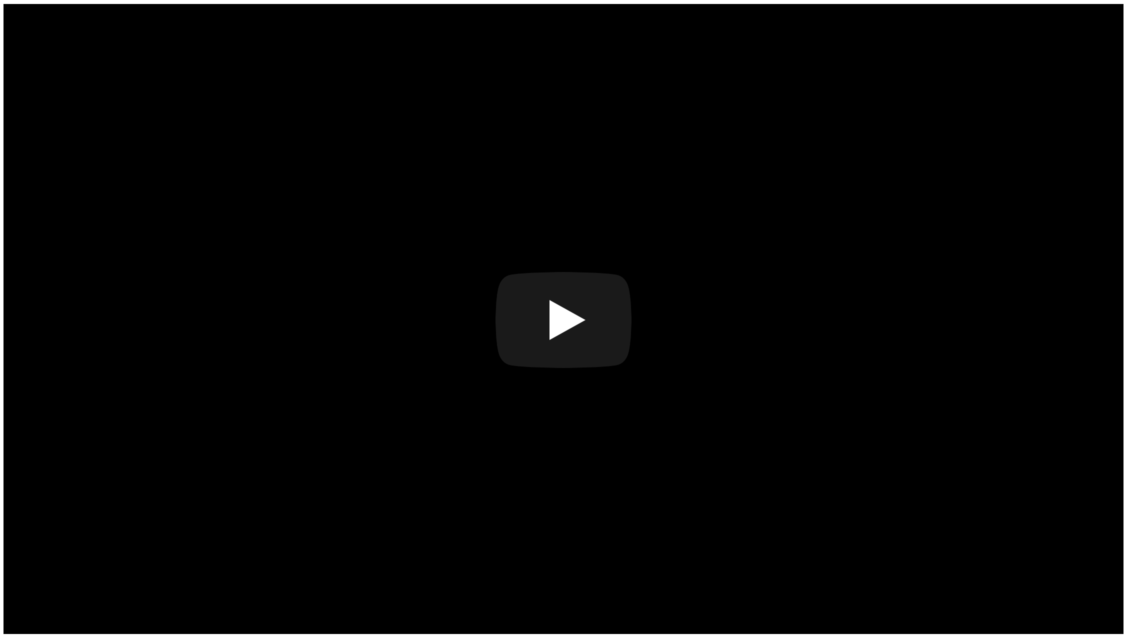
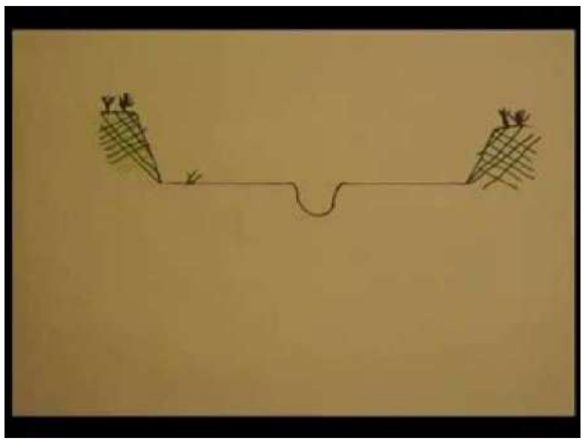


Figure 11.7.1: Landsat image of lower Mississippi River floodplain

Many fluvial landforms occur in a floodplain near a meandering stream. A **floodplain** is the broad, mostly flat area next to a meandering river that is regularly flooded. A stream creates its floodplain as the channel meanders back and forth over thousands, even millions of years. Regular flooding contributes to creating the floodplain by eroding uplands next to the floodplain. The stream channels are confined by small **natural levees** that have been built up over many years of regular flooding. Natural levees can isolate flow from contributing channels from immediately reaching the main channel on the floodplain. The smaller isolated streams, called **yazoo streams**, will flow parallel to the main trunk stream until there is an opening in the levee to allow for a belated confluence [23].





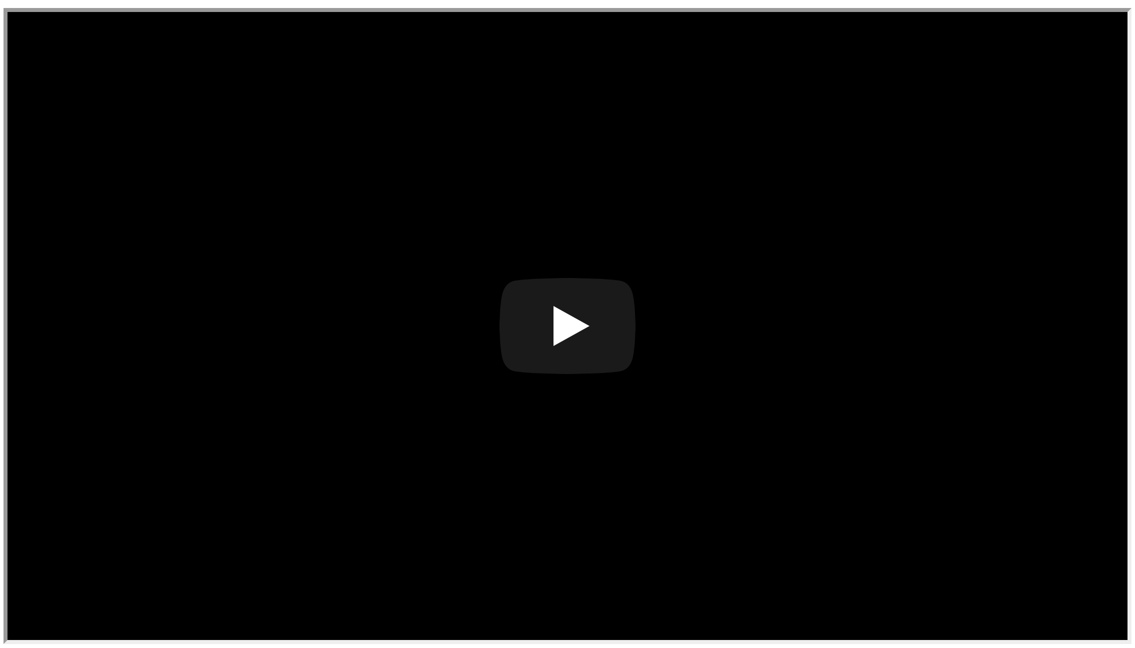
The location and width of floodplains naturally vary, however, humans build artificial levees on flood plains to limit flooding. Sediment that breaches the levees during flood stage is called **crevasse splays** delivering silt and clay into the floodplain. Floodplains are nutrient-rich from the fine-grained deposits and thus often make good farmland. Floodplains are also easy to build on due to their flat nature, however, when floodwaters crest over human-made levees, the levees quickly erode with potentially catastrophic impacts. Because of the good soils, farmers regularly return after floods and rebuild year after year.

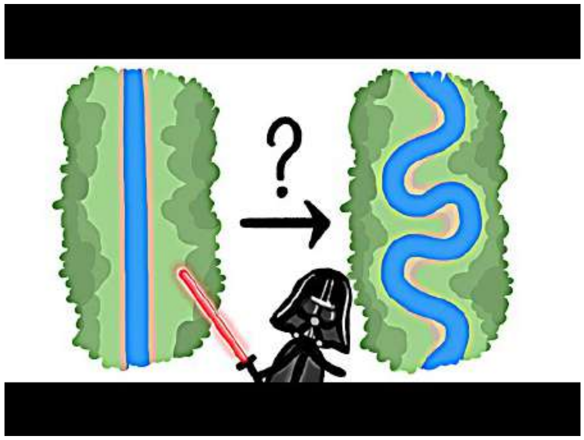
Figure 11.7.1: Point bar and cut bank on the Cirque de la Madeleine in France.

Meandering rivers create additional landforms as the channel migrates within the floodplain. Meandering rivers erode side-to-side because the highest velocity water having the most capacity to erode is located on the outside of the bend. Erosion of the outside of the bend of a stream channel is called a **cut bank** and the meander extends its loop by this erosion. The **thalweg** of the stream is the deepest part of the stream channel. In the straight parts of the channel, the thalweg and highest velocity are in the center of the channel. But at the bend of a meandering stream, the thalweg moves to the cut bank. Opposite the cut bank on the inside bend of the channel is the lowest stream velocity and therefore becomes an area of deposition call a **point bar**.

Figure 11.7.1: Meander nearing cutoff on the Nowitna River in Alaska

Through erosion on the outsides of the meanders and deposition on the insides, the channels of meandering streams move back and forth across their floodplain over time. Sometimes on very broad floodplains with very low gradients, the meander bends can become so extreme that they cut across themselves at a narrow neck (see figure). The former channel becomes isolated from streamflow and forms an **oxbow lake** seen on the right of the figure. Eventually, the oxbow lake fills in with sediment and becomes a wetland and eventually a **meander scar**. Stream meanders can migrate and form oxbow lakes in a relatively short amount of time. Where stream channels form geographic and political boundaries, this shifting of channels can cause conflicts.





Deltas

Figure 11.7.1: Location of the Mississippi River drainage basin and Mississippi River delta.

When a stream reaches a low energy body of water such as a lake or some parts of the ocean, the velocity slows and the bedload and suspended load sediment come to rest, forming a **delta**. If wave erosion from the water body is greater than deposition from the river, the deposition will not occur and a delta will not form. The largest and most famous delta in the United States is the Mississippi River delta formed where the Mississippi River flows into the Gulf of Mexico. The Mississippi River drainage basin is the largest in North America, draining 41% of the contiguous U.S. [24]. Because of the large drainage area, the river carries a large amount of sediment that is supplied to the delta. The Mississippi River is a major shipping route and human engineering has ensured that the channel no longer meanders significantly within the floodplain. In addition, the river has been artificially straightened so that it meanders less and is now 229 km shorter than it was before humans began engineering it [24]. Because of these restraints, the delta is now solely focused in one area and thus has created a "bird's foot" pattern. These two NASA images of the delta (follow the link) show how the shoreline has retreated and the land was inundated with water while deposition of sediment was located at end of the delta. These images have changed over a 25 year period from 1976 to 2001. These are stark changes illustrating sea level rise and land subsidence from the compaction of peat due to the lack of sediment resupply [25].

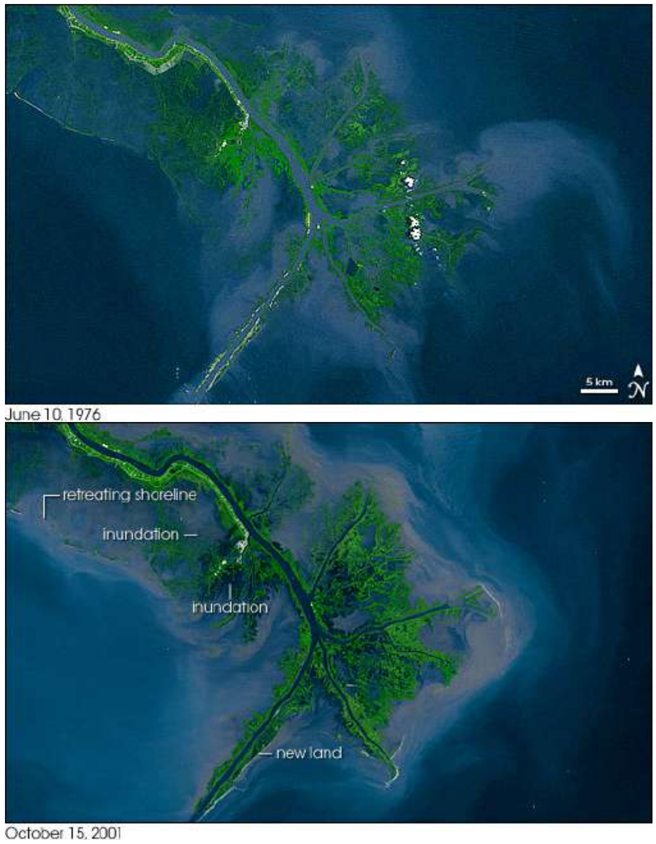


Figure 11.7.1: Before (left) vs. After (right)

The formation of the Mississippi River delta started about 7500 years ago when postglacial sea level stopped rising. In the past 7000 years, prior to anthropogenic modifications, the Mississippi River delta had several lobes that were sequentially created by the river, abandoned for a shorter route to the Gulf of Mexico, then reworked by the ocean waves of the Gulf of Mexico [26]. After the lobes were abandoned by the river, isostatic depression and compaction of the sediments caused basin subsidence (e.g. the mass and compaction of the new sediments caused the land to sink).

Figure 11.7.1: Delta in Quake Lake Montana. The deposition of this delta began in 1959 when the Madison River was dammed by the landslide caused by the 7.5 magnitude earthquake.

A clear example of how deltas form came from an unlikely source, an earthquake. During the 1959 Madison Canyon 7.5 magnitude earthquake in Montana, a large landslide dammed the Madison River forming Quake Lake [27], which is still there today. A small tributary stream that once flowed into the Madison River now flows into Quake Lake where a delta has been forming since. This a modern example of a Gilbert-type delta, which is a delta composed primarily of coarse material actively eroded from the mountainous upthrown block to the north.

Deltas represent stream deposits protruding into a quiet water body and can be further categorized as wave-dominated or tide-dominated. Wave-dominated deltas occur where the tides are small and wave energy dominates. An example is the Nile River delta in the Mediterranean Sea that has the classic shape like the Greek character (Δ) from which the landform is named. A tide-dominated delta is when ocean tides are powerful and influence the shape of the delta. For example, Ganges-Brahmaputra Delta in the Bay of Bengal (near India and Bangladesh) is the world’s largest delta and mangrove swamp called the Sundarban.

Sundarban Delta in Bangladesh, a tide-dominated delta of the Ganges River Tidal forces creates linear segments in the delta shoreline by ocean intrusion into the delta deposits. This delta also holds the world’s largest mangrove swamp, and incidentally is the only place where the Bengal tiger still actively hunts humans as prey.

Figure 11.7.1: Nile Delta showing its classic “delta” shape.

Special Topic: Ancient Deltas in Lake Bonneville

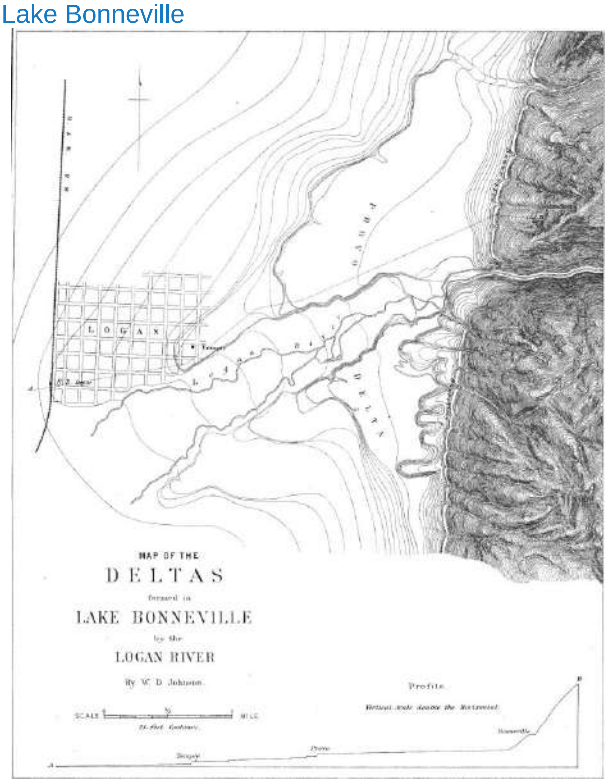


Figure 11.7.1: Map of the Logan Delta

Lake Bonneville was a large, pluvial lake that occupied the western half of Utah and parts of eastern Nevada from about 30,000 to 12,000 years ago [30]. The lake filled to a maximum elevation as great as approximately 5100 feet above the mean sea level, covering the basins, leaving the mountains exposed, many as islands. The presence of the lake allowed for deposition of fine-grained lake mud and silt, as well as coarse gravels entrained by mountain streams that lost their sediment load and energy to the open water of the lake. The lake's average surface elevation varied over its existence. The variations in lake level were controlled by regional climate and a catastrophic failure of Lake Bonneville's main outlet, Red Rock Pass [31]. Extended periods of time where the lake level remained stable caused wave-cut terraces that can be seen today on the flanks of many mountains in the region and allowed for the development of large deltaic deposits at the mouths of major canyons in Salt Lake, Cache, and other valleys. As the lake regressed to its remnant, the Great Salt Lake, the rivers that created the deltaic deposits incised stream valleys through the same deposits.

Figure 11.7.1: Deltaic deposits of Lake Bonneville near Logan, Utah.

Entrenched Meanders

Figure 11.7.1: Entrenched meander of the Colorado River, downstream of Page, Arizona.

In some rare cases, uplift will occur on a low-gradient landscape with a meandering river. This effectively increases the gradient of the stream causing it to erode downward instead of side-to-side. An example of this process is where the Colorado River and other streams crossed the Colorado Plateau as meandering streams. As the Colorado Plateau has uplifted over the past several million years, the Colorado River has incised into the flat-lying rocks of the plateau by hundreds of feet.

Figure 11.7.1: The Rincon is an abandoned meaner loop on the entrenched Colorado River in Lake Powell.

The entrenched meanders continue to experience lateral erosion. The Rincon on the entrenched Colorado River at Lake Powell is an incised oxbow lake. Another excellent example of entrenched meanders is Goosenecks State Park, Utah, where the San Juan River is deeply entrenched into the Colorado Plateau.

To summarize, an **entrenched channel** occurs when a meandering channel rapidly down cuts due to a drop in base level. This causes the original meandering shape to be preserved within a deeply entrenched channel. This channel type is rare worldwide but is common in the Colorado Plateau region which is a broad flat area near four-corners where Utah, Colorado, Arizona, and New Mexico meet. For example, the Green, Colorado, and San Juan Rivers famously form entrenched channels.

Figure 11.7.1: The San Juan River has incised meanders into the flat Colorado Plateau at Goosenecks State Park, southeastern Utah.

Terraces

Stream terraces are remnants of older floodplains located above the existing floodplain and river. Like entrenched meanders, stream terraces form when uplift occurs or base level drops and streams erode downward, leaving behind their old floodplains. In other cases, stream terraces can form from extreme flood events associated with retreating glaciers. A classic example of multiple stream terraces is along the Snake River in Grand Teton National Park in Wyoming [32, 33].


Figure 11.7.1: Terraces along the Snake River, Wyoming.


Figure 11.7.1: Terraces in Glen Roy, Scotland

References

6. Charlton R (2007) Fundamentals of fluvial geomorphology. Taylor & Francis

20. Cirrus Ecological Solutions (2009) Jordan River TMDL. Utah State Division of Water Quality

21. United States Geological Survey (1967) The Amazon: Measuring a Mighty River. United States Geological Survey

22. Brush LM Jr (1961) Drainage basins, channels, and flow characteristics of selected streams in central Pennsylvania. pubs.er.usgs.gov

23. Fairbridge RW (1968) Yazoo rivers or streams. In: Geomorphology. Springer Berlin Heidelberg, pp 1238–1239

24. Turner RE, Rabalais NN (1991) Changes in Mississippi River water quality this century. Bioscience 41:140–147

25. Törnqvist TE, Wallace DJ, Storms JEA, et al (2008) Mississippi Delta subsidence primarily caused by compaction of Holocene strata. Nat Geosci 1:173–176

26. Galloway WE, Whiteaker TL, Ganey-Curry P (2011) History of Cenozoic North American drainage basin evolution, sediment yield, and accumulation in the Gulf of Mexico basin. Geosphere 7:938–973

27. Myers WB, Hamilton W (1964) The Hebgen Lake, Montana, earthquake of August 17, 1959. US Geol Surv Prof Pap 435:51

30. Oviatt CG (2015) Chronology of Lake Bonneville, 30,000 to 10,000 yr B.P. Quat Sci Rev 110:166–171

31. Gilbert GK (1890) Lake Bonneville. United States Geological Survey, Washington, D.C.

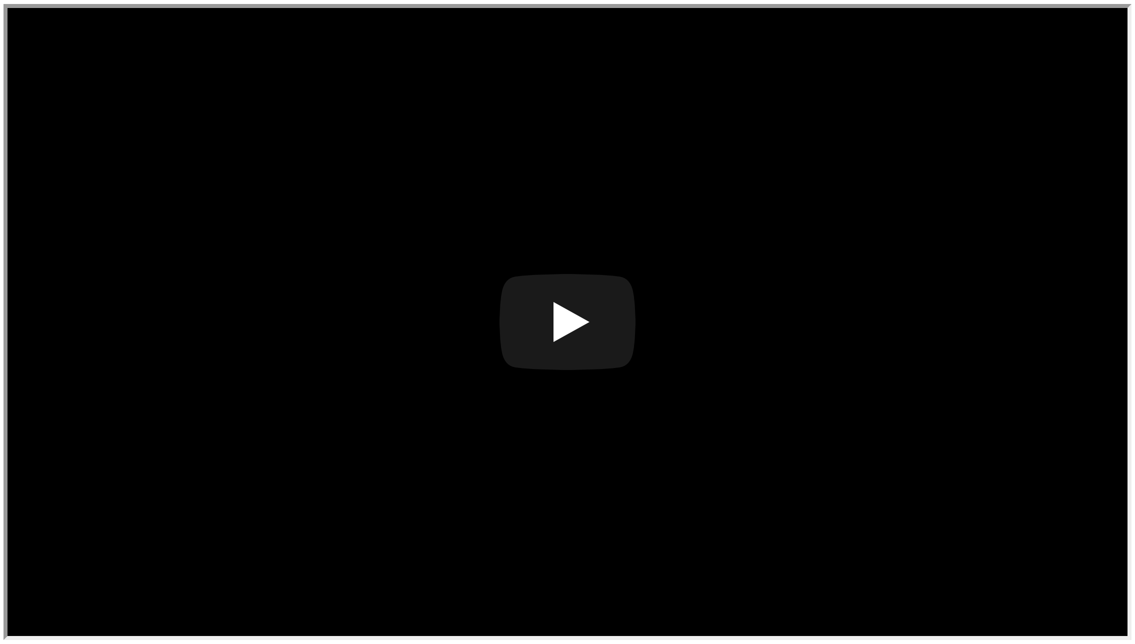
32. Reed JC, Love D, Pierce K (2003) Creation of the Teton landscape: a geologic chronicle of Jackson Hole and the Teton Range. pubs.er.usgs.gov

33. Marston RA, Mills JD, Wrazien DR, et al (2005) Effects of Jackson lake dam on the Snake River and its floodplain, Grand Teton National Park, Wyoming, USA. Geomorphology 71:79–98

Contributions and Attributions

11.8: Groundwater

Groundwater is an important source of freshwater. It can be found in all places under the ground but is limited by extractable quantity and quality.

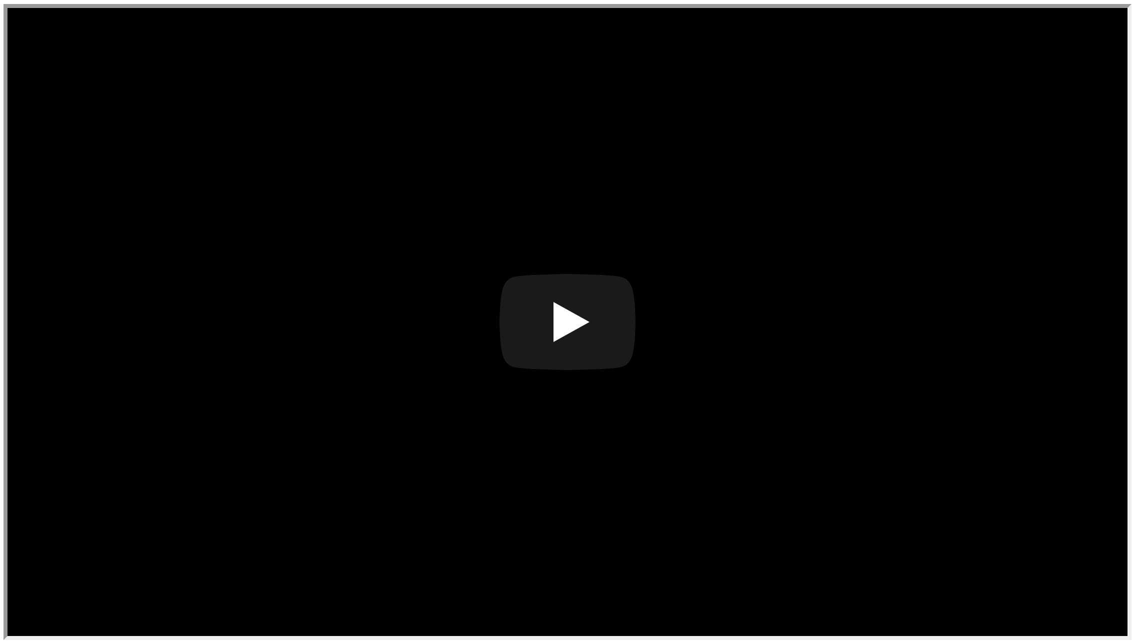


Properties

Permeability and Porosity

Most rocks are not entirely solid and contain a certain amount of open space between grains or crystals, known as pores. **Porosity** is a measure of the open space in rocks –expressed as the percentage of open space that makes up the total volume of the rock or sediment material. Porosity can occur as primary porosity, which represents the original pore spaces in the rock (e.g. space between sand grains, vesicles in volcanic rocks), or secondary porosity which occurs after the rock forms (e.g. fractures, dissolved portions of rock). Lithification of unconsolidated sediments will reduce porosity because it compacts grains and adds cement (see chapter 5.3). Water trapped in the unconnected pores of the rock during the processes of deposition and lithification is called **connate water**.

Permeability is a measure of the interconnectedness of pores in a rock or sediment. The connections between pores allows for that material to transmit water. A combination of a place to put water (porosity) and the ability to move water (permeability) makes a good **aquifer**—a rock unit or sediment that contains extractable groundwater. Well-sorted sediments have higher porosity because there are not smaller sediment particles filling in the spaces between the larger particles. Clays generally have very high porosity, but the pores are poorly connected, thereby causing low permeability.



While permeability is important as a measure of the ability to transmit fluids, it is generally not the most commonly used descriptor among geologists for this property. **Hydraulic conductivity** is another common measure of the connectedness of pore spaces and is a function of both permeability and fluid properties. Because it considers fluid properties, hydraulic conductivity is used by both petroleum geologists and hydrogeologists to describe the production capability of oil reservoirs and aquifers. A high hydraulic conductivity indicates a rapid transmission of fluid through an aquifer. Unconsolidated gravels, highly fractured and dissolved rocks, and well-sorted sandstones have high hydraulic conductivities.

Aquifers and Confining Layers

An **aquifer** is a geologic material capable of delivering water in usable quantities. Geologic material includes any rock or sediment. In order for a geologic material to be considered an aquifer, it must be at least partially saturated, where its open spaces are filled with water, and be permeable, i.e. able to transmit water. For drinking water aquifers, the water must also be **potable**. Aquifers can vary dramatically in scale, from spanning several formations to being limited to a small area on the side of a hill. Aquifers adequate for water supply are both permeable and porous.

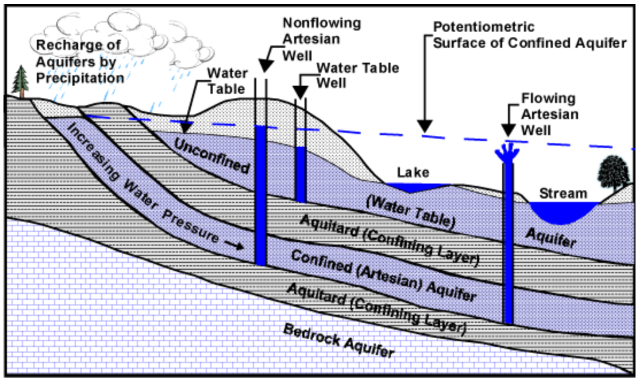


Figure 11.8.1: Potentiometric surface and water table in an aquifer system.

A good aquifer will provide a sufficient quantity of water to meet demand. The quantity of water that an aquifer can hold and transmit is governed by its physical properties. Most simply, the aquifer's porosity and permeability (defined above) are variables that govern its hydraulic conductivity and **storativity**.

A **confining layer** is a layer of low permeability geologic material that restricts the flow of water to or from the aquifer. Confining layers include **aquicludes** (a.k.a. aquifuges), which are so impermeable that no water travels through them, and **aquitards**, which significantly decrease the speed at which water travels through them due to their low permeability.

Groundwater Flow

From the Surface into the Ground

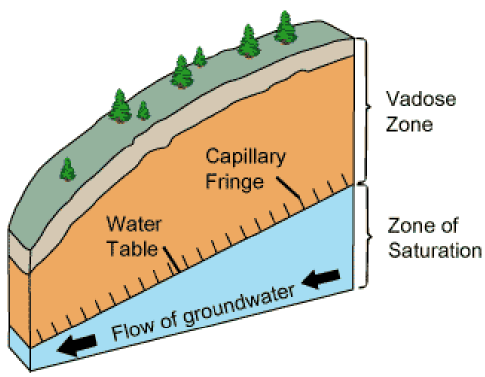


Figure 11.8.1: Zone of saturation.

When surface water infiltrates or seeps into the ground, it usually enters the unsaturated zone (a.k.a. vadose zone, a.k.a. zone of aeration). The **vadose zone** is the volume of material between the land surface and the zone of saturation, which consists of geologic materials in which the pore spaces are not completely filled with water [34]. Plants' roots inhabit the upper vadose zone. In the vadose zone, fluid pressure in the pores is less than atmospheric pressure. Below the vadose zone is the capillary fringe. The capillary fringe is the usually thin zone below the vadose zone where the pores are completely filled with water (saturation), but the fluid pressure is less than atmospheric pressure. The pores in the capillary fringe are filled because of capillary action, described in the Properties of Water section above. Below the capillary fringe is the saturated zone (a.k.a. phreatic zone), where the pores are completely saturated and the fluid in the pores is at or above atmospheric pressure [34]. The interface between the capillary fringe and the saturated zone marks the location of the water table.

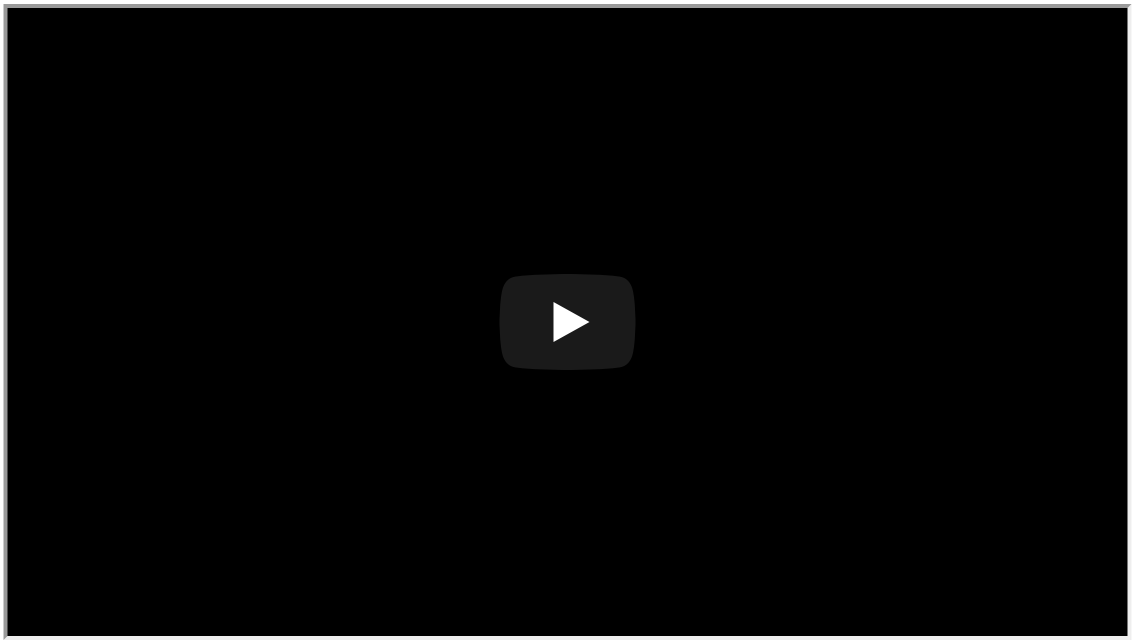
Water Table

Wells are conduits that extend into the ground with openings to the aquifers, to extract from, measure, and sometimes add water to the aquifer. Wells are generally the way that geologists and hydrologists measure the depth to groundwater from the land surface as well as withdraw water from aquifers.

Water is found throughout porosity in sediments and bedrock. The **water table** is the area at which the pores are fully saturated with water. The most simple case of a water table is when the aquifer is unconfined, meaning it does not have a confining layer above it. Confining layers can pressurize aquifers by trapping water that is recharged at a higher elevation underneath the confining layer, allowing for a potentiometric surface higher than the top of the aquifer, and sometimes higher than the land surface. The

potentiometric surface represents the height that water would rise in a well penetrating the pressurized aquifer system. Breaches in the pressurized aquifer system, like faults or wells, can cause **springs** or **flowing wells**, also known as **artesian wells**.

The water table will generally mirror surface topography, though more subdued because hydrostatic pressure is equal to atmospheric pressure along the surface of the water table. If the water table intersects the ground surface the result will be water at the surface in the form of a gaining stream, spring, lake, or wetland. The water table intersects the channel for **gaining streams** which then gains water from the water table. The channels for **losing streams** lie below the water table, thus losing streams lose water to the water table. Losing streams may be seasonal during a dry season or **ephemeral** in dry climates where they may normally be dry and carry water only after rainstorms. Ephemeral streams pose a serious danger of flash flooding in dry climates.



Mentioned in the video is the USGS Groundwater Watch site. Here is the link to that site: groundwaterwatch.usgs.gov/

Geologists measure the height of the water table and potentiometric surface using wells. Graphs of the depth to groundwater over time are known as **hydrographs**. Hydrographs can indicate changes in the water table over time. The water level in a well can change very frequently (every minute), seasonally, and over long periods of time, and is controlled by many forces.

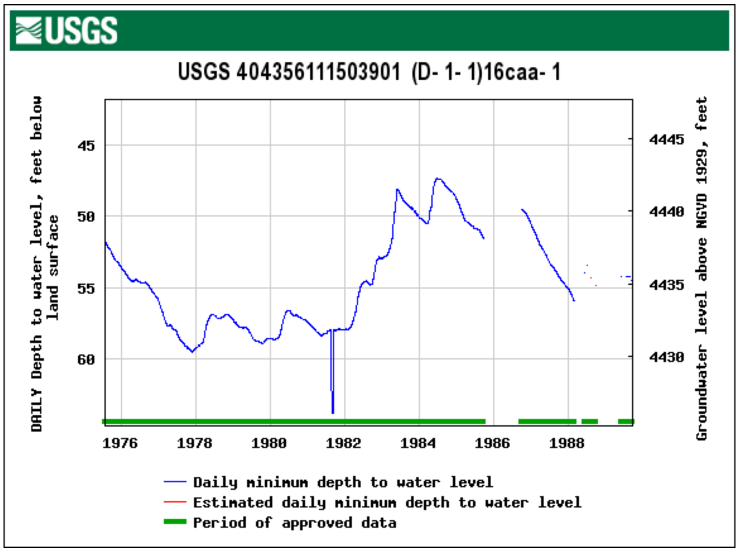


Figure 11.8.1: Example of a hydrograph.

Darcy's Law

Darcy's Law was an empirical relationship established by Henry Darcy in 1856 showing how to discharge through a porous medium is controlled by permeability, pressure, and cross-sectional area. In his experiment, Darcy used tubes of packed sediment with water running through them. The relationships described by Darcy's Law have close similarities to Fourier's law in the field of heat conduction, Ohm's law in the field of electrical networks, or Fick's law in diffusion theory. Darcy's Law provides a quantitative measure of hydraulic conductivity and discharge.

$$Q = K * A * \frac{\Delta h}{L} \tag{11.8.1}$$

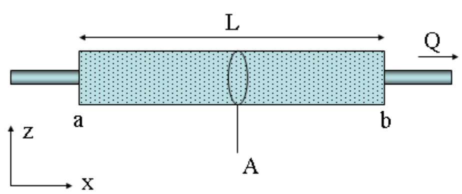


Figure 11.8.1: Pipe showing apparatus that would demonstrate Darcy's Law. Δh would be measured across L from a to b.

- Q = flow (volume/time)
- K = hydraulic conductivity (length/time)
- A = cross-sectional area of flow (area)
- Δh = change in pressure head (pressure difference)
- L = distance between pressure (h) measurements (length)
- Δh/L is commonly referred to as the hydraulic gradient

Cone of Depression

Pumping water from an aquifer lowers the water table or potentiometric surface around the well. In an unconfined aquifer, the water table is lowered as water is removed from the aquifer near the well. In a confined aquifer, the pressure around the well is reduced. The amount of change from before pumping to pumping level is termed drawdown. Drawdown is the greatest nearest the well, resulting in a concentric pattern of drawdown termed the **cone of depression**.

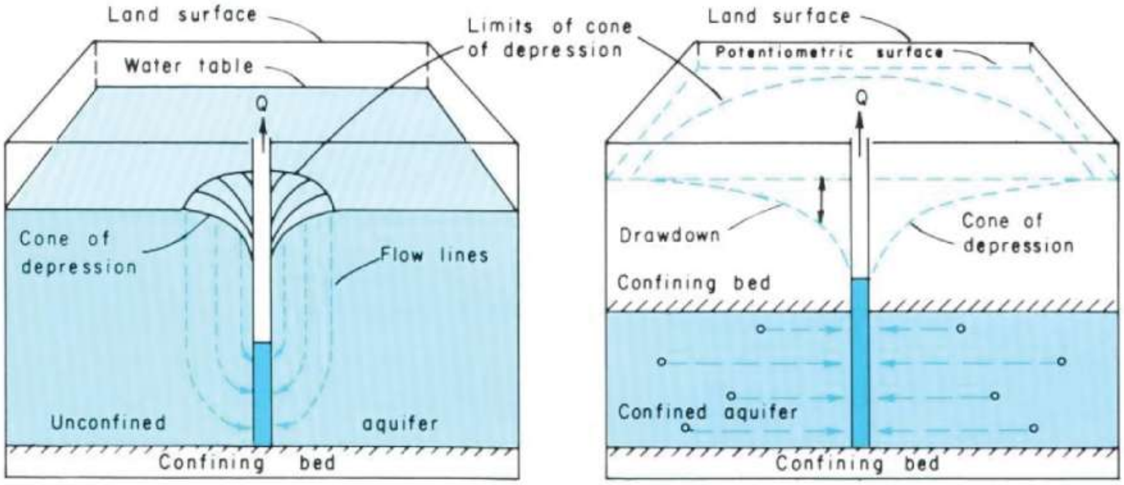


Figure 11.8.1: Cones of depression (Heath 1983).

When one cone of depression intersects another cone of depression or a barrier feature like an impermeable mountain block, drawdown is intensified. When a cone of depression intersects a recharge zone, the cone of depression is lessened.

Recharge

The **recharge** area is where surface water enters an aquifer through the process of infiltration. Recharge areas are generally the topographically highest location of an aquifer. They are characterized by losing streams and sediment or rock that allows infiltration into the subsurface. Recharge areas mark the beginning of groundwater flow paths.

In the Basin and Range region, the recharge zones for the unconsolidated aquifers of the valley areas are along the valley margins, near the foothills of the mountains. In Salt Lake Valley, as mountain streams leave the mountainous areas, they lose water to the gravel-rich deltaic deposits of ancient Lake Bonneville.

Recharge can be induced through the aquifer management practice of aquifer storage and recovery. Injection wells and infiltration galleries (basins) allow for humans to increase the rate of recharge into an aquifer system [35]. Injection wells pump water into an aquifer. Injection wells are regulated by state and federal governments to ensure that the injected water is not negatively impacting the quality or supply of the existing groundwater in the aquifer. Some aquifers are capable of storing significant quantities of water, allowing water managers to use the aquifer system like a surface reservoir. Water is stored in the aquifer during periods of low water demand and high water supply and later extracted during times of high water demand and low water supply.

