THE LITTLE BOOK OF GEOMORPHOLOGY:

EXERCISING THE PRINCIPLE OF CONSERVATION

Robert S. Anderson

Table of Contents

1. Introduction

There is no general theory of geomorphology. We cannot cast the subject in a single equation, or set of equations. As with geology, geomorphology is a tangle of physics, chemistry, biology and history. It is also geometry, as the geomorphology plays out in a complex geographic, topographic setting in which both the tectonic and climate processes responsible for driving evolution of the topography change in style and intensity. There is no grand quest for a Universal Law of Geomorphology. Our subject is often subdivided according to the geographic elements of the geomorphic system: hillslopes, rivers, eolian dunes, glaciers, coasts, karst, and so on. You can see this in the chapters of most textbooks on geomorphology.

That said, there is indeed order to the natural system on and near the earth's surface that in turn serves to connect these subdisciplines. Features in common among many geomorphic realms include:

- Surface materials most often move in one direction: downhill, downstream, downdrift, or downwind. (Note that prior to emergence on the surface, the motion is effectively vertical, in the reference frame of the surface, during near-surface exhumation.)
- Materials are transformed as they move through the system. Some of this is again unidirectional, this time meaning irreversible: large grains can be broken into small grains, but not the reverse; chemical reactions are for the most part permanent, resulting in solutional loss and change in mineralogy toward low temperature hydrous phases. Exceptions to this general statement do exist: duricrusts form by cementation of soils, and carbonate deposits accumulate.
- Motion is concentrative. Material gathers itself into more efficient streams (Shreve's dichotomy of hillslopes and streams (Shreve, 1979)) resulting in spatially-branching networks. This is one of several instances in which the surface processes lead to self-organization.
- If given enough time, the system evolves toward a state in which the material flow is adjusted to transport that supplied to it.

These features argue for some degree of universality in our subject, some degree of connection between the elements of landscape and our treatment of those elements. The most useful scientific principle that we will employ here is that of conservation: conservation of mass, of energy, and momentum. As Richard Feynman reminded us, keep track of the stuff in little boxes, and you can't go wrong. The principles I refer to here, and on which I dwell almost exclusively in this Little Book, are really derivatives of the statements that "mass is conserved" or "energy is conserved". These are indeed fundamental laws of physics. And physicists lean on them heavily. So also is momentum conserved, and angular momentum. So also is baryon number, strangeness, and so on. The fact that on the earth's surface the speeds involved are far less than the speed of light allows us to dodge the complexity introduced by Einstein; we can go right back to Newton. As we do not need to worry about the conversion of mass to energy, mass is conserved perfectly.

Here is a general word statement capturing the essence of conservation:

Rate of change of something = rate of input – rate of loss +/- sources or sinks of that something

Of course, it is not always obvious what that "something" is in a particular problem. Our first task is therefore to decide what to balance. For example, we must decide whether to craft the balance for sediment mass or sediment volume, whether to balance the number of radionuclides, or their concentration per unit volume. In addition, we must decide what region of the world we choose to craft this balance. Is it just a little 2D box? Is it whole hillslope, is it a vertical pillar, or is it a 3D box across which this something might be transported across all sides? Formally, we are choosing the "control volume" for the problem. We will see that these choices will depend on the system, and on the questions we are asking of it.

Our next task is to transform this word statement into a mathematical one. For the most part this follows easily if (and only if) we have been careful about setting up the word problem. The mathematical equations we derive will look quite similar, at least at some level. That this is the case allows us to make progress in solving them. While this is not a book about solving the problems in their most general form, I will work a few of the simplest cases.

The art comes in just how to set up the problems… and as always with art, it is practice that allows you to get better. This little book is aimed at providing that practice by showing many examples of how to set up problems in quantitative geomorphology. By assembling many examples between two covers not too widely spaced, I hope that you will see the connections between the problems, and by seeing the connections be able to approach new problems that are in one or another sense analogous to those I introduce on these pages. We will see that the mathematics provides the bridge not only between the geomorphic systems I discuss, but between our science and others – physics, from which we unabashedly borrow, but also chemistry and biology.

The principle of conservation that I exercise here should serve the same purpose that the more fundamental laws of physics serve for the physicist. By being able to lean so heavily, or confidently, on these principles, we can cast our problems formally so that we can then focus on what we do not know, and what we must know about the system in order to make progress on it. My hope is that it will promote the formulation of better geomorphic experiments, and exploration of the world in search of the natural experiments.

I realize that this all sounds fuzzy at the moment. Let me be more concrete. The mathematical statement that most frequently arises from writing down the statement of conservation carefully sounds something like these words:

The rate of change of some quantity at a site = the spatial gradient in the rate of transport of that quantity + the rate of birth of this quantity at the site - the rate of destruction at the site.

Any progress beyond this point in the analysis of our system will then require that we come to grips with what sets the transport rate (we need that before we can evaluate its spatial gradient), and what sets the birth rate or rate of destruction (or death) of the quantity. As this still sounds a little fuzzy, let me illustrate with a biological example. Say our system is a band of sheep on a hillside in a mountain pasture (Figure 1). For economic reasons we are interested only in live sheep. The sheep are wandering from left to right. The change of the number of sheep on the hillside, which we might measure by repeatedly photographing the hillside, will depend largely on the difference between the number of sheep that leave the hillside to the right between times of measurement and those that drift into the hillside from the left; any difference in these numbers will result in accumulation or loss of sheep form the hillslope. But to this we must add any newly born lambs, and subtract any sheep that die (by whatever means – old age, coyotes, shepherds getting hungry). To make more progress than this, we would need more knowledge about the biology of sheep. The rate of births, for example, should be dependent upon the concentration of ewes in the band (the rams are usually separated from the ewes except in the breeding season). Prediction of the death rate would require knowledge of the age structure of the band and some actuarial statistics for sheep. It would also involve the concentration of predators at the site, the number of coyotes or bears or wolves or eagles per unit area of landscape. So you see that by writing down the conservation equation we are automatically forced to focus on the other issues of transport and of sources and sinks, in this case involving specific biological processes.

I will take the ovine example one step further to ask a different question. Can we predict, instead of the number of sheep, the distribution of the ages of the sheep? In this case we might alter the conservation equation to one in which we conserve sheep of each of many age brackets. One would then have to account for the number of sheep that leave each age class, by aging a year, the number of sheep that enter the age bracket by promotion from the next lower age bracket, and the number of sheep that die in each age bracket (the actuarial table again). We would end up with a set of equations that need to be solved, with explicit links between them. If we knew enough about sheep biology, we might be able to predict the age structure of the band of sheep. Or, once we had carefully formulated the set of equations, we could turn the question around and, given the age structure of the band of sheep, determine some of these fundamental biological processes.

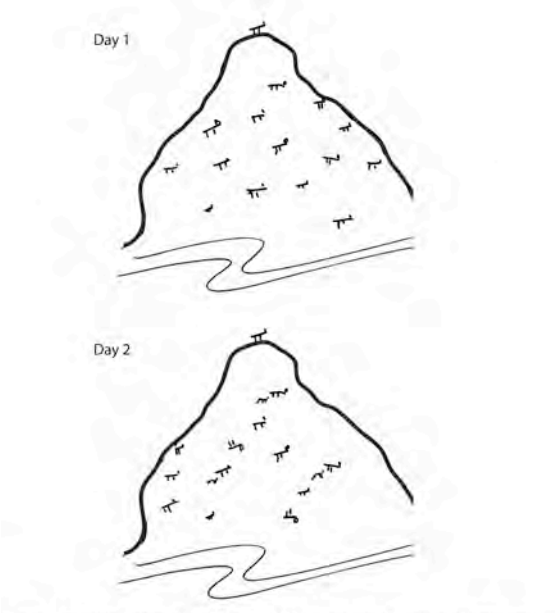


Figure 1. Conservation of sheep. On this hillslope pasture, sheep can migrate onto and off of the pasture, and they can both be born and die. The rate of change in the number of sheep depends upon the differences in rates of migration across the boundaries, and upon the birth and death rates. In this instance, in the 1 day interval between measurements, 2 sheep die, 3 are born, 4 migrate off the right hand side of the slope, and 2 migrate in from the left. The rate of change of number of sheep is therefore $dS/dt = [(2-4) + 3-2]/day = -1/day$. A simple count of the total live sheep in each frame confirms this: 15 on Day 1 and 14 on Day 2.