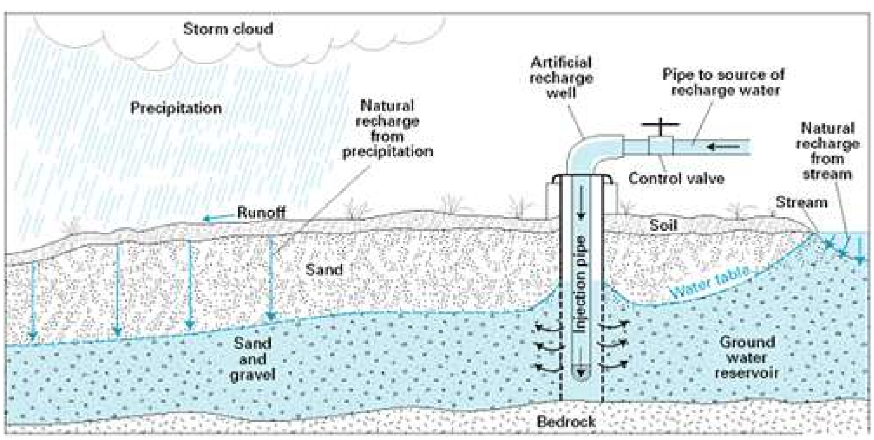


Figure 11.8.1: Different ways an aquifer can be recharged.

Discharge

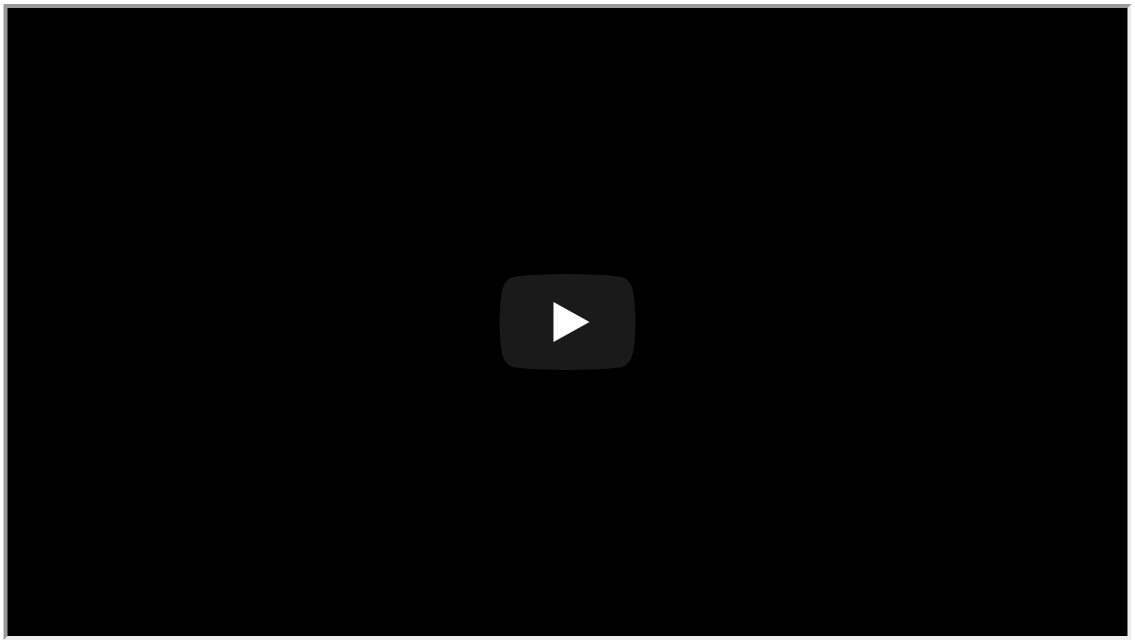
Discharge areas are where groundwater emerges at the land surface. Groundwater emerges at the land surface when the potentiometric surface or water table intersects the land surface. These areas are characterized by springs, flowing (artesian) wells, gaining streams, and playas. Discharge areas mark the end of groundwater flow paths. In the Basin and Range of the western

United States, discharge zones are typically in the middle of the valley basins, where playa lakes, springs, and gaining streams signify groundwater emerging at the land surface.

Groundwater Mining and Subsidence

Groundwater as a Limited Resource

Like other natural resources on our planet, the quantity of fresh and potable water is finite. The only natural source of water on land is from the sky in the form of precipitation. Because of a slow rate of travel, limited recharge areas, and intensifying extraction and demand, in many places groundwater is being extracted faster than it is being replenished. When groundwater is extracted faster than recharge can renew it, groundwater levels (potentiometric surfaces) decline and areas of discharge can diminish or dry up completely. Regional pumping-induced groundwater decline is known as **groundwater mining** or groundwater overdraft. Groundwater mining can lead to dry wells, reduced spring and streamflow, and subsidence. Where there is a continual lowering of the water table in an area, e.g. extraction by pumping is greater than replenishment by precipitation, groundwater mining is happening.



Subsidence

Figure 11.8.1: Example of an earth fissure in Cedar City.

In many places, water actually helps hold up the skeleton of the aquifer by the water pressure exerted on the grains in an aquifer. If pore pressure decreases because of groundwater mining, the aquifer can compact, causing the surface of the ground to sink. Areas especially susceptible to this effect are aquifers made of unconsolidated sediments. Unconsolidated sediments with multiple layers of clay and other fine-grained material are at higher risk because clay can compact considerably when drained of water [36, 37].

Figure 11.8.1: The pole shows subsidence from groundwater pumping for irrigation in the San Joaquin Valley of California.

In many cases, the amount of compaction in one area will be greater than the amount of compaction in an adjacent area. The different amounts of compaction in areas that are next to each other can cause the land to offset and develop cracks and fissures.

Subsidence from groundwater mining has been documented in southwestern Utah, notably Cedar Valley, Iron County, Utah. Groundwater levels have declined more than 100 feet in certain parts of Cedar Valley, causing earth fissures and measurable amounts of land subsidence.

The photo shows documentation of subsidence from the pumping of groundwater for irrigation in the Central Valley in California. The dates marked on the pole show land elevation in the past.

References

34. Freeze AR, Cherry JA (1979) Groundwater, 1 edition. Prentice Hall

35. Reese RS (2014) Review of Aquifer Storage and Recovery in the Floridan Aquifer System of Southern Florida

36. Galloway D, Jones DR, Ingebritsen SE (1999) Land subsidence in the United States. U.S. Geological Survey

37. Knudsen T, Inkenbrandt P, Lund W, et al (2014) Investigation of land subsidence and earth fissures in Cedar Valley, Iron County, Utah: Utah Geological Survey

Contributions and Attributions

11.9: Water Contamination

Types of Contamination

Water can be contaminated by various human activities or by existing natural features, like mineral-rich geologic formations. Agricultural activities, industrial operations, landfills, animal operations, and small and large scale sewage treatment processes, among many other things, all can potentially contribute to contamination. As water runs over the land or infiltrates into the ground, it dissolves material left behind by these potential contaminant sources. There are three major groups of contamination: inorganic chemicals, organic chemicals, and biological agents. Small sediments that cloud the water, causing **turbidity**, are also an issue with some wells but it is not considered contamination. The risks and type of remediation for a contaminant depend on the type of chemicals present.

Point source pollution can be attributed to a single, definable source, while **nonpoint source** pollution is from multiple dispersed sources. Point sources include waste disposal sites, storage tanks, sewage treatment plants, and chemical spills. Nonpoint sources are dispersed and indiscreet, where the whole of the contribution of pollutants is harmful, but the individual components do not have harmful concentrations of pollutants. A good example of nonpoint pollution are residential areas, where lawn fertilizer on one person’s yard may not contribute much pollution to the system, but the combined effect of many residents using a fertilizer can lead to significant nonpoint pollution. Other nonpoint sources include nutrients (nitrate and phosphate), herbicides, pesticides contributed by farming, nitrate contributed by animal operations, and nitrate contributed by septic systems.

Organic chemicals are common pollutants. They consist of strands and rings of carbon atoms, usually connected by covalent bonds. Other types of atoms, like chlorine, and molecules, like hydroxide (OH^-), are attached to the strands and rings. The number and arrangement of atoms will decide how the chemical behaves in the environment, its danger to humans or ecosystems, and where the chemical ends up in the environment. The different arrangements of carbon allow for tens of thousands of organic chemicals, many of which have never been studied for negative effects on human health or the environment. Common organic pollutants are herbicides and pesticides, pharmaceuticals, fuel, and industrial solvents and cleansers.

Organic chemicals include surfactants (cleaning agents) and synthetic hormones associated with pharmaceuticals, which can act as endocrine disruptors. Endocrine disruptors mimic hormones and can cause long-term effects in developing sexual reproduction systems in developing animals. Only very small quantities of endocrine disruptors are needed to cause significant changes in animal populations.

An example of organic chemical contamination is the Love Canal, in Niagara Falls, New York. From 1942 to 1952, the Hooker Chemical Company disposed of over 21,000 tons of chemical waste, including chlorinated hydrocarbons, into a canal and covered it with a thin layer of clay. Chlorinated hydrocarbons are a large group of organic chemicals that have chlorine functional groups, most of which are toxic and carcinogenic to humans. The company sold the land to the New York School Board, who developed it into a neighborhood. After residents began to suffer from serious health ailments and pools of oily fluid started rising into residents’ basements, the neighborhood had to be evacuated. This site became a U.S. Environmental Protection Agency **Superfund Site**, a site with federal funding and oversight to ensure its cleanup.

Inorganic chemicals are another set of chemical pollutants. They can contain carbon atoms, but not in long strands or links. Inorganic contaminants include chloride, arsenic, and nitrate (NO^3). Nutrients can be from geologic material, like phosphorus-rich rock, but are most often sourced from fertilizer and animal and human waste. Untreated sewage and agricultural runoff concentrate nitrogen and phosphorus which are essential for the growth of microorganisms. Nutrients like nitrate and phosphate in surface water can promote the growth of microbes, like blue-green algae (cyanobacteria), which in turn use oxygen and create toxins (microcystins and anatoxins) in lakes [38]. This process is known as eutrophication.

Metals are common inorganic contaminants. Lead, mercury, and arsenic are some of the more problematic inorganic groundwater contaminants. Bangladesh has a well-documented case of arsenic contamination from natural geologic material dissolving into the groundwater. Acid mine drainage can also cause significant inorganic contamination. See the Energy and Mineral Resources Chapter for a description of acid mine drainage.

Salt, typically sodium chloride, is a common inorganic contaminant. It can be introduced into groundwater from natural sources, such as evaporite deposits like the Arapien Shale of Utah, or from anthropogenic sources like the salts applied to roads in the winter to keep ice from forming. Salt contamination can also occur from saltwater intrusion, where cones of depression around fresh groundwater pumping near ocean coasts induce the encroachment of saltwater into the freshwater body.

Another common groundwater contaminant is biological, which includes harmful bacteria and viruses. A common bacteria contaminant is *Escherichia coli (E. coli)*. Generally, harmful bacteria are not present in groundwater unless the source of groundwater is closely connected with a contaminated surface source, such as a septic system. Karst is especially susceptible to this form of contamination because water moves relatively quickly through the dissolved conduits of limestone. Bacteria can also be used for remediation (see below).

Table: Groundwater Contaminants.

Remediation

Remediation is the act of cleaning contamination. Biological remediation usually consists of using specific strains of bacteria to break down a contaminant into safer chemicals. This type of remediation is usually used on organic chemicals but also works on reducing or oxidizing inorganic chemicals like nitrate. Phytoremediation is a type of bioremediation that uses plants to absorb the chemicals over time.

Chemical remediation uses the introduction of chemicals to remove the contaminant or make it less harmful. One example is reactive barriers, a permeable wall in the ground or at a discharge point that chemically reacts with contaminants in the water. Reactive barriers made of limestone can increase the pH of acid mine drainage, making the water less acidic and more basic, which removes dissolved contaminants by precipitation into a solid form.

Physical remediation consists of removing the contaminated water and either treating it (aka pump and treat) with filtration or disposing of it. All of these options are technically complex, expensive, and difficult, with physical remediation typically being the most costly.

Reference

38. U.S. Environmental Protection Agency (2014) Cyanobacteria/Cyanotoxins

Contributions and Attributions

11.10: Karst

Figure 11.10.1: Steep karst towers in China left as remnants as limestone is dissolved away by acidic rain and groundwater.

Karst refers to landscapes and hydrologic features created by the dissolution of limestone. Karst can be found anywhere where there are limestone and other soluble subterranean substances like salt deposits. The dissolution of limestone creates features like sinkholes, caverns, disappearing streams, and towers.

Figure 11.10.1: Sinkholes of the McCauley Sink in Northern Arizona, produced by the collapse of Kaibab Limestone into caverns caused by the solution of underlying salt deposits.

The dissolution of underlying salt deposits has caused sinkholes to form in the Kaibab Limestone on the Colorado Plateau in Arizona.

Figure 11.10.1: Sinkhole from the collapse of the surface into an underground cavern that appeared in the front yard of this home in Florida.

Karst forms when natural water, in combination with carbon dioxide, creates carbonic acid and dissolves calcite (calcium carbonate) in limestone. Remember that CO_2 in the atmosphere dissolves readily in the water droplets that form clouds from which precipitation comes in the form of rain and snow. Thus precipitation is slightly acidic with carbonic acid.

Water + Carbon Dioxide Gas equals Carbonic Acid in Water:

$$\mathrm{H_2O + CO_2 \longrightarrow H_2CO_3} \tag{11.10.1}$$

Solid Calcite + Carbonic Acid in Water Dissolved equals Calcium Ion + Dissolved Bicarbonate Ion:

$$\mathrm{CaCO_3 + H_2CO_3 \longrightarrow Ca^{2\,+} + 2\,HCO_3^-} \tag{11.10.2}$$

Figure 11.10.1: Mammoth hot springs, Yellowstone National Park.

After the slightly acidic water dissolves the calcite, changes in temperature or gas content in the water can cause the water to redeposit the calcite in a different place as tufa (travertine), often deposited by a spring or in a cave. Speleotherms are secondary deposits, typically made of travertine, deposited in a cave. Travertine speleotherms form by water dripping through cracks and dissolution openings in caves and evaporating, leaving behind the travertine deposits. Speleotherms commonly occur in the form of stalactites, when extending from the ceiling, and stalagmites, when extending from the floor.

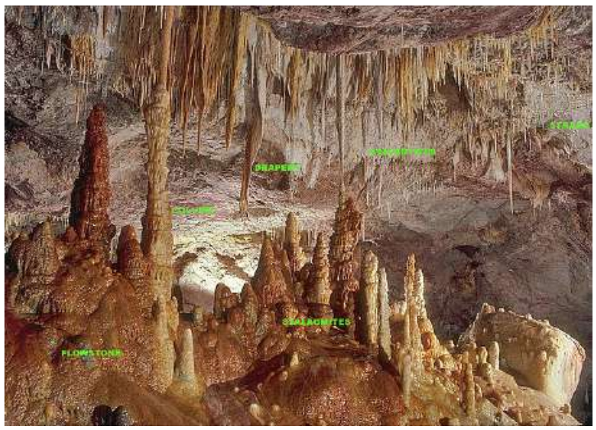


Figure 11.10.1: Varieties of speleotherms.

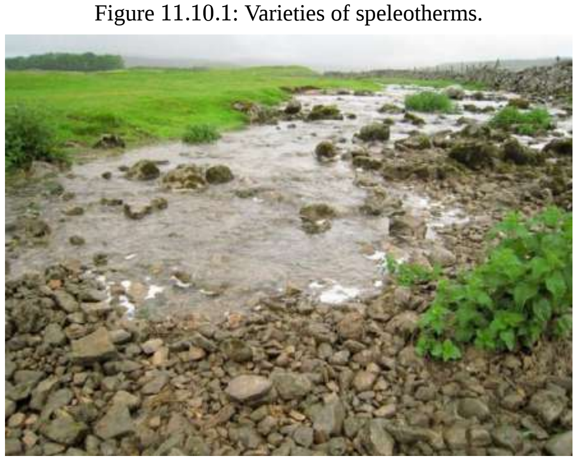


Figure 11.10.1: This stream disappears into a subterranean cavern system to re-emerge a few hundred yards downstream.

Meteoric (surface) water enters the karst system through sinkholes, losing streams, and disappearing streams. Changes in base level can cause rivers running over limestone to dissolve the limestone and sink into the ground. As the water continues to dissolve its way through the limestone, it can leave behind intricate networks of caves and narrow passages. Often dissolution will follow and expand fractures in the limestone. Water exits the karst system as springs and rises. In mountainous terrane, dissolution can extend all the way through the vertical profile of the mountain, with caverns dropping thousands of feet.

Contributions and Attributions

This page is licensed under a CC BY-NC-SA license and was authored, remixed, and/or curated by Chris Johnson, Matthew D. Affolter, Paul Inkenbrandt, & Cam Mosher. A detailed versioning history of the edits to source content is available upon request.

CHAPTER OVERVIEW

12: COASTLINES

Coastlines are the great interface between the 29% of earth's surface that is land and 71% of earth that is covered by the oceans. Therefore, it is the longest visible boundary on earth. To understand the processes that take place at this interface, we must first consider the energetic action at this boundary; namely, waves. The importance of this interface is seen in the study of ancient shorelines, and particularly for natural resources, a process called sequence stratigraphy.

12.1: Waves and Wave Processes

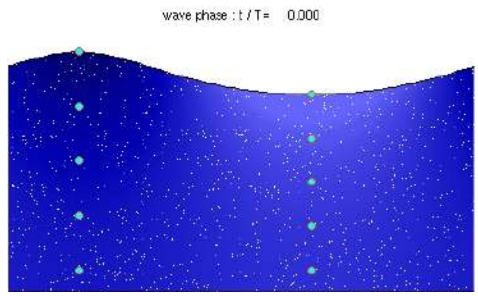


Figure 12.1.1: Particle motion within a wind-blown wave.

Waves are created when wind blows over the surface of the water. Energy is transferred from wind to the water by friction and carried in the upper part of the water by waves. Waves move across the water surface with individual particles of water moving in circles, the water moving forward with the crest and moving backward in the trough. This can be demonstrated by watching the movement of a cork or some floating object as a wave passes.

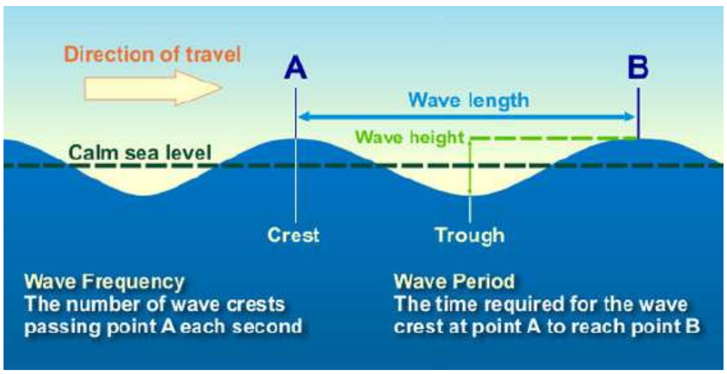


Figure 12.1.1: Aspects of water waves, labeled.

Important terms to understand in the operation of waves include: **Wave crest** is the highest point of the wave; the **trough** is the lowest point of the wave. **Wave height** (equal to twice **wave amplitude**) is the vertical distance from the trough to the crest and depends on the amount of energy carried in the wave. Even before reaching shore, wave height increases with increasing wave energy. **Wavelength** is the horizontal distance between adjacent wave crests or corresponding features of the wave. **Wave velocity** is the speed by which a wave crest moves forward, which is also related to the energy carried by the wave. **Wave period** is the time interval it takes for adjacent wave crests to pass a given point.

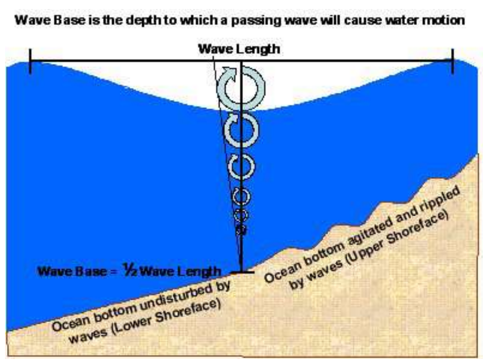


Figure 12.1.1: Diagram describing the wave base.

The circular motion of water particles diminishes with depth and is negligible at about one-half wavelength, an important dimension to remember in connection with waves. The vertical reach of waves in the water is called the **wave base**. Looking at incoming waves at a beach, one will appreciate that most ocean waves have a wavelength on the order of a few tens of feet. Thus, seawater is typically disturbed by a wave motion to a depth of a few tens of feet. This is known as a **fair-weather wave base**. In strong storms such as hurricanes, both wave length and the maximum depth of water disturbance increase dramatically. The effective depth to which waves can erode sediment is thus called a **storm wave base,** which is approximately 300 feet [1].

Waves are generated by wind blowing across the ocean surface. The amount of energy imparted to the water depends on the wind velocity and the distance across which the wind is blowing. This distance is called **fetch**. Waves striking a shore were typically generated hundreds of miles from the coast by storms and may have been traveling across the ocean for days.

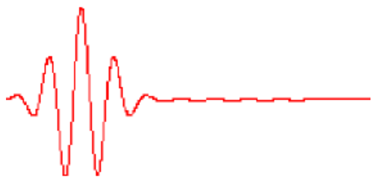


Figure 12.1.1: Model of a wave train moving with dispersion.

Winds blowing in a relatively constant direction generate waves moving in that direction. Such a group of approximately parallel waves traveling together is called a **wave train**. As wave trains spread from different areas of generation, they may move in different directions and carry different amounts of energy. Interaction of these different wave trains produces the choppy sea surface seen in the open ocean. Also of interest is that many wave lengths are produced in a given wave train from a fetch region. Longer waves travel at a faster velocity than shorter wavelengths thus there is a sorting of wavelengths that takes place during travel of the wave train with the longer waves arriving first at a distant shore. This is a process called **wave dispersion**.

Behavior of Waves Approaching Shore

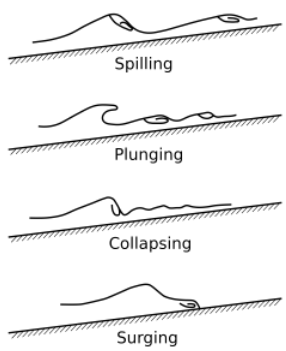


Figure 12.1.1: Types of breakers

On the open sea, waves generally appear choppy because wave trains from many directions are interacting with each other. Where crests converge with other crests (called constructive interference) they add together producing peaks, a process referred to as wave amplification. Constructive interference of troughs produces hollows. Where crests converge with troughs, they cancel each other out (called destructive interference). As waves approach the shore and begin to make frictional contact with the seafloor (i.e., water depth is a half wavelength or less) they begin to slow down, but the energy carried by the wave remains the same so they build up higher. Remember that the water moves in a circular motion as the wave passes, with the water that feeds each circle being drawn from the trough in front of the advancing wave. As the wave encounters shallower water at the shore, there is eventually insufficient water in front of the wave to supply a complete circle, and the crest pours over creating a **breaker**.

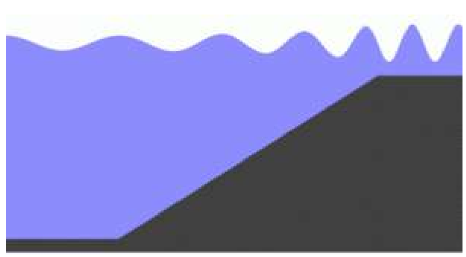


Figure 12.1.1: All waves, like tsunamis, slow down as they reach shallow water. This causes the wave to increase in height.

A special type of wave is generated by any energetic event affecting the seafloor, such as earthquakes, submarine landslides, and volcanic eruptions. Such waves are called **tsunamis** and, in the case of earthquakes, are created when a portion of the seafloor is suddenly elevated by movement in the crustal rocks below that are involved in the earthquake. The water is suddenly lifted and a wave train spreads out in all directions from the mound carrying enormous energy and traveling very fast (hundreds of miles per hour). Tsunamis may pass unnoticed in the open ocean because the wavelength is very long and the wave height is very low. But as the wave train approaches the shore, each wave makes contact with the shallow seafloor, friction increases, and the wave slows down. Wave height builds up and the wave strikes the shore as a wall of water a hundred or more feet high. The massive wave may sweep inland well beyond the beach. This is called the tsunami runup, which destroys structures far inland. Tsunamis deliver a catastrophic blow to observers at the beach as the water in the trough in front of it is drawn back toward the tsunami wave, exposing the seafloor. Curious and unsuspecting people on the beach may run out to see exposed offshore sea life only to be overwhelmed when the breaking crest hits.

Reference

1. Rich, J. L. THREE CRITICAL ENVIRONMENTS OF DEPOSITION, AND CRITERIA FOR RECOGNITION OF ROCKS DEPOSITED IN EACH OF THEM. *Geol. Soc. Am. Bull.* **62**, 1–20 (1951).

Contributions and Attributions

12.2: Shoreline Features

Many different erosional and depositional features exist in the high energy of the coast. The **coast** or **coastline** includes all parts of the land-sea boundary area that are directly affected by the sea. This includes land far above high tide and well below normal wave base. But the **shore** or **shoreline** itself is the direct interface between water and land that migrates with the tides and with deposition and erosion of sediment. Processes at the shoreline are called **littoral** processes.

Shoreline Zones

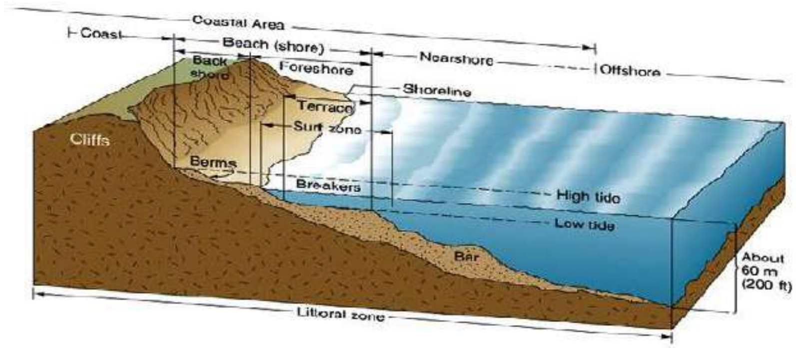


Figure 12.2.1: Diagram of zones of the shoreline.

Shoreline zones can be viewed by looking at the beach profile, which is divided into four primary zones – offshore, nearshore, foreshore, and backshore. The **offshore** is below any shoreline-derived process but is still geologically active due to cascading sands of turbidites and deeper currents (with deposits called contourites). The **nearshore** is affected by the waves, i.e., that part of the shore where water depth is a half wavelength or less. The width of this zone thus depends on the maximum wavelength of the approaching wave train and with the slope of the seafloor. The nearshore area, when looking at rocks deposited in this zone, is typically called the **shoreface**, and is broken into two segments: upper shoreface, which is affected by everyday wave action, typically consisting of finely-laminated and cross-bedded sand, and lower shoreface, the area only moved by storm waves, which has hummocky cross-stratified sand. The **surf zone** is where the waves break. The area (mostly overlapping the surf zone) that is periodically wet and dry, because of wave action and tides, is called the **foreshore**, which is made of planer-laminated, well-sorted sand.

The **beach face** is where the swash of the breaking wave runs up and the backwash flows back down. Above the beach face are low ridges called **berms**. During the summer in North America, when most people visit the beach, the zone of footprints and beach umbrellas is the **summer berm**. Wave energy is typically lower in the summer, which allows sand to be piled onto the beach. Behind the summer berm is commonly a low ridge of sand representing the **winter berm**. Beachgoers walk across this winter berm from the parking lot to the summer berm where they park their items. Higher winter storm energy moves the summer berm sand off the beach and piles it in the near from which it will be replaced next year as it is moved back onto the summer berm. There may be a zone of dunes behind the berms representing sand blown behind the beach by onshore winds. This area behind the berms that is always above the ocean in normal conditions is known as the **backshore**. Shorelines are actually simplified for this discussion; they are a dynamic and geologically-complicated place.

Refraction, Longshore Currents, and Longshore Drift

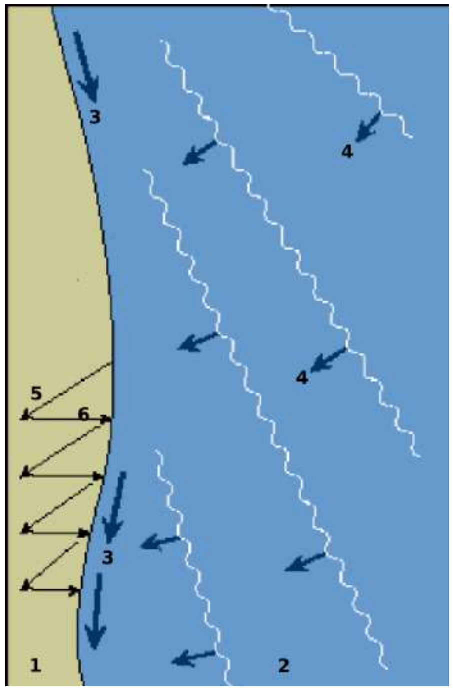


Figure 12.2.1: Longshore Drift. 1=beach, 2=sea, 3=longshore current direction, 4=incoming waves, 5=swash, 6=backwash

As waves enter shallower water, they slow down. Waves usually approach the shoreline at an angle, with one end of the waves of the train slowing down first. This causes the waves to bend toward the beach. Such bending of the waves as they enter shallower water is called **wave refraction**, which produces the appearance from the beach that waves are approaching the beach face generally straight on, parallel to the beach. However, refracted waves on average approach the shoreline at somewhat of an angle creating a slight difference between the swash as it moves up the beach face and the backwash as it flows back down. This results in a net movement of the water along the beach creating a current called the **longshore current**. Sand stirred up by waves in the surf zone is thus moved along the shore by **longshore drift**. Longshore drift along both the west and east coasts of North America moves sand on average from north to south.

Figure 12.2.1: Farewell Spit, New Zealand

Longshore drift can be carried down the coast until it reaches a bay or inlet where it begins to deposit in the quieter water. Here, a **spit** begins to form. As the spit grows, it may extend across the mouth of the bay forming a **baymouth bar**. Where the bay or inlet serves as the anchorage for boats, such spit growth and baymouth bars are a severe inconvenience. Communities thus affected attempt measures to keep their harbor open.

Figure 12.2.1: Jetties near Carlsbad, California. Notice the left jetty is loaded with sand, while the right jetty is lacking sand. This is due to the longshore drift going left to right.

One means to do this is a **jetty**, often built of concrete or large stones, forming a long barrier to deflect the sand away from the harbor mouth or other ocean waterways in which transport is desired. If the jetty does not succeed in deflecting the sand far enough out, it may continue to flow along the shore, building a spit around the end of the jetty. A more expensive but effective method is then to dredge the sand from the growing spit, put it on barges, and deliver it back to the drift downstream of the harbor opening. An even more expensive (but effective) means is to install large pumps and pipes to draw in the sand upstream of the harbor, pump it through pipes, and discharge it back into the drift downstream of the harbor mouth. Because natural processes are at work continuously, human efforts to mitigate inconveniences are sometimes not equal to the task or require ongoing modifications. The community of Santa Barbara, California, tried several methods to keep their harbor open before settling on pumps and piping [2].

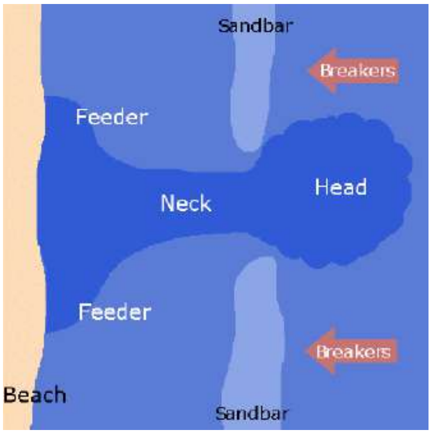


Figure 12.2.1: Animation of rip currents.

Another coastal phenomenon related to longshore currents is the presence of **rip currents**. These involve the nearshore configuration of the seafloor and/or the arrival of wave trains straight onto the shore**.** In areas where wave motion pushes water directly toward the beach face, or the shape of the nearshore seafloor refracts and focuses the water movement toward a point on the beach, the water piling up there must find an outlet back to the sea. The outlet is provided by relatively narrow rip currents that carry the water directly away from the beach. Swimmers caught in such currents find themselves being carried out to sea. They may attempt to return to shore by swimming directly against the current. This is generally a fruitless effort because they tire against the strong current. A better solution is to ride it out to where it dissipates, then swim around it and return to the beach or swim laterally, parallel to the beach, until out of the current, then return to the beach. Awareness of the presence of rip currents with a plan is the key or avoid them altogether.

Emergent and Submergent Coasts

Figure 12.2.1: Island Arch, a sea arch in Victoria, Australia.

Coastlines that have a relative fall in sea level, either caused by tectonics or sea level change, are called **emergent**. Where the shoreline is rocky, perhaps with a sea cliff, waves refracting around headlands attack the rocks behind the point of the headland.

Figure 12.2.1: This tombolo, called “Angel Road,” connects the stack of Shodo Island, Japan.

They may cut out the rock at the base forming a sea arch which may collapse to isolate the point as a **stack.** Rocks behind the stack may be eroded away and sand eroded from the point collects behind it forming a **tombolo**, a sand strip that connects the stack to the shoreline. Where sand supply is low, wave energy may erode a **wave-cut platform** across the surf zone, exposed as a bare rock with tidal pools at low tide. Wave energy expended at the base of a sea cliff may cut a **wave notch**.

Figure 12.2.1: Wave notches carved by Lake Bonneville, Antelope Island, Utah.

Sea cliffs tend to be persistent features as the waves cut away at their base and higher rocks calve off by mass wasting. If the coast is emergent, these erosional features may be elevated relative to the wave zone. Wave-cut platforms become marine **terraces**, perhaps with remnant sea cliffs inland from them.

Figure 12.2.1: Landsat image of Chesapeake Bay, eastern United States. Note the barrier islands parallel to the coastline.

Tectonic subsidence or sea level rise produces a **submergent** coast. Features associated with the submergent coasts include estuaries, bays and river mouths flooded by the higher water. **Fjords** are former glacial valleys now flooded by post-Ice Age sea level rise (see chapter 14). Elongated bodies of sand called **barrier islands** form parallel to the shoreline from the old beach sands, often isolated from the mainland by lagoons behind them [3]. The formation of barrier islands is controversial; some workers believe as above that barrier islands were formed by rising sea level as the ice sheets melted after the last ice age. Accumulation of spits and far offshore bar formations are also mentioned as possible formation hypotheses for barrier islands.

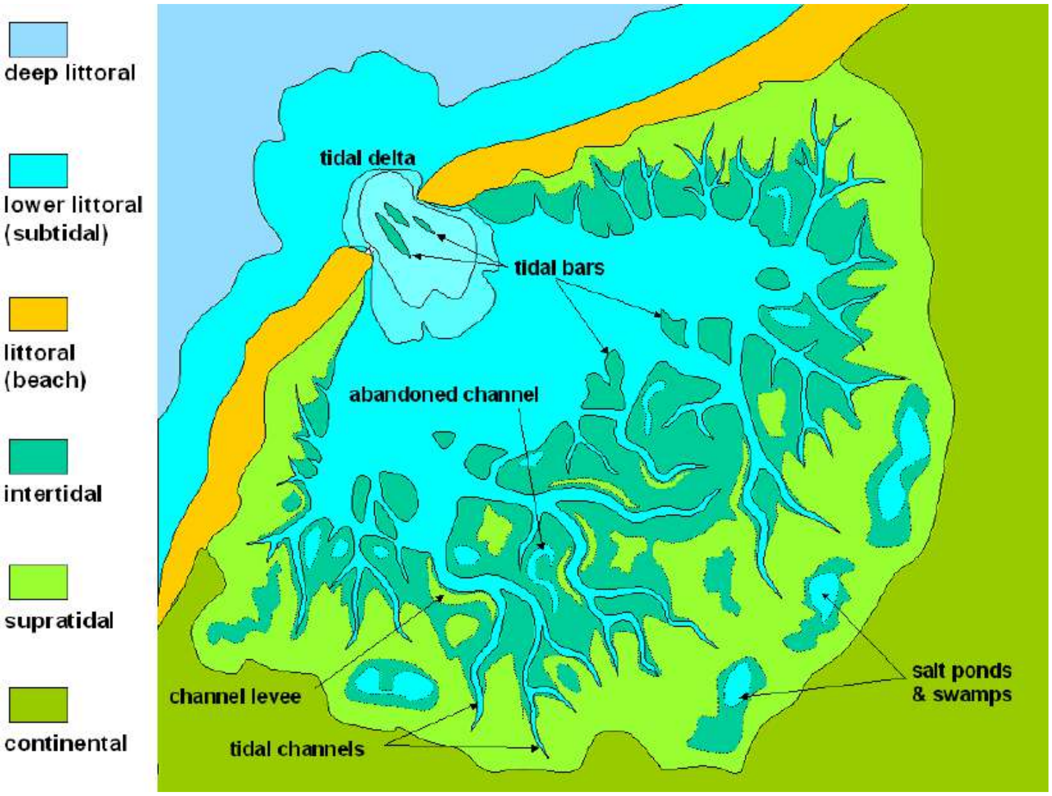


Figure 12.2.1: General diagram of a tidal flat and associated features.

Tidal flats or mudflats form where tides alternately flood and expose low areas along the coast. Combinations of symmetrical ripple marks, asymmetrical ripple marks from tidal currents, and mud cracks from drying form on these flats. An example of ancient tidal flat deposits is exposed in the Precambrian strata found in the central part of the Wasatch Mountains of Utah. These ancient deposits provide an example of applying Hutton's Uniformity Principle. The presence of features common on modern tidal flats prompts the interpretation that these ancient deposits were formed in a similar environment. There were shorelines, tides, and shoreline processes acting at that time, yet the age of the ancient rocks indicates that there were no land plants to hold products of mechanical weathering in place so rates of erosion would have been different. The Uniformity Principle must be applied with some knowledge of the context of the application.

Typically tidal flats are broken into three different sections, which may be abundant or absent in each individual tidal flat. Barren zones are areas with strong, flowing water and coarser sediment, with ripples and cross-bedding common. Marshes are vegetated with common sand and mud. Salt pans are the finest-grained parts of the tidal flats, with silty sediment, mud cracks, and are less often submerged [4].

Figure 12.2.1: Kara-Bogaz Gol lagoon, Turkmenistan.

Lagoons are locations where spits, barrier islands, or other features have partially cut off a body of water from the ocean. **Estuaries** are a (typically vegetated) type of lagoon where freshwater is flowing into the area as well, making the water **brackish** (between salt and freshwater). However, terms like a lagoon, estuary, and even bay are often loosely used in place of one another

[5]. Lagoons and estuaries are certainly transitional between terrestrial and marine geologic environments, where littoral, lacustrine, and fluvial processes can overlap.

Human Impact on Coastal Beaches

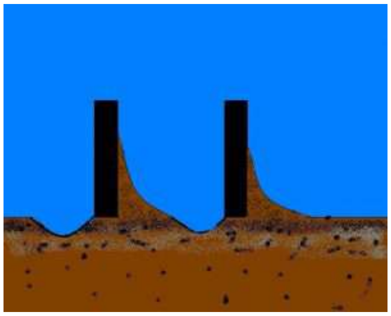


Figure 12.2.1: Groins gathering sediment from longshore drift.

Coasts are prime real estate land that attracts the development of beach houses, condominiums, and hotels. This kind of interest and investment leads to ongoing efforts to manage the natural processes in coastal areas. Humans who find longshore drift is removing sand from their beaches often use **groins** (also spelled groyne) in an attempt to retain it.

Figure 12.2.1: Groin system on a coast in Virginia

Similar but smaller than jetties, groins are bits of wood or concrete built across the beach perpendicular to the shoreline at the downstream end of one's property. Unlike jetties, they are used to preserve sand on a beach, rather than to divert it from an area. Sand erodes on the downstream side of the groin and collects against the upstream side. Every groin thus creates the need for another one downstream. The series of groins along a beach develops a scalloped appearance for the shoreline.

Sand for longshore drift and beaches comes from rivers flowing to the oceans from inland areas. Beaches may become starved of sand if sediment carried by streams and rivers is trapped behind dams. To mitigate this, **beach replenishment** may be employed where sand is hauled in from other areas by trucks or barges and dumped on the depleted beach. Unfortunately, this can disrupt the ecosystem that exists along the shoreline by exposing native creatures to foreign sandy material and foreign microorganisms and can even bring in foreign objects that impact humans on replenished beaches. Visitors to one replenished east coast beach found munitions and metal shards in the sand which had been brought from abandoned test ranges from which the sand had been dredged [6].

Figure 12.2.1: A tombolo has formed behind the breakwater at Venice, CA

Another approach to reduce erosion or provide protected areas for boat anchoring is the construction of a **breakwater**, an offshore structure against which the waves break, leaving calmer waters behind it. Unfortunately, this means that waves can no longer reach the beach to keep the longshore drift of sand moving. The drift is interrupted, the sand is deposited in the quieter water, and the shoreline builds out forming a tombolo behind the breakwater, eventually covering the structure with sand [7]. The image shows this result at the breakwater constructed by the city of Venice, California in an attempt to create a quiet water harbor. The tombolo behind the breakwater is now acting as a large groin in the beach drift.

Submarine Canyons

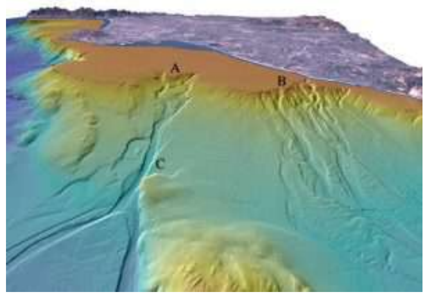


Figure 12.2.1: Submarine canyons off of Los Angeles. A=San Gabriel Canyon, B=Newport Canyon. At point C, the canyon is 815 m wide and 25 m deep.

Submarine canyons are narrow and deep canyons located in the marine environment on continental shelves. They typically form at the mouths of large landward river systems, both by cutting down into the continental shelf during times of low sea level and also by continual material slumping or flowing down from the mouth of the river or a delta. Underwater currents rich in sediment pass through the canyons, erode them and drain onto the ocean floor. Steep delta faces and underwater flows of sediments are released down the continental slope as underwater landslides, called **turbidity flows**. The erosive action of this type of flow continues to cut the canyon and eventually fan-shaped deposits develop on the ocean floor beyond the continental slope [8]. See chapter 5.5.2 for more information.

References

2. Runyan, K. & Griggs, G. Implications of Harbor Dredging for the Santa Barbara Littoral Cell. in *California and the World Ocean '02* 121–135 (American Society of Civil Engineers, 2005).

3. Davis, R. A., Jr. & Fitzgerald, D. M. *Beaches and coasts*. (John Wiley & Sons, 2009).

4. Reineck, H.-E. & Singh, I. B. *Depositional sedimentary environments: with reference to terrigenous clastics*. (Springer Science & Business Media, 2012).

5. Davis, R. A. *The evolving coast*. (Scientific American Library New York, 1997).

6. Greene, P., Follett, G. & Henker, C. Munitions and dredging experience on the United States coast. *Mar. Technol. Soc. J.* **43**, 127–131 (2009).

7. Jackson, N. L., Harley, M. D., Armaroli, C. & Nordstrom, K. F. Beach morphologies induced by breakwaters with different orientations. *Geomorphology* **239**, 48–57 (2015).

8. Normark, W. R. & Carlson, P. R. Giant submarine canyons: Is size any clue to their importance in the rock record? *Geological Society of America Special Papers* **370**, 175–190 (2003).

Contributions and Attributions

12.3: Current and Tides

Water in the ocean, when moving, can move via waves, currents, and tides. Waves have been discussed in chapter 12.1, and this section will focus on the other two. Currents in the ocean are driven by persistent global winds blowing over the surface of the water and water density. They are part of the Earth's heat engine in which solar energy is absorbed by the ocean water (remember the specific heat of water). The absorbed energy is distributed by ocean currents.

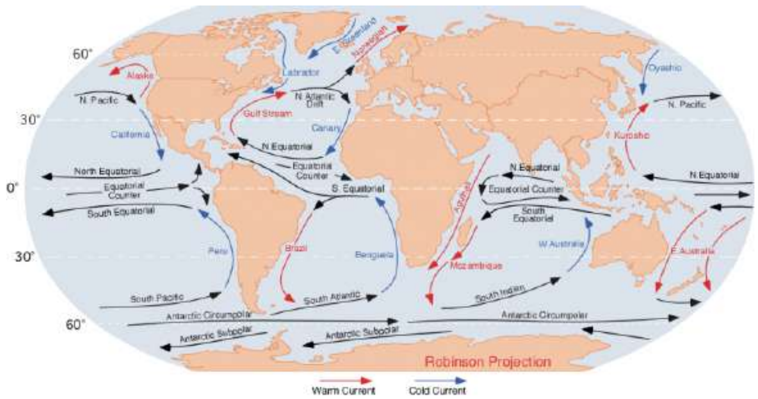


Figure 12.3.1: World ocean currents.

Surface Currents

In the above figure, notice the large sub-circular currents in the northern and southern hemispheres in the Atlantic, Pacific, and Indian Oceans. These are driven by prevailing atmospheric circulation and are called **gyres** [9] and rotate clockwise in the northern hemisphere and counterclockwise in the southern hemisphere because of the Coriolis Effect (see Chapter 13). Currents flowing from the equator toward the poles tend to be narrow as a result of the Earth's rotation and carry warm water poleward along the east coasts of adjacent continents. These are called western boundary currents, and they are key contributors to the local climate. The Gulf Stream and the Kuroshio currents in the northern hemisphere and the Brazil, Mozambique, and Australian currents in the southern hemisphere are such western boundary currents. Currents returning cold water toward the equator tend to be broad and diffuse along the western coasts of adjacent landmasses. These warm and cold currents affect nearby lands making them warmer or colder than other areas at equivalent latitudes. For example, the warm Gulf Stream makes Northern Europe much milder than similar latitudes in Northeastern Canada and Greenland. Another example is the cool Humboldt Current flowing north along the west coast of South America. This cold current limits evaporation in the ocean and contributes to the arid condition of the Atacama Desert [10].

Deep Currents

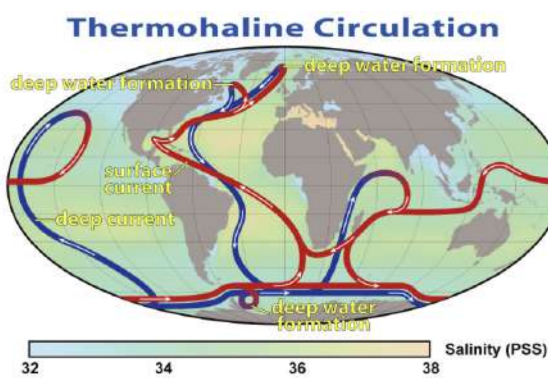


Figure 12.3.1: Global thermohaline circulation. PSS=practical salinity units.

Whether an ocean current moves horizontally or vertically depends on its density. The density of seawater is determined by factors such as temperature and salinity. Evaporation and the influx of freshwater from rivers also affect salinity and therefore the density of seawater. As the western boundary currents cool, some of the cool, denser water sinks to become the deep water of the oceans. Movement of this deep water is called the **thermohaline circulation** (*thermo* refers to temperature, *haline* refers to salinity) and connects the deep waters of all the world's oceans. This can be best illustrated by the Gulf Stream. After the warm water within the current promotes much evaporation and the heat dissipates, the resulting water is much colder and saltier. As the denser water reaches the North Atlantic and Greenland, it begins to descend and becomes a deepwater current. This worldwide (connected) shallow and deep ocean circulation is sometimes referred to as the **global conveyor belt** [11].

Tides

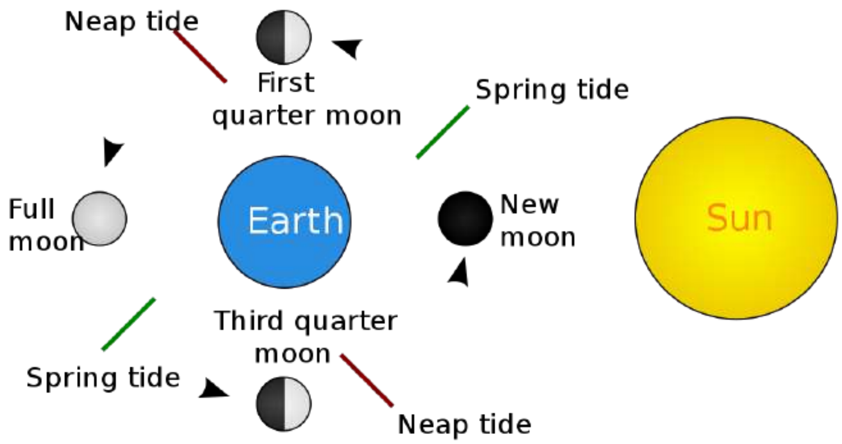


Figure 12.3.1: Diagram showing tides, in relation to the sun and moon.

The gravitational effects of the sun and moon on the oceans create **tides,** the rising and lowering of sea level during the day [12]. The earth rotates daily within the gravity fields of the moon and sun. Although the sun is much larger and exerts a more powerful gravity, its great distance from earth means that the gravitational influence of the moon on tides is dominant. The magnitude of the tide at a given location, the difference between high and low tide (the tidal range), depends primarily on the configuration of the moon and sun with respect to the earth. When the sun, moon, and earth line up with each other at full moon or new moon, the tidal range is at a maximum. This is called **spring tide.** Approximately two weeks later when the moon and sun are at right angles with the earth, the tidal range is lowest. This is called a **neap tide.**

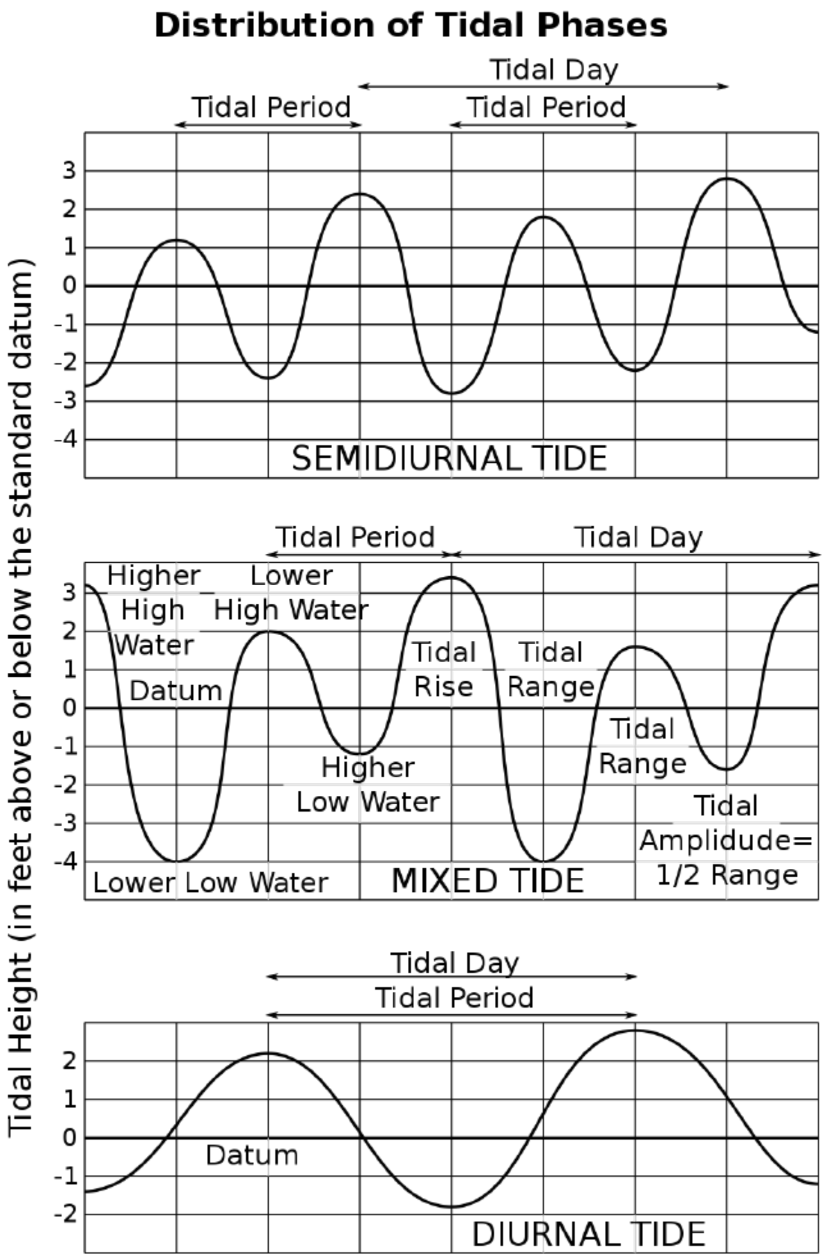


Figure 12.3.1: Different tide types

The earth rotates within the tidal envelope so we experience the rising and ebbing of the tide on a daily basis. Tides are measured at coastal locations and these measurements and tidal predictions based on them are published for those who depend on this information (e.g. this NOAA website) [13]. The rising and falling of the tides (tidal patterns) as experienced at a given shore location are of three types, diurnal, semidiurnal, and mixed.

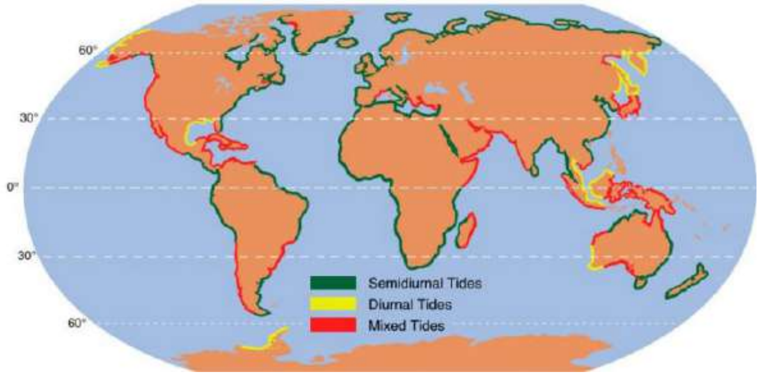


Figure 12.3.1: Global tide types

Diurnal tides go through a complete cycle once in each tidal day. Keeping in mind that the moon orbits around the Earth, a t**idal day** is the amount of time the Earth rotates to the same location of the Moon above the Earth, which (considering the movement of the bodies) consists of slightly longer than 24 hours. **Semidiurnal tides** go through the complete cycle twice in each tidal day with the tidal range typically showing some inequality in each cycle. **Mixed tides** are a combination of diurnal and semidiurnal patterns and show two tidal cycles per tidal day, but the relative amplitudes of each cycle and their highs and lows vary during the tidal month with a diurnal overprint. The pattern at a given shore location and the times of arrival of tidal phases are complicated and determined by the bathymetry (depth) of the ocean basins and continental obstacles in the way of the tidal envelope within which the earth rotates. Local tidal experts use tidal charts (indicated in the example above) based on daily observations to make forecasts for expected tides for the next few days.

Typical tidal ranges are on the order of 3 feet. Extreme tidal ranges occur where the tidal wave enters a restrictive zone. An example is the English Channel between Great Britain and the European continent where tidal ranges of 25 to 32 feet have been observed. The earth's highest tidal ranges occur at the Bay of Fundy, the funnel-like bay between Nova Scotia and New Brunswick, Canada, where the average range is nearly 40 feet and extremes of around 60 feet have been observed. At these locations of extreme tidal range, a person who ventures out onto the seafloor exposed during ebb tide may not be able to outrun the advancing water during flood tide. Additional information on tides can be found at this NOAA website.

References

9. Colling, A. *Ocean Circulation*. (Butterworth-Heinemann, 2001).

10. Munk, W. H. On the wind-driven ocean circulation. *J. Meteorol.* **7**, 80–93 (1950).

11. Stommel, H. & Arons, A. B. On the abyssal circulation of the world ocean—I. Stationary planetary flow patterns on a sphere. *Deep Sea Research (1953)* **6**, 140–154 (1959).

12. Stewart, R. H. *Introduction to physical oceanography*. (Texas A & M University Texas, 2008).

13. Schwiderski, E. W. On charting global ocean tides. *Rev. Geophys.* **18**, 243–268 (1980).

Contributions and Attributions

CHAPTER OVERVIEW

13: DESERTS

Approximately 30 percent of Earth's terrestrial surface is desert. Deserts are defined as locations of low precipitation. While temperature extremes are often associated with deserts, they do not define them. The lack of moisture, including the lack of humidity and cloud cover, allow temperature extremes to occur. The sun's energy is more absorbed by the Earth's surface without cloud cover, and nighttime cooling is more drastic without cloud cover and humidity to absorb the emitted heat, so temp

13.1: Prelude to Deserts

Figure 13.1.1: A playa filled with evaporite minerals (such as gypsum) erodes and forms ripple-covered dunes in White Sands National Monument, New Mexico.

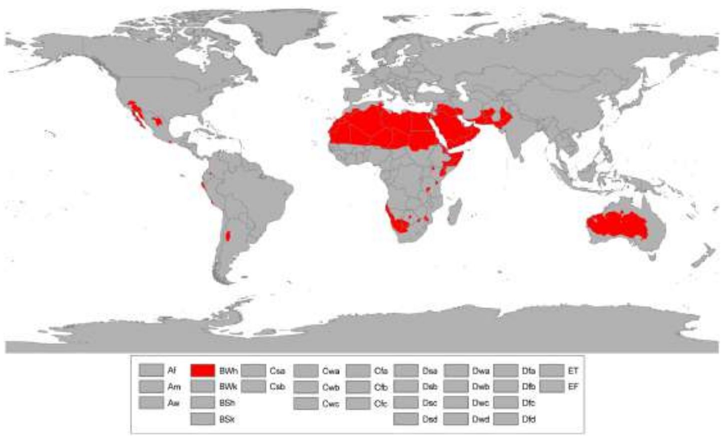


Figure 13.1.1: World hot deserts (Koppen BWh)

The location of climates on Earth's surface is not random. Jungles, tundras, and deserts have scientific explanations for their locations. Approximately 30 percent of Earth's terrestrial surface is desert. Deserts are defined as locations of low precipitation. While temperature extremes are often associated with deserts, they do not define them. The lack of moisture, including the lack of humidity and cloud cover, allow temperature extremes to occur. The sun's energy is more absorbed by the Earth's surface without cloud cover, and nighttime cooling is more drastic without cloud cover and humidity to absorb the emitted heat, so temperature extremes are common in deserts.

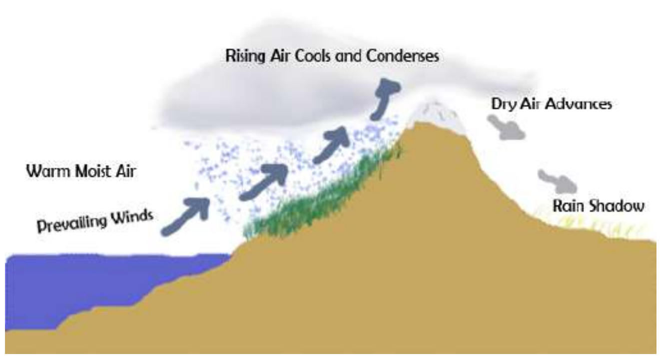


Figure 13.1.1: Diagram of rain shadow.

Deserts tend to occur at latitudes of around 30° and at the poles, both north and south, driven by circulation and prevailing wind patterns in the atmosphere. At approximately 30° north and south of the equator, sinking air produces **trade wind deserts** like the Sahara and the Outback of Australia [1]. **Rainshadow deserts** are produced where prevailing winds with moist air dries as it is forced to rise over mountains.

Figure 13.1.1: In this image from the ISS, the Sierra Nevada Mountains are perpendicular to prevailing westerly winds, creating a rain shadow to the east (down in the image). Note the dramatic decrease in snow on the Inyo Mountains.

The Western Interior Desert of North America and the Atacama Desert of Chile (the driest warm desert on earth) are examples of rain shadow deserts. Finally, **polar deserts**, such as the vast areas of the Antarctic and Arctic are covered by sinking cold air which is usually too cold to hold much moisture. Though covered with ice and snow, the average annual precipitation is very low, with Antarctica being Earth's driest continent.

References

1. Walker, A. S. *Deserts: geology and resources*. (Government Printing Office, 1996).

13.2: The Origin of Deserts

Atmospheric Circulation

The engine that drives circulation in the atmosphere and oceans is solar energy which is determined by the average position of the sun over the earth's surface. Direct light provides uneven heating depending on latitude and angle of incidence, with high solar energy in the tropics, and little or no energy at the poles. Atmospheric circulation and geographic location are the primary causal agents of deserts.

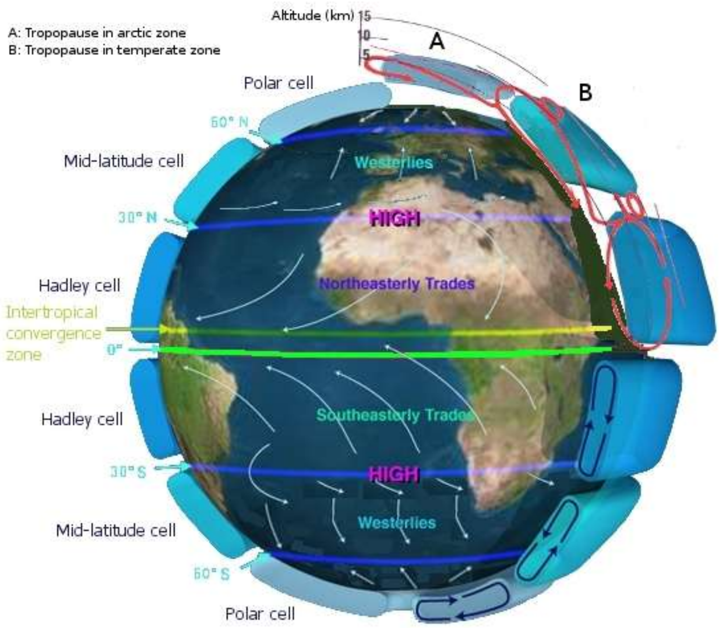


Figure 13.2.1: Generalized atmospheric circulation

The figure shows the generalized circulation of the atmosphere. There are three generalized circulating cells of rising and sinking air called the **Hadley Cell**, the **Ferrel** or **Midlatitude Cell**, and the **Polar Cell**. Solar energy falling on the equatorial belt heats the air and causes it to rise. The rising air cools and its contained moisture falls back on the tropics as rain. The drier air then continues to spread toward the north and south where it collides with the Ferrel Cell and they sink back at about 30 degrees north and south latitudes. This sinking drier air creates belts of predominant high pressure along which desert conditions prevail in what are called the "horse latitudes." These belts of predominantly high pressure have air that descends along these belts and flows either north to become the westerlies or south to become the trade winds. These circulation cells in the atmosphere rising in the tropics and polar regions and sinking in the horse latitudes produce the desert belts along the horse latitudes at approximately 30 degrees north and south of the equator [2]. Note the arrows indicating general directions of winds in the latitude zones. The trade winds are predominant in the tropics and the westerlies in the mid-latitudes.

Figure 13.2.1: USGS Map of the Great Basin Desert.

Other deserts have other atmospheric phenomena to owe (at least part of) their origin, like the desert of Utah, Nevada, and the surrounding areas called the **Great Basin Desert** [3]. This desert, while having some sinking air effects due to global circulation, is also a rain shadow desert produced as moist air from the Pacific rises by orthographic lifting over the Sierra Nevada (and other) Mountains and loses moisture from previous condensation and precipitation on the rainy side of the range(s).

Figure 13.2.1: Map of the Atacama desert (yellow) and surrounding related climate areas (orange).

One of the driest places on earth is the **Atacama Desert** of northern Chile [4]. This is a strip along the west coast of South America, west of the Andes, lying north of 30 degrees south latitude, at the southern edge of the trade wind belt. Warm moist air moves west across the Amazon basin and rises over the Andes where it loses moisture, its precipitation falling on the rain forest side of the mountains. Once over the mountains, it descends onto the Atacama where it meets air cooled by the cold Peru (Humboldt) ocean current flowing north along the coast. This is considered to be the driest (non-polar) place on earth with locations in the Atacama having not received any precipitation for periods of years [5].

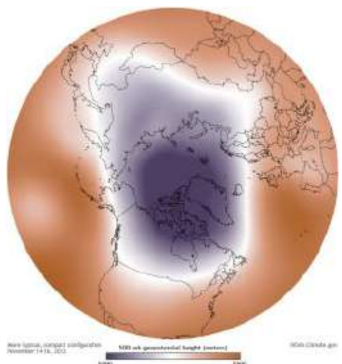
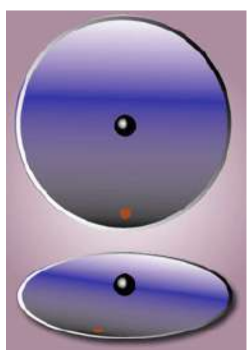


Figure 13.2.1: The polar vortex of mid-November, 2013. This cold, descending air (shown in purple) is characteristic of polar circulation.

Referring again to the figure above, note that the polar regions are also predominantly high-pressure areas of descending cold and dry air. Another circulation cell occurs there known as the **Polar Cell** [6]. Here, air not only descends convectively because it is cold, but cold air can hold much less moisture than warm air, and thus the driest and coldest places on Earth are the polar deserts. Antarctica is not only currently the driest land on Earth today, but any land that occupies the poles in Earth history should always be dry.

Coriolis Effect



In a non-rotating Earth, air would rise at the equator, sink at the poles, creating one circulation cell. However, as noted above, Earth has three cells. Why? As objects move on a rotating sphere, an effect called the **Coriolis Effect** occurs which causes a deflection in the motion. In the northern hemisphere, this deflection is to the right; in the southern hemisphere, it is to the left. This has two consequences on masses of air (and water) moving on the earth. The lower air in the Hadley Cells moves toward the equator over the earth's surface. This air is deflected to the right in the northern hemisphere and to the left in the southern hemisphere creating the **trade winds** that carried European explorers to South America and the Caribbean. The midlatitude cells move surface air north toward the pole in the northern hemisphere (and south in the southern hemisphere) from the horse latitudes which is deflected again to the right (or left in the southern) producing the zone of **westerlies**. High above the Earth, the rising air from the equator would attach to the sinking air at the poles, but again, it is deflected, causing instead sinking air at 30° and rising at 60°. This splits the circulation into three cells instead of one.

To understand the Coriolis Effect, first consider motion along the meridians (the lines connecting the poles and running north-south). The earth rotates toward the east, i.e. everything on Earth moves at an eastward speed depending on its latitude. At a given latitude, objects possess a certain momentum of that motion depending on the length of the radius from its latitude to the rotational axis of the Earth. In the northern hemisphere, if the object is moving north, it has greater momentum toward the east than other objects at the new more northerly latitude. If it moves south, it has less momentum than other objects at that new more southerly latitude. It, therefore, tends to move to the right compared to fixed locations at that new latitude. The opposite happens in the southern hemisphere.

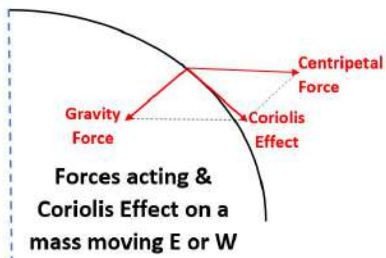
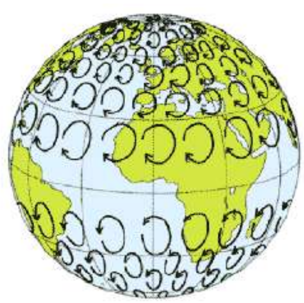
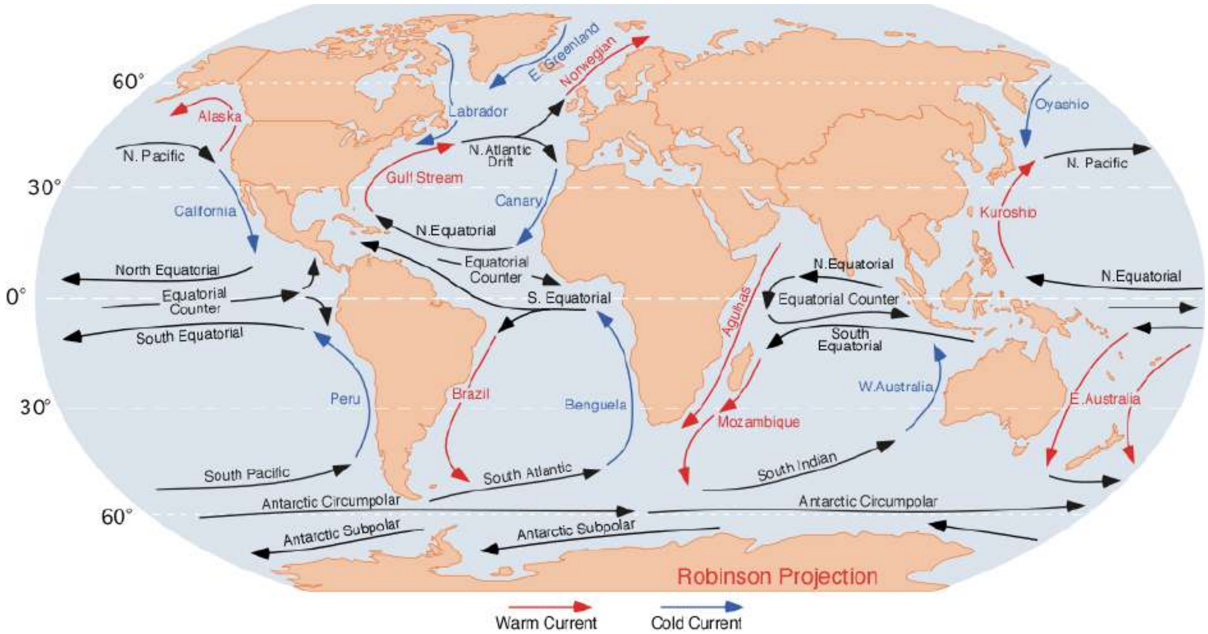


Figure 13.2.1: Forces acting on a mass moving East-West in the Northern Hemisphere on the rotating Earth that produce the Coriolis Effect

Now consider motion in an east-west direction again thinking of the momentum imparted by the radius from the Earth's rotational axis. The centripetal effect of Earth's rotation causes objects on the earth to tend to be forced outward perpendicular to the rotational axis, the centripetal effect or "force." Since gravity holds things on the earth's surface (the gravitational force points toward the Earth's center, perpendicular to the surface), they do not actually fly outward of course. But considering the components of force involved acting on a mass at or above the Earth's surface, the centripetal component is perpendicular to the Earth's spin axis. The component of gravity is perpendicular to the earth's surface pointing toward the Earth's center. If the object is moving eastward, the speed of the object adds to the earth's rotational speed and the centripetal effect is enhanced, thus the net effect of gravity and the centripetal component parallel to the surface causes deflection to the right (left in the southern hemisphere). If the object is moving west, its speed subtracts from the rotational speed and reduces the centripetal effect. Deflection is again to the right (left in the southern hemisphere). At any direction of motion, the meridional and the centripetal effects combine, thus no matter which direction an object moves on the rotating earth, there is a tendency for deflection to the right in the northern hemisphere (left in the southern).



The objects of greatest interest in geoscience that are affected by the Coriolis Effect on earth are air and water masses. Since wind patterns, especially prevailing patterns, cause ocean currents, then water masses feel it as well. In reality, any object moving on the earth experiences it. For example, the Coriolis Effect must be taken into account by artillerymen calculating the trajectory of artillery shells for accuracy in hitting targets over long distances.



The Coriolis effect creates large sub-circular rotating currents called **gyres** in the oceans, turning clockwise in the northern hemisphere and counterclockwise in the southern under the Coriolis Effect. These currents bring cold water along the west coasts of both North and South America contributing to the drier climates of the Atacama and Central and Southern California. The Coriolis Effect acting on both the atmosphere and ocean is a major contributor to climate and weather on the earth.

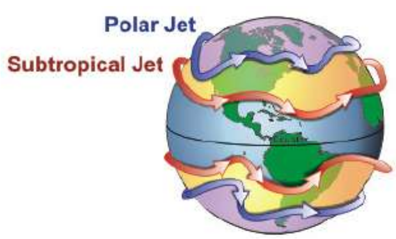


Figure 13.2.1: Earth’s two Jet Streams. The stronger Polar Jet is associated with low pressure. When the Polar Jet moves from its normal average location of 60°, it brings low pressure to desert regions near 30°. This is the main, but not only, cause of precipitation in mid-latitudes.

An application of the Coriolis Effect can be seen on the TV weather report. High-pressure systems are typically shown by a large “H” and indicate dry conditions, and low-pressure systems by a large “L” indicating clouds or precipitation. Air flows outward from a high and because of the Coriolis Effect, it rotates clockwise (to the right). It flows inward to a low and again turns to the right, rotating counterclockwise. Of course weather reports in the Southern Hemisphere show the opposite. Another interesting realization from the Coriolis Effect and the Zone of Westerlies is that weather systems tend to move from west to east across both North America and the southern part of South America. The high pressures and low pressures that exist due to uneven heating of the atmosphere and the Coriolis Effect create the high and low pressures on the weather map. The chaotic nature of the atmosphere (and fast-moving flows like the Jet Stream) makes these high and low pressures constantly and consistently move. This is important, because at 30 degrees, without this movement, low pressure would *never* exist! This means rain would never arrive. Even in the driest parts of this zone, like the Atacama, it rains on occasion. High pressure normally exists here, but just not all the time. These air movements, both prevailing and sporadic, are thus important in understanding climate and its geological implications.

References

2. Hadley, G. Concerning the cause of the general trade-winds: By Geo. Hadley, Esq; FRS. *Philosophical Transactions* **39**, 58–62 (1735).

3. Grayson, D. K. *The desert's past: a natural prehistory of the Great Basin*. (Smithsonian Inst Pr, 1993).

4. Hartley, A. J. & Chong, G. Late Pliocene age for the Atacama Desert: Implications for the desertification of western South America. *Geology* **30**, 43–46 (2002).

5. Collado, G. A., Valladares, M. A. & Méndez, M. A. Hidden diversity in spring snails from the Andean Altiplano, the second highest plateau on Earth, and the Atacama Desert, the driest place in the world. *Zool. Stud.* **52**, 50 (2013).

6. Littell, E. & Littell, R. S. *Littell's Living Age*. (T.H. Carter & Company, 1846).

Contributions and Attributions

13.3: Desert Weathering and Erosion

Figure 13.3.1: Weathering and erosion of Canyonlands National Park have created a unique landscape, including arches, cliffs, and spires.

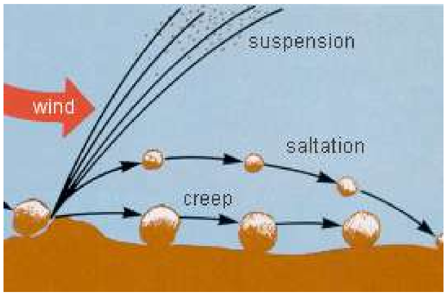
Weathering takes place in desert climates by the same means as other climates, only at a slower rate. This is besides the higher temperatures, which typically spur faster weathering. Water is the main agent of weathering, and lack of water slows weathering. Precipitation occurs in deserts, only less than in other climatic regions. Chemical weathering proceeds more slowly in deserts compared to more humid climates because of the lack of water. Even mechanical weathering is slowed, because of a lack of runoff and even a lack of moisture to perform ice wedging. However, when precipitation does occur, often in the form of flash floods, a large amount of mechanical weathering can happen quite quickly.

Figure 13.3.1: Newspaper rock, near Canyonlands National Park, has many petroglyphs carved into the desert varnish.

One unique weathering product of deserts is **desert varnish**. Also known as desert patina or rock rust, they are thin dark brown layers of clays and iron and manganese oxides that form on very stable surfaces within arid environments. The exact cause of the material is still unknown, though cosmogenic and biologic mechanisms have been proposed.

Figure 13.3.1: A dust storm (haboob) hits the Mongolian Gobi.

While water is still the dominant agent of erosion in most desert environments, wind is a notable agent of weathering and erosion in many deserts. This includes suspended sediment traveling in **haboobs**, or dust storms, that frequent deserts. Deposits of windblown dust are called **loess**. Loess deposits cover wide areas of the midwestern United States, much of it from dust that melted out of the ice sheets during the last ice age [7]. Lower energy than water, wind transport nevertheless moves sand, silt, and dust [8]. As noted in chapter 11, the load carried by a fluid (like air) is distributed among bedload and suspended load. As with water, in wind these components depend on wind velocity.



Sand size material moves by a process called **saltation** in which sand grains are lifted into the moving air and carried a short distance where they drop and splash into the surface dislodging other sand grains which are then carried a short distance and splash dislodging still others [8].

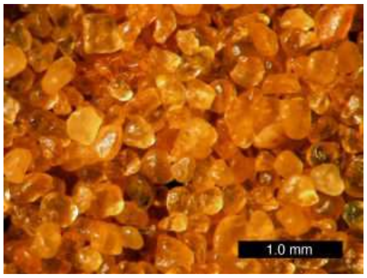


Figure 13.3.1: Enlarged image of frosted and rounded windblown sand grains

Since saltating sand grains are constantly impacting other sand grains, windblown sand grains are commonly pretty well rounded with frosted surfaces. Saltation is a cascading effect of sand movement creating a zone of windblown sand up to a meter or so above the ground. This zone of saltating sand is a powerful erosive agent in which bedrock features are effectively sandblasted. The fine-grained suspended load is effectively removed from the sand and the surface carrying silt and dust in haboobs. Wind is thus an effective sorting agent separating sand and dust-sized (≤70 μm) particles [9]. When wind velocity is high enough to slide or roll materials along the surface, the process is called **creep**.

One extreme version of sediment movement was shrouded in mystery for years: **Sliding stones**. Also called sailing stones and sliding rocks, these are large moving boulders along flat surfaces in deserts, leaving trails. This includes the famous example of the Racetrack Playa in Death Valley National Park, California. For years, scientists and enthusiasts attempted to explain their movement, with little definitive results [10; 11]. In recent years, several experimental and observational studies have confirmed that thin layers of ice allow the stones to move with high winds providing propulsive energy [12; 13]. These studies include measurements of actual movement, as well as re-creations of the conditions, with resulting movement in the lab.

Figure 13.3.1: A yardang in Bolivia

Figure 13.3.1: Ventifact from Mojave Desert near Barstow, CA

The zone of saltating sand is an effective agent of erosion through sand abrasion. A bedrock outcrop which has such a sandblasted shape is called a **yardang** [14]. Rocks and boulders lying on the surface may be blasted and polished by saltating sand. When predominant wind directions shift, multiple sandblasted and polished faces may appear. Such polished desert rocks are called **ventifacts** [15].

Figure 13.3.1: Blowout in Texas

In places with sand dunes, clumps of vegetation often anchor sediment that has accumulated on the desert surface. Yet, saltation from winds may be sufficient to move or remove materials not anchored by vegetation. This causes a bowl-shaped depression in the sand called a **blowout** [16].

References

7. Muhs, D. R. & Bettis, E. A. Quaternary loess-paleosol sequences as examples of climate-driven sedimentary extremes. *Special Papers-Geological Society of America* 53–74 (2003).

8. Bagnold, R. A. The physics of blown sand and desert dunes. *Methum, London, UK* 265 (1941).

9. Shao, Y. *Physics and Modelling of Wind Erosion*. (Springer Science & Business Media, 2008).

10. Stanley, G. M. Origin of playa stone tracks, Racetrack Playa, Inyo County, California. *Geol. Soc. Am. Bull.* **66**, 1329–1350 (1955).

11. Clements, T. Wind-blown rocks and trails on Little Bonnie Claire Playa, Nye County, Nevada. *J. Sediment. Res.* **22**, (1952).

12. Norris, R. D., Norris, J. M., Lorenz, R. D., Ray, J. & Jackson, B. Sliding rocks on Racetrack Playa, Death Valley National Park: first observation of rocks in motion. *PLoS One* **9**, e105948 (2014).

13. Kletetschka, G. *et al.* Sliding stones of Racetrack Playa, Death Valley, USA: The roles of rock thermal conductivity and fluctuating water levels. *Geomorphology* **195**, 110–117 (2013).

14. Hedin, S. A. *Central Asia and Tibet*. **1**, (Hurst and Blackett, limited, 1903).

15. Laity, J. E. Landforms, landscapes, and processes of aeolian erosion. in *Geomorphology of Desert Environments* (eds. Parsons, A. J. & Abrahams, A. D.) 597–627 (Springer Netherlands, 2009).

16. Livingstone, I. & Warren, A. *Aeolian geomorphology: an introduction*. (Longman, 1996).

Contributions and Attributions

13.4: Desert Landforms

Figure 13.4.1: Aerial image of an alluvial fan in Death Valley

In deserts like those of the American Southwest, streams draining mountains flow through canyons and emerge into adjacent valleys. As the stream emerges from the narrow canyon and spreads out, and with a lower slope angle and slower speeds [17] and no longer constrained by the canyon walls [18], it drops its coarser load. As the channel fills with this conglomeratic material, the stream is deflected around it. This process causes the stream to be deflected back and forth, developing a system of radial distributaries and constructing a fan-shaped feature call an **alluvial fan**, similar to a delta made by a river entering a body of water [18].