This is neither as far-fetched, nor as irrelevant to our quest as geomorphologists, as it might sound. Biologists perform such an analysis all the time in order to understand and to model the demography of a population. To bring it back to geomorphology, this method is being exploited to modernize lichenometry (e.g., Loso et al., 2007). These authors document the age distribution of a lichen population, from which they deduce the rule set that controls the demography, really the population evolution, of lichen, which in turn is required before we make use of the distribution of lichen to deduce the age of the surface on which they are growing.

I have strayed into biological examples. But I hope you can see that the same strategy can be employed to organize our thoughts about what sets the size distribution of grains in a river or along a coast, for example. Big grains can become little grains by comminution (or breaking down), while little grains cannot stick themselves back together. Speaking mathematically, the big grains become a source for small grains and show up as a source term in the conservation equation of little grains. Once small enough, small grains may be lost from the system, as they get wafted out onto the shelf from the beach system, or go into suspension and no longer participate in the shape or grain size make-up of a river bed or a beach. We can organize these ideas and the processes they represent into conservation equations, again a linked set of equations, that must be simultaneously solved to predict the grainsize evolution of a system.

Let me provide one more concrete example. Imagine that you are asked to assess the efficiency of a wood stove. You design a special room in which to place the wood stove, you load in some wood, and you light the fire. The room must be very well insulated, so that the only way heat can enter or leave the room is through one inlet and one outlet.

Then you install measurement devices to monitor the heat that comes into the room through the air inlet, the heat that leaves through the one outlet, the stovepipe, and the change in the heat in the room itself. I think you would agree that the change in the heat in the room must be equivalent to the input of heat from the outside (through the inlet), minus the heat that leaves the room through the stovepipe, plus the heat produced by the stove (Figure 2). This is a statement of conservation of heat.

The word picture is simply:

Change of heat in the room = input of heat from the outside – outlet of heat from the room + heat generated in the room

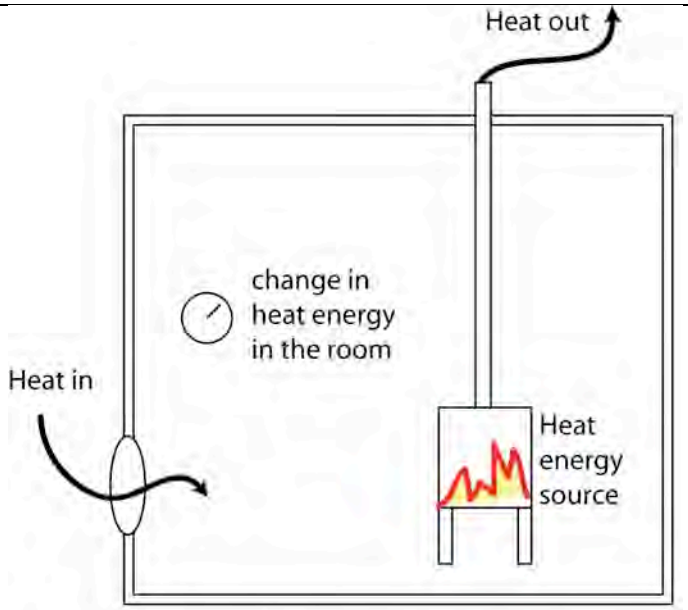


Figure 2. Heat energy balance in a calorimetry room designed to test the efficiency of wood stoves. One must measure the heat coming into the room through the inlet, the heat leaving through the stovepipe, and the change in heat in the room to be able to deduce the strength of the energy source.

The experiment I have described mirrors the manner in which the efficiency of various wood stoves is actually measured. My first brush with the principle of conservation came as an undergraduate student employed to help run such experiments designed by Jay Shelton, then a physics professor at Williams College (see Shelton and Shapiro (1976) for descriptions of wood burning stoves and their proper operation). He simply turned the equation around to solve for the heat produced by the stove, placing all the measurable quantities on the right hand side of the equation:

Heat generated in the room = change of heat in the room – input of heat + outlet of heat

Knowing how much wood was placed in the stove (we weighed it), and its energy content per unit mass (which depended upon its water content), we could then divide the heat produced in the stove by the energy consumed to determine what fraction of the energy went into warming up the room.

Efficiency = heat generated in the room (by the stove) / heat content of the wood

As an aside, one could place any heat source in the room. The most interesting examples were human beings, of which the college had many. It was through such experiments that Jay was able to determine that a typical college student, at rest, puts out energy at a rate of about 100 Joules per second, or 100 watts. So if you ever need to know what you are worth as a producer of energy, think about a 100-watt light bulb.

I have organized the book to move from general to geomorphic. I start with problems that are not purely geomorphic, but that are both pertinent and parallel in their development. Two involve heat, and one introduces a dating method. I start with a heat problem in which the conduction of heat dominates, a problem relevant to understanding the temperature profile in the earth's crust, which in turn is directly pertinent to the thickness of permafrost. But while I stop there with applications of the equations we derive (this is after all a Little Book), the same equation is the backbone to thinking quantitatively about the variations in temperature to which near-surface rock and soil is subjected, which in turn influences the rates of chemical and mechanical breakdown of rock. I then turn to another heat problem, the global climate system, in which the uneven distribution of radiation forces the world's oceans and atmosphere to redistribute heat from the equatorial latitudes toward the poles. Next, the dating problem provides an example of local sources and sinks, here the production of new nuclides from cosmic rays, and the decay of these nuclides to daughter products. This is the essence of many atomic dating methods. The example provided is but a brief introduction to a blossoming field of cosmogenic nuclide applications in the earth sciences. I introduce other applications in the weathering problem and in the coastal system.

I then turn to geomorphic systems that will sound more familiar. I start with lakes, and ask what combination of precipitation and evaporation sets a particular lake level. The section on glaciers deals with the expected pattern of ice discharge in a valley, and allows us to predict the expected glacier length given a prescribed climate (by which here we mean the spatial pattern of local mass balance). I treat then the simplest case of in situ weathering to predict the pattern of degree of weathering in a column of accumulating sediment (e.g., loess or a floodplain). Here I take one baby step toward linking equations for different minerals: as the feldspars are the source of the clays, we must track the concentrations of both. We then allow transport of the sediment by addressing the expected shapes of a hillslope. Here we revisit the problem first worked by G. K. Gilbert to explain why hilltops are typically convex up. They are rounded, not spiky. I then pour water (rain) on a hillside and ask what sets the spatial pattern of overland flow of water on the slope. This is a prerequisite to answering how such an overland flow of water might modify the landscape. Finally, I address two problems involving sediment transport. The first addresses what might form and allow the translation of bedforms like dunes and ripples. In the second, I go to the beach, and ask what sets the spatial pattern of sediment transport (longshore drift) in the coastal or littoral zone.

I summarize briefly by wrapping back to find the common thread among the problems treated. While I do not solve all of the equations derived, I hint how one might go about exploring their behavior. Again, I emphasize that one can proceed in an organized manner, starting with steady solutions, 1D solutions, and proceeding to transient cases of

interest. Given that the thread I pull is a mathematical one, and that mathematics is a highly evolved discipline with a long history of solving equations of the sort we derive, we can then borrow heavily on past experience in math and other fields that have already accessed this deep history.

Finally, let me conclude this introduction with a caveat. This is not an encyclopedic book on geomorphology. It cannot be so and remain small. It is meant as a tutorial on how to approach geomorphic problems. There is not room to embellish by showing many real-world examples. We cannot even treat all of the geomorphic processes. I have left out, for example, the incision of bedrock by rivers, the formation of caves, the tectonic processes responsible for raising rock above sealevel, and so on. I am indeed working on a larger book to be published by Cambridge Press that will address all of these processes, titled Geomorphology: the Mechanics and Chemistry of Landscapes, co-authored with Suzanne Anderson. As this will be out soon, I needn't apologize for the incompleteness of the little book you have before you. Think of it as an extended abstract in which I try to capture the essence of our subject.

Further reading

For examples of how analogous problems are set up in other fields, I recommend the following textbooks.

Turcotte and Schubert's *Geodynamics* has for 20 years been a source of inspiration for earth scientists. While they focus on geophysics, this has broadened in the most recent edition to include some geomorphic topics. But it is the geophysics, and particularly the accessible treatment of heat and elastic problems that I most treasure.