Figure 13.4.1: Bajada along Frisco Peak in Utah

Alluvial fans continue to grow and may eventually coalesce with neighboring fans to form an apron of alluvium along the mountain front called a **bajada** [19].

Figure 13.4.1: Inselbergs in Mojave Desert

As the mountains erode away and the debris accumulates first in alluvial fans, then bajadas, the mountains eventually are buried in their own erosional debris. Such residual buried mountain remnants are called **inselbergs** [20], “island mountains,” as first described by the German geologist Wilhelm Bornhardt.

Figure 13.4.1: Satellite image of desert playa surrounded by mountains

Where the desert valley is an enclosed basin, i.e. streams entering it do not drain out but the water is removed by evaporation, a dry lake bed is formed called a **playa.**

Figure 13.4.1: Dry wash (or ephemeral stream)

Playas are among the flattest of all landforms. Such a lake may cover a large area and be only a few inches deep, and that only after a heavy thunderstorm. Playa lakes and desert streams that flow only after rainstorms are called **intermittent** or **ephemeral**. Drainage basins of ephemeral streams gather water from large areas and ephemeral channels may suddenly fill with water from storms many miles away and not even visible at that location plus, lack of organic matter and soil structure in arid regions inhibits infiltration and adds to runoff.

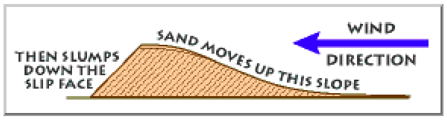
Figure 13.4.1: Flash flood in a dry wash

Such high-volume ephemeral lows may be non-channelized and move as sheet flows. Such **flash floods** are a major factor in desert deposition and a serious concern for desert travelers who need to pay attention to regional weather. Water is less able to infiltrate because the flow compacts the surface, plants are less common to slow flows, and soils in deserts can become more hydrophobic. Water typically runs off as **sheet wash** to stream channels called **arroyos** or a dry wash that may be dry part or most of the year. Dry ephemeral channels can fill quickly, creating a mass of water and debris that charges down the channel, possibly even overflowing the banks of the arroyo. People entering such channels or camping by them have been swept away by sudden flash floods.

Sand

Figure 13.4.1: Sahara Desert erg

While deserts are defined by dryness, not sand, the popular conception of a typical desert is a sand sea called an **erg**. An erg is a broad area of desert covered by a sheet of fine-grained sand often blown by aeolian forces (wind) into dunes [21]. Probably the best-known erg is the Empty Quarter (Rub' al Khali) of Saudi Arabia, but other ergs exist in Colorado (Great Sand Dunes National Park), Utah (Little Sahara Recreation Area), New Mexico (White Sands National Monument), and California (parts of Death Valley National Park). It is not only deserts that form dunes; the high supply of sand can form ergs anywhere, even as far north as 60° in Saskatchewan at the Athabasca Sand Dunes Provincial Park. Coastal ergs on the shores of lakes and oceans also do exist and can be found in places like Oregon, Michigan, and Indiana.



The way dunes form creates an internal feature called **cross-bedding**. As the wind blows up the windward side of the dune, it carries sand to the dune crest depositing layers of sand parallel to the windward (or "stoss") side. The sand builds up the crest of the dune and pours over the top until the leeward (downwind or slip) face of the dune reaches the **angle of repose,** the maximum angle which will support the sand pile. Dunes are unstable features and move as the sand erodes from the stoss side and continues to drop down the leeward side covering previous stoss and slip-face layers and creating the cross-beds. Mostly, these are reworked over and over again, but occasionally, the features are preserved in a depression, then lithified. Shifting wind directions and abundant sand sources create chaotic patterns of cross-beds like those seen in the fossil ergs represented by the Navajo Sandstone and Zion National Park of Utah.

Figure 13.4.1: Cross beds in the Navajo Sandstone at Zion National Park

In the Mesozoic, Utah was covered by a series of ergs, thickest in Southern Utah. Perhaps the best known of these sandstone formations is the Navajo Sandstone. The Navajo forms the dramatic cliffs and spires in Zion National Park and covers a large part of the Colorado Plateau. It is exposed beneath the Entrada Sandstone in Arches National Park, a later series of sand dunes in which the conditions of the lithified rock allowed the formation of arches.

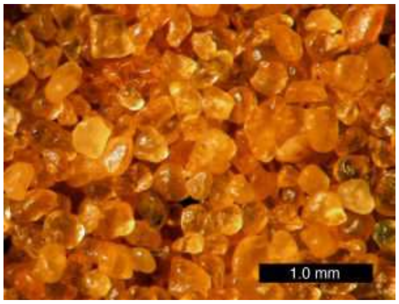


Figure 13.4.1: Enlarged image of frosted and rounded windblown sand grains from Coral Pink Sand Dunes.

As the cement that holds the grains together in these modern sand cliffs disintegrates and the freed grains gather at the base of the cliffs and move down the washes, sand grains may be recycled and redeposited. These great sand ergs may represent ancient quartz sands recycled many times, just passing now through another cycle. One example of this is Coral Pink Sand Dunes State Park in Southwestern Utah, which is sand that is eroded from the Navajo Sandstone forming new dunes.

Dune Types

Figure 13.4.1: NASA image of a barchan dune field in coastal Brazil

Dunes are complex features formed by a combination of wind direction and sand supply, in some cases interacting with vegetation. There are several types of dunes representing the variable of wind direction, sand supply, and vegetative anchoring. **Barchan dunes** or **crescent dunes** form where sand supply is limited and there is a fairly constant wind direction. Barchans move downwind and develop a crescent shape with wings on either side of a dune crest. Barchans are known to actually move over homes, even towns.

Figure 13.4.1: Satellite image of longitudinal dunes in Egypt

Longitudinal dunes or **linear dunes** form where sand supply is greater and the wind is variable around a dominant direction, in a back-and-forth manner. They may form ridges tens of meters high lined up with the predominant wind directions.

Figure 13.4.1: Parabolic dunes, Cape Cod

Parabolic dunes form where vegetation anchors parts of the sand and unanchored parts blowout. Parabolic dune shape is similar to barchan dunes but usually reversed, and it is determined more by the anchoring vegetation than a strict parabolic form.

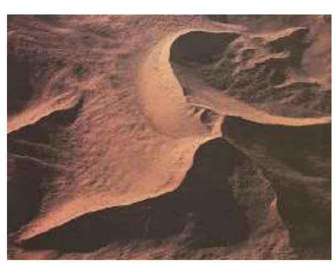


Figure 13.4.1: Star dune in Sahara

Star dunes form where the wind direction is variable in all directions. Sand supply can range from limited to quite abundant. It is the variation in wind direction that forms the star.

References

17. Hooke, R. L. Processes on arid-region alluvial fans. *J. Geol.* **75**, 438–460 (1967).

18. Boggs, S. J. *Principles of sedimentology and stratigraphy*. (Pearson, 2011).

19. Easterbrook, D. J. *Surface processes and landforms*. (Pearson College Division, 1999).

20. King, L. C. Canons of landscape evolution. *Geol. Soc. Am. Bull.* **64**, 721–752 (1953).

21. Wilson, I. G. Desert Sandflow Basins and a Model for the Development of Ergs. *Geogr. J.* **137**, 180–199 (1971).

Contributions and Attributions

13.5: The Great Basin and the Basin and Range

Figure 13.5.1: Map of the Great Basin

The Great Basin is the largest area of interior drainage in North America, meaning there is no outlet to the ocean and all precipitation remains in the basin or is evaporated. It covers western Utah, most of Nevada, and extends into eastern California, southern Oregon, and southern Idaho. Streams in the Great Basin gather runoff and groundwater discharge and deliver it to lakes and playas within the basin. A subregion within the Great Basin is the **Basin and Range** which extends from the Wasatch Front in Utah west across Nevada to the Sierra Nevada Mountains of California. The basins and ranges referred to in the name are horsts and grabens, formed by normal fault blocks from the crustal extension, as discussed in chapter 2 and chapter 9. The lithosphere of the entire area has stretched by a factor of about 2, meaning from end to end, the distance has doubled over the past 30 million years or so. This has created the bowl-like shape of the region, which creates an overall internal drainage, and countless sub-drainages in individual basins. Each of these is lined by alluvial sediments leading into playa or lacustrine depositional environments. Even without the arid conditions, there would be these types of deposits, with lacustrine becoming more common in place of playa. This most recently occurred with pluvial lakes that formed during the last glacial maximum (see chapter 14.4.3).

Figure 13.5.1: Typical Basin and Range scene. Ridgecrest, CA sits just east of the southern Sierra Nevada Mountains.

The desert of the Basin and Range extends from about 35° to near 40° and has a rain shadow effect created by westerly winds from the Pacific rising and cooling over the Sierras, depleted of moisture by precipitation on the western side. The result is relatively dry air descending across Nevada and western Utah. A journey from the Wasatch Front southwest to the Pacific Ocean will show stages of desert landscape evolution from the young fault blocks of Utah with sharp peaks and alluvial fans at the mouths of canyons, through older landscapes in Southern Nevada with bajadas along the mountain fronts, to the oldest landscapes in the Mojave Desert of California with subdued inselbergs sticking up through a sea of old bajadas. These landscapes illustrate the evolutionary stages of desert landscape development.

Desertification

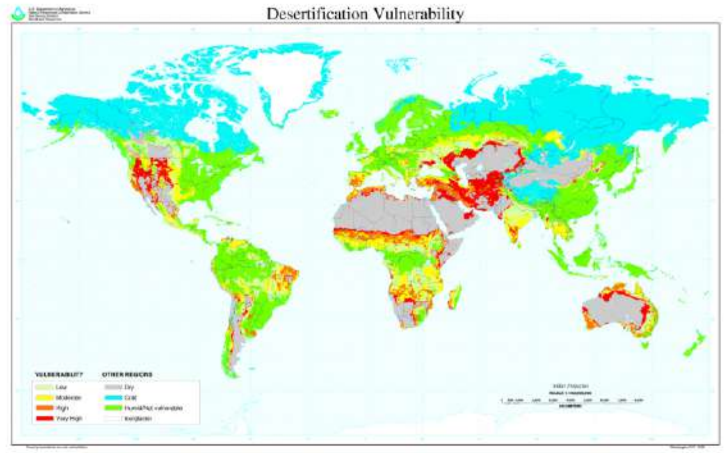


Figure 13.5.1: World map showing desertification vulnerability

Previously arable and usable land may be turned into desert by climate change and the activities of humans, such as poor farming practices, livestock overgrazing, and overuse of available water. This is a process called **desertification** and it is a serious problem worldwide [22]. Plants and soil types that are non-arid specifically help groundwater infiltration and water retention. Adding aridity to an area converts these soils and plants to be less effective in retaining water, and via a **positive feedback loop** (meaning that the processes feed on themselves promoting an increasing spiral). This only increases the aridity and spreads the desert further. The figure shows areas of the world and their vulnerability to desertification. Note the red and orange areas in the western and midwestern United States. The Dust Bowl of the 1930s is a classic example of human-caused desertification. Sometimes there is a conflict between what is known to prevent desertification and what an individual farmer feels he needs to do to make a living. Mitigating the desertification process includes both societal steps and individual education on alternatives.

Reference

22. Geist, H. *The causes and progression of desertification*. (Ashgate Aldershot, 2005).

Contributions and Attributions

CHAPTER OVERVIEW

14: GLACIERS

The hydrosphere, liquid water, is the single most important agent of erosion and deposition. The cryosphere, the solid state of water in the form of ice also has its own unique erosional and depositional features. Large accumulations of year-round ice on the land surface are called glaciers. Masses of ice floating on the ocean as sea ice or icebergs are not glaciers, although they may have had their origin in glaciers.

14.1: Prelude to Galciers

The hydrosphere, liquid water, is the single most important agent of erosion and deposition. The cryosphere, the solid state of water in the form of ice also has its own unique erosional and depositional features. Large accumulations of year-round ice on the land surface are called **glaciers**. Masses of ice floating on the ocean as sea ice or icebergs are not glaciers, although they may have had their origin in glaciers.

Figure 14.1.1: Avalanche Lake in Glacier National Park, Montana is an example of a glacially-carved cirque basin.

Glaciers cover about 10% of the surface of the Earth, and are powerful erosional agents that sculpt the planet's surface. Glaciers form when more snow accumulates over a long span of time than melts and eventually turns into ice. This usually occurs in mountainous areas that have both cold temperatures and high precipitation. But snow can also accumulate and turn into ice in extremely cold low lying areas such as Greenland and Antarctica. This chapter focuses on types of glaciers, how glaciers function, erosional and depositional landforms created by glaciers, and how glaciers are connected to past climates and modern day climate change.

14.2: Types of Glaciers

Figure 14.2.1: Glacier in the Bernese Alps.

There are two general types of glaciers – alpine glaciers and ice sheets. **Alpine glaciers** form in mountainous areas either at high elevations or near cool and wet coastal areas like the Olympic Peninsula of Washington. A common type of alpine glacier is a **valley glacier** which is confined to a long, narrow valley located in mountainous areas especially at higher latitudes (closer to either the north or south pole). Most alpine glaciers are located in the major mountain ranges of the world such as the Andes, Rockies, Alps, and Himalayas.

Figure 14.2.1: Greenland ice sheet.

The other major glacier type is **ice sheets** (also called **continental glaciers**). These are thick accumulations of ice that occupy a large geographical area. The main ice sheets on the earth today are located on Greenland and Antarctica. The Greenland Ice Sheet has an extensive surface area and thickness up to 3,300 meters (10,800 feet or two miles) and has a volume estimated at nearly 3 million cubic kilometers (~102 billion cubic feet) [1].

The Antarctic Ice Sheet is much larger and covers almost the entire continent. The thickest parts of this massive ice sheet are over 4,000 meters thick (>13,000 feet or 2.5 miles) and its weight depresses the Antarctic bedrock to below sea level in many places beneath the ice [2]. The Antarctic Ice Sheet contains the most ice as illustrated by the figure below comparing cross-sectional views of both ice sheets.

Former ice sheets, present during the last glacial maximum event (also known as the last ice age) in North America, are called the Laurentide Ice Sheet.

Figure 14.2.1: Thickness of Greenland ice sheet in meters (Source: Eric Gaba).

Figure 14.2.1: Maximum extent of Laurentide Ice Sheet

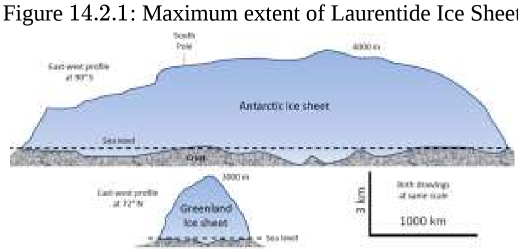


Figure 14.2.1: Cross-sectional view of both Greenland and Antarctic ice sheets drawn to scale for size comparison (Source: Steve Earle)

References

1. Bamber, J. L., Layberry, R. L. & Gogineni, S. P. A new ice thickness and bed data set for the Greenland ice sheet: 1. Measurement, data reduction, and errors. *J. Geophys. Res.* **106**, 33773–33780 (2001).

2. Lythe, M. B. & Vaughan, D. G. BEDMAP: A new ice thickness and subglacial topographic model of Antarctica. *J. Geophys. Res.* **106**, 11335–11351 (2001).

Contributions and Attributions

14.3: Glacier Formation and Movement

Glaciers form when accumulating snow compresses into firn and eventually turns into ice. In some cases, perennial snow accumulates on the ground and lasts all year. This makes a snowfield and not a glacier since it is a thin accumulation of snow. Snow and glacial ice actually have a fair amount of void space (porosity) that traps air. As the snow settles, compacts, and bonds with underlying snow, the amount of void space diminishes. When the snow gets buried by more snow, it compacts into granular **firn** (or névé) with less air and it begins to resemble ice more than snow. Continual burial, compression, and recrystallization make the firn denser and more ice-like. Eventually, the accumulated snow turns fully to ice, however, small air pockets remain trapped in the ice and form a record of the past atmosphere.

Figure 14.3.1: Glacial crevasses.

Figure 14.3.1: Cravasse on the Easton Glacier in the North Cascades

As the ice accumulates, it begins to flow downward under its own weight. An early study of glacier movement conducted in 1948 on the Jungfraufirn Glacier in the Alps installed hollow vertical rods in the ice and measured the tilt over two years. The study found that the top part was fairly rigid and the bottom part flowed internally. A P-T diagram of ice shows that ice actually melts under pressure (one of the unique properties of water) so ice at the base of a typical glacier is actually melting. About half of the overall glacial movement was from sliding on a film of meltwater along the bedrock surface and a half from internal flow [3]. These studies show that the ice near the surface (about the upper 165 feet [50 meters] depending on location, temperature, and flow rate) is rigid and brittle [4]. This upper zone is the **brittle zone**, the portion of the ice in which ice breaks when it moves to form large cracks along the top of a glacier called **crevasses**. These crevasses can be covered and hidden by a snow bridge and thus are a hazard for glacier travelers.

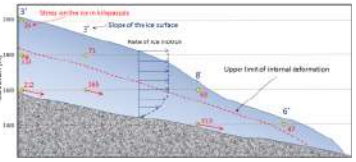


Figure 14.3.1: Cross-section of a valley glacier showing stress (red numbers) increases with depth under the ice. The ice will deform and flow where the stress is greater than 100 kilopascals, and the relative extent of that deformation is depicted by the red arrows. Downslope movement is shown with blue arrows. The upper ice above the red dashed line does not flow but is pushed along en masse. (Source: Steve Earle)

Below the brittle zone, there is so much weight of the overlying ice (typically exceeding 100 kilopascals-approximately 100,000 times atmospheric pressure) that it no longer breaks when force is applied to it but rather it bends or flows. This is the **plastic zone** and, within this zone, the ice flows. The plastic zone represents the great majority of the ice of a glacier and often contains a fair amount of sediment from as large as boulders and as small as silt and clay which act as grinding agents. The bottom of the plastic zone slides and grinds across the bedrock surface and represents the zone of erosion.

References

3. Gerrard, J. A. F., Perutz, M. F. & Roch, A. Measurement of the velocity distribution along a vertical line through a glacier. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **213**, 546–558 (1952).

4. Nye, J. F. The Distribution of Stress and Velocity in Glaciers and Ice-Sheets. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **239**, 113–133 (1957).

Contributions and Attributions

14.4: Glacial Budget

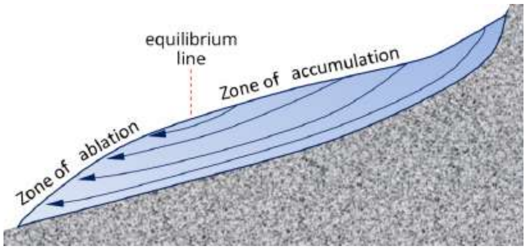


Figure 14.4.1: Cross-sectional view of an alpine glacier showing internal flow lines, zone of accumulation, snow line, and zone of melting.

Glaciers gain mass during the winter as snow accumulates. During summer the snow melts. The glacier is like a bank account, if more money is coming in (snow accumulating in winter) than going out (snow melting in summer), then the bank account grows. The glacial budget works in a similar way. The **glacial budget** describes how ice accumulates and melts on a glacier which ultimately determines whether a glacier advances or retreats. The balance of accumulating ice (**zone of accumulation**) is weighed against melting ice (**zone of melting** or **zone of ablation**), and whichever is greater determines whether the glacier will advance or retreat. In the zone of accumulation, the rate of annual snowfall is greater than the rate of melting. In other words, not all of the snow that falls each winter melts during the following summer, and the ice surface is always covered with snow. In the zone of melting or ablation, more ice melts then accumulates as snow during the year. The equilibrium line (or **snowline**, also called the **firnline**) marks the boundary between the zones of accumulation and ablation. Below the equilibrium line, in the zone of melting, bare ice is exposed because last winter's snow has all melted; above that line, the ice is still mostly covered with snow from last winter. The position of the equilibrium line changes from year to year as a function of the balance between snow accumulation in the winter and snowmelt during the summer. More winter snow and less summer melting obviously favors the advance of the equilibrium line (and of the glacier's leading-edge (or **terminus**), but of these two variables, it is the summer melt that matters most to a glacier's budget. Cool summers promote glacial advance and warm summers promote glacial retreat [5].

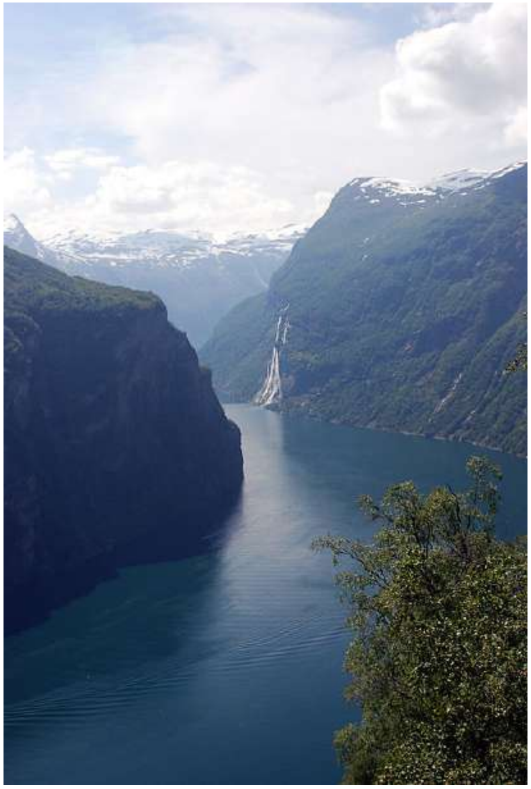


Figure 14.4.1: Fjord

If warmer summers promote glacial retreat, then overall climate warming over many decades and centuries causes the glacier to melt and retreat significantly. Since the global climate has been warming due to human burning of fossil fuels [6], this warming is likely causing the ice sheets to melt (or lose mass) at an increasing rate over years and decades rather than over centuries [7]. This means that as time goes on, the glaciers are melting faster and contributing more to rising sea-level than expected based on the previous history.

When ice sheets start to melt, such as those in Antarctica and Greenland, their flow into the ocean speeds up eventually creating floating ice sheets. The edges of the glacier or its extension as floating ice break off in a process called **calving**. In cases like these, the end of the glacier in the fjord may retreat but it will also lose thickness or deflate [8; 9]. A **fjord** is a narrow ocean-flooded valley with steep walls that were carved by a recent glacier. The retreating glacier or glaciers may add to sea level, and this increased sea level can also add to the flooding of the former glacially-carved valleys. Glacial retreat and deflation are well-illustrated in the 2009 TED Talk by James Balog.

Sorry: we can't play video on this browser. Please make sure it's up to date and that Flash 11.1 or higher is installed. Load this talk on ted.com

References

5. Earle, S. *Physical geology OER textbook*. (BC Campus OpenEd, 2015).

6. Stocker, T. F. *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*. (Cambridge University Press, 2014).

7. Velicogna, I. Increasing rates of ice mass loss from the Greenland and Antarctic ice sheets revealed by GRACE. *Geophys. Res. Lett.* **36**, L19503 (2009).

8. Nick, F. M., Vieli, A., Howat, I. M. & Joughin, I. Large-scale changes in Greenland outlet glacier dynamics triggered at the terminus. *Nat. Geosci.* **2**, 110–114 (2009).

9. Jensen, T. S., Box, J. E. & Hvidberg, C. S. A sensitivity study of annual area change for Greenland ice sheet marine terminating outlet glaciers: 1999–2013. *J. Glaciol.* **62**, 72–81 (2016).

Contributions and Attributions

14.5: Glacial Landforms

Glacial landforms are of two kinds, erosional and depositional landforms. Erosional landforms are formed by removing material. The internal pressure and movement within glacial ice cause some melting and glaciers to slide over bedrock on a thin film of water. Glacial ice also contains a large amount of sediments such as sand, gravel, and boulders. Together, the movement plucks off bedrock and grinds the bedrock producing a polished surface and fine sediment called rock flour as well as other poorly-sorted sediments. Characteristic depositional landscapes are produced when the ice melts and retreats and leaves behind these sediments with distinct shapes and compositions. Because glaciers are among the earliest geological processes studied by geologists whose studies were in Europe, the terminology applied to glaciers and glacial features contains many terms from European languages.

Erosional Glacial Landforms

Both alpine and continental glaciers erode bedrock and create erosional landforms. Glaciers are heavy masses of abrasive ice that grind over the surface. Elongated grooves are created by fragments of rock embedded in the ice at the base of a glacier scraping along the bedrock surface called **glacial striations**. In addition, rock flour as fine grit in the ice can polish a hard granite or quartzite bedrock to a smooth surface called **glacial polish**. This figure illustrates these glacial landforms.

Figure 14.5.1: A polished and striated rock in Yosemite.

Figure 14.5.1: Glacial striations in Mt. Rainier National Park.

Following is a description of erosional landforms produced by alpine glaciers. Since glaciers are typically much wider than rivers of similar length, and since they tend to erode more at their bases than their sides, they transform former V-shaped stream valleys into broad valleys that have relatively flat bottoms and steep sides with a distinctive "U" shape [5]. Little Cottonwood Canyon near Salt Lake City, Utah was occupied by a large glacier that extended down to the mouth of the canyon and into Salt Lake Valley [10]. Today, that canyon is the location of many erosional landforms including the U-shaped valley, as well as polished and striated rock surfaces. In contrast, river-carved canyons have a V-shaped profile when viewed in cross-section. Big Cottonwood Canyon, its neighbor to the north, has that V-shape in its lower parts, indicating that its glacier did not extend clear to its mouth but was confined to its upperparts.

Figure 14.5.1: The U-shape of the Little Cottonwood Canyon, Utah, as it enters into the Salt Lake Valley.

Figure 14.5.1: V-shaped Big Cottonwood Canyon, carved by a stream.

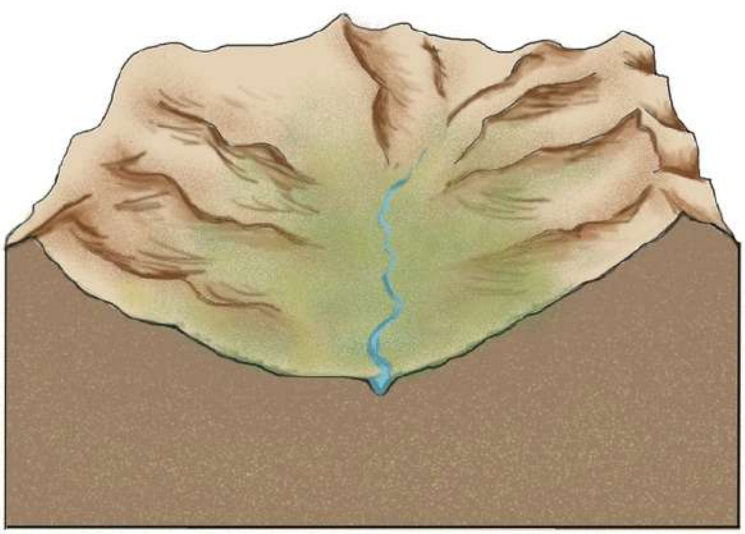


Figure 14.5.1: Formation of a glacial valley. Glaciers change the shape of the valley from a "V" shape to a "U".

When two U-shaped valleys are adjacent to each other, the ridge between them can be carved into a sharp ridge called an **arête**. Since glaciers erode a broad valley, the arêtes are left behind with steep walls separating them. At the head of a glacially carved valley is a bowl-shaped feature called a **cirque** representing where the head of the glacier is eroding against the mountain by plucking rock away from it and the weight of the thick ice is eroding out a bowl. After the glacier is gone, the bowl at the bottom of the cirque is often occupied by a lake called a **tarn.** Headward cirque erosion by three or more mountain glaciers produces **horns,** which are steep-sided, spire-shaped mountains with pronounced cirques on three or more sides. Low points along arêtes or between horns (also mountain passes) are termed **cols**. When a smaller tributary glacier intersects a larger trunk glacier, the smaller glacier erodes down less. Therefore, once the ice has been removed, the tributary valley is left as a **hanging valley**, sometimes with a waterfall. Where the trunk glacier straightens and widens the former V-shaped valley and erodes the ends of side ridges, the result is a steep triangle-shaped cliff called a **truncated spur**.

Figure 14.5.1: Cirque with Upper Thornton Lake in the North Cascades National Park, Washington. Photo by Walter Siegmund.

Figure 14.5.1: An example of a horn, Kinnerly Peak, Glacier National Park, Montana. Photo by USGS.

Figure 14.5.1: Bridalveil Falls in Yosemite National Park, California, is a good example of a hanging valley.

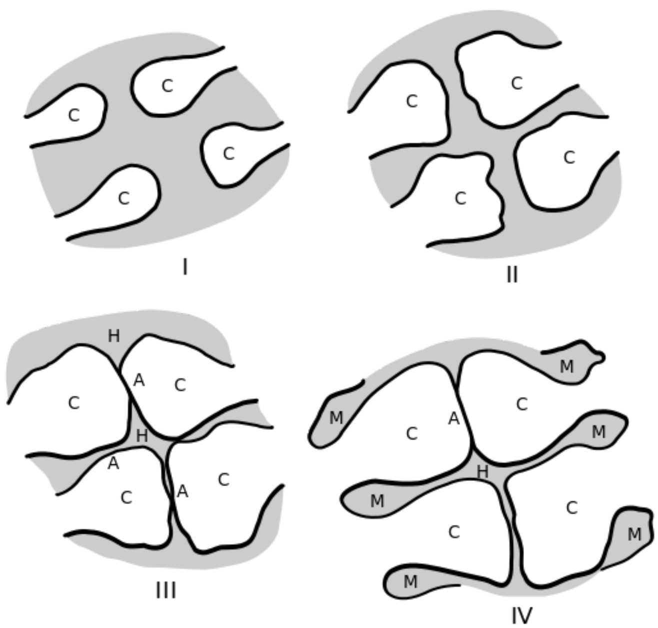


Figure 14.5.1: Key: **A** = Arête, **C** = Cirque, **H** = Horn, **M** = Monument, Preglacial surface shaded

Stage Number	Stage Name	Description
I	Grooved Upland	Early-stage. Only cirques (C) developed at this point.
II	Early-Fretted Upland	Cirque (C) development moves headward, very much like a stream and headward erosion. Still no ridge-like glacial features.

Stage Number	Stage Name	Description
III	Mature-Fretted Upland	Classic mountain glacial landscape. Horns (H) and Arêtes (A) have developed.
IV	Monumented Upland	Final stage of mountain glacial erosion. Erosion so extensive that monuments (M) of the mountain and knife-like ridges are some of the only features remaining.

Depositional Glacial Landforms

Figure 14.5.1: Boulder of diamictite of the Mineral Fork Formation, Antelope Island, Utah, United States.

Sediment is deposited by glaciers in both alpine and continental environments. Since ice is responsible for a large amount of erosion, there is a lot of sediment in glacial ice. When sediment is left behind by a melting glacier, it is called **till** and is characteristically poorly sorted with grain sizes ranging from clay and silt to subrounded pebbles and boulders, possibly striated. Many of the depositional landforms described in this section are composed of till. Lithified rocks of this type are sometimes referred to as **tillites**, though that implies an interpretation of glacial origin. A more objective and descriptive term is **diamictite**, meaning a rock with a wide range of clast sizes.

Material carried by the glacier is called **moraine**, which is an accumulation of glacial till produced by the grinding and erosive effects of a glacier. In valley glaciers, moraine also includes material falling on the sides of the glacier by mass wasting from the valley walls. The glacier acts like a conveyor belt, carrying sediment inside and on the ice and depositing it at the end and sides of the glacier. Because ice in the glacier is always flowing downslope, the deposited moraines build up even if the end of the glacier doesn't advance. The type of moraine depends on its location. A **terminal moraine** is a ridge of unsorted till at the maximum extent of a glacier or the farthest extent of a glacier [11]. When glaciers retreat in episodes, they may leave behind deposits called **recessional moraines**. The recessional moraines look similar to terminal moraines but are formed when the glacier retreat pauses. Moraines located along the side of a glacier are called **lateral moraines** and mostly represent material that fell on the sides of the glacier from the mass wasting of the valley walls. When two tributary glaciers join together, the two lateral moraines combine to form a **medial moraine**. Behind the terminal and recessional moraines is a veneer (or thin sheet) of till on top of bedrock called the **till sheet** (or **ground moraine**).

Figure 14.5.1: Lateral moraines of Little Cottonwood and Bell canyons, Utah.

Figure 14.5.1: Medial moraines where tributary glaciers meet. At least seven tributary glaciers from upstream have joined to form the trunk glacier flowing on out of the upper left of the picture.

In addition to moraines, as glaciers melt they leave behind other depositional landforms. The intense grinding process creates a lot of silt. Streams of water melting from the glacier carry this silt (along with sand and gravel) and deposit it in front of the glacier in an area called an **outwash plain**. In addition, when glaciers retreat, they may leave behind large boulders of a type of rock that doesn't match the local bedrock. These are called **glacial erratics**. When continental glaciers melt, large blocks of ice can be left behind to melt within the impermeable till and can create a depression called a **kettle** that can be later filled with surface water like a **kettle lake**. As glaciers melt, the meltwater flows over the ice surface until it descends into crevasses, perhaps finding channels within the ice or continuing to the base of the glacier into channels along the bottom. Such streams located under continental glaciers carry sediment in a sinuous channel within or under the ice, similar to a river. When the ice recedes, the sediment remains as a long sinuous ridge known as an **esker**. Meltwater descending down through the ice or along the margins of the ice may deposit mounds of sediment that remain as hills called **kames**.

Figure 14.5.1: A small group of Ice Age drumlins in Germany.

Also common in continental glacial areas of New York state and Wisconsin are drumlins. A **drumlin** is an elongated asymmetrical drop-shaped hill with its steepest side pointing upstream to the flow of ice and streamlined side (low angle side) pointing in the direction the ice is flowing.

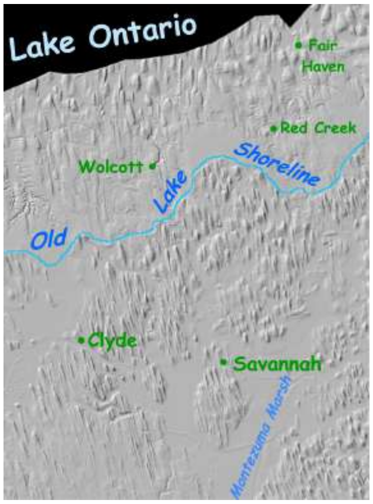
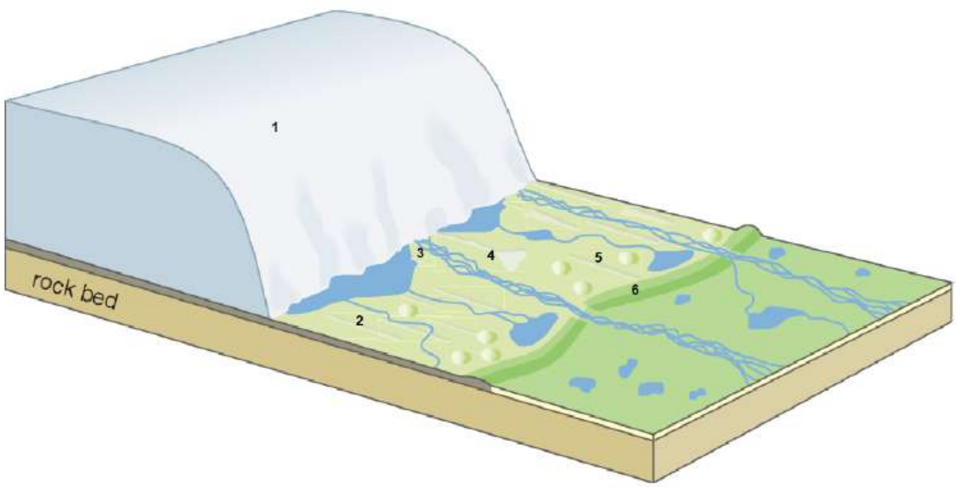


Figure 14.5.1: Drumlin field in Wayne County, New York. Their shape indicates the glacier moved in a NNW to SSE direction.

Drumlins can occur in great numbers in drumlin fields. The origin of drumlins is still debated and some leading ideas are the incremental accumulation of till under the glacier, large catastrophic meltwater floods located under the glacier, and surface deformation by the weight of the overlying glacial ice [12; 13].



Number	Name	Note
1	Receding Glacier	Sheet of ice with a terminus that is moving back with time.
2	Ground Morraine	N/A
3	Esker	N/A
4	Esker	
5	Drumlin	N/A

Number	Name	Note
6	Terminal Morraine	

Glacial Lakes

Figure 14.5.1: Tarn in a cirque.

Lakes are common features in alpine glacial environments. A lake that is confined to a glacial cirque is known as a **tarn** such as Silver Lake near Brighton Ski resort located in Big Cottonwood Canyon or Avalanche Lake in Glacier National Park. Tarns are common in areas of alpine glaciation because the thick ice that carves out a cirque also typically hollows out a depression in bedrock that after the glacier is gone, fills with precipitated water.

Figure 14.5.1: Paternoster lakes

In some cases, recessional moraines may isolate a series of basins within a glaciated valley, and the resulting chain of lakes is called **paternoster lakes**.

Figure 14.5.1: Satellite view of Finger Lakes region of New York

Lakes that occupy long glacially carved depressions are known as **finger lakes** [5].

In areas of continental glaciation, the crust is depressed isostatically by crustal loading from the weight of thick glacial ice. Basins are formed along the edges of continental glaciers (except for those that cover entire continents like Antarctica and Greenland), and these basins fill with glacial meltwater forming **proglacial lakes**. The classic example of a proglacial lake is Lake Agassiz, located mostly in Manitoba, Canada, with Lake Winnipeg serving as the remnant of the lake. Many such lakes, some of them huge, existed at various times along the southern edge of the Laurentide Ice Sheet.

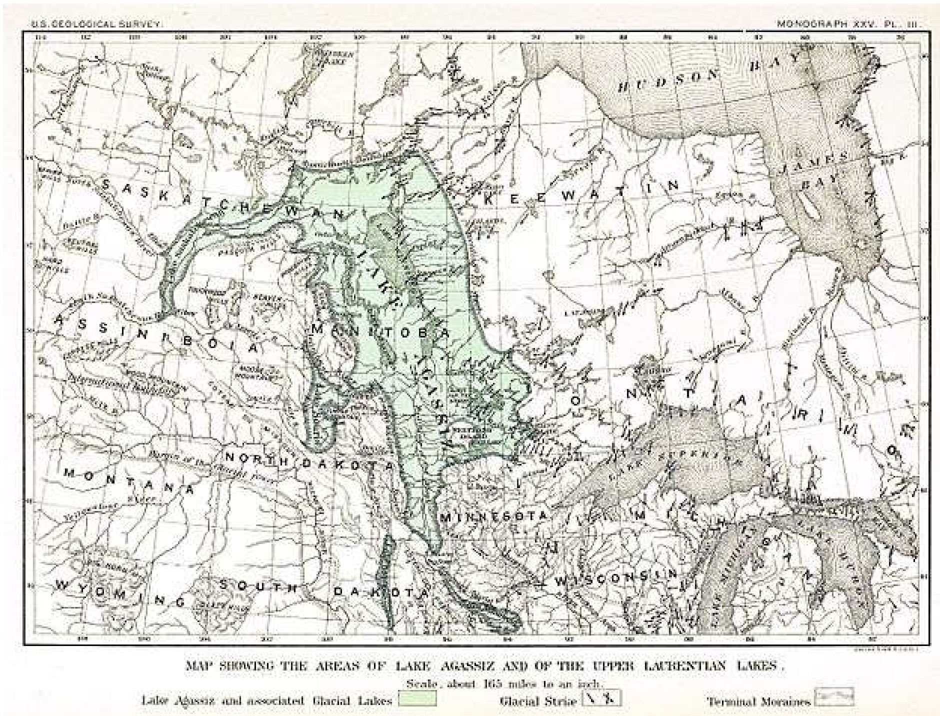


Figure 14.5.1: Extent of Lake Agassiz

Other proglacial lakes formed when glaciers dammed rivers causing the valley to be flooded. The classic example is Glacial Lake Missoula, which formed in Idaho and Montana when the Clark Fork River was dammed by the ice sheet. During the latter part of the last glaciation about 18,000 years ago, the ice holding back Lake Missoula was breached enough to allow some of the lake water to start flowing out, which escalated into a massive and rapid outflow (over days to weeks) during which much of the volume of the lake drained along the valley of the Columbia River to the Pacific Ocean. The ice dam and the lake then formed again. It is estimated that this process of massive flooding happened at least 25 times over that span. In many cases, the rate of outflow was equivalent to the discharge of all of Earth's current rivers combined.

Figure 14.5.1: View of Channeled Scablands in eastern Washington showing huge potholes and massive erosion

The landscape produced by these massive floods is preserved in the Channelled Scablands of Idaho, Washington, and Oregon [5].

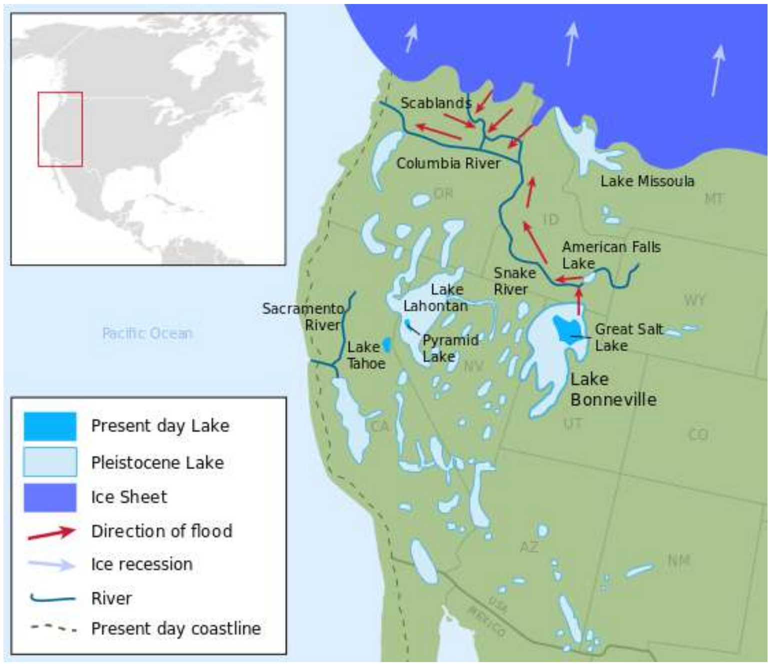


Figure 14.5.1: Pluvial lakes in the western United States

During the last glaciation, most of the western United States had a cooler and wetter climate. Due to less evaporation and more precipitation, many large lakes formed in the basins of the Basin and Range Province called **pluvial lakes**. Two of the largest pluvial lakes were Lake Bonneville and Lake Lahontan. Lake Bonneville occupied much of western Utah and eastern Nevada (Figure from Miller et al 2013) and filled Salt Lake Valley, which is densely urbanized today, with water hundreds of feet deep. During the last glaciation, the level of the lake fluctuated many times leaving several pronounced old shorelines in the western portion of Utah including Salt Lake Valley identified by wave-cut terraces. Lake level peaked around 18,000 years ago [14] and spilled over Red Rock Pass in Idaho into the Snake River causing rapid erosion and a very large flood that rapidly lowered the lake level and scoured the land in Pocatello Valley, Snake River Plain, and Twin Falls Idaho. The flood ultimately flowed into the Columbia River across part of the scablands area and had an incredible discharge of about 4,750 km^3 [15] which is about the volume of Lake Michigan. This would be as if one of the Great Lakes completely emptied within days. Lake Lahontan existed at about the same time mostly in northwestern Nevada.

Figure 14.5.1: The great lakes

The five great lakes in the upper Midwest of North America occupy five basins carved by the ice sheet in a large depression during the Ice Age and were exposed as the ice retreated about 14,000 years ago. They form a naturally interconnected body of freshwater that drains into the Atlantic through the St. Lawrence River. Emergent coastline features are forming on the lakes due to isostatic rebound since the ice retreated (see Chapter 12).

References

5. Earle, S. *Physical geology OER textbook.* (BC Campus OpenEd, 2015).

10. Laabs, B. J. C. *et al.* Chronology of latest Pleistocene mountain glaciation in the western Wasatch Mountains, Utah, U.S.A. *Quat. Res.* **76**, 272–284 (2011).

11. Lewis, R. S. & Stone, J. R. Late Quaternary stratigraphy and depositional history of the Long Island Sound Basin: Connecticut and New York. *J. Coast. Res.* 1–23 (1991).

12. Englert, R. G. *et al.* Quantifying Eroded Sediment Volume during Drumlin Formation in Simcoe County. *Cartographica: The International Journal for Geographic Information and Geovisualization* **50**, 172–178 (2015).

13. Maclachlan, J. C. & Eyles, C. H. Quantitative geomorphological analysis of drumlins in the Peterborough drumlin field, Ontario, Canada. *Geogr. Ann. Ser. A. Phys. Geogr.* **95**, 125–144 (2013).

14. Miller, D. M., Oviatt, C. G. & Mcgeehin, J. P. Stratigraphy and chronology of Provo shoreline deposits and lake-level implications, Late Pleistocene Lake Bonneville, eastern Great Basin, USA. *Boreas* **42**, 342–361 (2013).

15. O'Connor, J. E. Hydrology, hydraulics, and geomorphology of the Bonneville flood. *Geological Society of America Special Papers* **274**, 1–84 (1993).

Contributions and Attributions

14.6: Ice Age Glaciations

A **glaciation** (or ice age) occurs when the Earth's climate is cold enough that large ice sheets grow on continents. There have been four major, well-documented glaciations in Earth's history: one during the Archean-early Proterozoic (~2.5 billion years ago), another in late Proterozoic (~700 million years ago), another in the Pennsylvanian (323 to 300 million years ago), and the most recent Pliocene-Quaternary glaciation (Chapter 15). A minor glaciation that occurred around 440 million years ago in modern-day Africa is also mentioned by some authors. The best-studied glaciation is, of course, the most recent. The Pliocene-Quaternary glaciation is a series of many glacial cycles, possibly 18, during the last 2.5 million years. There is especially strong evidence for eight glacial advances within the last 420,000 years as recorded in the Antarctic ice core record [16]. The last of these, known in popular media as "The Ice Age" but known by geologists as the last glacial maximum, reached its height between 26,500 and 19,000 years ago [10; 17]. Follow this link to an **infographic** that illustrates the glacial and climate changes over the last 20,000 years ending with the human influences since the Industrial Revolution.

Causes of Glaciations

Why do glaciations occur? The causes include both long-term and short-term factors. In the geologic sense, long-term means a scale of 10's to 100's of millions of years and short-term means a 100 to 200,000-year scale. Ideas about long-term causes of glaciations over geologic time include the positioning of continents near poles by plate tectonics and the Wilson Cycle and changes in ocean circulation due to re-positioning of the continents such as the closing of the Panama Strait. Short-term factors are more recognizable for the most recent Pliocene-Quaternary Glaciation and are most relevant to today's anthropogenic climate change, but may have taken place in the earlier glaciations.

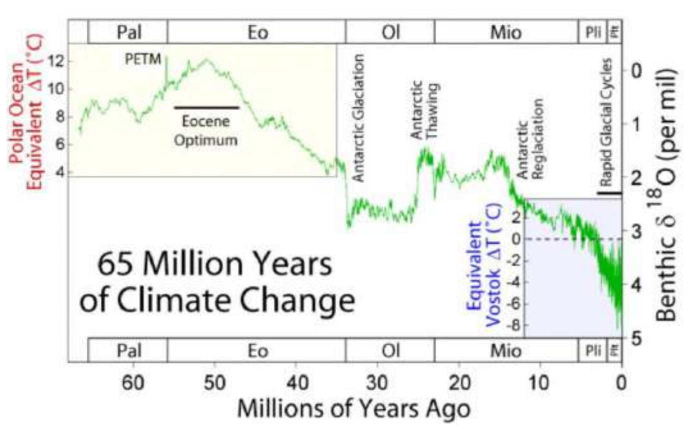
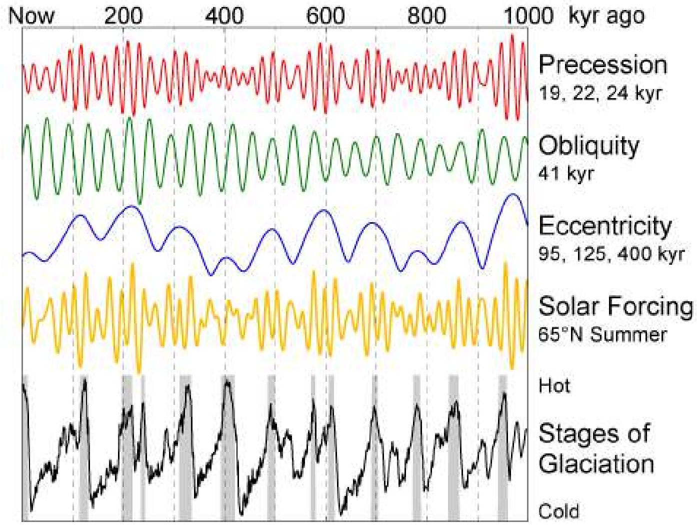
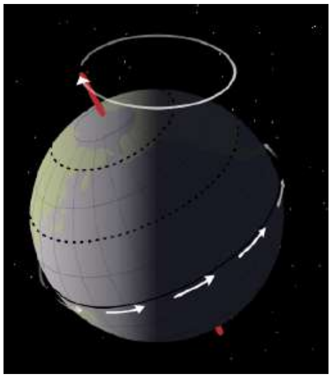
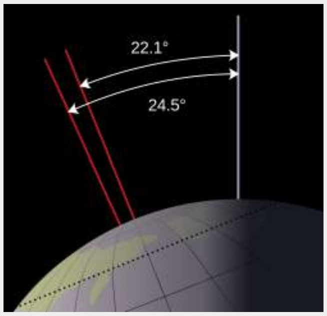
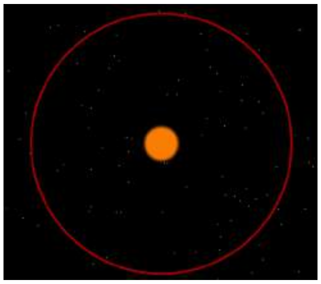


Figure 14.6.1: Atmospheric CO_2 has declined during the Cenozoic from a maximum in the Paleocene-Eocene to its dramatic rise since the Industrial Revolution.

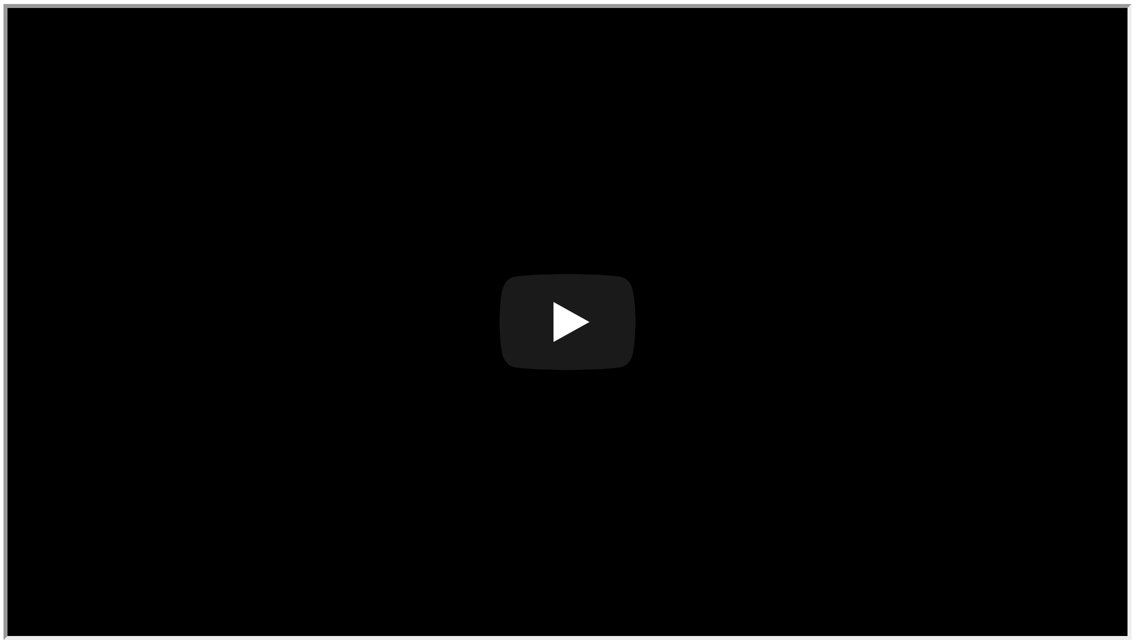
Short-term causes of glacial fluctuations are attributed to cycles in the rotational axis of the Earth and in Earth-Sun relations due to variations in the earth's orbit called **Milankovitch Cycles**. These cycles affect the amount of incoming solar radiation, and changes in carbon dioxide in the atmosphere. During the Cenozoic, carbon dioxide levels steadily decreased from a maximum in the Paleocene causing a gradual climatic cooling [18]. As the climate cooled, the effects of the Milankovitch Cycles began to influence climate with regular cycles of warming and cooling. Milankovitch Cycles are three orbital changes named after the Serbian astronomer Milutin Milankovitch. The three orbital changes are the wobbling of Earth's axis called **precession** with a span of 21,000 years, the angle of Earth's axis called **obliquity** with a span of about 41,000 years, and variations of the distance from the sun in Earth's orbit around the sun referred to as **eccentricity** with a span of 93,000 years [19]. These orbital changes created a glacial-interglacial cycle of 41,000 years from 2.5 to 1.0 million years ago and a longer cycle of ~100,000 years from 1.0 million years ago to today (for detail, see this **chart**). The combination of these three Milankovitch Cycles changes the angle at which the sun's energy strikes the surface of the earth near the poles and the amount of energy (insolation) received by Earth (for detail, see this **chart**). As the climate cooled during the Cenozoic Era, the subtle changes in energy received by the planet were expressed as a warmer and cooler climate cycle, thus the glacial-interglacial cycles.

This chart illustrates the effect of the Milankovitch Cycles.



Name	Illustration
Precession	
Obliquity	
Eccentricity	

This video summarizes ice ages: their characteristics and causes.



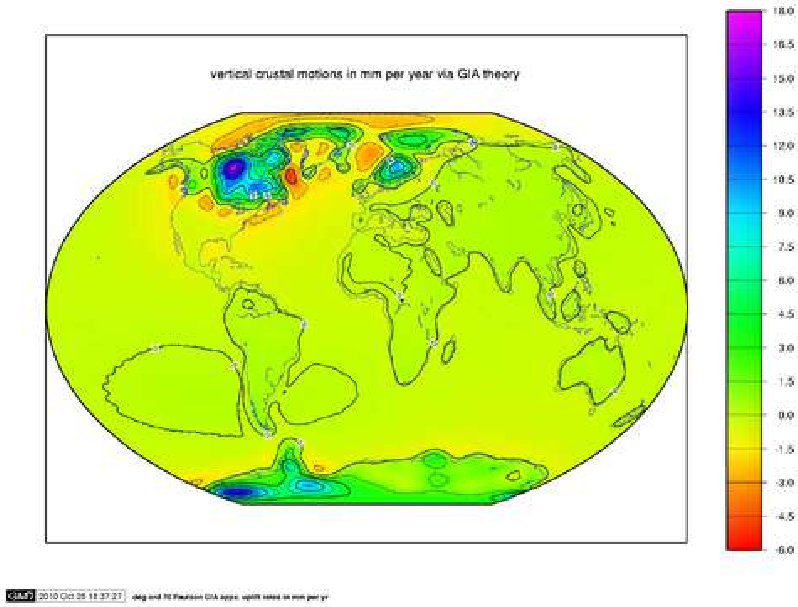
Sea-Level Change and Isostatic Rebound

Since glaciers are ice located on land (not floating in the ocean), when glaciers melt and retreat two things happen, sea-level rises globally and the land rises locally due to isostatic rebound. Melting glacial water runs off into the ocean and sea-level worldwide will rise. For example, since the last glacial maximum about 19,000 years ago [17] sea-level has risen about 400 feet (125 meters) [20]. An overall global change in sea level is called **eustatic** sea-level change. More water in the ocean causes a eustatic sea-level rise. Another important factor causing eustatic sea-level rise is thermal expansion. According to basic physics, thermal expansion occurs when a solid, liquid, or gas expands in volume when the temperature increases. Watch this **30-second video** demonstrating thermal expansion with the classic brass ball and ring experiment. About half of the eustatic sea-level rise during the last century has been the result of thermal expansion, the rest from the melting of glaciers [21; 22].

However, tectonics and isostatic rebound can move the land up and down. The change of sea level as it relates to a more local continental landscape is called relative sea-level change. The relative sea-level change includes both the vertical movement of the eustatic sea-level and the vertical movement of land, so that the sea-level change is measured *relative* to the land. Therefore, if the land rises a lot and sea-level rises only a little, then the sea-level would appear to go down.

The lithosphere can move vertically as a result of two main processes, tectonics, and isostatic rebound. Tectonic uplift occurs when tectonic plates collide as discussed in the Plate Tectonics chapter. **Isostasy** describes the equilibrium that exists for the earth's lithosphere, where denser lithosphere "sinks" lower on top of the asthenosphere and less dense lithosphere "floats" higher on the asthenosphere. **Isostatic rebound** is when some weight is removed from the continental lithosphere causing it to "float" higher on the asthenosphere. Erosion can remove this weight very slowly or the relatively rapid removal of glaciers can remove a lot of weight in a short amount of time. Melting glaciers removes weight from the continental lithosphere causing it to rise or "rebound" from being previously depressed. Most glacial isostatic rebound is occurring where glaciers recently melted (19,000 years ago) such as Canada and Scandinavia. Glacial isostatic rebound causes the relative sea-level to fall or rise less quickly as seen from the land. Isostatic rebound also occurred in Utah when the water from Lake Bonneville was removed [23]. Isostatic rebound is still taking place wherever Ice Age ice or water bodies were present on continental surfaces. Its effects can be seen in terraces forming on river floodplains that cross these areas.

This map shows the rates of vertical crustal movement worldwide. Note that the greatest upward movement occurs in regions affected by recent glaciation, isostatic rebound. Also, note that crustal depression has also occurred in adjacent regions as subcrustal material displaced by isostatic lowering from the weight of the ice has flowed back in under the rebound.



Rates of isostatic rebound worldwide, greatest in regions of recent glaciation.

References

10. Laabs, B. J. C. *et al.* Chronology of latest Pleistocene mountain glaciation in the western Wasatch Mountains, Utah, U.S.A. *Quat. Res.* **76**, 272–284 (2011).

16. Augustin, L. *et al.* Eight glacial cycles from an Antarctic ice core. *Nature* **429**, 623–628 (2004).

17. Clark, P. U. *et al.* The Last Glacial Maximum. *Science* **325**, 710–714 (2009).

18. McInerney, F. A. & Wing, S. L. The Paleocene-Eocene Thermal Maximum: A Perturbation of Carbon Cycle, Climate, and Biosphere with Implications for the Future. *Annu. Rev. Earth Planet. Sci.* **39**, 489–516 (2011).

19. Hays, J. D., Imbrie, J. & Shackleton, N. J. Variations in the Earth's Orbit: Pacemaker of the Ice Ages. *Science* **194**, 1121–1132 (1976).

20. Fleming, K. *et al.* Refining the eustatic sea-level curve since the Last Glacial Maximum using far- and intermediate-field sites. *Earth Planet. Sci. Lett.* **163**, 327–342 (1998).

21. Church, J. A. *et al.* Changes in sea level. , *in: J.T Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. Van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.): Climate Change 2001: The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel* 639–694 (2001).

22. Wigley, T. M. L. & Raper, S. C. B. Thermal expansion of sea water associated with global warming. *Nature* **330**, 127–131 (1987).

23.Max D Crittenden. New Data on the Isostatic Deformation of Lake Bonneville. *GEOLOGICAL SURVEY PROFESSIONAL PAPER 454-E* (1963).

Contributions and Attributions

CHAPTER OVERVIEW

15: GLOBAL CLIMATE CHANGE

This chapter describes the systems involved in regulating Earth's temperature, its climate, geologic evidence of past climate changes, and the role humans have on today's climate change. It is critically important to be aware of the geologic context of climate change processes if we want to understand anthropogenic (human-caused) climate change.

15.1: Global Climate Change

This chapter describes the systems involved in regulating Earth's temperature, its climate, geologic evidence of past climate changes, and the role humans have on today's climate change. It is critically important to be aware of the geologic context of climate change processes if we want to understand anthropogenic (human-caused) climate change. First, this awareness increases the understanding of how and why our activities are causing present-day climate change, and second, it allows us to distinguish between natural and anthropogenic processes in the climate record in the past.

Figure 15.1.1: The "Blue Marble," a picture of our planet from the 1972 Apollo 17 mission, shows that our planet is a finite place with many interacting systems. While the exact photographer is unknown, it was most likely taken by the first (and only) geologist on the moon: Harrison "Jack" Schmitt.

In science, a system is a group of objects and processes that interact, such as the rock cycle. **Earth System Science** is the study of how earth systems (geosphere, atmosphere, hydrosphere, cryosphere, and biosphere) interact and change in response to natural cycles and new human-driven forces. Changes in one earth system affect other systems. A significant part of this chapter introduces various processes from different earth systems and discusses how they influence each other and impact the global climate. For example, global temperature largely changes based on the composition of atmospheric gases (atmosphere), circulation of the ocean (hydrosphere), and characteristics of the land surface (geosphere, cryosphere, and biosphere).

In order to understand climate change, it is important to distinguish between climate and weather. **Weather** is the temperature and precipitation patterns occurring in the short-term such as right now or later this week. **Climate** is the temperature and precipitation patterns and range of variability averaged over the long-term for a particular region (Chapter 13.1). Thus, a single cold winter does not mean that the entire globe is cooling—indeed, the cold winters in the US of 2013 and 2014 took place while the rest of the Earth was experiencing record warm winter temperatures. To avoid these generalizations, many scientists use a 30-year average as a good baseline [1]. Therefore, climate change refers to slow changes in temperature and precipitation patterns over the long-term for a particular area or the Earth as a whole.

Reference

1. Gutro, R. NASA - What's the Difference Between Weather and Climate? (2005). Available at: http://www.nasa.gov/mission_pages/noaa-n/climate/climate_weather.html. (Accessed: 28th September 2016)

15.2: Earth’s Temperature

Because the Moon doesn’t have much of an atmosphere, daytime temperatures on the moon are around 224℉ and nighttime temperatures are around -298℉. That is an astonishing 522 degrees of change between the light-side and dark-side of the Moon [2]. This section describes how Earth’s atmosphere is involved in regulating the Earth’s temperature.

Earth’s Energy Budget

Solar radiation arriving at Earth from the Sun is relatively uniform. Energy (or heat) radiates from the Earth’s surface and lower atmosphere back to space. This flow of incoming and outgoing energy is the Earth’s energy budget. For Earth’s temperature to be stable over long stretches of time, incoming energy and outgoing energy have to be equal on average so that the energy budget at the top of the atmosphere balances. About 29 percent of the incoming solar energy arriving at the top of the atmosphere is reflected back to space by clouds, atmospheric particles, or reflective ground surfaces like sea ice and snow. About 23 percent of incoming solar energy is absorbed in the atmosphere by water vapor, dust, and ozone. The remaining 48 percent passes through the atmosphere and is absorbed at the surface. Thus, about 71 percent of the total incoming solar energy is absorbed by the Earth system [3].

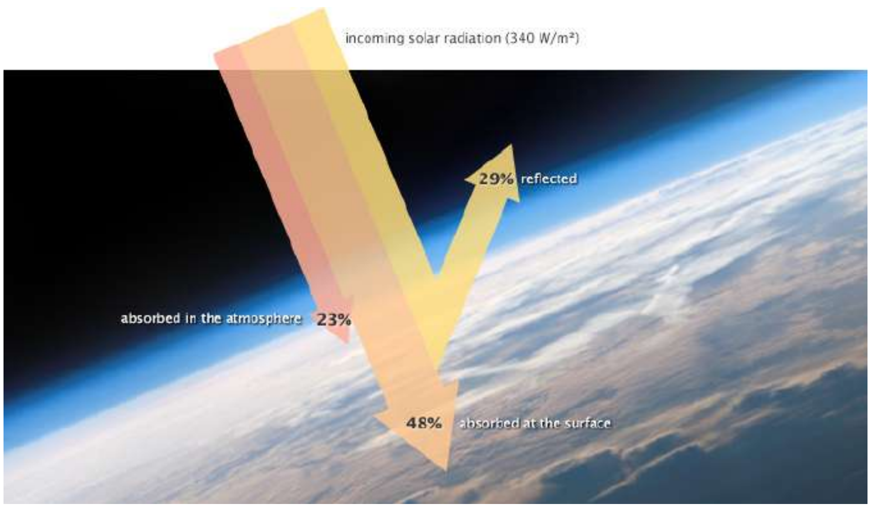


Figure 15.2.1: Incoming solar radiation filtered by the atmosphere.

When this energy reaches Earth, the atoms and molecules making up the atmosphere and surface absorb the energy and they increase in temperature. If this material could only absorb energy, then the temperature of the Earth would be like the water level in a sink with no drain where the faucet runs continuously. The sink would eventually overflow. However, the temperature does not infinitely rise because the Earth is not just absorbing sunlight. The Earth’s surface is also radiating thermal energy (heat) back into the atmosphere. If the temperature of the Earth rises, the planet emits an increasing amount of heat to space and this is the primary mechanism that prevents Earth from continually heating [3].

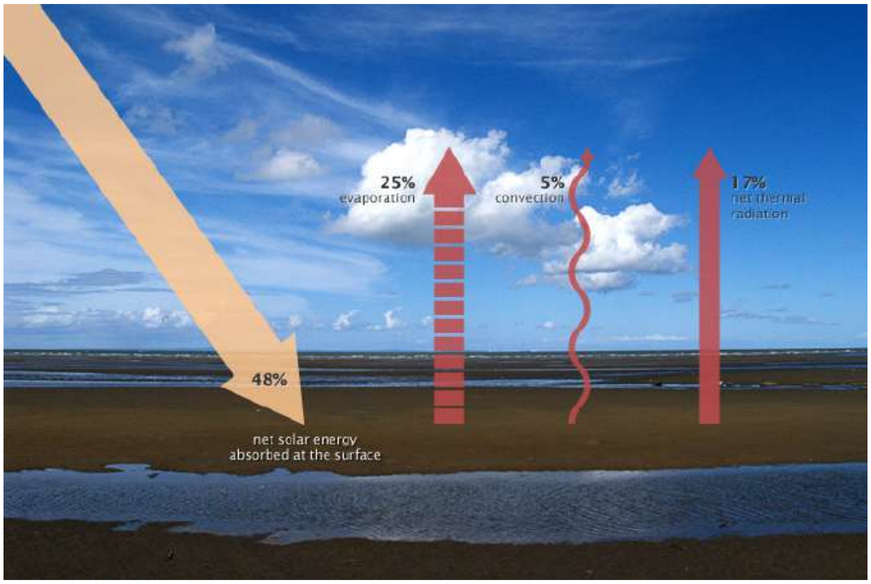


Figure 15.2.1: Some of the thermal infrared energy (heat) radiated from the surface into the atmosphere is trapped by gasses in the atmosphere.

Greenhouse gases act like a giant blanket for Earth. The more greenhouse gases in the atmosphere, then the more outgoing heat will be retained by Earth and the less of this thermal infrared energy (heat) dissipates to space. The greenhouse effect is discussed in more detail in the next section.

Factors that can affect the Earth's energy budget are not limited to greenhouse gases. Increases in solar irradiance (more solar energy) can increase the energy received by the earth. However, increases associated with this are very small [3; 4; 5]. In addition, less ice and snow covers the land and the Arctic Sea increases the amount of sunlight absorbed by land and water (see animation below). The reflectivity of the Earth's surface is called **albedo**. Furthermore, aerosols (dust particles) produced from burning coal, diesel engines, and volcanic eruptions can reflect more incoming solar radiation and actually cool the planet. The effect of anthropogenic aerosols is weak on the climate system but anthropogenic production of greenhouse gases is not weak. Thus, the net effect is warming due to more anthropogenic greenhouse gases associated with fossil fuel combustion [6; 7; 8].

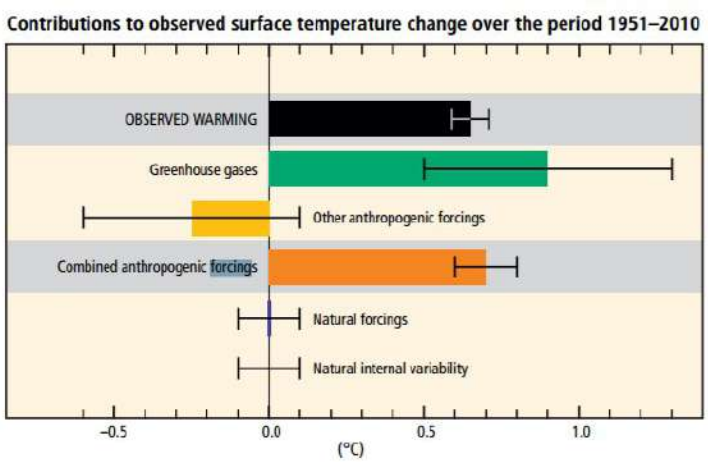


Figure 15.2.1: Net effect of factors influencing warming.

An effect that changes the planet can trigger feedback mechanisms that amplify or suppress the original effect. A **positive feedback** mechanism is when the output or effect enhances the original stimulus or cause. Thus, it increases the effect later. For example, the loss of sea ice at the North Pole makes that area less reflective (reduced albedo). This allows the surface air and ocean to absorb more energy in an area that was once covered by sea ice [3]. Another example is the melting permafrost. **Permafrost** is permanently frozen soil located near the high latitudes, mostly in the Northern Hemisphere. As the climate warms, more permafrost thaws and the thick deposits of organic matter are exposed to oxygen and begin to oxidize (or decay). This oxidation process releases carbon dioxide and methane which in turn causes more warming which melts more permafrost, etc.

A **negative feedback** mechanism occurs when the output or effect reduces the original stimulus or cause [3]. For example, in the short term, more carbon dioxide (CO_2) is expected to cause forest canopies to grow and absorb more CO_2. An example of the long term is increased carbon dioxide (CO_2) in the atmosphere is expected to cause more carbonic acid and chemical weathering, resulting in transport of dissolved bicarbonate and other ions to the oceans which then become stored in sediment.

Composition of Atmosphere

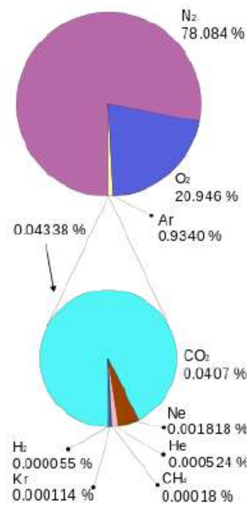


Figure 15.2.1: Composition of the atmosphere

The composition of the atmosphere is a key component of the regulation of the planet's temperature. The atmosphere is 78% nitrogen (N_2), 21% oxygen (O_2), 1% argon (Ar), and less than 1% for all other gases known as trace components. The trace components include carbon dioxide (CO_2) water vapor (H_2O), neon, helium, and methane. Water vapor is highly variable, mostly based on region, but has been estimated to be about 1% of the atmosphere [9]. The trace gases include several important greenhouse gases, which are the gases responsible for warming and cooling the plant. On a geologic scale, the source of atmospheric CO_2 is volcanoes and the sink for CO_2 is the weathering process that buries CO_2 in sediments. Biological processes both add and subtract CO_2 from the atmosphere [10].

Greenhouse gases trap heat in the atmosphere and warm the planet. They have little effect on incoming solar radiation (which is shortwave radiation) but absorb some of the outgoing infrared radiation (longwave radiation) that is emitted from Earth, thus keeping it from being lost to space. More greenhouse gases in the atmosphere absorb more longwave heat and make the planet warmer.

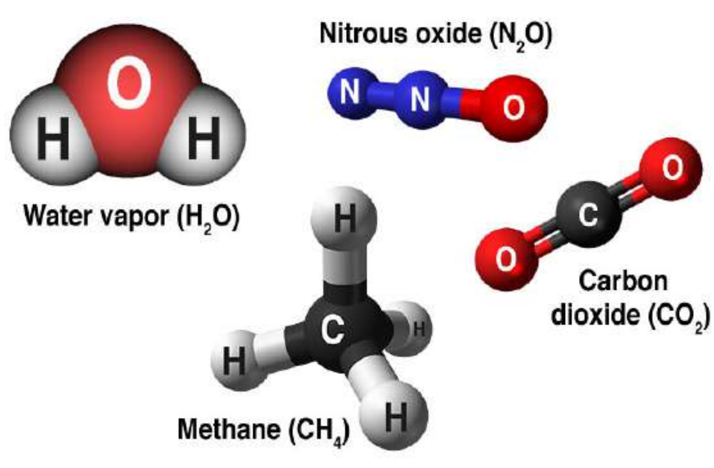


Figure 15.2.1: Common greenhouse gases

The most common greenhouse gases are water vapor (H_2O), carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O). Water vapor is the most abundant greenhouse gas but its abundance in the atmosphere does not change much over time. Carbon dioxide is much less abundant than water vapor, but carbon dioxide is being added to the atmosphere by human activities such as burning fossil fuels, land-use changes, and deforestation. Further, natural processes such as volcanic eruptions add carbon dioxide [3], but at an insignificant rate compared to anthropogenic contributions.

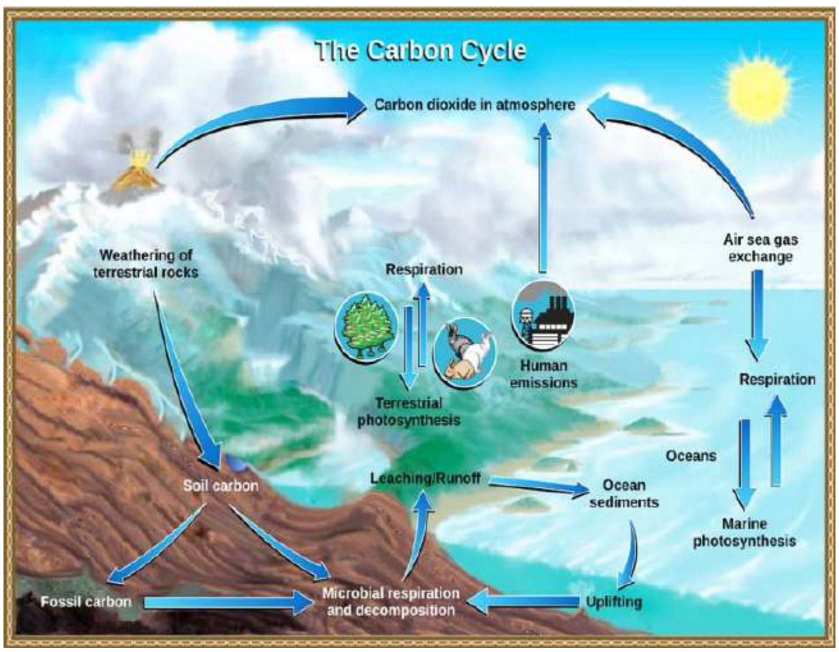
There are two important reasons why carbon dioxide is the most important greenhouse gas. First, carbon dioxide has a long residence time in the atmosphere (meaning that it does not go away for hundreds of years). Second, most of the additional carbon dioxide is "fossil" in origin. That means that it is released by burning fossil fuels. For example, coal is a fossil fuel. Coal is made from plant material created by photosynthesis millions of years ago and stored in the ground. Photosynthesis takes sunlight plus carbon dioxide and creates the carbohydrates of plants. This occurs over millions of years, as a slow process accumulating fossil carbon in rocks and sediments. When we burn coal, we instantaneously release the stored solar energy and the fossil carbon dioxide that took millions of years to accumulate in the first place.

Carbon Cycle

Earth has two important carbon cycles. One is the biological one, wherein living organisms—mostly plants—consume carbon dioxide from the atmosphere to make their tissues through photosynthesis, and then, after they die, that carbon is released back into the atmosphere when they decay over several years or decades [11]. The following is the general equation for photosynthesis.

$$CO_2 + H_2O + sunlight \rightarrow sugar + O_2$$

The second is the geologic carbon cycle. A small portion of this biological-cycle carbon becomes buried in sedimentary rocks during the slow formation of coal, as tiny fragments and molecules in organic-rich shale, and as the shells and other parts of marine organisms in limestone. This then becomes part of the geological carbon cycle, a cycle that actually involves a majority of Earth's carbon, but one that operates only very slowly [11].



Carbon cycle.

The following is a list of storage reservoirs for the geological carbon cycle.

- Organic matter from plants is stored in peat, coal, and permafrost for thousands to millions of years.
- Weathering of silicate minerals converts atmospheric carbon dioxide to dissolved bicarbonate, which is stored in the oceans for thousands to tens of thousands of years.
- Dissolved carbon is converted by marine organisms to calcite, which is stored in carbonate rocks for tens to hundreds of millions of years.
- Carbon compounds are stored in sediments for tens to hundreds of millions of years; some end up in petroleum deposits.
- Carbon-bearing sediments are transferred by subduction to the mantle, where the carbon may be stored for tens of millions to billions of years.
- During volcanic eruptions, carbon dioxide is released back to the atmosphere, where it is stored for years to decades [11].

During much of Earth's history, the geological carbon cycle has been balanced, with carbon being released by volcanism at approximately the same rate that it is stored by the other processes. Under these conditions, the climate remains relatively stable. During some times of Earth's history, that balance has been upset. This can happen during prolonged stretches of greater than average volcanism. One example is the eruption of the Siberian Traps at around 250 million years ago, which appears to have led to strong climate warming over a few million years. A carbon imbalance is also associated with significant mountain-building events. For example, the Himalayan Range has been forming since about 40 Ma and over that time — and still today — the rate of weathering on Earth has been enhanced because those mountains are so high and the range is so extensive. The weathering of these rocks — most importantly the hydrolysis of feldspar — has resulted in the consumption of atmospheric carbon dioxide and transfer of the carbon to the oceans and to ocean-floor carbonate minerals. The steady drop in carbon dioxide levels over the past 40 million years, which contributed to the Pleistocene glaciations, is partly attributable to the formation of the Himalayan Range. Another, non-geological form of carbon-cycle imbalance is happening today on a very rapid time scale. We are in the process of extracting vast volumes of fossil fuels (coal, oil, and gas) that were stored in rocks over the past several hundred million years, and converting these fuels to energy and carbon dioxide. By doing so, we are changing the climate faster than has ever happened in the past [11].

Greenhouse Effect

The **greenhouse effect** is a natural process by which the atmosphere warms surface temperatures. Without an atmosphere, Earth would have huge fluctuations in temperature between day and night like the moon. Daytime temperatures would be hundreds of degrees Fahrenheit above normal and nighttime temperatures would be hundreds of degrees below normal. The greenhouse effect occurs because of the presence of greenhouse gases in the atmosphere.

The greenhouse effect is named after a similar process that warms a greenhouse or a car on a hot summer day. Sunlight passes through the glass of the greenhouse or car, reaches the interior, and changes into heat. The heat radiates upward and gets trapped by the glass windows. The greenhouse effect for the Earth can be explained in three steps.

Step 1: Solar radiation from the sun is composed of mostly ultraviolet (UV), visible light, and infrared (IR) radiation. Components of solar radiation include parts with a shorter wavelength than visible light, like ultraviolet light, and parts of the spectrum with longer wavelengths, like IR and others. Some of the radiation gets absorbed, scattered, or reflected by the atmospheric gases but about half of the solar radiation eventually reaches the Earth's surface.