Carslaw and Jaeger's *Conduction of Heat in Solids* is the bible on the topic of heat transport. It is something we all ought to have on our shelves for reference whenever a heat problem or something analogous to one (for example any diffusion problem) is approached.

Bird, Stewart, and Lightfoot's *Transport Phenomena*, while targeted at engineers, is a classic treatise on how to approach problems of heat and mass transport. The authors provide as well excellent atomic-level treatments of the origin of viscosity in fluids.

Geoff Davies's *Dynamic Earth* provides an avenue into thinking about the whole earth as a system. As his aim is to explore the modes of motion in the earth's mantle, he must treat both thermal and fluid flow problems. These are intricately connected in the convection process, either driven from the bottom, giving rise to plumes, or from the top from the sinking of cool slabs. He treats each of these problems at several levels, from beginner to advanced.

David Furbish, in his *Fluid Physics in Geology: An Introduction to Fluid Motions on Earth's Surface and Within its Crust* provides a formal introduction to many earth science problems involving fluid flow. Many of these are geomorphic in character, but the book is more general than that.

Hornberger, Raffensperger, Wiberg and Eshelman, in their *Elements of Physical Hydrology*, provide a well-illustrated, organized view of the hydrologic system. They begin many problems with a simple approach meant to develop the intuition of the reader.

John Harte's *Consider a Spherical Cow* is a favorite for entrance into formal treatment of environmental problems. The approach emphasizes scaling, and box models in which the principles of concern in this book are employed. I also have it on my shelf for its appendices, which like those in Turcotte and Schubert provide excellent summaries of important physical constants and other "useful numbers".

I have also taken inspiration from Kerry Emanuel's *Divine Wind*, a book about the science and history of hurricanes. I appreciate the manner in which he has not shied away from talking about the science in some detail, and about how the science is accomplished, in a book that is clearly targeted at a broad audience. He has also paid homage to the art such events have inspired, through photography, poetry and prose. I find it a wonderful introduction to how the atmosphere works, as illuminated through the window of the most terrible storms. We need more such books.

2. The geotherm, permafrost and the lithosphere

Sorted circles, Spitzbergen. Naturally occurring self-organization of the beach sediment in the active layer above permafrost. Coarse grains are sorted to the rims, leaving interior of the circles fine-grained. Scale of circles: roughly 2 m diameter.

In this first chapter I ask what dictates the temperature profile within the earth's crust. While this is not strictly geomorphology, it is an appropriate place to start because heat problems serve as analogs for many other geomorphic problems, as we will see. In addition, knowledge of the temperature profile within the near-surface earth is fundamental to the interpretation of exhumation histories from thermochronology, for the prediction of the depth of permafrost, and may be used to deduce the temperature history of the earth's surface. Indeed, many near-surface processes are driven by changes in temperature, or have rates that are modulated by temperature. The rates of chemical reactions, for example, are strongly dependent on temperature. This section is designed

to serve as an introduction to the basic physics of heat conduction in this region of the earth. It will serve as well as an introduction to diffusion problems.

Consider the box depicted in Figure 1. This is called the control volume for the problem. While I have chosen a parallelepiped for its shape, with faces perpendicular to the Cartesian coordinates x, y and z, I could have chosen any shape as long as it is a closed surface. Our goal is to craft an equation for the evolution of temperature in the volume.

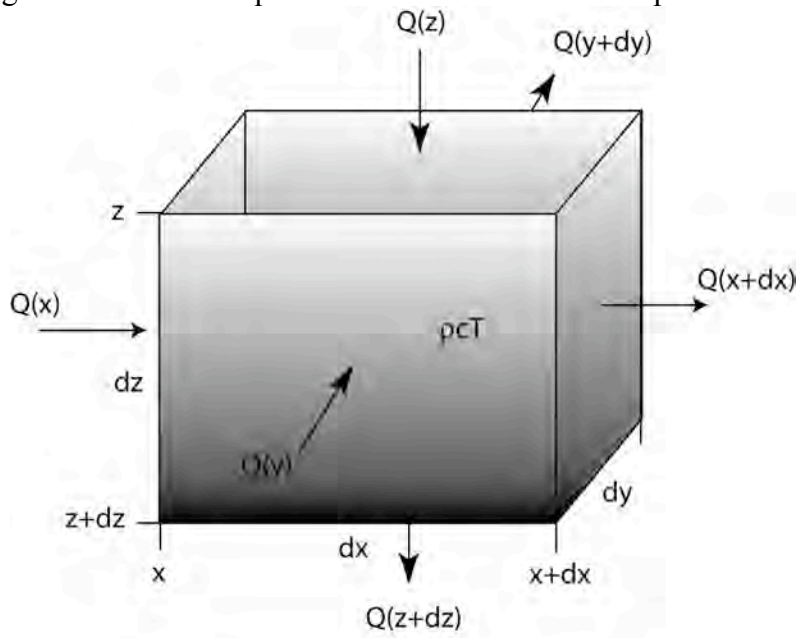


Figure 1. Control volume for setting up conservation of heat energy. Fluxes are shown in and out of each side of the box. Note z is taken to be downward, as we are addressing heat flux in the upper crust of the Earth.

We begin by writing an equation for the conservation of heat within this parcel of the earth. Here we take the material to be a solid, and do not allow it to be in motion (although this assumption could be relaxed later). The word picture for this statement of conservation is:

Rate of change of heat = rate of heat input – rate of heat loss +/- sources or sinks of heat

Each term will have units of energy per unit time. The heat in the box is simply the temperature of the box (which we can think of as the concentration of heat per unit mass), times the volume of the box, times the mass per unit volume (the density), times the thermal heat capacity (energy per unit mass per degree of temperature): $H = \rho cTdxdydz$. The rate of change of this quantity with time is therefore the left hand side of this equation. We use the partial derivative of the quantity with respect to time to represent the rate of change: $\frac{\partial(\rho cTdxdydz)}{\partial t}$, all other variables fixed, in other words, at a fixed point in space (x,y,z). Now we need an expression for the rate at which heat is coming across the left hand side of the box. The heat moving across a boundary or wall of the box can be written as the product of a heat flux, here denoted Q, with the area of the wall.

Throughout the book I will reserve the use of the term "flux" to mean the rate at which some quantity (here heat) moves *per unit area per unit time*. The inputs and losses of heat across left and right walls will therefore look like $Q_x dxdz$. The transport of heat across the left wall is the product of $Q_x(x)$ (which reads "heat flux in the x-direction, evaluated at the location x") with the area of the left wall, $dy\ dz$, while that crossing the right hand wall is the product of $Q_x(x+dx)$, or the heat flux in the x direction evaluated at $x+dx$ with the area of the right wall, $dy\ dz$. Terms relating to the fluxes of heat across the other sides of the box are analogous. Our equation then becomes:

$$\frac{\partial(\rho c T dxdydz)}{\partial t} = Q_x(x)dydz - Q_x(x+dx)dydz + Q_y(y)dxdz - Q_y(y+dy)dxdz + Q_z(z)dxdy - Q_z(z+dz)dxdy \quad (2.1)$$

Up to this point we have made very few assumptions. The only item missing is any heat that is produced within the box, which could happen by radioactive decay of elements, or strain heating of a fluid, or the change of phase of the material. We will ignore all of these for the present. Now let's simplify this a little. We can hold the volume of material in this solid constant through time, allowing us to pull the dx, dy and dz out of the partial derivative with respect to time. For now I also assume that the density and the heat capacity of the material do not change over the temperature range of interest. We may then divide both sides by $\rho\ c\ dx\ dy\ dz$, and the equation simplifies to one of temperature change:

$$\frac{\partial T}{\partial t} = -\left(\frac{1}{\rho c}\right)\left\{\frac{\left[Q_x|(x+dx) - Q_x|(x)\right]}{dx} + \frac{\left[Q_y|(y+dy) - Q_y|(y)\right]}{dy} + \frac{\left[Q_z|(z+dz) - Q_z|(z)\right]}{dz}\right\} \quad (2.2)$$

The last step involves the simple recognition that if we were to shrink the size of our box so that dx tends toward zero, the terms in brackets on the right hand side are the spatial derivatives of the heat flux, $\partial Q_x/\partial x, \partial Q_y/\partial y, \partial Q_z/\partial z$, and the final equation for the conservation of heat becomes

$$\frac{\partial T}{\partial t} = -\left(\frac{1}{\rho c}\right)\left\{\frac{\partial Q_x}{\partial x} + \frac{\partial Q_y}{\partial y} + \frac{\partial Q_z}{\partial z}\right\} \quad (2.3)$$

In doing this we are effectively using the definition of a derivative. This equation says that the temperature in a region will rise if there is a negative gradient in the flux of heat across it, and vice versa (Figure 2). For example, if there is a positive gradient in the x direction, then more heat is leaving out the right hand side of the box than is arriving through the left hand side of the box, and the heat content and hence the temperature in the box ought to decline.