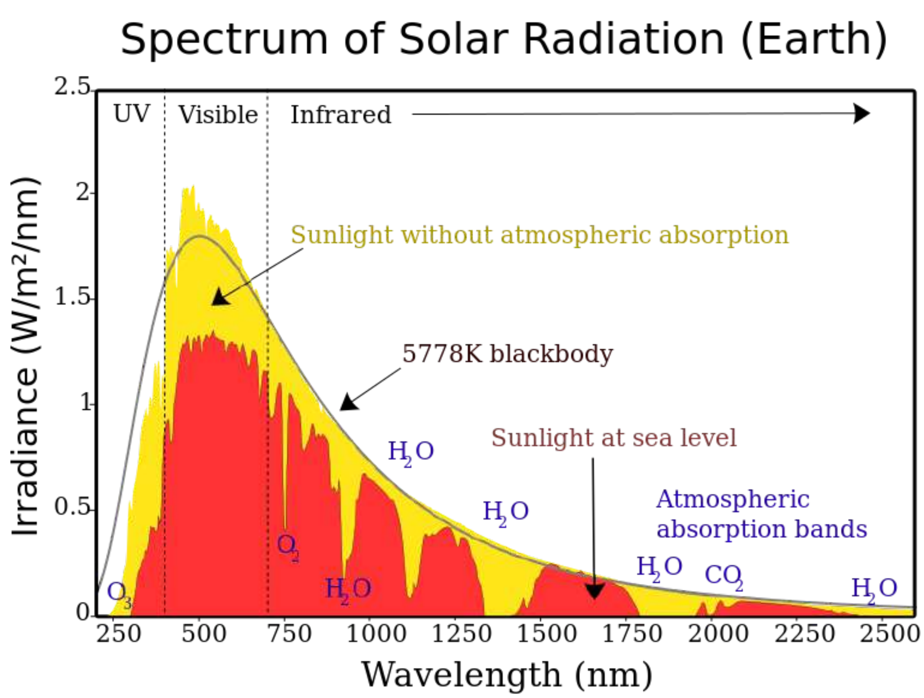


Figure 15.2.1: Incoming radiation absorbed, scattered, and reflected by atmospheric gases.

Step 2: The visible, UV, and IR radiation, that reaches the surface converts to heat energy. Most students have experienced sunlight warming a surface such as a paved surface, a patio, or deck. When this occurs, the warmer surface thus emits more thermal radiation, which is a type of IR radiation. So, there is a conversion from visible, UV, and IR to just thermal IR. This thermal IR is what we experience as heat. If you have ever felt the heat radiating from a fire or a hot stovetop, then you have experienced thermal IR.

Step 3: Thermal IR radiates from the earth's surface back into the atmosphere. But since it is thermal IR instead of UV, visible, or regular IR, this thermal IR gets trapped by greenhouse gases. In other words, the sun's energy leaves the Earth at a different wavelength than it enters, so, the sun's energy is not absorbed in the lower atmosphere when energy is coming in, but rather when the energy is going out. The gases that typically do this blocking on Earth include carbon dioxide, water vapor, methane, and nitrous oxide. More greenhouse gases in the atmosphere result in more thermal IR being trapped. Explore this external link to an interactive animation on the greenhouse effect from the National Academy of Sciences.

References

2. Wolpert, S. New NASA temperature maps provide a 'whole new way of seeing the moon'. (2009). Available at: http://newsroom.ucla.edu/releases/new-nasa-temperature-maps-provide-102070. (Accessed: 23rd February 2017)

3. Lindsey, R. Climate and Earth's Energy Budget: Feature Articles. (2009). Available at: http://earthobservatory.nasa.gov. (Accessed: 14th September 2016)

4. Fröhlich, C. & Lean, J. The Sun's total irradiance: Cycles, trends and related climate change uncertainties since 1976. *Geophys. Res. Lett.* **25**, 4377–4380 (1998).

5. Lean, J., Beer, J. & Bradley, R. Reconstruction of solar irradiance since 1610: Implications for climate change. *Geophys. Res. Lett.* **22**, 3195–3198 (1995).

6. Pachauri, R. K. *et al. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (IPCC, 2014).

7. Oreskes, N. The scientific consensus on climate change. *Science* **306**, 1686–1686 (2004).

8. Levitus, S. *et al.* Anthropogenic warming of Earth's climate system. *Science* **292**, 267–270 (2001).

9. North Carolina State University. Composition of the Atmosphere. (2013). Available at: http://climate.ncsu.edu/edu/k12/.AtmComposition. (Accessed: 12th September 2016)

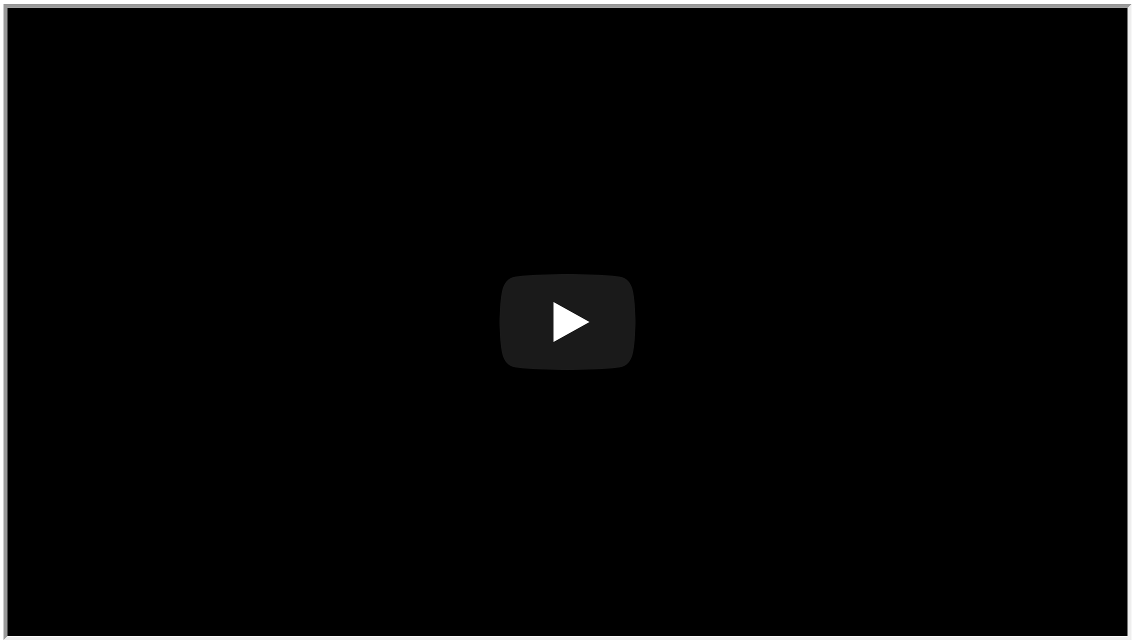
10. Lacis, A. A., Hansen, J. E., Russell, G. L., Oinas, V. & Jonas, J. The role of long-lived greenhouse gases as principal LW control knob that governs the global surface temperature for past and future climate change. *Tellus B Chem. Phys. Meteorol.* **65**, 19734 (2013).

11. Earle, S. *Physical geology OER textbook*. (BC Campus OpenEd, 2015).

Contributions and Attributions

15.3: Evidence of Recent Climate Change

While climate has changed many times in the past (see chapter 14.5.1 and chapter 15.3), the scientific consensus is that human activity is causing the climate to change today more rapidly [11; 7]. While this seems like a new idea, it has been suggested for more than 75 years [12]. This section describes the evidence that scientists agree is most likely a result of **anthropogenic climate change**, or, human-caused climate change. For more information, watch this six-minute video on climate change by two professors at a North Carolina State University.



Global Temperature Rise

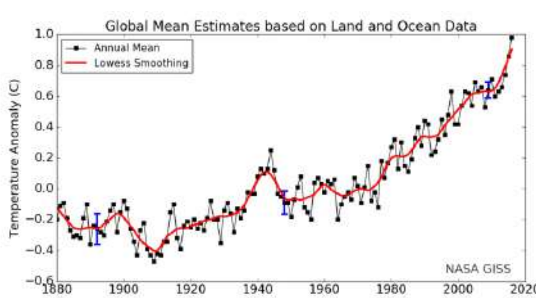


Figure 15.3.1: Land-ocean temperature index, 1880 to present, with a base time 1951-1980. The solid black line is the global annual mean and the solid red line is the five-year lowess smooth. The blue uncertainty bars (95% confidence limit) account only for incomplete spatial sampling.

Since 1880, average global surface temperatures have trended upward and most of that warming has occurred since 1970 (see this NASA animation). Since the ocean is absorbing a lot of the additional trapped heat, surface temperatures include both land surface and ocean temperatures [13]. Changes in land surface or ocean surface temperatures can be expressed as temperature **anomalies**. A temperature anomaly is the difference in average temperature measurement from a predetermined datum. This datum is the average temperature of a particular date range, for example, 1951 to 1980. Another common datum is the last century (1900-2000). Therefore, an anomaly of 1.25 ℃ for 2015 (last century datum) means that the average temperature for 2015 was 1.25 ℃ greater than the 1900-2000 average. In 1950, the temperature anomaly was -0.28 ℃, so this is -0.28 ℃ lower than the 1900-2000 average [3]. These temperatures are annual average surface temperatures.

This video figure shows worldwide temperature changes since 1880. The more blue, the cooler; the more yellow and red, the warmer.

In addition to a rising average surface land temperature, the ocean has absorbed a lot of the heat (remember that the specific heat of water is unusually large). With oceans covering about 70% of the earth's surface, there is a lot of opportunities to absorb energy. The ocean has been absorbing about 80% to 90% of the additional heat added due to human activities. As a result, the top 2,300 feet of the ocean has increased in temperature 0.3℉ since 1969 (external link to this 3-minute video by NASA JPL on heat capacity of the ocean) [3]. The reason the ocean has warmed less than the atmosphere, while still taking on most of the heat, is due to the very high specific heat of water, which means that water can absorb a lot of energy for a small temperature increase. In contrast, the atmosphere needs less energy to increase its temperature.

Some scientists suggest that anthropogenic greenhouse gases do not cause global warming since surface temperatures have not increased very much between 1998 and 2013, while greenhouse gas concentrations have continued to increase during that time period. However, since the oceans are absorbing most of the heat, decade-scale circulation changes (similar to La Niña) in the ocean push warmer water deeper under the surface [14; 15; 16]. Once the absorption and circulation of the ocean is accounted for and the heat added back into surface temperatures, then the temperature increases become apparent as shown in the above figure. Furthermore, this ocean heat storage is temporary, as reflected in the record-breaking warm years of 2014-2016. Indeed, with this temporary ocean storage effect, 15 of the first 16 years of the 21st century have been the hottest in recorded history.

Carbon Dioxide

Anthropogenic greenhouse gases, mostly carbon dioxide (CO_2), have increased since the industrial revolution when the burning of fossil fuels dramatically increased. These levels are unprecedented in the last 800,000-year earth history as recorded in geologic sources such as ice cores. Carbon dioxide has increased by 40% since 1750 and the rate (or speed) of increase has been the fastest during the last decade [3; 6]. For example, since 1750, 2040 gigatons of CO_2 have been added to the atmosphere, about 40% have remained in the atmosphere while the remaining 60% have been absorbed into the land (by plants and soil) or the oceans [6]. Indeed, during the lifetime of most young adults, the total atmosphere has increased by 50 ppm or 15%.

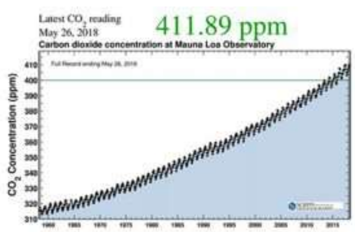
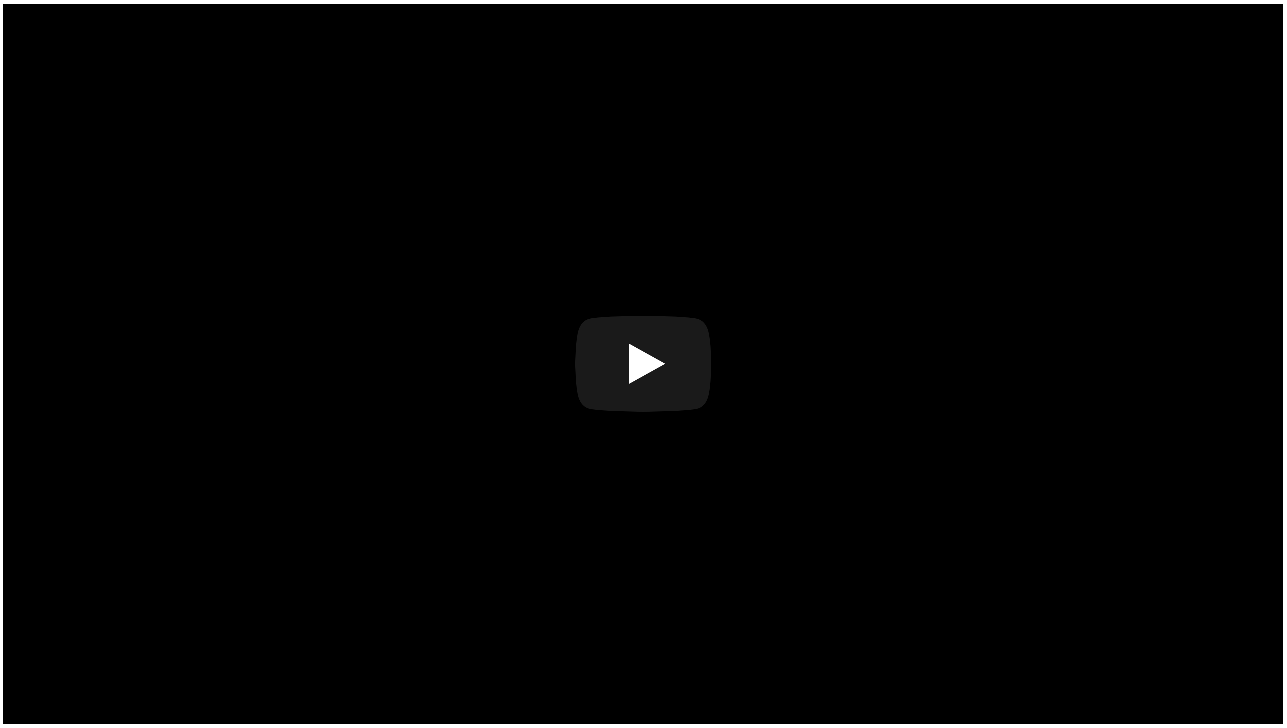
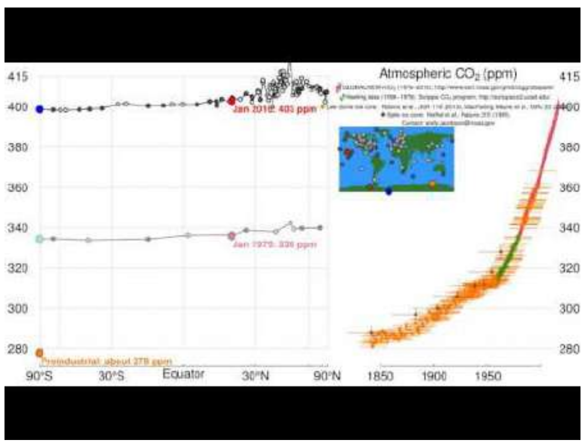


Figure 15.3.1: The Keeling Curve showing increasing atmospheric CO_2 since 1958. Note that its increase is exponential, not linear!

Charles Keeling, an oceanographer with Scripps Institution of Oceanography in San Diego, California was the first person to make regular measurements of atmospheric CO_2. Using his methods, constant measurements of CO_2 in the atmosphere have been made at the Mauna Loa Observatory on Hawaii since 1958. These measurements are published regularly by NASA at this website: scripps.ucsd.edu/programs/keelingcurve/. Go there now to see the very latest measurement. Keeling's measured values have been posted in a curve of increasing values called the Keeling Curve. This curve varies annually up and down from summer (when the plants in the Northern Hemisphere are using CO_2) to winter when the plants are dormant, but shows a steady increase over the past several decades. This curve increases exponentially, not linearly indicating that the rate of increase of CO_2 is itself increasing!

The following video shows how atmospheric CO_2 has varied recently and also over the last 800,000 years as determined by many CO_2 monitoring stations (shown on the insert map). It is also instructive to watch the CO_2 variation of the Keeling portion of the video by latitude. This shows that most of the human sources of CO_2 are in the Northern Hemisphere.





Melting Glaciers and Shrinking Sea Ice

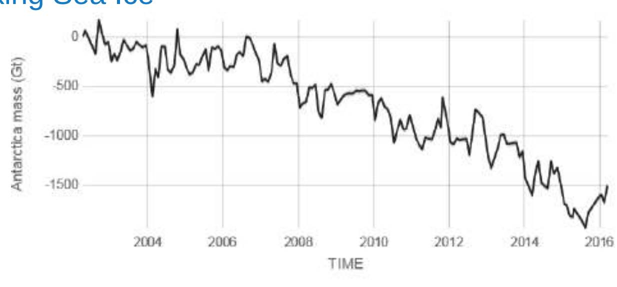
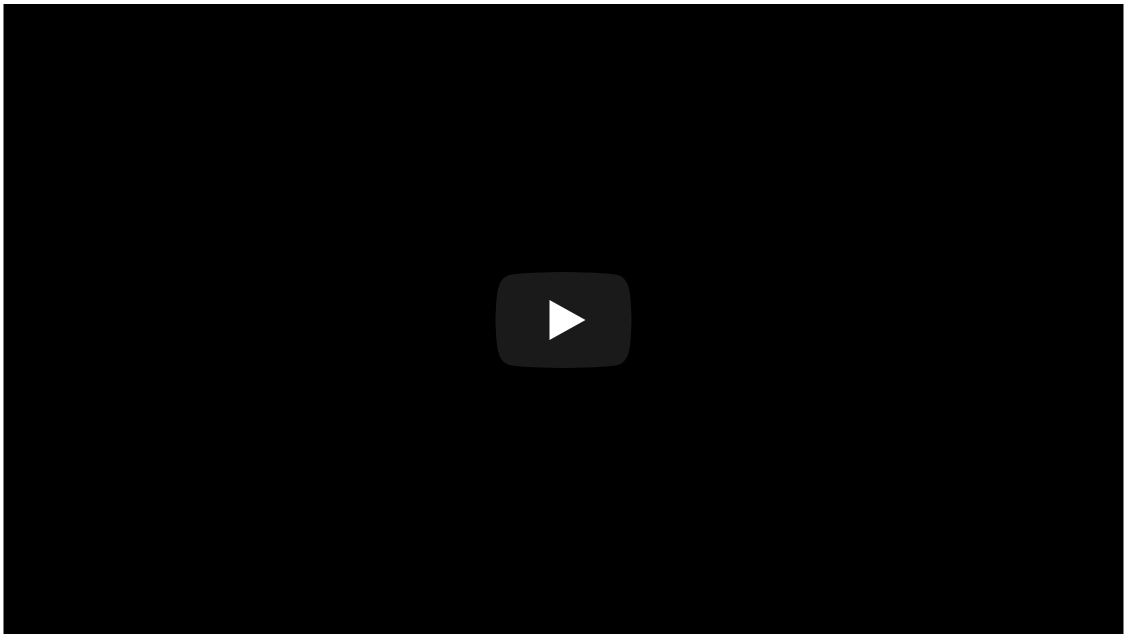
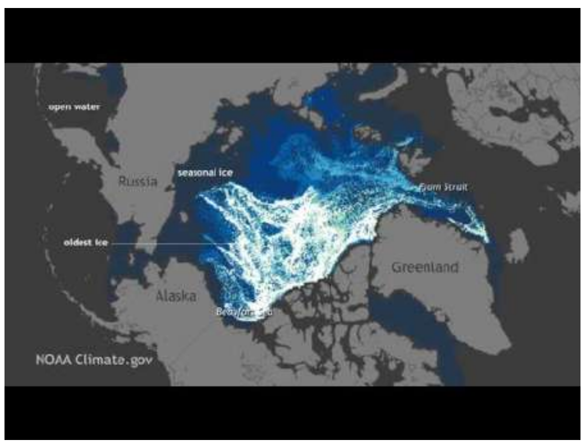


Figure 15.3.1: Decline of Antarctic ice mass from 2002 to 2016

Glaciers are ice on top of the land. Alpine glaciers, ice sheets, and sea ice are all melting. Explore melting glaciers at NASA's interactive Global Ice Viewer). Satellites have recorded that Antarctica is melting at 118 gigatons per year and Greenland is melting at 281 gigatons per year (1 gigaton is over 2 trillion pounds). Almost all major alpine glaciers are shrinking, deflating, and retreating and the rate of ice mass loss is unprecedented (never observed before) since the 1940's when quality records for most began. Before anthropogenic warming, glacial activity was variable with some retreating and some advancing [17]. The extent of spring snow cover has decreased. In addition, the extent of sea ice is shrinking. Sea ice is ice floating in the ocean (not on land like a glacier). Most sea ice is at the North Pole which is only occupied by the Arctic Ocean and sea ice [3; 6]. Below, the NOAA animation shows how perennial sea ice has declined from 1987 to 2015. The oldest ice is white and the youngest (seasonal) ice is dark blue. The amount of old ice has declined from 20% in 1985 to 3% in 2015.

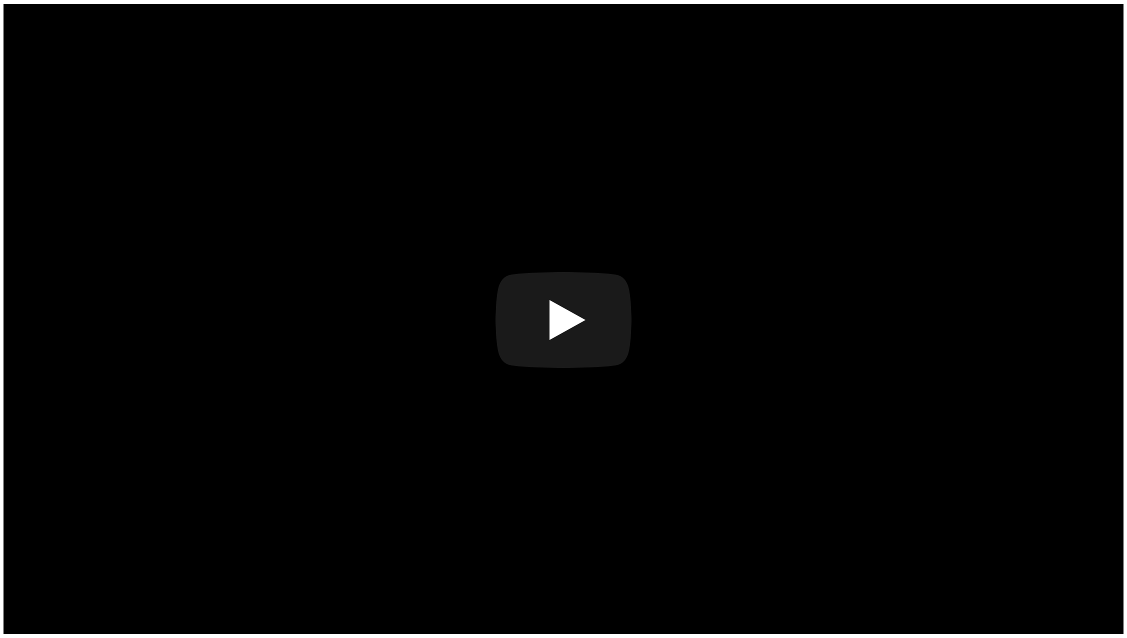




Rising Sea-Level

Sea-level is rising 3.4 millimeters (0.13 inches) per year and has risen 0.19 meters (7.4 inches) from 1901 to 2010. This is thought largely to be from both the melting of glaciers and thermal expansion. Thermal expansion means that as objects such as solids, liquids, and gases heat up, they expand in volume. Since 1970, the melting of glaciers and thermal expansion account for 75% of the sea-level rise [6].

Classic video demonstration (30 seconds) on thermal expansion with brass ball and ring (North Carolina School of Science and Mathematics).



Ocean Acidification

Since 1750, about 40% of the new anthropogenic carbon dioxide has remained in the atmosphere. The remaining 60% gets absorbed by the ocean and vegetation. Therefore, the ocean has absorbed about 30% of new anthropogenic carbon dioxide. When carbon dioxide gets absorbed in the ocean, it creates carbonic acid which makes the ocean more acidic which has an impact on marine organisms that secrete calcium carbonate shells. Recall that hydrochloric acid reacts by effervescing with limestone rock made of calcite, which is calcium carbonate. Ocean acidification associated with climate change has been linked to the thinning of the carbonate walls of some sea snails (pteropods) and small protozoan zooplankton (foraminifera) and declining growth rates of corals [6]. Small animals like protozoan zooplankton are an important component in the marine ecosystem. Acidification combined with warmer temperature and lower oxygen levels is expected to have severe impacts on marine ecosystems and human-used fisheries, possibly affecting our ocean-derived food sources [6].

Video

Extreme Weather Events

Occurrence and intensity of extreme weather events such as hurricanes, precipitation, and heatwaves are increasing [3; 6]. Since the 1980s, hurricanes, which are generated from warm ocean water, have increased in frequency, intensity, and duration and connections to a warmer climate are likely. Since 1910, average precipitation has increased by 10% in the contiguous United States, and much of this increase is associated with heavy precipitation events like storms [18]. However, the distribution is not even and more precipitation is projected for the northern United States while less precipitation is projected for the already dry southwest [3].

Further, heatwaves have increased and rising temperatures are already affecting crop yields in northern latitudes [6]. Increased heat allows for greater moisture capacity in the atmosphere, increasing the potential for more extreme events [19].

Reference

3. Lindsey, R. Climate and Earth's Energy Budget: Feature Articles. (2009). Available at: http://earthobservatory.nasa.gov. (Accessed: 14th September 2016)

6. Pachauri, R. K. *et al. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (IPCC, 2014).

7. Oreskes, N. The scientific consensus on climate change. *Science* **306**, 1686–1686 (2004).

11. Earle, S. *Physical geology OER textbook*. (BC Campus OpenEd, 2015).

12. Callendar, G. S. The artificial production of carbon dioxide and its influence on temperature. *Q.J.R. Meteorol. Soc.* **64**, 223–240 (1938).

13. Hansen, J., Sato, M., Kharecha, P. & others. Earth's energy imbalance and implications. *Atmospheric* (2011).

14. Foster, G. & Rahmstorf, S. Global temperature evolution 1979–2010. *Environ. Res. Lett.* **6**, 044022 (2011).

15. Kosaka, Y. & Xie, S.-P. Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature* **501**, 403–407 (2013).

16. Easterling, D. R. & Wehner, M. F. Is the climate warming or cooling? *Geophys. Res. Lett.* **36**, (2009).

17. Zemp, M. *et al.* Historically unprecedented global glacier decline in the early 21st century. *J. Glaciol.* **61**, 745–762 (2015).

18. Karl, T. R. & Knight, R. W. Secular trends of precipitation amount, frequency, and intensity in the United States. *Bull. Am. Meteorol. Soc.* **79**, 231–241 (1998).

19. Santer, B. D. *et al.* Identification of human-induced changes in atmospheric moisture content. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 15248–15253 (2007).

Contributions and Attributions

15.4: Prehistoric Climate Change

Over Earth history, the climate has changed a lot. For example, during the Mesozoic Era, the Age of Dinosaurs, the climate was much warmer and carbon dioxide was abundant in the atmosphere. However, throughout the Cenozoic Era (65 Million years ago to today), the climate has been gradually cooling. This section summarizes some of these major past climate changes.

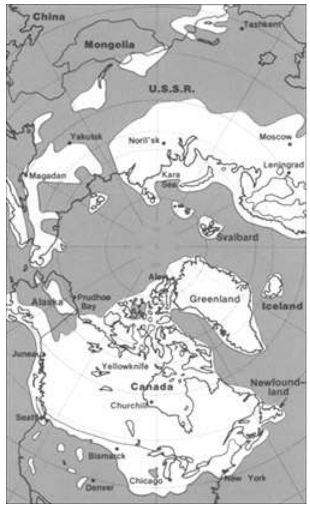


Figure 15.4.1: Maximum extent of Laurentide Ice Sheet

Past Glaciations

Through geologic history, the climate has changed slowly over millions of years. Before the most recent Pliocene-Quaternary glaciation, there were three other major glaciations [20]. The oldest, known as the Huronian, occurred toward the end of the Archean-early Proterozoic (~2.5 billion years ago). The major event of that time, the great oxygenation event (Chapter 8), is most commonly associated with the cause of that glaciation. The increased oxygen is thought to have reacted with the potent greenhouse gas methane, causing cooling [21].

The end of the Proterozoic (about 700 million years ago) had another glaciation, known as the **Snowball Earth hypothesis** [22]. Glacial evidence has been interpreted in widespread rock sequences globally and even has been linked to low-latitude glaciation [23]. Limestone rock (usually formed in tropical marine environments) and glacial deposits (usually formed in cold climates) are often found together from this time in regions all around the world. In Utah, Antelope Island in the Great Salt Lake has interbedded limestone and glacial deposits (diamictites) interpreted to be formed by continental glaciation [24]. The idea of the controversial Snowball Earth hypothesis is that a runaway albedo effect (ice and snow reflecting solar radiation) might cause the complete freezing of land and ocean surfaces and a collapse of biological activity. The ice-covered earth would only melt when carbon dioxide from volcanoes reached high concentrations, due to the inability for carbon dioxide to enter the then-frozen ocean. Some studies estimated carbon dioxide was 350 times higher than today's concentrations [22]. The complete freezing [25] and the extent of the freezing [26] has come into question.

Glaciation also occurred in the Paleozoic, most notably with the Karoo Glaciation of the Pennsylvanian (323 to 300 million years ago). This also was caused by an increase of oxygen and a subsequent drop in carbon dioxide, most likely produced by the evolution and rise of land plants [27].

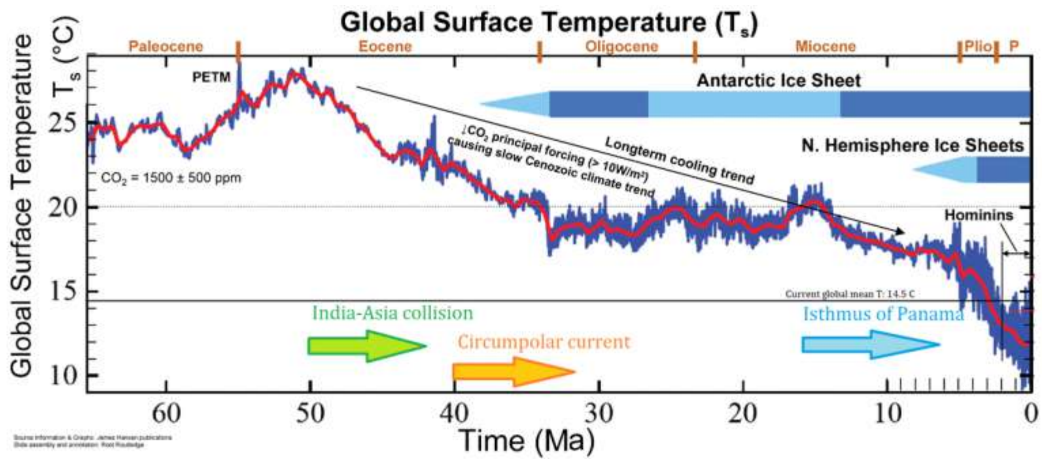


Figure 15.4.2: Global average surface temperature over the past 70 million years.

During the Cenozoic Era (the last 65 million years), the climate started out warm and gradually cooled to today. This warm time is called the **Paleocene-Eocene Thermal Maximum** and Antarctica and Greenland were ice-free during this time. Since the Eocene, tectonic events during the Cenozoic caused persistent and significant planetary cooling. For example, the collision of the Indian Plate with the Asian Plate created the Himalaya Mountains increasing weathering and erosion rates. An increased rate of weathering of silicate minerals, especially feldspar, consumes carbon dioxide from the atmosphere and therefore reduces the greenhouse effect, resulting in long-term cooling [28].

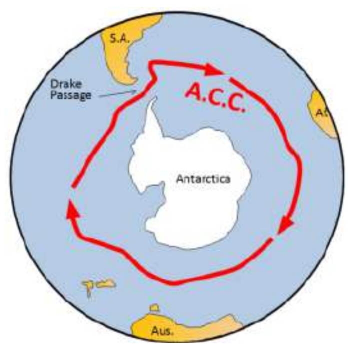


Figure 15.4.1: The Antarctic Circumpolar Current

At about 40 Ma, the narrow gap between the South American Plate and the Antarctica Plate widened, resulting in the opening of the Drake Passage. This allowed for the unrestricted west-to-east flow of water around Antarctica, the Antarctic Circumpolar Current, which effectively isolated the southern ocean from the warmer waters of the Pacific, Atlantic, and Indian Oceans. The region cooled significantly, and by 35-million-year ago (Oligocene) glaciers had started to form on Antarctica [29].

At around 15 Ma, subduction-related volcanism between Central and South America created the Isthmus of Panama that connected North and South America. This prevented water from flowing between the Pacific and Atlantic Oceans and reduced heat transfer from the tropics to the poles. This created a cooler Antarctica and larger Antarctic glaciers. The expansion of that ice sheet (on land and water) increased Earth’s reflectivity (albedo), a positive feedback loop of further cooling: more reflective glacial ice, more cooling, more ice, and so on [30; 31].

By 5 million years ago (Pliocene Epoch), ice sheets had started to grow in North America and northern Europe. The most intense part of the current glaciation is the last 1 million years of the Pleistocene Epoch. The Pleistocene has significant temperature variations (through a range of almost 10°C) on time scales of 40,000 to 100,000 years, and corresponding expansion and contraction of ice sheets. These variations are attributed to subtle changes in Earth’s orbital parameters called **Milankovitch cycles** [32; 33], which are explained in more detail in the chapter on glaciers. Over the past million years, the glaciation cycles have been approximately every 100,000 years [34] with many glacial advances in the last 2 million years (Lisiecki and Raymo, 2005) [35].

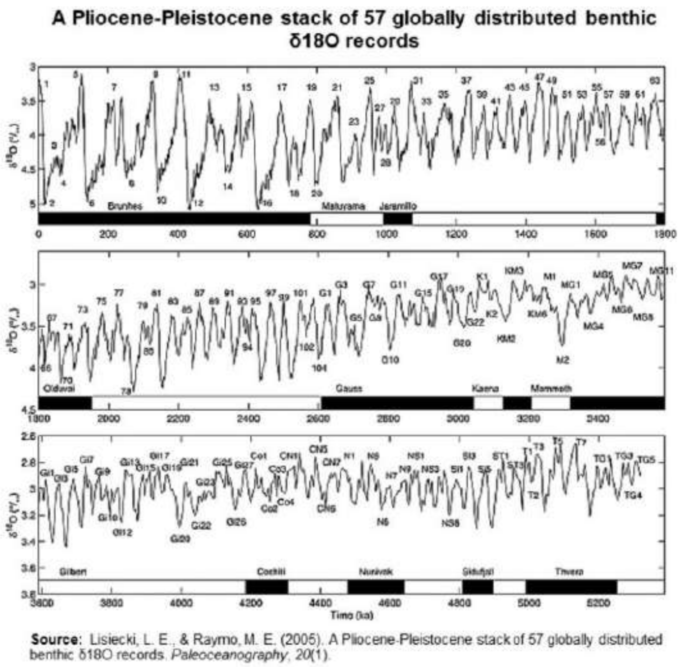


Figure 15.4.3: A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. X-axis is time in thousands of years (ka) so 200 is actually 200,000. (Source: Lisiecki and Raymo, 2005)

Warmer portions of climate within an ice age are called **interglacials**, with brief versions called **interstadials**. These warming upticks are related to variations in Earth's climate like Milankovitch cycles. In the last 500,000 years, there have been 5 or 6 interglacials, with the most recent belonging to our current time, the Holocene.

Two of the more recent climate swings demonstrate the complexity of the changes: the Younger Dryas and the Holocene Climatic Optimum. These events are more recent and yet have conflicting information. The Younger Dryas cooling is widely recognized in the Northern Hemisphere [36], though the timing of the event (about 12,000 years ago) does not appear to be equal everywhere [37]. It also is difficult to find in the Southern Hemisphere [38]. The Holocene Climatic Optimum is the warming around 6,000 years ago [39], though it was not universally warmer, and probably not as warm as current warming [40], and not at the same time everywhere [41].

Proxy Indicators of Past Climates

How do we know about past climates? Geologists use proxy indicators to understand past climate. A **proxy indicator** is a biological, chemical, or physical signature preserved in the rock, sediment, or ice record that acts like a "fingerprint" of something in the past [42]. Thus they are an indirect indicator of something like climate. For ancient glaciations from the Proterozoic and Paleozoic, there are rock formations of glacial sediments such as the diamictite (or tillite) of the Mineral Fork Formation in Utah. This dark rock has many fine-grained components plus some large out-sized clasts like a modern glacial till [43; 44].

For climate changes during the Cenozoic Era (the last 65 Ma), there is a detailed chemical record from the coring of deep-sea sediments as part of the Ocean Drilling Program. Studies of deep-sea sediment use stable carbon and oxygen isotopes obtained from the shells of deep-sea benthic foraminifera that have settled on the ocean floor over millions of years. Oxygen isotopes are a proxy indicator of deep-sea temperatures and continental ice volume [45].

Sediment Cores – Stable Oxygen Isotope

Figure 15.4.4: Sediment core from the Greenland continental slope (Source: Hannes Grobe)

Oxygen isotopes are an indicator of past climate. The two main stable oxygen isotopes are ^{16}O and ^{18}O. They both occur in water (H_2O) and in the calcium carbonate ($CaCO_3$) shells of foraminifera as the oxygen component of both of those molecules. The most abundant and lighter isotope is ^{16}O. Since it is lighter, it evaporates more easily from the ocean's surface as water vapor, which later turns to clouds and precipitation on the ocean and land.

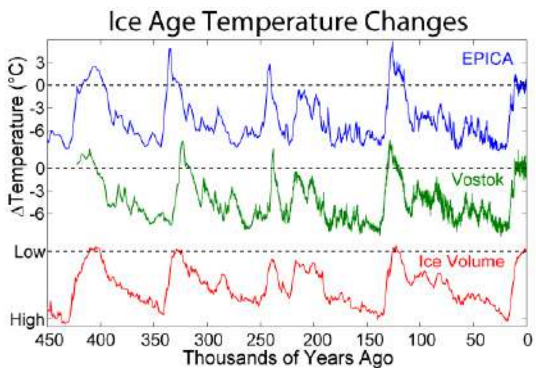


Figure 15.4.5: Antarctic temperature changes during the last few glaciations compared to global ice volume. The first two curves are based on the deuterium (heavy hydrogen) record from ice cores (EPICA Community Members 2004, Petit et al. 1999). The bottom line is ice volume based on oxygen isotopes from a composite of deep-sea sediment cores (Lisiecki and Raymo 2005).

During geologic times when the climate is cooler, more of this precipitation is locked onto land in the form of glacial ice. Consider the giant ice sheets, more than a mile thick, that covered a large part of North America during the last ice age only 14,000 years ago. During glaciation, the glaciers effectively lock away more ^{16}O, thus the ocean water and foraminifera shells become enriched in ^{18}O. Therefore, a ratio of ^{18}O to ^{16}O in calcium carbonate shells of foraminifera is an indicator of past climate. The sediment cores from the Ocean Drilling Program record a continuous accumulation of sediment.

Sediment Cores – Boron Isotopes and Acidity

Boron-isotope ratios in ancient planktonic foraminifera shells in deep-sea sediment cores have been used to estimate the pH (acidity) of the ocean over the past 60 million years. Ocean acidity is a proxy for past atmospheric CO^2 concentrations. In the early Cenozoic, around 60 million years ago, CO^2 concentrations were over 2,000 ppm and started falling around 55 to 40 million years ago possibly due to reduced CO^2 outgassing from ocean ridges, volcanoes, and metamorphic belts and increased carbon burial due to uplift of the Himalaya Mountains. By the Miocene (about 24 million years ago), CO^2 levels were below 500 ppm [46] and by 800,000 years ago CO^2 levels didn't exceed 300 ppm [47].

Carbon Dioxide Concentrations in Ice Cores

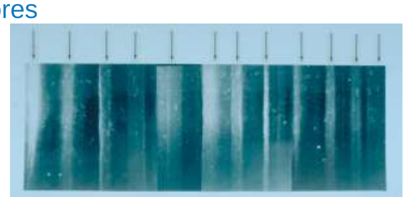


Figure 15.4.6: 19 cm long section of ice core showing 11 annual layers with summer layers (arrowed) sandwiched between darker winter layers. (Source: US Army Corps of Engineers)

For the more recent Pleistocene climate, there is a more detailed and direct chemical record from coring into the Antarctic and Greenland ice sheets. Snow accumulates on these ice sheets and creates yearly layers. Ice cores have been extracted from ice sheets covering the last 800,000 years. Oxygen isotopes are collected from these annual layers and the ratio of ^{18}O to ^{16}O is used to determine temperature as discussed above. In addition, the ice traps small atmospheric gas bubbles as the snow turns to ice.

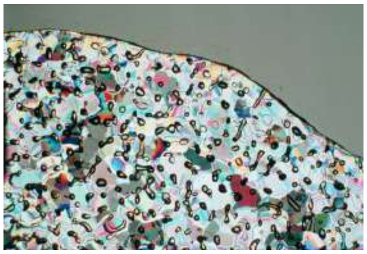


Figure 15.4.7: Antarctic ice showing hundreds of tiny trapped air bubbles from the atmosphere thousands of years ago. (Source: CSIRO)

Small pieces of this ice are crushed and the ancient air extracted into a mass spectrometer that can detect the chemistry of the ancient atmosphere. Carbon dioxide levels are recreated from these measurements. Over the last 800,000 years, the maximum carbon dioxide concentration during warm times was about 300 ppm and the minimum during cold stretches was about 170 ppm [46; 47; 48]. The carbon dioxide content of earth's atmosphere is currently over 400 ppm.

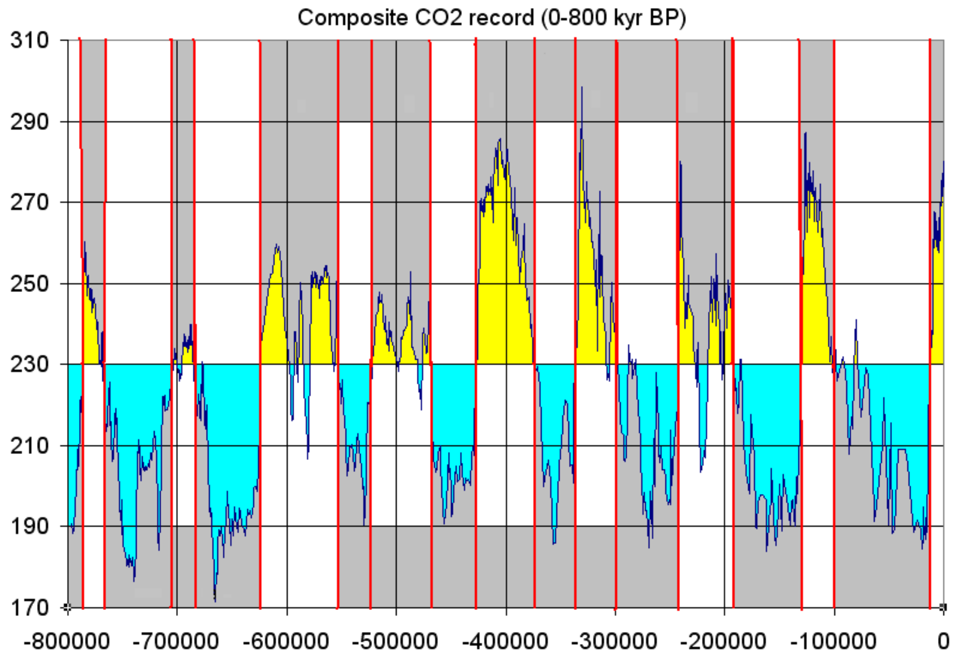
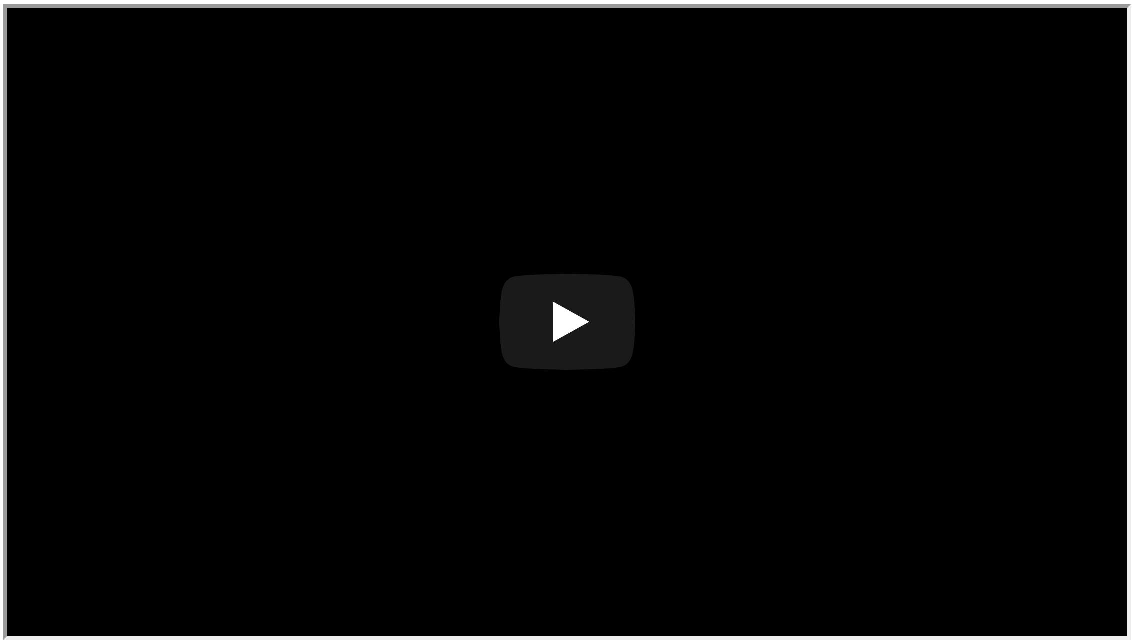


Figure 15.4.8: Composite carbon dioxide record from the last 800,000 years based on ice core data from EPICA Dome C Ice Core.

Oceanic Microfossils

Microfossils, like foraminifera, diatoms, and radiolarians, can be used to interpret past climate records. In sediment cores, different species of microfossils are found in different layers. Groups of these microfossils are called assemblages. One assemblage consists of species that lived in cooler ocean water (in glacial times) and another assemblage found at a different level in the same sediment core is made of warmer water species [49].



Tree Rings

Figure 15.4.9: Tree rings form every year. Rings that are farther apart are from wetter years and rings that are closer together are from dryer years.

Every year a tree will grow one ring with a light section and dark section. The rings vary in width. Since trees need a lot of water to survive, narrower rings indicate colder and drier climates. Since some trees can be several thousands years old, we can use their rings for regional paleoclimatic reconstructions. Further, dead trees such as those used in Puebloan ruins can be used to extend this proxy indicator, which showed long term droughts in the region and why their villages were abandoned.

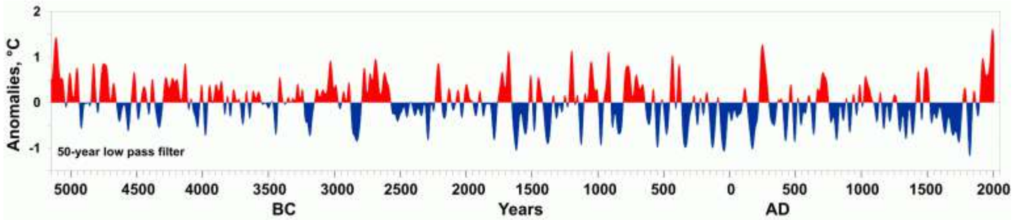


Figure 15.4.10: Summer temperature anomalies for the past 7000 years (Source: R.M.Hantemirov)

Pollen

Flowering plants produce pollen grains. Pollen is distinctive when viewed under a microscope. Sometimes pollen can be preserved in lake sediments that accumulate every year. Coring of lake sediments can reveal ancient pollen. Fossil pollen assemblages are groups of pollen from multiple species such as spruce, pine, and oak. Through time (via the sediment cores and radiometric age-dating techniques), the pollen assemblage will change revealing the plants that lived in the area at the time. Thus the pollen assemblages are an indicator of past climate since different plants will prefer different climates [50]. For example, in the Pacific Northwest east of the Cascades, a region close to the border of grasslands and forest, a study tracked pollen over the last 125,000 years covering the last two glaciations. As shown in the figure (Fig. 2 from reference Whitlock and Bartlein 1997 [51]), pollen assemblages with more pine tree pollen are found during glaciations and pollen assemblages with less pine tree pollen are found during interglacial times [51].

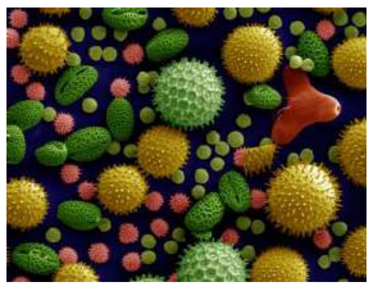


Figure 15.4.12: Scanning electron microscope image of modern pollen with false color added to distinguish plant species. (Source: Dartmouth Electron Microscope Facility, Dartmouth College)

Other Proxy Indicators

Paleoclimatologists study many other phenomena to understand past climates such as human historical accounts, human instrument record from the recent past, lake sediments, cave deposits, and corals.

References

20. Deynoux, M., Miller, J. M. G. & Domack, E. W. *Earth's Glacial Record*. (Cambridge University Press, 2004).

21. Kopp, R. E., Kirschvink, J. L., Hilburn, I. A. & Nash, C. Z. The Paleoproterozoic snowball Earth: a climate disaster triggered by the evolution of oxygenic photosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 11131–11136 (2005).

22. Hoffman, P. F., Kaufman, A. J., Halverson, G. P. & Schrag, D. P. A neoproterozoic snowball earth. *Science* **281**, 1342–1346 (1998).

23. Schopf, J. W. & Klein, C. Late Proterozoic Low-Latitude Global Glaciation: the Snowball Earth. in *The Proterozoic biosphere : a multidisciplinary study* (eds. Schopf, J. W. & Klein, C.) 51–52 (Cambridge University Press, 1992).

24. Doelling, H. H. *et al.* Geology of Antelope Island. *Davis County, Utah: Utah Geological and Mineral Survey Open-File Release* **144**, 82 (1988).

25. Allen, P. A. & Etienne, J. L. Sedimentary challenge to Snowball Earth. *Nat. Geosci.* **1**, 817–825 (2008).

26. Eyles, N. & Januszczak, N. 'Zipper-rift': a tectonic model for Neoproterozoic glaciations during the breakup of Rodinia after 750 Ma. *Earth-Sci. Rev.* (2004).

27. Berner, R. A. The carbon cycle and carbon dioxide over Phanerozoic time: the role of land plants. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **353**, 75–82 (1998).

28. Kump, L. R., Brantley, S. L. & Arthur, M. A. Chemical Weathering, Atmospheric CO2, and Climate. *Annu. Rev. Earth Planet. Sci.* **28**, 611–667 (2000).

29. Lagabrielle, Y., Goddéris, Y., Donnadieu, Y., Malavieille, J. & Suarez, M. The tectonic history of Drake Passage and its possible impacts on global climate. *Earth Planet. Sci. Lett.* **279**, 197–211 (2009).

30. Kuipers Munneke, P. *et al.* A new albedo parameterization for use in climate models over the Antarctic ice sheet. *J. Geophys. Res.* **116**, D05114 (2011).

31. Curry, J. A., Schramm, J. L. & Ebert, E. E. Sea Ice-Albedo Climate Feedback Mechanism. *J. Clim.* **8**, 240–247 (1995).

32. Milankovitch, M. Mathematische klimalehre und astronomische theorie der klimaschwankungen. (1930).

33. Roe, G. In defense of Milankovitch. *Geophys. Res. Lett.* **33**, L24703 (2006).

34. Abe-Ouchi, A. *et al.* Insolation-driven 100,000-year glacial cycles and hysteresis of ice-sheet volume. *Nature* **500**, 190–193 (2013).

35. Lisiecki, L. E. & Raymo, M. E. A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. *Paleoceanography* **20**, (2005).

36. Carlson, A. E. The younger Dryas climate event. (2013).

37. Fairbanks, R. G. The age and origin of the “Younger Dryas climate event” in Greenland ice cores. *Paleoceanography* **5**, 937–948 (1990).

38. Kaplan, M. R. *et al.* Glacier retreat in New Zealand during the Younger Dryas stadial. *Nature* **467**, 194–197 (2010).

39. Ganopolski, A., Kubatzki, C., Claussen, M., Brovkin, V., V. & Petoukhov, V., V. The influence of vegetation-atmosphere-ocean interaction on climate during the mid-holocene. *Science* **280**, 1916–1919 (1998).

40. Hewitt, C. D. & Mitchell, J. F. B. A fully coupled GCM simulation of the climate of the mid-Holocene. *Geophys. Res. Lett.* **25**, 361–364 (1998).

41. He, Y. *et al.* Asynchronous Holocene climatic change across China. *Quat. Res.* **61**, 52–63 (2004).

42. Weissert, H. Deciphering methane’s fingerprint. *Nature* **406**, 356–357 (2000).

43. Ojakangas, R. W. & Matsch, C. L. Upper Precambrian (Eocambrian) Mineral Fork Tillite of Utah: A continental glacial and glaciomarine sequence. *GSA Bulletin* **91**, 495–501 (1980).

44. Christie-Blick, N., Ojakangas, R. W. & Matsch, C. L. Upper Precambrian (Eocambrian) Mineral Fork Tillite of Utah: A Continental Glacial and Glaciomarine Sequence: Discussion and Reply. *The Geological Society of America Bulletin* **93**, 184–187 (1982).

45. Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**, 686–693 (2001).

46. Pearson, P. N. & Palmer, M. R. Atmospheric carbon dioxide concentrations over the past 60 million years. *Nature* **406**, 695–699 (2000).

47. Lüthi, D. *et al.* High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature* **453**, 379–382 (2008).

48. Oak Ridge National Laboratory. 800,000-year Ice-Core Records of Atmospheric Carbon Dioxide (CO2). (2008). Available at: http://cdiac.ornl.gov/trends/co2/ice_core_co2.html. (Accessed: 14th September 2016)

49. Cunningham, W. L., Leventer, A., Andrews, J. T., Jennings, A. E. & Licht, K. J. Late Pleistocene–Holocene marine conditions in the Ross Sea, Antarctica: evidence from the diatom record. *The Holocene* **9**, 129–139 (1999).

50. Webb, T. & Thompson, W. Is vegetation in equilibrium with climate? How to interpret late-Quaternary pollen data. *Vegetatio* **67**, 75–91 (1986).

51. Whitlock, C. & Bartlein, P. J. Vegetation and climate change in northwest America during the past 125 kyr. *Nature* **388**, 57–61 (1997).

Contributions and Attributions

15.5: Anthropogenic Causes of Climate Change

As shown in the previous section, prehistoric changes in climate have been very slow. Climate changes typically occur slowly over many millions of years. The climate changes observed today are rapid and largely human-caused. Evidence shows that climate is changing, but what is causing that change? Scientists have suspected since the late 1800s that human-produced (anthropogenic) changes in atmospheric greenhouse gases would likely cause climate change, as changes in these gases have been the case every time in the geologic past. By the middle 1900s, systematic measurements began which confirmed that human-produced carbon dioxide was accumulating in the atmosphere and other earth systems, like forests and the oceans. By the end of the 1900s and into the early 2000's the Theory of Anthropogenic Climate Change was solidified as evidence from thousands of ground-based studies and continuous satellite measurements of land and ocean mounted in number revealing the expected temperature increase. Theories evolve and transform as new data and new techniques become available, but they represent the state of thinking for that field. The Theory of Anthropogenic Climate Change is that humans are causing most of the current changes to climate by burning fossil fuels such as coal, oil, and natural gas. This section summarizes the scientific understanding of anthropogenic climate change.

Scientific Consensus

The overwhelming majority of climate studies indicate that human activity is causing rapid changes to the climate, which will cause severe environmental damage. There is a strong scientific consensus on the issue. Studies published in peer-reviewed scientific journals show that 97 percent of climate scientists agree that climate warming is from human activities [52]. There is no alternative explanation for the observed link between human-produced greenhouse gas emissions and the changing modern climate. Most leading scientific organizations endorse this position, including the U.S. National Academy of Science which was established in 1863 by an act of Congress under President Lincoln. Congress charged the National Academy of Science "with providing independent, objective advice to the nation on matters related to science and technology" [53]. Therefore, the National Academy of Science is the leading authority when it comes to policy advice related to scientific issues.

One way we know that the increased greenhouse gas emissions are from human activities is with isotopic fingerprints. For example, fossil fuels have a ratio of stable carbon-13 to carbon-12 ($^{13}C/^{12}C$) that is different from today's stable carbon ratio in the atmosphere. Studies have been using isotopic carbon signatures to identify anthropogenic carbon in the atmosphere since the 1980s. Isotopic records from the Antarctic Ice Sheet show stable isotopic signature from ~1000 AD to ~1800 AD and a steady isotopic signature gradually changing since 1800 followed by rapid change after 1950. These changes show the atmosphere having a carbon isotopic signature increasingly more similar to that of fossil fuels [42; 54].

Anthropogenic Sources of Greenhouse Gases

Anthropogenic emissions of greenhouse gases have increased since pre-industrial times due to global economic growth and population growth. Atmospheric concentrations of the leading greenhouse gas, carbon dioxide, are at unprecedented levels that haven't been observed in at least the last 800,000 years [6]. The pre-industrial level of carbon dioxide was at about 278 parts per million (ppm). As of 2016, carbon dioxide was, for the first time, above 400 ppm for the entirety of the year. Measurements of atmospheric carbon at the Mauna Loa Carbon Dioxide Observatory show a continuous increase since 1957 when the observatory was established from 315 ppm to over 410 ppm in 2017. The daily reading today can be seen at Daily CO_2. Based on the ice core record over the past 800,000 years, carbon dioxide ranged from about 185 ppm during ice ages to 300 ppm during warm times [52]. View the data-accurate NOAA animation below of carbon dioxide trends over the last 800,000 years.

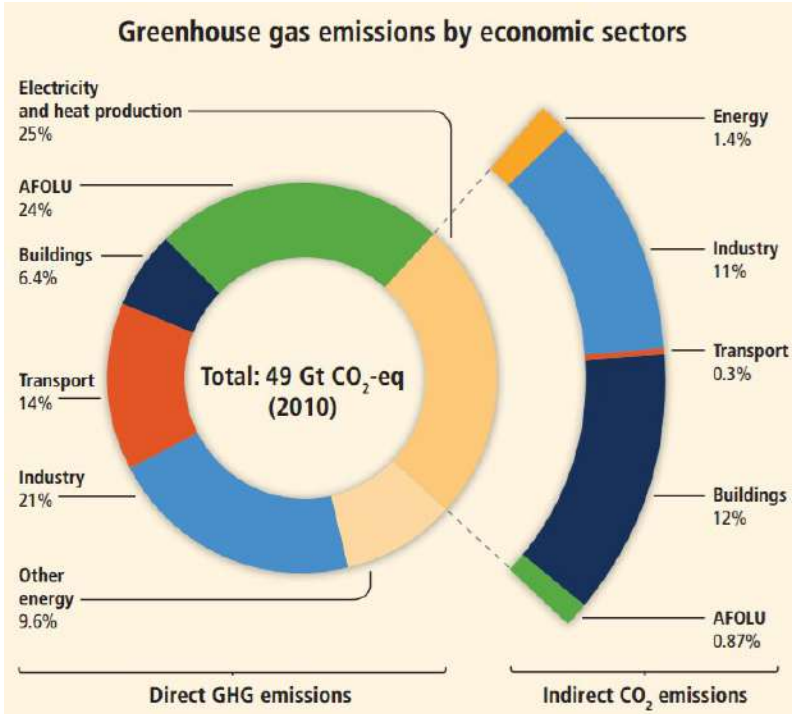


Figure 15.5.1: Total anthropogenic greenhouse gas (GHG) emissions from economic sectors in 2010. The circle shows the shares of direct GHG emissions (in % of total anthropogenic GHG emissions) from five economic sectors in 2010. The pull-out shows how shares of indirect CO2 emissions (in % of total anthropogenic GHG emissions) from electricity and heat production are attributed to sectors of final energy use. AFOLU is agriculture, forestry, and other land use (Source: Pachauri et al. 2014).

What is the source of these anthropogenic greenhouse gas emissions? Fossil fuel combustion and industrial processes contributed 78% of all emissions since 1970. Sectors of the economy responsible for most of this include electricity and heat production (25%); agriculture, forestry, and land use (24%); industry (21%); transportation including automobiles (14%); other energy production (9.6%); and buildings (6.4%) [6]. More than half of greenhouse gas emissions have occurred in the last 40 years (Figure 1.5 p.45 of [6] and 40% of these emissions have stayed in the atmosphere. Unfortunately, despite scientific consensus, efforts to mitigate climate change require political action. Despite the growing amount of climate change concern, mitigation efforts, legislation, and international agreements have reduced emissions in some places, yet the continued economic growth of the less developed world has increased global greenhouse gas emissions. In fact, the time between 2000 and 2010 saw the largest increases since 1970 [6].

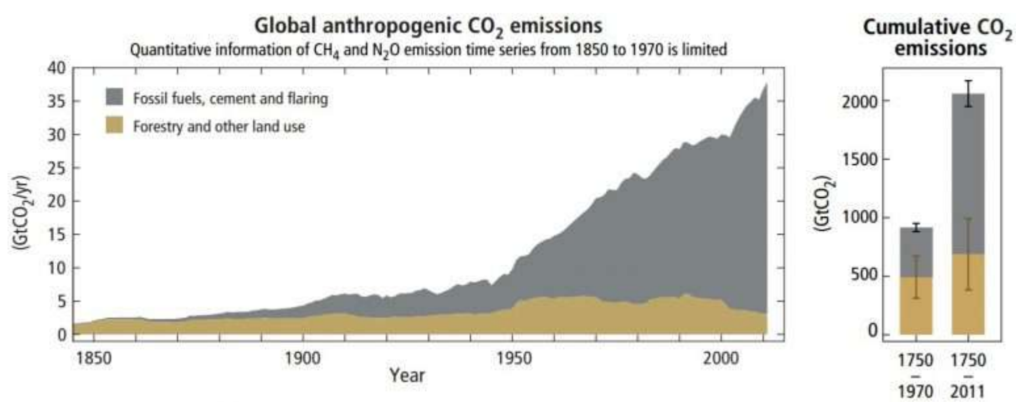


Figure 15.5.1: Annual global anthropogenic carbon dioxide (CO_2) emissions in gigatonne of CO_2-equivalent per year ($GtCO_2$/yr) from fossil fuel combustion, cement production and flaring, and forestry and other land use, 1850–2011. Cumulative emissions and their uncertainties are shown as bars and whiskers.

References

6. Pachauri, R. K. *et al. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (IPCC, 2014).

42. Weissert, H. Deciphering methane's fingerprint. *Nature* **406**, 356–357 (2000).

52. National Oceanic and Atmospheric Administration. ESRL Global Monitoring Division - Global Greenhouse Gas Reference Network. (2005). Available at: http://www.esrl.noaa.gov/gmd/ccgg/trends/history.html. (Accessed: 14th September 2016)

53. National Academy of Sciences. Mission Statement. (2016). Available at: http://www.nasonline.org/about-nas/mission/. (Accessed: 3rd August 2016)

54. Francey, R. J., Allison, C. E., Etheridge, D. M., Trudinger, C. M. & others. A 1000-year high precision record of δ13C in atmospheric CO2. *Tellus B Chem. Phys. Meteorol.* (1999).

Contributions and Attributions

CHAPTER OVERVIEW

16: ENERGY AND MINERAL RESOURCES

An important use of geologic knowledge is locating economically valuable materials for use in society. All items we use can come from only three sources: they can be farmed, hunted or fished, or they can be mined. without mining, modern civilization would not exist. Geologists are essential in the process of mining.

16.1: Prelude to Energy and Mineral Resources

Figure 16.1.1: The Latrobe Gold Nugget, as seen on display in the London Natural History Museum, is 717 grams and displays the rare cubic form of native gold. Most gold, even larger nuggets, grow in confined spaces where the euhedral nature of the mineral is not seen.

This text has discussed pioneers in the scientific study of geology like James Hutton and Charles Lyell, but the first “geologists” were the hominids who picked up stones, beginning the stone age. Maybe stones were first used as curiosity pieces, maybe as weapons, but ultimately, they were used as tools. This was the Paleolithic Period, the beginning of the study of geology and it goes back 2.6 million years ago to east Africa [1].

Figure 16.1.1: A Mode 1 Oldowan tool used for chopping

In modern times, an important use of geologic knowledge is locating economically valuable materials for use in society. All items we use can come from only three sources: they can be farmed, hunted or fished, or they can be mined. At the turn of the Twentieth Century, speculation was rampant that food supplies would not keep pace with world demand, and artificial fertilizers would need to be developed [2]. The ingredients for fertilizers are mined: nitrogen from the atmosphere using the Haber process [3], potassium from the hydrosphere (lakes or oceans) by evaporation, and phosphorus from the lithosphere (minerals like apatite from phosphorite rock, found in Florida, North Carolina, Idaho, Utah, and around the world). Thus, without mining, modern civilization would not exist. Geologists are essential in the process of mining.

References

1. Semaw, S. *et al.* 2.6-Million-year-old stone tools and associated bones from OGS-6 and OGS-7, Gona, Afar, Ethiopia. *J. Hum. Evol.* **45**, 169–177 (2003).

2. Laylin, J. K. *Nobel Laureates in Chemistry, 1901-1992*. (Chemical Heritage Foundation, 1993).

3. Haber, F. The synthesis of ammonia from its elements Nobel Lecture, June 2, 1920. *Reson* **7**, 86–94 (2002).

16.2: Mining

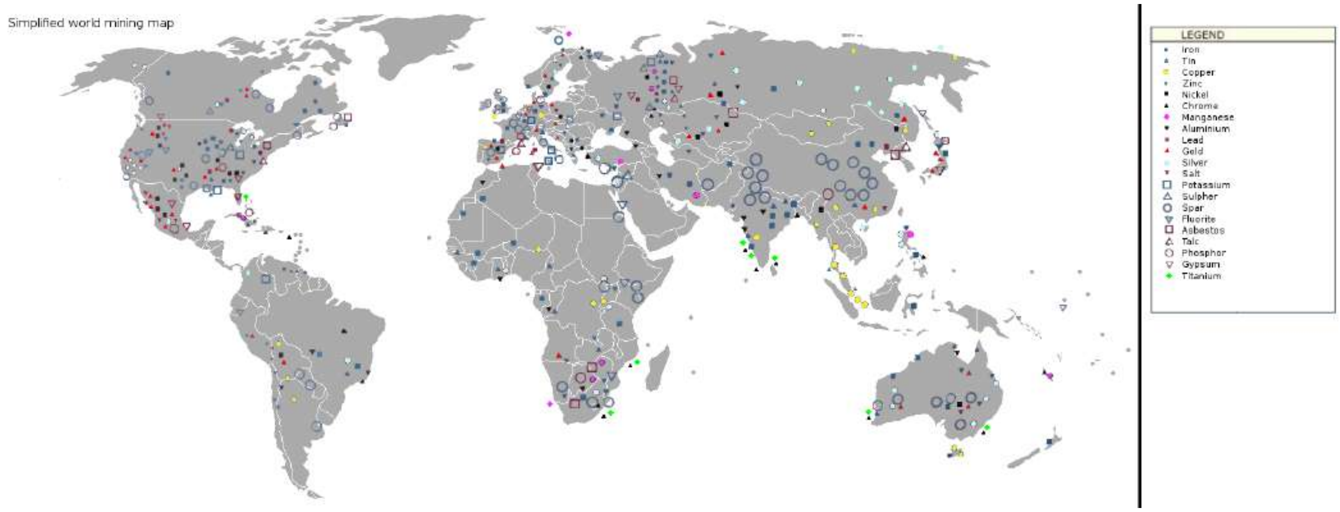


Figure 16.2.1: Map of world mining areas.

Mining is defined as the extraction, from the Earth, of valuable material for societal use. Usually, this includes solid materials (e.g. gold, iron, coal, diamond, sand, and gravel), but can also include fluid resources such as oil and natural gas. Modern mining has a long relationship with modern society. The oldest evidence of mining, with a concentrated area of digging into the Earth for materials, has a history that may go back 40,000 years to the hematite (used as red dye) of the Lion Cave in Swaziland [4]. Resources extracted by mining are generally considered to be nonrenewable.

Renewable vs. Nonrenewable Resources

Resources generally come in two major categories: renewable, which can be reused over and over, or replicate over the course of a short (less than a human life span) time, and nonrenewable, which cannot.

Figure 16.2.1: Hoover Dam provides hydroelectric energy and stores water for southern Nevada.

Renewable resources are items that are present in our environment which can be exploited and replenished. Some of the more common energy sources in this category are linked with green energy resources because they are associated with environmental impacts that are relatively small or easily remediated. Solar energy is the energy that comes from fusion within the Sun, which radiates electromagnetic energy. This energy reaches the Earth both constantly and consistently and should continue to do so for about 5 billion more years [5]. Wind energy is maybe the oldest form of renewable energy, used in sailing ships and windmills. Both solar and wind-generated energy are variable on the Earth's surface. These limitations may be offset through the use of energy storing devices such as batteries or electricity exchanges between producing sites. The heat of the Earth, known as geothermal, can be viable anywhere if drilling goes deep enough. In practice, it is more useful where heat flow is great, such as volcanic zones or regions with thinner crust [6]. Hydroelectric dams provide energy by allowing water to fall through the dam activating turbines that produce energy. Ocean tides can also be a reliable source of energy. All of these types of renewable resources can provide the energy that powers society. Other renewable resources that are not directly energy-related are plant and animal matter, which are used for food, clothing, and other necessities.

Figure 16.2.1: Natural, octahedral shape of diamond.

Nonrenewable resources cannot be replenished at a sustainable rate. They are finite within a human lifetime. Many nonrenewable resources are chiefly a result of planetary, tectonic, or long-term biologic processes, and include items such as gold, lead, copper, diamonds, marble, sand, natural gas, oil, and coal. Most nonrenewable resources are utilized for their concentration of specific elements on the periodic table. For example, if society needs sources of iron (Fe), then it is the exploration geologist who will search for iron-rich deposits that can be economically extracted. Non-renewable resources may be abandoned when other materials become cheaper or serve their purpose better. For example, abundant coal is available in England, but the availability of North Sea oil and natural gas (at a lower cost and lower environmental impact) led to the decrease in coal usage.

Ore

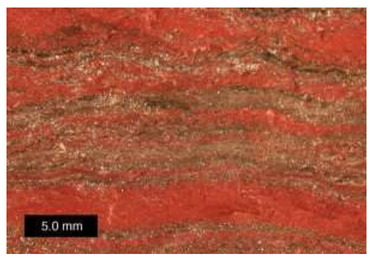


Figure 16.2.1: Banded-iron formations are an important ore of iron (Fe).

The elements of the periodic table are found within the materials that make up the Earth. However, it is rare for the amount of the element to be concentrated to the point where the extraction and processing of the material into usable product becomes profitable. Any place where the amount of valuable material is concentrated is a geologic and geochemical anomaly. If the material can be mined at a profit, the body constitutes an **ore** deposit. Typically, the term ore is used for only metal-bearing minerals, though the concept of ore as a non-renewable resource can be applied to valuable concentrations of fossil fuels, building stones, and other non-metal deposits, even groundwater. The term “natural resource” is more common than ore for these types of materials.

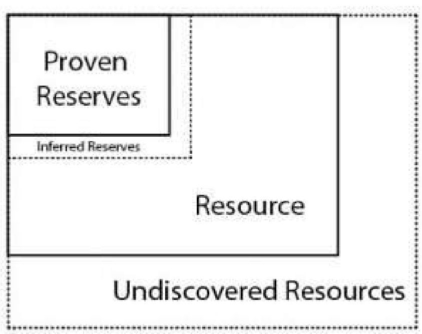


Figure 16.2.1: Diagram illustrating the relative abundance of proven reserves, inferred reserves, resources, and undiscovered resources. (Source: Chris Johnson)

It is implicit that the technology to mine is available, economic conditions are suitable, and political, social and environmental considerations are satisfied in order to classify a natural resource deposit as ore. Depending on the substance, it can be concentrated in a narrow vein or distributed over a large area as a low-concentration ore. Some materials are mined directly from bodies of water (e.g. sylvite for potassium; water through desalination) and the atmosphere (e.g. nitrogen for fertilizers). These differences lead to various methods of mining, and differences in terminology depending on the certainty. **Ore mineral resource** is used for an indication of ore that is potentially extractable, and the term **ore mineral reserve** is used for a well defined (proven), profitable amount of extractable ore.

Cumulative Production	IDENTIFIED RESOURCES			UNDISCOVERED RESOURCES	
	Demonstrated		Inferred	Probability Range (or)	
	Measured	Indicated		Hypothetical	Speculative
ECONOMIC	Reserves		Inferred Reserves		
MARGINALLY ECONOMIC	Marginal Reserves		Inferred Marginal Reserves		
SUB-ECONOMIC	Demonstrated Subeconomic Resources		Inferred Subeconomic Resources		

Other Occurrences	Includes nonconventional and low-grade materials

Figure 16.2.1: McKelvey diagram showing different definitions for different degrees of concentration and understanding of mineral deposits.

Mining Techniques

Figure 16.2.1: Bingham Canyon Mine, Utah. This open-pit mine is the largest man-made removal of rock in the world.

The style of mining is a function of technology, social license, and economics. It is in the best interest of the company extracting the resources to do so in a cost-effective way. Fluid resources, such as oil and gas, are extracted by drilling wells. Over the years, drilling has evolved into a complex discipline in which directional drilling can produce multiple bifurcations and curves originating from a single drill collar at the surface. Using geophysical tools like seismic imaging, resources can be pinpointed and extracted efficiently.

Figure 16.2.1: A surface coal mine in Wyoming.

Solid resources are extracted by two principal methods, of which there are many variants. **Surface mining** is the practice of removing material from the outermost part of the Earth. **Open-pit mining** is used to target shallow, broadly disseminated resources. Typically, the pit progressively deepens through additional mining cuts to extract the ore, and the walls of the pit are as steep as can safely be managed. A steep wall means there is less waste (non-valuable) rock or overburden to remove and is an engineering balance between efficient mining and mass wasting. Occasionally landslides do occur, including a very large landslide that occurred in the Bingham Canyon mine in 2013. These events are costly and dangerous, though careful monitoring gave the Bingham Canyon mine ample warning time. **Strip mining** and **mountaintop mining** are surface mining techniques also used for resources that cover large areas, especially layered resources like coal. In this case, an entire mountaintop or rock layer is removed to gain access to the ore below. The environmental impacts of surface mining are usually greater due to the larger surface disturbance footprint [7].

Figure 16.2.1: Underground mining in Estonia of Oil Shale.

Underground mining is often used for higher-grade, more localized, or very concentrated resources. Some ore minerals are mined underground by introducing chemical agents that dissolve the target mineral followed by solution extraction and subsequent precipitation in a surface operation, but more often a mining shaft/tunnel (or a large network of these shafts and tunnels) is dug to access the material. Whether mining occurs underground or from Earth's surface is dictated by ore deposit depth, geometry, land-use policies, economics, strength of the surrounding rock, and physical access to the ore to be mined. For example, deeper deposits might require removal of too much material, it may be too dangerous or impractical to remove, or it may be too expensive to remove the entire overburden. These factors may prevent materials from being mined from the surface, and cause a project to be mined underground. Also, places where the mining footprint can not be large may force underground mining to occur. The method of mining and whether mining is feasible depends on the price of the commodity and the cost of available technology to remove it and deliver it to market. Thus mines and the towns that support them come and go as the price of the commodity varies. Technological advances and market demands may reopen mines and revive ghost towns.

Concentrating and Refining

Figure 16.2.1: A phosphate smelting operation in Alabama, 1942.

All ore minerals are mixed with less desirable components called **gangue**. The process of physically separating gangue minerals from ore-bearing minerals is called **concentrating**. Separating the desired element from a host mineral by chemical means (including heating in the presence of other minerals) is called **smelting**. Finally, taking a metal such as copper and removing other trace metals such as gold or silver is done through the process of **refining**. Typically, this is done one of three ways:

1. Items can either be mechanically separated and processed based on the unique physical properties of the ore mineral, like recovering placer gold based on its high density
2. Items can also be heated to chemically separate desired components, like refining crude oil into gasoline
3. Items can be smelted, in which controlled chemical reactions unbind metals from the minerals they are contained in, such as when copper is taken out of chalcopyrite ($CuFeS_2$).

Mining, concentrating, smelting and refining processes require enormous amounts of energy. Continual advances in metallurgy and mining practices aim to develop ever more energy-efficient and environmentally benign processes and practices.

References

4. Vogel, J. C. Groningen radiocarbon dates IX. *Radiocarbon* **12**, 444–471 (1970).

5. Schröder, K.-P. & Connon Smith, R. Distant future of the Sun and Earth revisited. *Mon. Not. R. Astron. Soc.* **386**, 155–163 (2008).

6. Duffield, W. A. Volcanoes, geothermal energy, and the environment. *Volcanoes and the Environment* 304 (2005).

7. Palmer, M. A. *et al.* Science and regulation. Mountaintop mining consequences. *Science* **327**, 148–149 (2010).

Contributions and Attributions

16.3: Fossil Fuels

Figure 16.3.1: Coal power plant in Helper, Utah.

Fossils fuels are extractable sources of stored energy created by ancient ecosystems. The natural resources that typically fall under this category are coal, oil (petroleum), and natural gas. This energy was originally formed via photosynthesis by living organisms such as plants, phytoplankton, algae, and cyanobacteria. Sometimes this is known as fossil solar energy since the energy of the sun in the past has been converted into the chemical energy within a fossil fuel. Of course, as the energy is used, just like respiration from photosynthesis that occurs today, carbon can enter the atmosphere, causing climate consequences (see ch. 15). Fossil fuels account for a large portion of the energy used in the world.

Figure 16.3.1: Modern coral reefs and other highly-productive shallow marine environments are thought to be the sources of most petroleum resources.

The conversion of living organisms into hydrocarbon fossil fuels is a complex process. As organisms die, decomposition is hindered, usually due to rapid burial, and the chemical energy within the organisms' tissues is added to surrounding geologic materials. Higher productivity in the ancient environment leads to a higher potential for fossil fuel accumulation, and there is some evidence of higher global biomass and productivity over geologic time [8]. Lack of oxygen and moderate temperatures seem to enhance the preservation of these organic substances [9; 10]. Heat and pressure that is applied after burial also can cause transformation into higher quality materials (brown coal to anthracite, oil to gas) and/or migration of mobile materials [11].

Oil and Gas

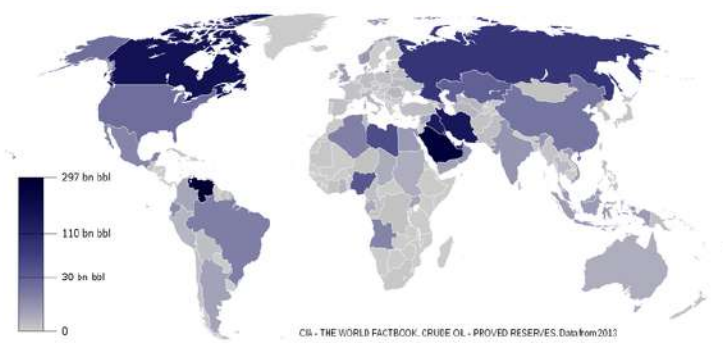


Figure 16.3.1: World Oil Reserves in 2013. Scale in billions of barrels.

Petroleum, with the liquid component commonly called **oil** and gas component called **natural gas** (mostly made up of methane), is principally derived from organic-rich shallow marine sedimentary deposits [12]. As the rock (which is typically shale, mudstone, or limestone) lithifies, the oil and gas leak out of the **source rock** due to the increased pressure and temperature, and migrate to a different rock unit higher in the rock column. Similar to the discussion of good aquifers in chapter 11, if the rock is sandstone, limestone, or other porous and permeable rock, then that rock can act as a **reservoir** for the oil and gas.

Figure 16.3.1: A structural or anticline trap. The red on the image represents pooling petroleum. The green layer would be a non-permeable rock, and the yellow would be a reservoir rock.

A **trap** is a combination of a subsurface geologic structure and an impervious layer that helps block the movement of oil and gas and concentrates it for later human extraction [13; 14]. The development of a trap could be a result of many different geologic situations. Common examples include an anticline or domal structure, an impermeable salt dome, or a fault-bounded stratigraphic block (porous rock next to non-porous rock). The different traps have one thing in common: they pool the fluid fossil fuels into a configuration in which extraction is more likely to be profitable. Oil or gas in strata outside of a trap renders extraction is less viable.

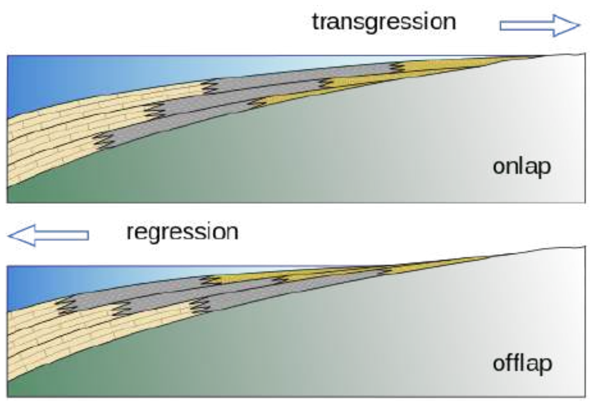


Figure 16.3.1: The rising sea levels of transgressions create onlapping sediments, regressions create offlapping.