An alternative derivation using the Taylor series expansion. Before moving on, I want to show you that we could have arrived at this equation by taking a slightly different mathematical route. Back up to equation 3.1. We need expressions for the fluxes of heat, Q, at positions $x+dx$, $y+dy$ and $z+dz$. This time we will use a method of estimating these values that employs what is known as a Taylor series expansion. As seen in Figure 3.2, the first estimate of Q_x at $x+dx$ would be simply that at x, in other words $Q_x(x+dx) = Q_x(x)$. A second, better estimate would entail projecting the slope of the function Q over the distance dx. This would yield $Q_x(x+dx) = Q_x(x) + (dQ_x/dx * dx)$. You can see from

the figure that this ought to yield a better estimate. Better yet, you might take into account the curvature of the function. A Taylor series expansion formalizes this process of estimation:

$$Q(x+dx) = Q(x) + \frac{\partial Q}{\partial x}dx + \frac{1}{2}\frac{\partial^2 Q}{\partial x^2}(dx)^2 + \ldots \frac{1}{n!}\frac{\partial^n Q}{\partial x^n}(dx)^n \qquad (2.4)$$

Note that as we take smaller and smaller elements, as we shrink our control volume, the terms with dx^2, dx^3 and so on, will become very small. These are called "higher order terms", and they may be neglected in the limit as dx->0.

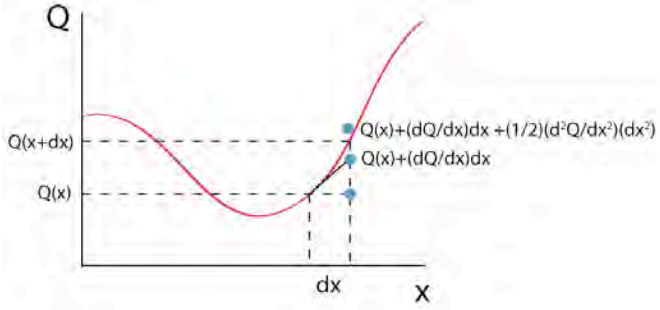


Figure 3.2 Taylor series expansion. Estimation of the value of the function Q at a position x+dx. The first guess would be the value of the function at a nearby point x (lower blue dot). The second estimate acknowledges the rate of change of the function at x, and projects that rate of change over the interval dx (upper blue dot). A third estimate would include the curvature of the function at x, and so on. The Taylor series expansion includes all of these terms in the estimate. As the distance dx over which these projections are being made shrinks (dx->0), the role of the higher order terms declines, and it is sufficient to use simply the local gradient.

Neglecting these higher order terms, when this estimate and the comparable ones for $Q_y(y+dy)$ and $Q_z(z+dz)$ are inserted into equation 3.1, it becomes

$$\frac{\partial(\rho cTdxdydz)}{\partial t} = Q_x(x)dydz - \left(Q_x(x) + \frac{\partial Q_x}{\partial x}dx\right)dydz + Q_y(y)dxdz - \left(Q_y(y) + \frac{\partial Q_y}{\partial y}dy\right)dxdz$$
$$+Q_z(z)dxdy - \left(Q_z(z) + \frac{\partial Q_z}{\partial z}dz\right)dxdy \qquad (2.5)$$

Performing the subtractions yields

$$\frac{\partial(\rho cTdxdydz)}{\partial t} = -\frac{\partial Q_x}{\partial x}dxdydz - \frac{\partial Q_y}{\partial y}dydxdz - \frac{\partial Q_z}{\partial z}dzdxdy \qquad (2.6)$$

Dividing by $\rho cdxdydz$, we arrive again at equation 3.3. The two mathematical routes have converged at the same equation for conservation of heat.

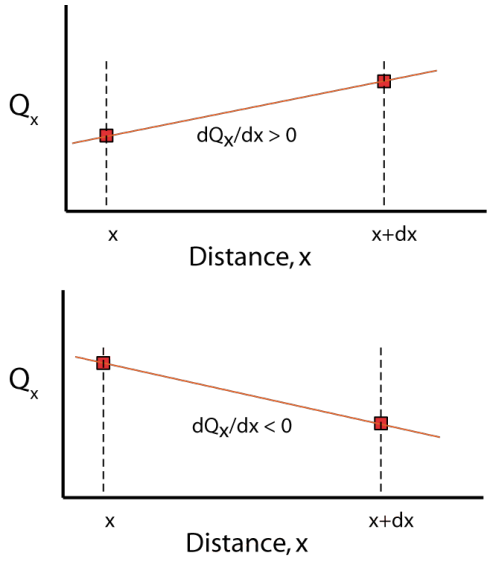


Figure 2. Opposing patterns of heat flux in the x-direction. Top: positive gradient of heat flux results in more hat leaving the element of space than arriving in it, which will cause the temperature in it to drop. Bottom: a negative gradient in heat flux will cause the temperature to rise.

We will from this point on assume that the heat problem is one-dimensional, and vertical. This will allow us to ignore gradients in heat flux in the horizontal dimensions. As we see below, this situation arises when the gradients in temperature are small in the horizontal relative to those in the vertical.

As yet, we have not identified the specific process by which heat is transported in the medium. There are three means of heat transport: conduction, advection, and radiation. Radiation requires that the medium of concern is transparent in the proper wavelengths, which rock and soil certainly are not. Advection requires that the medium, or another fluid passing through it, be in motion. For now we are assuming that this is not the case.

The principal means by which the near-surface rock is cooled is by conduction. Vibrational energy is traded off between adjacent atoms in such a way as to even out the energy, meaning that it flows from regions of high energy (temperature) to regions of low temperature. The result is a relationship that has become known as Fourier's law, where the flux of heat in a particular direction, say along the x-axis, is proportional to the local gradient of temperature in that direction, through a constant called the thermal conductivity, *k*.

$$Q_z = -k_z \frac{\partial T}{\partial z} \tag{2.7}$$

This heat flux has units of energy per unit area per unit time. The minus sign assures that the heat travels *down* thermal gradients, from hot toward cold. When we combine this expression for the heat flux with the 1-D equation for conservation of heat, we obtain

$$\frac{\partial T}{\partial t} = -\left(\frac{1}{\rho c}\right)\left\{\frac{\partial(-k\partial T / \partial z)}{\partial z}\right\} \tag{2.8}$$

This may be simplified to

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2} \tag{2.9}$$

where we have defined $\kappa = k/\rho c$ as the thermal diffusivity. This is the diffusion equation, in which the rate of change of temperature is tied to the local curvature of the temperature profile.

Now we need to give this equation some exercise: we seek solutions for specific situations. We start with the simplest case: the steady geotherm. This is the expected steady temperature profile in the earth's crust. In steady cases, we assume that any partial derivatives with respect to time are zero. In this case, the left hand side is therefore zero, leaving

$$\frac{d^2 T}{dz^2} = 0 \tag{2.10}$$

Note that this step has reduced the partial differential equation (PDE) to an ordinary differential equation (ODE); the partial symbol $\partial/\partial z$ has become d/dz. This equation says that the curvature is zero. A function that does not curve is a straight line. Our only task now is to assess which of all possible straight lines to choose. In order to solve for $T(z)$, we must integrate this equation twice. Each integration will require appeal to a boundary condition to solve for the constant of integration. The first integration results in an equation for the slope of the temperature profile, the gradient

$$\frac{\partial T}{\partial z} = c_1 \tag{2.11}$$

where c_1 is the constant of integration. Here we appeal to Fourier's Law to assess the gradient, c_1. Equation 3.4 suggests that the gradient is set by the ratio of the heat flux and the conductivity

$$\frac{\partial T}{\partial z} = -\frac{Q}{k} \tag{2.12}$$

As it is customary to define the near-surface heat flux as positive upward, despite having defined depth as positive downward, we use $Q_m = -Q$ to alter this equation to:

$$\frac{\partial T}{\partial z} = \frac{Q_m}{k} \tag{2.13}$$

where here Q_m is the mantle-derived heat flux. The second integration results in

$$T = \frac{Q_m}{k} z + c_2 \tag{2.14}$$

One can see that this second constant of integration has units of temperature. We use again a boundary condition, a place at the edge of the space of concern at which we know something about the system. Here it is the temperature. The most obvious choice is to take the mean annual ground surface temperature, $\overline{T}_s$, which we might know from the local climatology. That temperature is being set by completely different physics than the temperatures at depth. It is set by the energy balance at the earth's surface (We deal with this in the next chapter). Substituting $T(0) = \overline{T}_s$ suggests that $c_2 = \overline{T}_s$, resulting in the final equation for the temperature profile:

$$T = \bar{T}_s + \frac{Q_m}{k} z \tag{2.15}$$

This is the linear geotherm we expect in the steady state situation we have described. Combinations of thermal conductivity and mantle heat flux result in typical geothermal gradients on the order of 25-30°C/km.

Permafrost

Let's make use of this equation in cold regions. Consider what sets the base of the permafrost. By definition, permafrost is permanently frozen ground, in which the temperature does not rise above 0°C over several years. We can solve the above equation for the depth of the base of the permafrost by solving the equation for z, and setting the temperature to be $T = 0$°C. The solution is:

$$z_b = -\bar{T}_s \frac{k}{Q_m} = -\frac{\bar{T}_s}{dT / dz} \tag{2.16}$$

For example, if the temperature gradient is 25°C/km, and the mean annual temperature is -30°C, the base of the permafrost is 30/25 = 1.2 km. Everything between about 1 or 2 m (the base of the active layer that annually freezes and thaws) and 1200 m depths is frozen stiff, "permanently". Note that this material does not have to contain water; the temperature simply has to be such that any water that does exist is in the form of ice. In Figure 4 the temperature profiles for several mean annual surface temperatures are shown, along with the resulting relationship between mean annual temperature and the depth to the base of the permafrost.

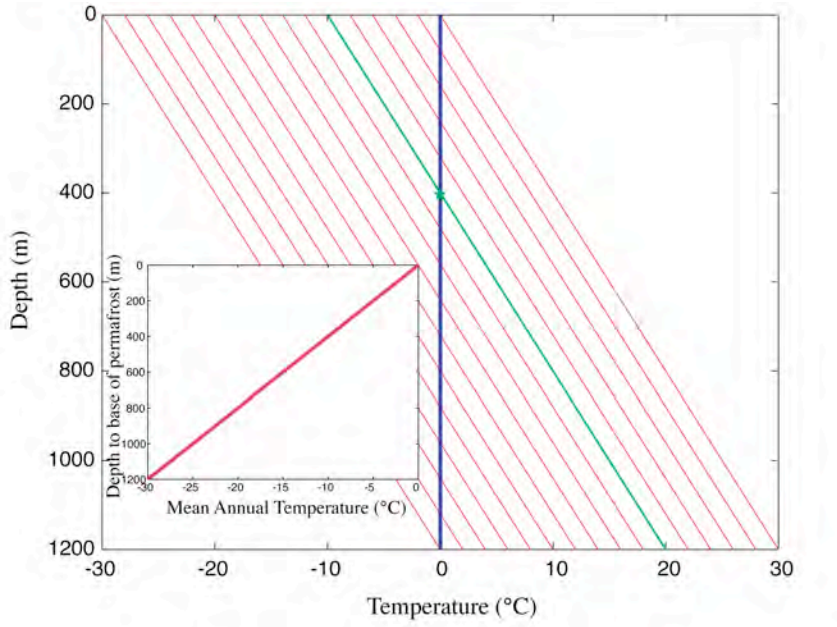


Figure 2. Mean annual temperature profiles in permafrost terrain for a range of mean annual surface temperatures (MAT) from -30 to 0°C. Temperature gradient, dT/dz, held fixed at a typical 25°C/km. Profile for MAT of-10°C is shown in green, with depth of permafrost (*) at 400 m. Depth of permafrost, z_b, is linearly dependent upon the mean annual surface temperature (inset).

While this theory works well to predict the steady state thermal profiles in permafrost terrain, the current state of the arctic is not steady. It is the departures of such measured profiles from steady state that have been used to document the history of warming of the ground surface temperatures through the last 150 years (e.g., Lachenbruch and Marshall, 1986; see Figure 5). These departures are the canaries in the coalmine of global warming.

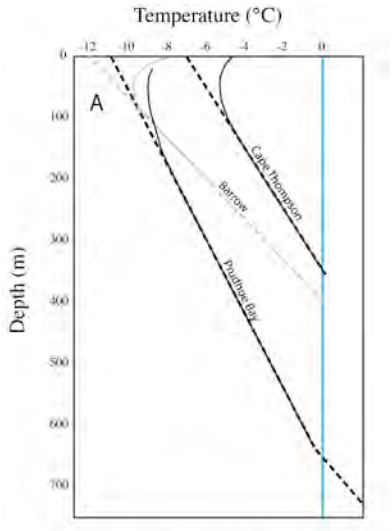


Figure x. Temperature profiles in north slope, Alaska. (as of early 1980s). (after Lachenbruch and Marshall, 1986)

This warming has apparently accelerated in the recent decades, as depicted by repeated measurements of the temperature profiles on the North Slope made by Gary Clow of the USGS (Figure 6). The importance of these records is that they are not proxies for some climate variable, as for example pollen is, or an isotope. They record temperature itself, and we know how heat moves about within permafrost. They are very direct measures, then, of the ground surface temperature that serves as the top boundary condition for the permafrost below.

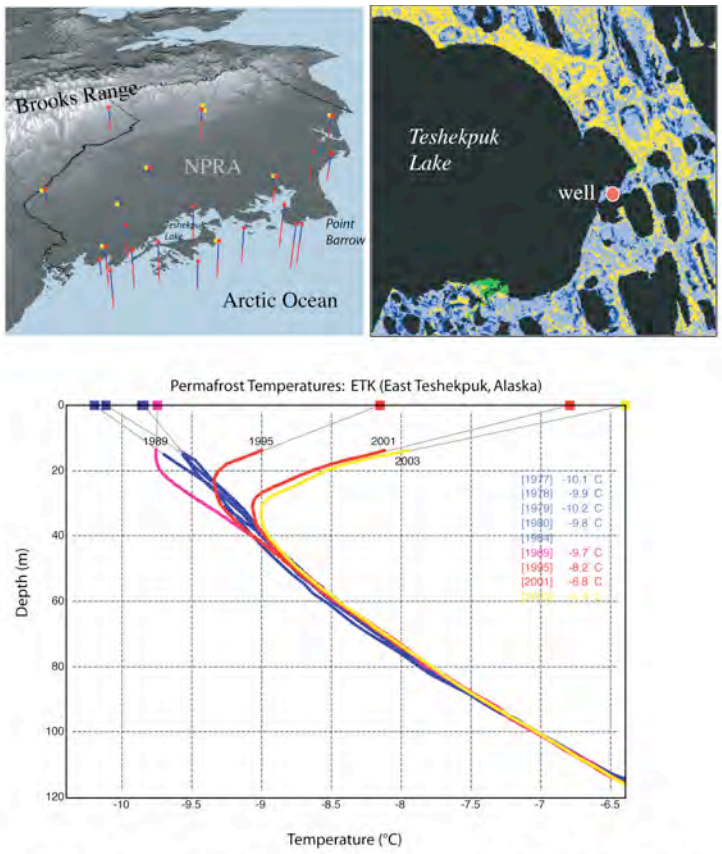


Figure 8.7 Permafrost temerature profiles over the last 3 decades measured in East Teshekpuk, on Alaska's north slope. DOI / GTN-P deep borehole array in the NPRA shown in top left, while false-color air photo of the lake Teshekpuk Lake shown in top right, with well site. Surface temperatures deduced from temperature profiles are shown as boxes at depth=0, and are listed on the right. Rapid warming of the surface since 1990 is evident. (from Gary Clow, USGS)

While significant ongoing research is designed to assess just how the near-surface air temperature translates into the temperatures at the base of the active layer, it is clear enough now that the air temperature must be rising significantly in many places around the globe.