A branch of geology that has grown from the desire to understand how changing sea level creates organic-rich shallow marine muds, carbonates, and sands in close proximity to each other are called **sequence stratigraphy** [15]. A typical shoreline environment has beaches next to lagoons next to coral reefs. Layers of beach sands and lagoonal muds and coral reefs accumulate into sediments that form sandstones, good reservoir rocks, next to mudstones next to limestones, both potential source rocks. As sea level either rises or falls, the location of the shoreline changes and the locations of sands, muds, and reefs with it. This places oil and gas producing rocks (like mudstones and limestones) next to oil and gas reservoirs (sandstones and some limestones). Understanding the interplay of lithology and ocean depth can be very important in finding new petroleum resources because using sequence stratigraphy as a model can allow predictions to be made about the locations of source rocks and reservoirs.

Tar Sands

Figure 16.3.1: Tar sandstone from the Miocene Monterrey Formation of California.

Conventional oil and gas (pumped from a reservoir) are not the only way to obtain hydrocarbons. The next few sections are known as unconventional petroleum sources, though, they are becoming more important as conventional sources increase in scarcity. **Tar sands**, or oil sands, are sandstones that contain petroleum products that are highly viscous (like tar), and thus, can not be drilled and pumped out of the ground, unlike conventional oil. The fossil fuel in question is bitumen, which can be pumped as a fluid only at very low rates of recovery and only when heated or mixed with solvents. Thus injections of steam and solvents, or direct mining of the tar sands for later processing can be used to extract the tar from the sands. Alberta, Canada is known to have the largest reserves of tar sands in the world [16].

Note

An energy resource becomes uneconomic once the total cost of extracting it exceeds the revenue which is obtained from the sale of extracted material

Oil Shale

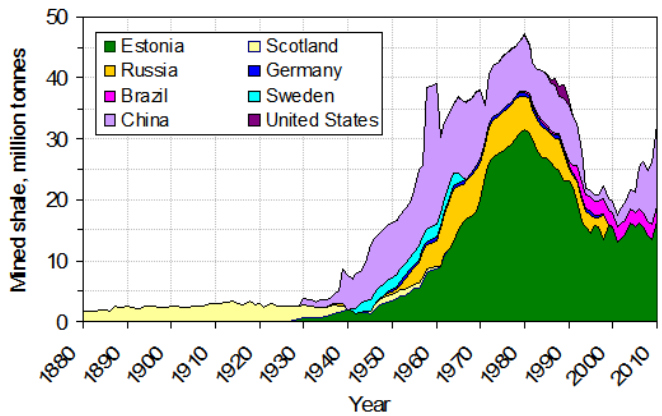


Figure 16.3.1: Global production of Oil Shale, 1880-2010.

Oil shale (or tight oil) is a fine-grained sedimentary rock that has a significant quantity of petroleum or natural gas. Shale is a common source of fossil fuels with high porosity but it has very low permeability. In order to get the oil out, the material has to be mined and heated, which, like with tar sands, is expensive and typically has a negative impact on the environment [17].

Fracking

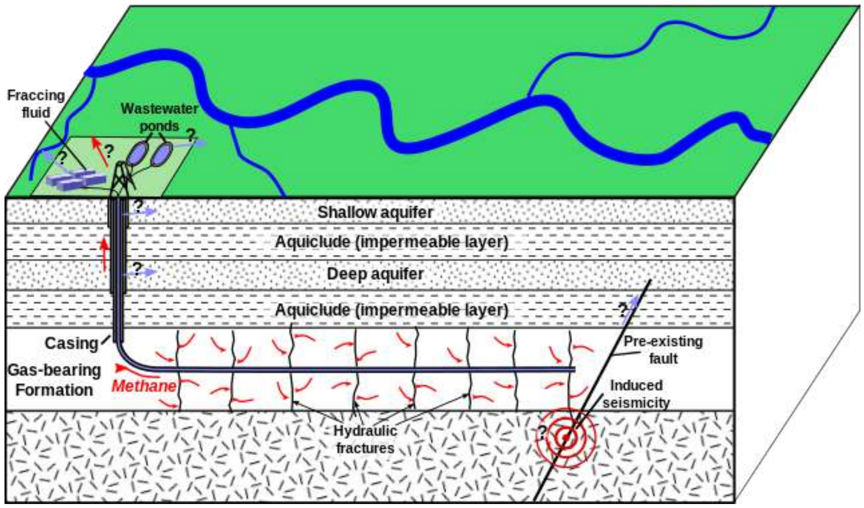


Figure 16.3.1: Schematic diagram of fracking.

Another process which is used to extract the oil and gas from shale and other unconventional tight resources is called **hydraulic fracturing**, better known as **fracking** [18]. In this method, high-pressure injections of water, sand grains, and added chemicals are pumped underground, creating and holding open fractures in the rocks, which aids in the release of the hard-to-access fluids, mostly natural gas. This is more useful in tighter sediments, especially shale, which has a high porosity to store the hydrocarbons but low permeability to transmit the hydrocarbons. Fracking has become controversial due to the potential for groundwater contamination [19] and induced seismicity [20] and represents a balance between public concerns and energy value.

Coal

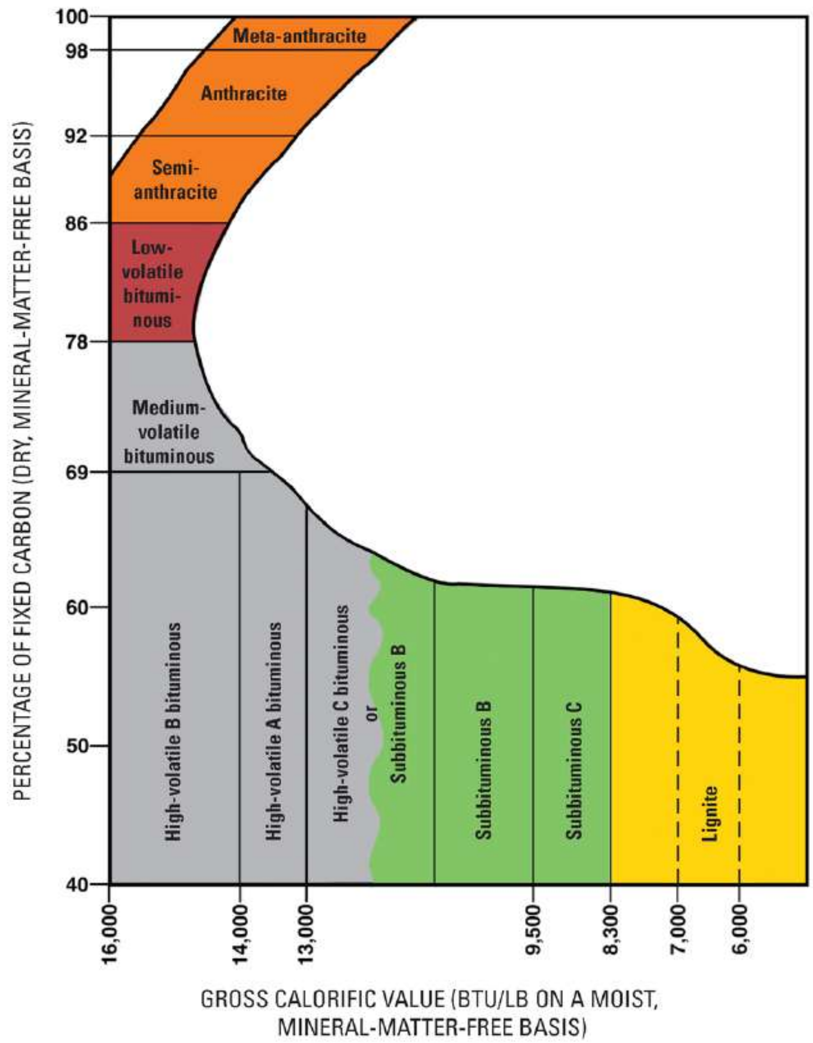


Figure 16.3.1: USGS diagram of different coal rankings.

Coal is the product of fossilized swamps [21], though some older coal deposits that predate terrestrial plants are presumed to come from algal buildups [22]. It is chiefly carbon, hydrogen, nitrogen, sulfur, and oxygen, with minor amounts of other elements [23]. As this plant material is incorporated into sediments, it undergoes a series of changes due to heat and pressure which concentrates fixed carbon, the combustible portion of the coal. In this sense, the more heat and pressure that coal undergoes, the greater is its fuel value and the more desirable is the coal. The general sequence of a swamp turning into the various stages of coal are:

Swamp => Peat => Lignite => Sub-bituminous => Bituminous => Anthracite => Graphite.

As swamp materials collect on the floor of the swamp, they turn to peat. As lithification occurs, peat turns to lignite. With increasing heat and pressure, lignite turns to sub-bituminous coal, bituminous coal, and then, in a process like metamorphism, anthracite. Anthracite is the highest metamorphic grade and most desirable coal since it provides the highest energy output. With even more heat and pressure driving out all the volatiles and leaving pure carbon, anthracite can turn to graphite.

Figure 16.3.1: Anthracite coal, the highest grade of coal.

Coal has been used by humans for at least 6000 years [23], mainly as a fuel source. Coal resources in Wales are often cited as a primary reason for the rise of Britain (and later, the United States) in the Industrial Revolution [24; 25; 26]. According to the US Energy Information Administration, the production of coal in the US has decreased due to cheaper prices of competing for energy sources and recognition of its negative environmental impacts, including increased very fine-grained particulate matter, greenhouse gases [27], acid rain [28], and heavy metal pollution [29]. Seen from this point of view, the coal industry is unlikely to revive.

References

8. Tappan, H. & Loeblich, A. R. Geobiologic Implications of Fossil Phytoplankton Evolution and Time-Space Distribution. *Geological Society of America Special Papers* **127**, 247–340 (1970).

9. Gordon, M., Jr, Tracey, J. I., Jr & Ellis, M. W. *Geology of the Arkansas bauxite region*. (1958).

10. Demaison, G. J. & Moore, G. T. Anoxic environments and oil source bed genesis. *Org. Geochem.* **2**, 9–31 (1980).

11. Tissot, B. Effects on prolific petroleum source rocks and major coal deposits caused by sea-level changes. *Nature* **277**, 463–465 (1979).

12. Pratt, W. E. *Oil in the Earth*. (University of Kansas Press, 1942).

13. Orton, E. *The Trenton Limestone as a Source of Petroleum and Inflammable Gas in Ohio and Indiana*. (U.S. Government Printing Office, 1889).

14. Dott, R. H. & Reynolds, M. J. Sourcebook for petroleum geology. (1969).

15. Vail, P. R. *et al.* Seismic stratigraphy and global sea level changes. *Seismic stratigraphy-applications to hydrocarbon exploration, edited by Payton, CE, Tulsa, American Association of Petroleum Geologists Memoir* **26**, 49–212 (1977).

16. Bauquis, P.-R. What future for extra heavy oil and bitumen: the Orinoco case. in **13**, 18 (1998).

17. Youngquist, W. Shale oil--The elusive energy. *Hubbert Center Newsletter* **4**, (1998).

18. Gandossi, L. An overview of hydraulic fracturing and other formation stimulation technologies for shale gas production. *Eur. Commisison Jt. Res. Cent. Tech. Reports* (2013).

19. Brown, V. J. Industry issues: Putting the heat on gas. *Environ. Health Perspect.* **115**, A76 (2007).

20. Kim, W.-Y. Induced seismicity associated with fluid injection into a deep well in Youngstown, Ohio. *J. Geophys. Res. [Solid Earth]* **118**, 3506–3518 (2013).

21. Taylor, E. L., Taylor, T. N. & Krings, M. *Paleobotany: The biology and evolution of fossil plants*. (Elsevier Science, 2009).

22. Mancuso, J. J. & Seavoy, R. E. Precambrian coal or anthraxolite; a source for graphite in high-grade schists and gneisses. *Econ. Geol.* **76**, 951–954 (1981).

23. Blander, M., Sinha, S., Pelton, A. & Eriksson, G. Calculations of the influence of additives on coal combustion deposits. *Argonne National Laboratory, Lemont, Illinois* 315 (2011).

24. Belloc, H. *The Servile State*. (T.N. Foulis, 1913).

25. McKenzie, H. & Moore, B. Social Origins of Dictatorship and Democracy. (1970).

26. Wrigley, E. A. *Continuity, Chance and Change: The Character of the Industrial Revolution in England*. (Cambridge University Press, 1990).

27. Quéré, C. L. *et al.* The global carbon budget 1959--2011. *Earth System Science Data* **5**, 165–185 (2013).

28. Barrie, L. A. & Hoff, R. M. The oxidation rate and residence time of sulphur dioxide in the Arctic atmosphere. *Atmos. Environ.* **18**, 2711–2722 (1984).

29. Crutzen, P. J. & Lelieveld, J. Human impacts on atmospheric chemistry. *Annu. Rev. Earth Planet. Sci.* **29**, 17–45 (2001).

Contributions and Attributions

16.4: Mineral Resources

Figure 16.4.1: Gold-bearing quartz vein from California.

Mineral resources, while principally nonrenewable, are generally placed in two main categories: **metallic** (containing metals) or **nonmetallic** (containing other useful materials). Most mining is focused on metallic minerals. A significant part of the advancement of human society has been developing the knowledge and technologies that yielded metal from the Earth and allowed the machines, buildings, and monetary systems that dominate our world today. The location and recovery of these metals have been a key facet of the study of geology since its inception. Every element across the periodic table has specific applications in human civilization. Metallic mineral mining is the source of many of these elements.

Types of Metallic Mineral Deposits

The number of ways that minerals and their associated elements concentrate to form ore deposits are too complex and numerous to fully review in this text. However, entire careers are built around them. Some of the more common types of these deposits are described, along with their associated elemental concentrations and world-class occurrences.

Magmatic Processes

Figure 16.4.1: Layered intrusion of dark chromium-bearing minerals, Bushveld Complex, South Africa

Crystallization and differentiation (see chapter 4) of a magmatic body can cause the concentration of certain minerals and elements. **Layered intrusion** (typically ultramafic to mafic) can be host to deposits that contain copper, nickel, platinum-palladium-rhodium, and chromium. The Stillwater Complex in Montana is an example of an economic layered mafic intrusion [30]. Associated deposit types can contain chromium or titanium-vanadium. The largest magmatic deposits in the world are the chromite deposits in the Bushveld Igneous Complex in South Africa [31]. Rocks of the Bushveld Igneous Complex have an areal extent larger than the state of Utah. The chromite occurs in layers, which resemble sedimentary layers, except this occurred within a crystallizing magma chamber.

Figure 16.4.1: This pegmatite from Brazil contains lithium-rich green elbaite (a tourmaline) and purple lepidolite (a mica).

Water and other volatiles that are not incorporated into mineral crystals while a magma crystallizes become concentrated around the margins of these crystallizing magmas. Ions in these hot fluids are very mobile and can form exceptionally large crystals. Once crystallized, masses of these large crystals are called **pegmatites** that form from the concentration of magma fluids near the end of crystallization when nearly the entire magma body has crystallized. In addition to minerals that are predominant in the main igneous mass, such as quartz, feldspar, and mica, pegmatite bodies may also contain very large crystals of unusual minerals that contain rare elements like beryllium, lithium, tantalum, niobium, and tin, as well as native elements like gold [32]. Such pegmatites are ores of these metals.

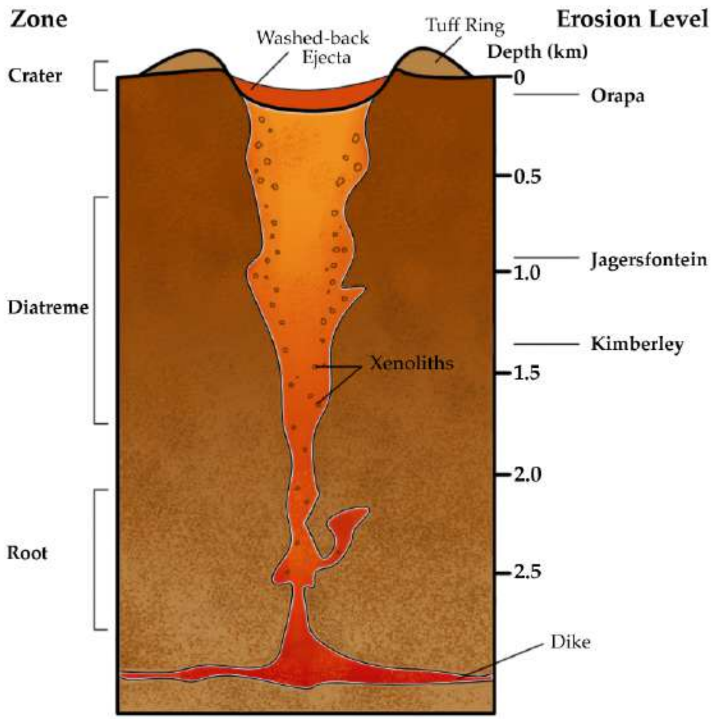


Figure 16.4.1: Schematic diagram of a kimberlite pipe.

An unusual magmatic process is a **kimberlite** pipe, which is a volcanic conduit that transports ultramafic magma from depths in the mantle to the surface. Diamonds, which are formed at great temperature and depth, are transported this way to locations where they can be mined. The process that emplaced these kimberlite (ultramafic) rocks is no longer common on Earth, and most of the known deposits are Archean [33].

Hydrothermal Processes

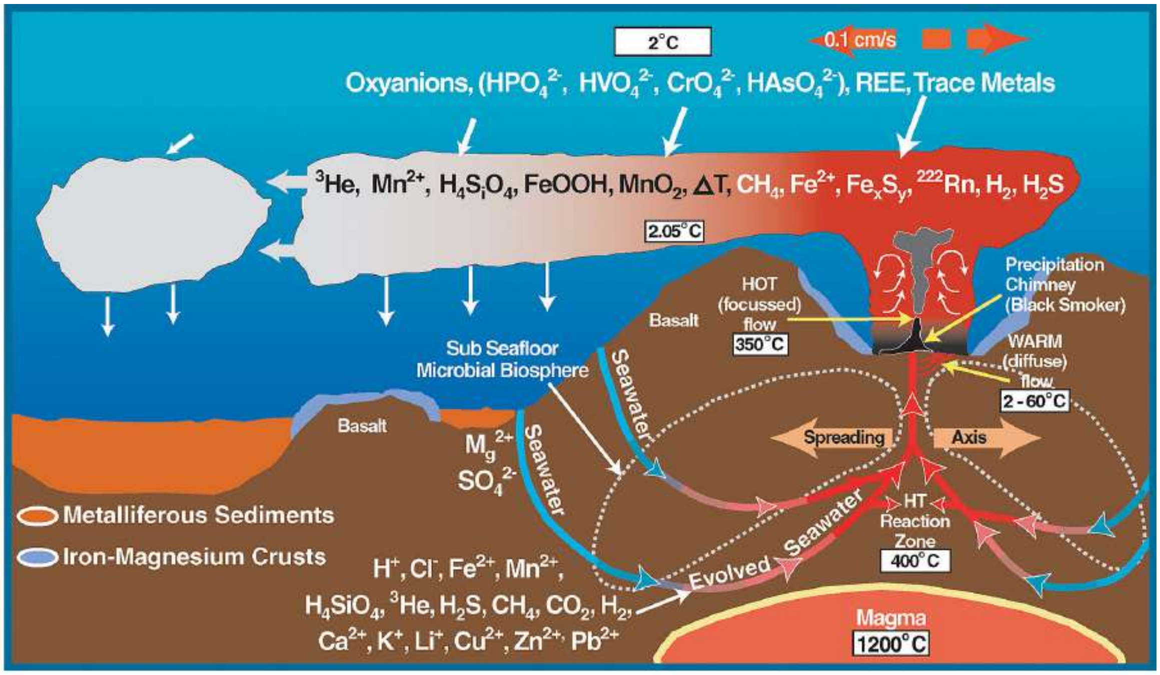


Figure 16.4.1: The complex chemistry around mid-ocean ridges.

Fluids rising from crystallizing magmatic bodies or heated by the geothermal gradient cause a wide range of geochemical reactions that can form a variety of mineral deposits. The most active hydrothermal process today produces **volcanogenic massive sulfide** (VMS) deposits, which form from black smoker activity near mid-ocean ridges all over the world, and commonly contain copper, zinc, lead, gold, and silver when found on the surface [34]. The largest of these deposits occur in Precambrian age rocks. The Jerome deposit in central Arizona is a good example.

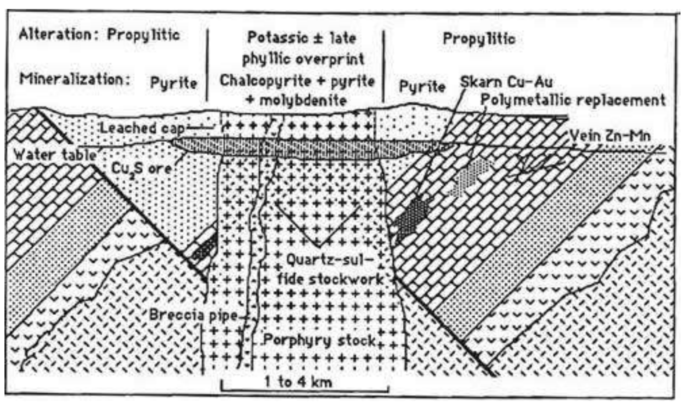


Figure 16.4.1: USGS schematic of a Porphyry copper deposit.

Another type of deposit which draws on heated water from magma is a **porphyry** deposit. This is not to be confused with the igneous texture porphyritic, although the name is derived from the porphyritic texture that is nearly always present in the igneous rocks in a porphyry deposit. Several types of porphyry deposits exist: porphyry copper, porphyry molybdenum, and porphyry tin. They are characterized by the presence of low-grade disseminated ore minerals closely associated with intermediate and felsic intrusive rocks over a very large area [35]. Porphyry deposits are typically the largest mines on Earth. One of the largest, richest, and possibly best-studied mines in the world is Utah's Bingham Canyon open-pit mine, which has had over 100 years of high production of several elements including copper, gold, molybdenum, and silver. Associated underground carbonate replacement deposits have produced lead, zinc, gold, silver, and copper [36]. Past open pit production at this mine was dominated by copper and gold from chalcopyrite and bornite. Gold occurs in minor quantities in the copper-bearing mineral, but the large scale of production makes Bingham Canyon one of the largest gold mines in the U.S. Future production may be more copper and molybdenum (molybdenite) from deeper underground mines.

Figure 16.4.1: The Morenci porphyry is oxidized toward its top (as seen as red rocks in the wall of the mine), creating supergene enrichment.

The majority of porphyry copper deposits owe their economic value to concentration by weathering processes occurring millions of years after the hosting intrusion called **supergene enrichment**. These occur once the hydrothermal event has ceased and the ore body has been uplifted, eroded, and exposed to oxidation [37]. When the upper pyrite-rich portion of the deposit is exposed to rain, pyrite in the oxidizing zone creates an extremely acid condition which dissolves copper out of copper minerals such as chalcopyrite, converting the chalcopyrite to iron oxides like hematite or goethite. The copper is carried downward in the solution until it arrives at the groundwater table and a reducing environment where the copper precipitates, converting primary copper minerals into secondary higher-copper content minerals. Chalcopyrite (35% Cu) is converted to bornite (63% Cu) and ultimately chalcocite (80% Cu). Without this enriched zone (2 to 5 times higher in copper content than the main deposit) most porphyry copper deposits would not be economic.

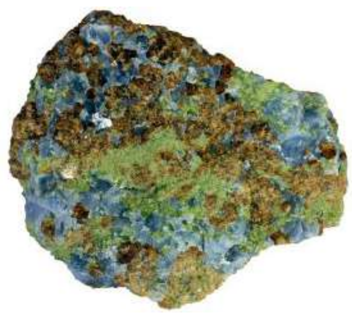


Figure 16.4.1: Garnet-augite skarn from Italy.

If limestone or other calcareous sedimentary rocks are present adjacent to the magmatic body, then another type of ore deposit called a **skarn** deposit can form. These metamorphic rocks form as magma-derived, highly saline metalliferous fluids react with carbonate rocks, creating calcium-magnesium-silicate minerals like pyroxene, amphibole, and garnet, as well as high-grade zones of iron, copper, and zinc minerals and gold [38]. Intrusions that are genetically related to the intrusion that made the Bingham Canyon deposit have also produced copper-gold skarns that were mined by the early European settlers in Utah [39; 40]. Metamorphism of iron and/or sulfide deposits commonly results in an increase in grain size that makes separation of gangue from the desired sulfide or oxide minerals much easier.

Figure 16.4.1: In this rock, a pyrite cube has dissolved (as seen with the negative "corner" impression in the rock), leaving behind small specks of gold.

Sediment-hosted disseminated gold deposits consist of low concentrations of microscopic gold as inclusions and disseminated atoms in pyrite crystals. These are formed via low-level hydrothermal reactions (generally in the realm of diagenesis) that occur in certain rock types, namely muddy carbonates and limey mudstones. This hydrothermal alteration is generally far-removed from a magma source but can be found in extended rocks with a high geothermal gradient. The earliest locally mined deposit of this type was the Mercur deposit in the Oquirrh Mountains of Utah where almost one million ounces of gold were recovered between 1890 and 1917. In the 1960s a metallurgical process using cyanide was developed for these types of low-grade ores. These deposits are also called **Carlin-type** deposits because the disseminated deposit near Carlin, Nevada is where the new technology was first applied and because the first definitive scientific studies were conducted there [41]. Gold was introduced by hydrothermal fluids which reacted with silty calcareous rocks, removing carbonate, creating additional permeability, and adding silica and gold-bearing pyrite in the pore space between grains. The Betze-Post mine and the Gold Quarry mine on the "Carlin Trend" are two of the largest of the disseminated gold deposits in Nevada. Similar deposits, but not as large, have been found in China, Iran, and Macedonia [42].

Non-Magmatic Geochemical Processes

Figure 16.4.1: Underground uranium mine near Moab, Utah.

Geochemical processes that occur at or near the surface without the aid of magma also concentrate metals, but to a lesser degree than hydrothermal processes. One of the main reactions is **redox** (short for reduction/oxidation) chemistry, which has to do with the amount of available oxygen in a system. Places where oxygen is plentiful, as in the atmosphere today, are considered oxidizing environments, while oxygen-poor environments are considered reducing. Uranium deposition is an example of redox mobilization. Uranium is soluble in oxidizing groundwater environments and precipitates as uraninite when reducing conditions are encountered. Many of the deposits across the Colorado Plateau (e.g. Moab, Utah) were formed by this method [43].

Redox reactions were also responsible for the creation of **banded iron formations** (BIFs), which are interbedded layers of iron oxide (hematite and magnetite), chert, and shale beds. These deposits formed early in the Earth's history as the atmosphere was becoming oxygenated. Cyclic oxygenation of iron-rich waters initiated the precipitation of the iron beds. Because BIFs are generally Precambrian in age, they are only found in some of the older exposed rocks in the United States, in the upper peninsula of Michigan and northeastern Minnesota [44].

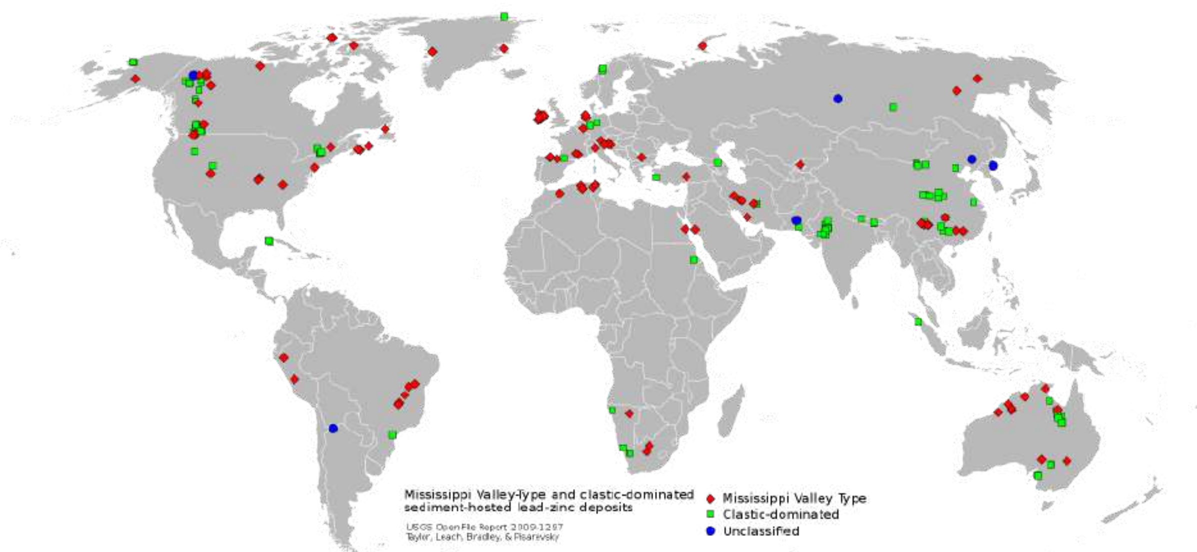


Figure 16.4.1: Map of Mississippi-Valley type ore deposits.

Deep, saline, connate fluids (trapped in the pore spaces), within sedimentary basins may be highly metalliferous. When expelled outward and upward during basin compaction, these fluids may form lead and zinc deposits in limestone by replacement or by filling open spaces (caves, faults) and in sandstone by filling pore spaces. The most famous of these are called **Mississippi Valley-type** deposits [44]. Also known as carbonate-hosted replacement deposits, they are large deposits of galena and sphalerite (lead and zinc ores) which form from fluids in the temperature range of 100 to 200°C. Although they are named for occurrences along the Mississippi River Valley in the United States, they are found worldwide.

Sediment-hosted copper deposits occurring in sandstones, shales, and marls are enormous in size and their contained resources are comparable to porphyry copper deposits. These were most-likely formed diagenetically by groundwater fluids in highly-permeable rocks [45]. Well-known examples are the Kupferschiefer in Europe, which has an areal coverage of >500,000 Km_2, and the Zambian Copper Belt in Africa.

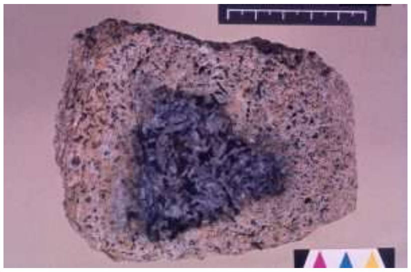


Figure 16.4.1: A sample of bauxite. Note the unweathered igneous rock in the center.

Deep and intense weathering of soils and mineral deposits exposed at the surface can result in the formation of surficial deposits. **Bauxite**, an ore of aluminum, is preserved in karst topography and laterites (soils formed in wet tropical environments) [46]. Aluminum concentrates in soils as feldspar and ferromagnesian minerals in igneous and metamorphic rocks undergo chemical weathering processes. Weathering of ultramafic rocks results in the formation of nickel-rich soils and weathering of magnetite and hematite in banded iron formation results in the formation of goethite, a friable mineral that is easily mined for its iron content.

Surficial Physical Processes

Figure 16.4.1: Lithified heavy mineral sand (dark layers) from a beach deposit in India.

At the earth's surface, the physical process of mass wasting or fluid movement concentrates high-density minerals by hydraulic sorting. When these minerals are concentrated in streams, rivers, and beaches, they are called **placer** deposits, whether in modern sands or ancient lithified rocks [47]. Native gold, native platinum, zircon, ilmenite, rutile, magnetite, diamonds, and other gemstones can be found in placers. Humans have mimicked this natural process to recover gold manually by gold panning and by mechanized means such as dredging.

Environmental Impacts of Metallic Mineral Mining

Figure 16.4.1: Acid mine drainage in the Rio Tinto, Spain.

The primary impact of metallic mineral mining comes from the mining itself, including disturbance of the land surface, covering of landscapes by tailings impoundments, and increased mass wasting by accelerated erosion [48]. In addition, many metal deposits contain pyrite, an uneconomic sulfide mineral placed on waste dumps, which may generate **acid rock drainage** (ARD) during weathering. In the presence of oxygenated water, sulfides such as pyrite react undergo complex reactions to release metal ions and hydrogen ions, lowering pH to highly acidic levels. Mining and processing of mined materials typically increase the surface area to volume ratio in the material, causing reactions to occur even faster than what would occur naturally. If not managed properly, these reactions may lead to acidification of streams and groundwater plumes that can carry dissolved toxic metals. In mines where limestone is a waste rock of carbonate minerals like calcite or dolomite are present, their acid-neutralizing potential helps reduce the likelihood of generating ARD. Although this is a natural process too, it is very important to isolate mine dumps and tailings from oxygenated water, both to prevent the dissolution of sulfides and subsequent percolation of the sulfate-rich water into waterways. Industry has taken great strides in preventing contamination in recent decades, but earlier mining projects are still causing problems with local ecosystems.

Nonmetallic Mineral Deposits

Figure 16.4.1: Carrara marble quarry in Italy, source to famous sculptures like Michelangelo's David.

While receiving much less attention, nonmetallic mineral resources (also known as industrial minerals) are just as vital to ancient and modern society as metallic minerals. The most basic of these is building stone. Limestone, travertine, granite, slate, and marble are common building stones and have been quarried for centuries. Even today, building stones from slate roof tiles to granite countertops are very popular. Especially-pure limestone is ground up, processed, and reformed as plaster, cement, and concrete. Some nonmetallic mineral resources are not mineral specific; nearly any rock or mineral can be used. This is generally called aggregate and is used in concrete, roads, and foundations. Gravel is one of the more common aggregates.

Evaporites

Figure 16.4.1: Salt-covered plain known as the Bonneville Salt Flats, Utah.

Evaporite deposits form in restricted basins, such as the Great Salt Lake or the Dead Sea, where evaporation of water exceeds the recharge of water into the basin [49]. As the waters evaporate, soluble minerals are concentrated and become supersaturated, at which point they precipitate from the now highly-saline waters. If these conditions persist for long stretches of time, thick deposits of rock salt and rock gypsum and other minerals can accumulate (see chapter 5).

Figure 16.4.1: Hanksite, $Na_{22}K(SO_4)_9(CO_3)_2Cl$, one of the few minerals that is considered a carbonate and a sulfate

Evaporite minerals like halite are used in our food as common table salt. Salt was a vitally important economic resource prior to refrigeration as a food preservative. While still used in food, now it is mainly mined as a chemical agent, water softener, or a de-icer for roads. Gypsum is a common nonmetallic mineral used as a building material, being the main component of drywall. It is also used as a fertilizer. Other evaporites include sylvite (potassium chloride) and bischofite (magnesium chloride), both of which are used in agriculture, medicine, food processing, and other applications. Potash, a group of highly soluble potassium-bearing evaporite minerals, is used as a fertilizer. In hyperarid locations, even rarer and more complex evaporites, like borax, trona, ulexite, and hanksite, are found and mined. They can be found in such localities as Searles Dry Lake and Death Valley, California, and in ancient evaporite deposits of the Green River Formation of Utah and Wyoming.

Phosphorous

Figure 16.4.1: Apatite from Mexico.

Phosphorus is an essential element that occurs in the mineral apatite, which is found in trace amounts in common igneous rocks. Phosphorite rock, which is formed in sedimentary environments in the ocean [50], contains abundant apatite and is mined to make fertilizer. Without phosphorus, life as we know it is not possible. Phosphorous is a major component of bone and a key component of DNA. Bone ash and guano are natural sources of phosphorus.

References

30. Boudreau, A. E. The Stillwater Complex, Montana – Overview and the significance of volatiles. *Mineralogical Magazine* **80**, 585–637 (2016).

31. Willemse, J. The geology of the Bushveld Igneous Complex, the largest repository of magmatic ore deposits in the world. *Economic Geology Monograph* **4**, 1–22 (1969).

32. London, D. & Kontak, D. J. Granitic Pegmatites: Scientific Wonders and Economic Bonanzas. *Elements* **8**, 257–261 (2012).

33. Arndt, N. T. Chapter 1 Archean Komatiites. in *Developments in Precambrian Geology* (ed. K.C. Condie) **11**, 11–44 (Elsevier, 1994).

34. Barrie, C. T. Volcanic — associated massive sulfide deposits: processes and examples in modern and ancient settings. (1999). Available at: https://www.researchgate.net/profile/Michael_Perfit/publication/241276560_Geologic_petrologic_and_geochemical_relationships_between_magmatism_and_massive_sulfide_mineralization_along_the_eastern_Galapagos_Spreading_Center/links/02e7e51c8707bbfe9c000000.pdf. (Accessed: 2nd July 2016)

35. Richards, J. P. Tectono-Magmatic Precursors for Porphyry Cu-(Mo-Au) Deposit Formation. *Econ. Geol.* **98**, 1515–1533 (2003).

36. Hawley, C. C. *A Kennecott Story: Three Mines, Four Men, and One Hundred Years, 1887-1997*. (University of Utah Press, 2014).

37. Ague, J. J. & Brimhall, G. H. Geochemical modeling of steady state fluid flow and chemical reaction during supergene enrichment of porphyry copper deposits. *Econ. Geol.* **84**, 506–528 (1989).

38. Einaudi, M. T. & Burt, D. M. Introduction; terminology, classification, and composition of skarn deposits. *Econ. Geol.* **77**, 745–754 (1982).

39. Bromfield, C. S., Erickson, A. J., Haddadin, M. A. & Mehnert, H. H. Potassium-argon ages of intrusion, extrusion, and associated ore deposits, Park City mining district, Utah. *Econ. Geol.* **72**, 837–848 (1977).

40. James, L. P. *Geology, ore deposits, and history of the Big Cottonwood Mining District, Salt Lake County, Utah*. (Utah Geological and Mineral Survey, Utah Department of Natural Resources, 1979).

41. Hofstra, A. H. & Cline, J. S. Characteristics and models for Carlin-type gold deposits. *Reviews in Economic Geology* **13**, 163–220 (2000).

42. Rui-Zhong, H., Wen-Chao, S., Xian-Wu, B., Guang-Zhi, T. & Hofstra, A. H. Geology and geochemistry of Carlin-type gold deposits in China. *Miner. Deposita* **37**, 378–392 (2002).

43. Lehmann, I. P', Publ. *Bur. Centr. Seism. Internat. Serie A* **14**, 87–115 (1936).

44. Klein, C. Some Precambrian banded iron-formations (BIFs) from around the world: Their age, geologic setting, mineralogy, metamorphism, geochemistry, and origins. *Am. Mineral.* **90**, 1473–1499 (2005).

45. Hitzman, M., Kirkham, R., Broughton, D., Thorson, J. & Selley, D. The Sediment-Hosted Stratiform Copper Ore System. *Econ. Geol.* **100th**, (2005).

46. Bárdossy, G. & Aleva, G. J. J. *Lateritic bauxites*. **27**, (Elsevier Science Ltd, 1990).

47. Cabri, L. J., Harris, D. C. & Weiser, T. W. Mineralogy and distribution of platinum-group mineral (PGM) placer deposits of the world. *Explor. Min. Geol.* **2**, 73–167 (1996).

48. Nuss, P. & Eckelman, M. J. Life cycle assessment of metals: a scientific synthesis. *PLoS One* **9**, e101298 (2014).

49. Gordon, W. A. Distribution by latitude of Phanerozoic evaporite deposits. *J. Geol.* **83**, 671–684 (1975).

50. Delaney, M. L. Phosphorus accumulation in marine sediments and the oceanic phosphorus cycle. *Global Biogeochem. Cycles* **12**, 563–572 (1998).

Contributions and Attributions

16.S: Summary

Summary

Energy and mineral resources are vital to modern society, and it is the role of a geologist to locate these resources for human benefit. As environmental concerns have become more prominent, the value of the geologist has not decreased, as they are still vital in locating and identifying the least intrusive methods of extraction.

Energy resources are general grouped as being renewable or nonrenewable. Geologists can aid in locating the best places to exploit renewables resources (e.g. locating a dam), but are commonly tasked with finding nonrenewable fossil fuels. Mineral resources are also grouped in two categories: metallic and nonmetallic. Minerals have a wide variety of processes that concentrate them to economic levels, and are usually mined via surface or underground methods.

Index

Glossary

Sample Word 1 | Sample Definition 1