Lithosphere

The alarmingly rapid warming of the permafrost is an example of an end-member case of instantaneous change in temperature of a surface bounding a half-space that was in steady state prior to the disturbance. The shorthand for this problem is "instantaneous warming or cooling of a half-space". In general, the surface could be either warmed or cooled relative to the interior; the mathematical solution for the resulting wave of thermal disturbance progressing into the interior is the same. You can imagine that because this disturbance propagates into the interior, there must be some region near the boundary within which the temperature has changes "significantly" from its prior state. We call this region a "thermal boundary layer". As we will see, the definition of "significant' change is somewhat arbitrary, but once we have defined it, the behavior of the boundary layer is quite understandable. Here I employ this notion to address a key element of the Earth – its lithosphere. We will see that the lithosphere is best thought of as a thermal boundary layer.

At a mid-ocean spreading ridge, two lithospheric plates slide away in opposite directions at speeds of several cm/year. Mantle beneath the ridge rises to replace that lost, and in so doing experiences decompression melting. The basaltic melt rises to produce oceanic crust, solidifying in either the surface or the subsurface. But the plate is not the crust. The plate includes the crust, but is largely that portion of the upper mantle that is cool enough to behave as an elastic solid on geologic timescales. In other words, the plate is a rheological boundary layer. The rheology of a substance is its response to forces (or more formally, to stresses, which are forces per unit area). An elastic rheology (the simplest solid) is one in which a finite strain of the material is produced by an imposed stress; the strain is recovered when the stress is released. In contrast a viscous rheology (the simplest fluid) is one in which the longer the stress is applied, the higher the strain. There is no single relationship between stress and strain. Instead, we find that the *rate* of strain is proportional to stress. In rocks, and for that matter in ice, we find that the rheology is critically dependent upon temperature. When either rock or ice is anywhere near its melting point, it behaves as a viscous material, while at much colder temperatures it behaves like an elastic solid. The cross-over temperature is sort of fuzzy, but call it 1200 to 1300°C. That part of the upper mantle that is cooler than 1200°C will therefore behave as a solid on geologic timescales – this is the lithospheric "plate". In assigning a thickness to the lithosphere, our task is now converted into determining the depth of the 1200°C isotherm. We have transformed the problem into a thermal problem.

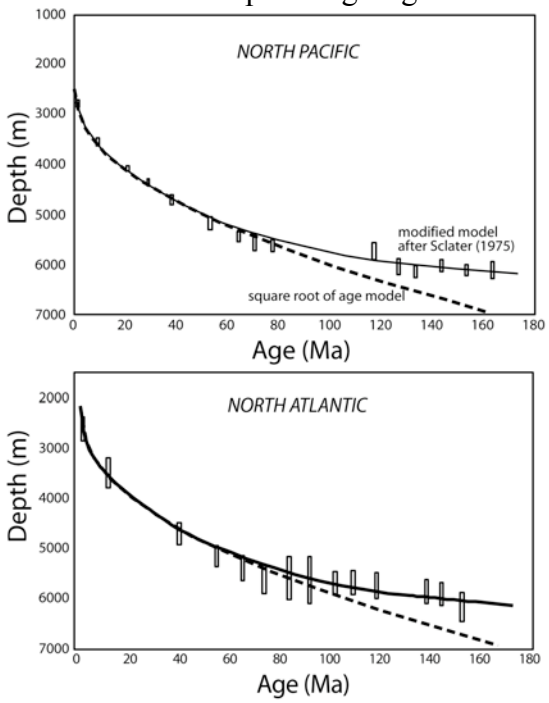
Beyond the region very close to the ridge, where groundwater circulation of fluids can extract heat efficiently from the thin cooling plate, the main means by which heat is transported in the lithosphere is conduction. The equations developed above therefore capture the physics. The system is by no means steady, however, meaning that we cannot make the simplifying assumptions that allowed us to develop the equation for the geotherm. The problem is however much like that involved in the warming of permafrost – only this time the temperature at the boundary has dropped rather than risen. Consider a column of crust and underlying upper mantle. This column is moving at half the spreading rate, but the entire column is moving. It is a solid. We can therefore think of the problem as being a 1D problem, in depth. At just before the clock starts, at time = 0-, the column is at the spreading ridge and is uniform in temperature. At time = 0, the top of the column is exposed to the base of the ocean, and thereafter this top surface is maintained at the temperature of the ocean floor. It IS the ocean floor. The column cools through time, and the distance to the 1200°C isotherm deepens. This is an instantaneous cooling problem. While the temperature structure within the lithosphere is somewhat daunting (it is something we call an error function, *erf*), the solution for the evolution of the depth to a specific isotherm (e.g. 1200°C) is surprisingly simple:

$$d = \eta\sqrt{\kappa t} \tag{2.17}$$

where η is a non-dimensional constant of order 1, κ is the thermal diffusivity, and t is the time since the parcel of lithosphere left the spreading age (the plate age). Given that plate speeds are pretty steady, we can transform time into distance from the spreading ridge through $x = u_{1/2}t$, where $u_{1/2}$ is the half spreading rate. We therefore predict that the profile of lithospheric thickness across a spreading basin is

$$d = \eta \sqrt{\frac{\kappa x}{u_{1/2}}} \tag{2.18}$$

The lithosphere should thicken as the square root of distance from the spreading ridge. How would we check this? The figure below shows the bathymetry of two major ocean basins with distance from the mid-ocean spreading ridges.



Oceanic bathymetry showing roughly square-root of time dependence.
(after Parsons and Sclater 1977)

This bathymetric profile reflects the thickening of lithosphere. The argument goes like this: the lithosphere thickens with time and distance as we have just discussed. By definition, the lithosphere is cooler than the mantle immediately below it (its mean temperature is about (1300-4)/2 ~ 650°C, which is therefore about 650°C cooler than the 1300°C mantle immediately beneath it). Cooler material is denser than hotter material. Isostatic compensation of this cooler denser lithosphere dictates that the thickness of low-density ocean be greater where the thickness of high-density lithosphere is greater. Hence the bathymetry should mimic the thickness profile of the lithosphere. The analysis of Parsons and Sclater (1977) has been updated by Stein and Stein (1992), but the essence remains: we can explain straight-forwardly the gross topography of 70% of the planet's surface.

References

Gold L. W. and A. H. Lachenbruch, 1973, Thermal conditions in permafrost—a review of North American literature. In: *Proceedings Second International Conference on Permafrost, Yakutsk, U.S.S.R., North American Contribution*, National Academy of Sciences, Washington, US, pp. 3–25.

Lachenbruch, A. H., and B. V. Marshall, 1986, Changing Climate: Geothermal Evidence

from Permafrost in the Alaskan Arctic, *Science* 234 (4777): 689-696.

Lachenbruch, A. H., J. H. Sass, and B. V. Marshall, 1982, Permafrost, heat-flow, and the geothermal regime at Purdue Bay, Alaska, *J. Geophys. Res.*, 87 (NB11): 9301-9316.

Parsons, B. and J. G. Sclater, 1977, An analysis of the variation of ocean floor bathymetry and heat flow with age. *J. Geophys. Res.* **82**, 803-827.

Stein, C. A. and S. Stein, 1992, A model for the global variation in oceanic depth and heat flow with lithospheric age. *Nature*, **359**, 123-129.

Turcotte, D. L. and G. Schubert, 2002, *Geodynamics*, 2nd edition, Cambridge University Press.

3. Gooshing of the mantle

Changes in topography over large distances can be caused by flow of the underlying mantle or lower crust. This reflects the mobility of these deep materials, and their behavior as viscous fluids over long timescales. They are driven from place to place by pressure gradients, driving transport from high toward low pressure. As the pressures at these depths are dictated by the thickness of the overlying crust and any other load (ice sheet, lake, ocean), the gradients in pressure are established by the gradients in the thickness of these overlying materials. If we can assess how fast the resulting changes in topography occur, we can deduce the viscosity of the material in motion. These phenomena intimately connect geomorphology to geophysics.

Glacial Isostatic Rebound

First let's take a crack at a classic problem in geomorphology, in which the surface of the earth is first pushed down by emplacement of a major load, and then rebounds when that load is taken off. The loads can be either a lake or an icesheet. I will use an ice sheet for now (Figure 1).

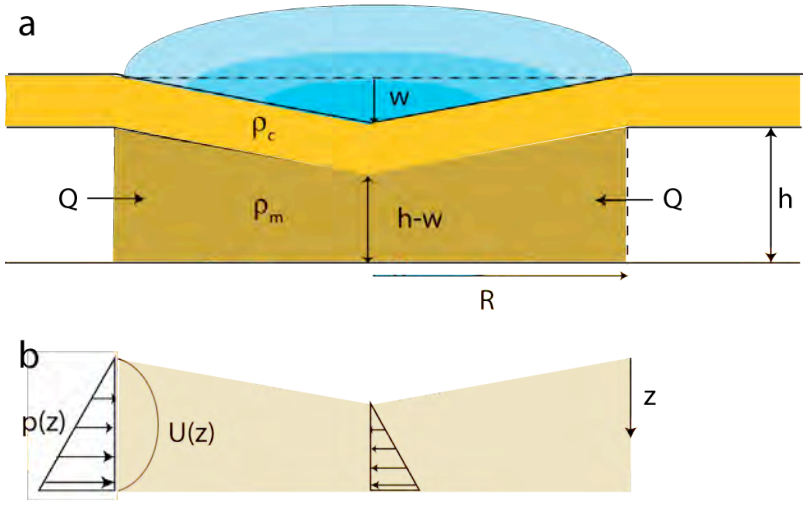


Figure x. Isotatic rebound. a) Schematic of glacial isostatic rebound problem, shown immediately after demise of a circular ice sheet of maximum radius R. The remaining deflection of the crust yet to rebound at this instant is w. The thickness of the low viscosity portion of the mantle is h, the base of which constitutes a no-flow boundary in this approximation. As mantle is transported across the circumference of the cylinder, the upper more rigid lid must rise, thereby reducing the deflection w. b) Pressure profiles in the mantle column are shown after subtraction of the equal pressures associated with the overlying uniform thickness crust. Imbalance of pressures across a channel element, resulting from the difference in heights of the mantle column at edge vs center of the deflection pattern, causes flow from high pressure to low pressure, resulting in a parabolic velocity profile, U(z). The integral of this flow profile is the mantle discharge, Q, shown in (a).

Our goal is to develop an equation for the evolution of the deflection *w* through time after the load is removed. We can test any rebound theory against data from regions once loaded by Last Glacial Maximum icesheets, such as that occupying the Fennoscandian shield, or the Laurentide icesheet over Hudson Bay (e.g., Andrews, 1968; Walcott, 1972; Peltier and Andrews, 1976).

The relevant conservation statement is simply:

The rate of change of mantle volume in the cylinder = inputs – losses

It is safe to assume that the mantle is incompressible, that it does not change in density with time. The problem then reduces to determining the rate of movement of mantle across the boundary; in the case of rebound, this will be inputs of mantle, while in the case of increasing ice load it will be loss of mantle. If we dictate that the cylinder not change in circumference over time, then the only way to accommodate changes in volume of the cylinder is by changing how thick it is. In other words,

$$\frac{d(\left[\pi R^2 h\right]-\left[\frac{1}{3}\pi R^2 w\right])}{dt}=\frac{d\left[-\frac{1}{3}\pi R^2 w\right]}{dt}=2\pi RQ \tag{3.1}$$

where R is the radius of the cylinder, h the thickness of the mantle channel in the absence of deflection, w the vertical deflection in the center of the loaded region, and Q is the volumetric discharge of the mantle either outward (during glacial loading) or inward (during rebound) across the edge of the cylinder. The term in brackets on the left side is the volume of mantle in the cylinder. The first term in square brackets is the volume of a filled cylinder; the second term in square brackets is the missing conical volume due to the depression of magnitude w. As neither R nor h changes in time, the leftmost term drops out when we take a time derivative. Q must have units of volume per unit length around the perimeter per unit time – or L^2/T.

What remains is to develop an equation for the discharge of mantle, Q, across the imaginary walls of our cylinder. For this we will appeal to the solution for viscous flow between two rigid plates. The velocity profile in such a flow is parabolic, the horizontal speed being

$$U(z)=-\frac{1}{2\mu}\frac{dp}{dx}\left(zh-z^2\right) \tag{3.2}$$

where z is the depth into the viscous channel, h its thickness, and dp/dx is the pressure gradient pushing the flow. Note that because zh is always greater than z^2, the flow is in the positive x direction only when the pressure gradient is negative. This acknowledges that fluid is pushed from high pressure toward low pressure. Our mantle discharge is the integral of this velocity profile

$$Q=\int_0^h U(z)dz=-\frac{1}{12\mu}\frac{dp}{dx}h^3 \tag{3.3}$$

We can now see the strong dependence on the thickness of the channel, but we have yet to come to grips with the pressure gradient. The pressure gradient arises from the difference in the thickness of the channel inside and outside the cylinder. In the rebound case, the thickness outside the channel is greater, and this extra thickness serves to push the mantle toward the thinner inner cylinder. We should use for the pressure the mean pressure in the viscous channel. Subtracting off the pressure associated with the overlying crust, and employing the mean value theorem, the mean pressure in the viscous channel becomes

$$\bar{p} = \frac{1}{h}\int_0^h p(z)dz = \frac{1}{h}\int_0^h (\rho_m g z)dz = \frac{1}{h}\left[\frac{\rho_m g z^2}{2}\right]_0^h = \frac{\rho_m g h}{2} \tag{3.4}$$

This pressure difference occurs over a length scale set by the radius of the cylinder. An estimate of the pressure gradient is therefore

$$\frac{dp}{dx} = \frac{d\left(\frac{1}{2}\rho_m g h\right)}{dx} = \frac{1}{2}\rho_m g \frac{dh}{dx} \cong \frac{1}{2}\rho_m g \frac{w}{R} \tag{3.5}$$

where the thickness change is represented by the maximum deflection w, and the length scale over which it changes is R (Figure 1). Insertion of this expression into the discharge equation, and insertion of the discharge equation into the conservation of volume equation yields

$$\frac{d(\frac{1}{3}\pi R^2 w)}{dt} = 2\pi R\left[-\frac{1}{12\mu}\frac{\rho_m g}{2R}h^3 w\right] \tag{3.6}$$

Simplifying this by dividing through by $\pi R^2/3$, none of which changes in time, results in an evolution equation for the deflection of the surface, w:

$$\frac{dw}{dt} = -\left[\frac{\rho_m g h^3}{4\mu R^2}\right]w = -\frac{w}{\tau} \tag{3.7}$$

The solution to this equation is an exponential decay function

$$w = w_0 e^{-t/\tau} \tag{3.8}$$

in which the characteristic time scale, τ, is the inverse of the expression in brackets in equation 1.7:

$$\tau = \frac{4\mu R^2}{\rho_m g h^3} \tag{3.9}$$

The time scale is strongly controlled by the viscosity of the material being gooshed across the cylinder walls, by the size of the depressed area, and by the thickness of the viscous channel. Assuming that the viscosity of the channel is 10^{18} Pa-s, the radius of the load 500 km, the thickness of the channel 100 km, the density of the mantle is 3000 kg/m^3 and g = 10 m/s^2, results in an expected e-folding timescale for rebound of about 1000 years. This is of the right order of magnitude, as dating of raised marine terraces in Fennoscandia suggest that timescales for rebound are about 4000 years (Figure 2).

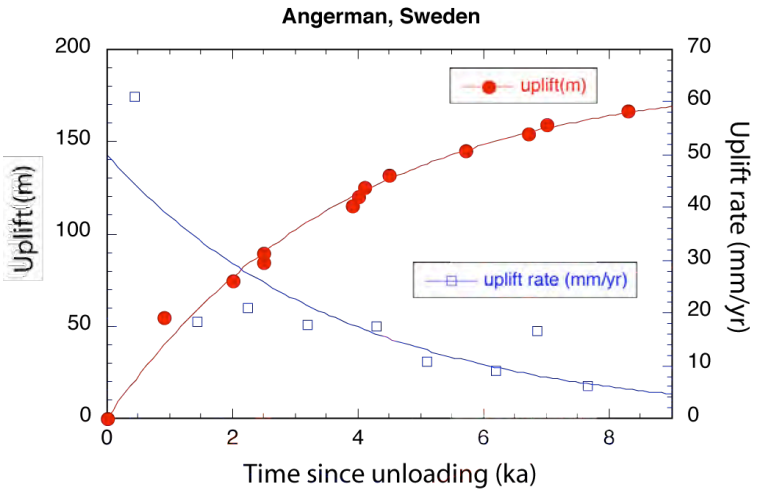


Figure x. Rebound of Angerman River area, Sweden, in the aftermath of the demise of the Fennoscandian Ice sheet. Data from dated shorelines. Boxes: uplift rates derived from differencing elevation and age data. The curves fit shown are the expected asymptotic exponential solution derived in the text for a site within the area of removed load. Same best-fit time scale for elevation data (3.8 ka) is used to fit rebound rate data. Note modern rate of uplift is about 5 mm/yr.

In this figure I have plotted not the deflection from the equilibrium elevation, but the uplift relative to modern sealevel. The least well known of the variables in the rebound time scale are the viscosity and the thickness of the viscous channel. There are obviously tradeoffs in choosing these, but that we are in the ballpark is encouraging, given this simple formulation of the problem. I could easily find the missing factor of 4 by increasing the viscosity to 4×10^{18} Pa-s, or by changing the channel thickness to 63 km.

Hurricane storm surge

I note that there is an oceanic analog in the storm surge associated with a hurricane. The ocean surface is mounded up in a dome centered on the eye of the storm for two reasons (see Kerry Emanuel's wonderful book on hurricanes). First, the low pressure of the atmosphere in the storm center, which itself generates the winds of the storm, creates a pressure gradient in the ocean that drives water toward the storm center. Second, the winds generate a current that piles water up as the storm encounters shallow water. The latter effect generates the majority of the storm surge and produces the asymmetry of the surge (to the right of where the eye hits the land in the northern hemisphere). These surges can be up to 20 meters! We can estimate the magnitude of the pure pressure effect (Figure 3).

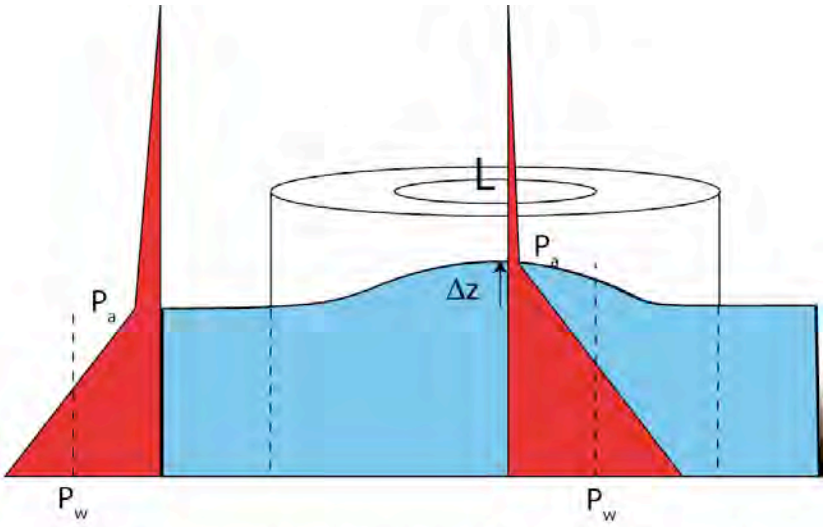


Figure 3. Mounding of the ocean surface beneath the eye of a hurricane. Low atmospheric pressure centered on the eye of the storm produces a pressure gradient in the ocean that pushes water from outside the storm toward the eye. Conservation of water requires that the amount of water increase within the cylinder (dashed margins) beneath the storm, resulting in rise of the water surface. The magnitude of the rise Δz is set by the requirement that the mean pressure in the water columns inside and outside the storm be equal. Under these conditions the pressure gradient driving flow across the cylinder vanishes. Here the pressure profiles in the atmosphere and the water are sketched illustrating this condition of equilillbrium. Note that the gradient in pressure in the air column differs, due to different density of the hot rising air in the storm center vs outside the storm, while that within the water column is the same.

The mean atmospheric pressure is 1013 millibars, or 1.013×10^5 Pa. Given that the formula for pressure beneath a fluid of a given uniform density, ρ, is $P = \rho g H$, we can estimate the thickness of this layer, H. For atmosphere at surface density of ρ=1.22 kg/m^3, this corresponds to a layer 8464 m thick. To generate the same pressure, water at ρ=1000 kg/m^3 would have to be H = 10.3 m thick. It is pure coincidence that the pressure exerted by the atmosphere is worth about a 10 m column of water. But it sure is convenient. Given that the lowest pressures associated with category 5 hurricanes are about 900 millibars, or about 9/10 of an atmosphere, the pressure drop corresponds to 1/10 of the column. This translates into 1/10th of 10 m, or about 1 m of water. So to balance the pressures in the column of water beneath the eye *vs* those in a comparable column exterior to a large storm, the water must dome up by about a meter. That this is only a small fraction of the 10 m typical of a storm surge associated with such monster storms implies that the remainder of the storm surge must come form the wind-driven currents.

Topographic oozing of the Tibetan Plateau

In a nifty article Marin Clark and Wicki Royden proposed that the smooth topographic ramp leading from SE China up to the Tibetan Plateau resulted from the oozing of hot lower crustal material from beneath the plateau (Clark and Royden, 2000). This ramp contrasts sharply with the abrupt topographic front of the Himalayas that bounds the plateau to the south. The physics of the problem is identical to that I have just introduced in the glacial rebound problem.

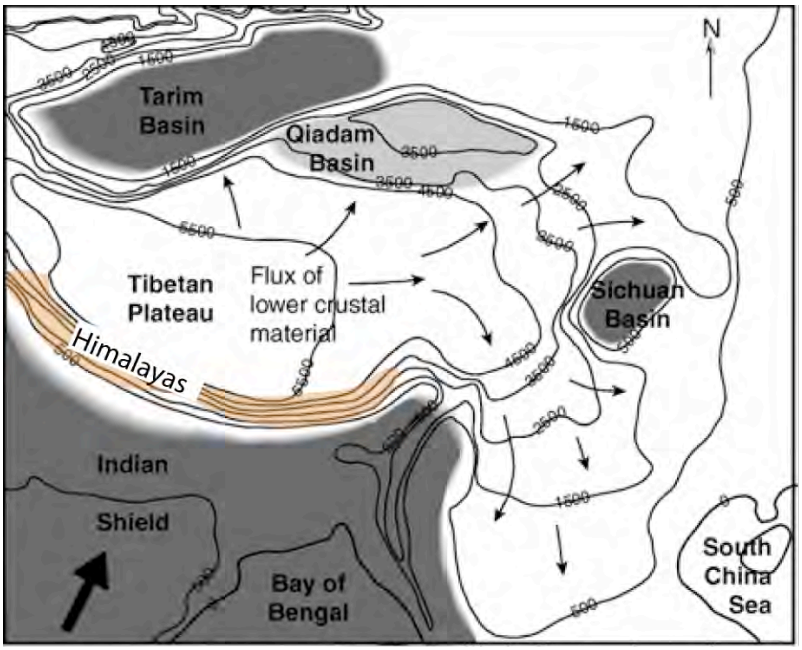


Figure 5. Map of Tibetan Plateau and surrounding lowlands, showing proposed direction of transport of lower crustal material from beneath plateau. Note the broad topographic ramp to the SE, and its contrast with the abrupt topographic front imposed by the Himalayas to the south of the plateau. (from Clark and Royden, 2000, Figure 5)

The analysis of Clark and Royden follows and simplifies an earlier model of the deformation of the Tibetan plateau as a whole (Shen et al., 2001). Viscous lower crust (in this case) is being driven down a channel by a pressure gradient. The chief difference is that the channel through which the lower crust is being driven is unbounded – material is not being gooshed into or out of a cylinder but into a slot of effectively infinite length. I emphasize that whether the hypothesis is right or wrong (and there is indeed some support for it in river incision along this ramp (Schoenbohm et al., 2006)), their treatment can be understood using the same physics we are exploring in this section.

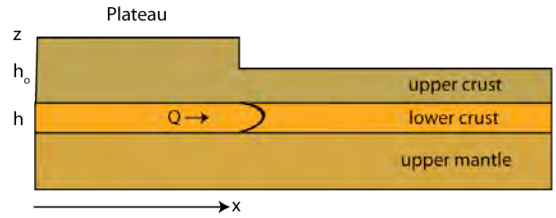


Figure 4. Set-up for flow of lower crustal material from beneath plateau. Change of thickness of the low viscosity channel, h, will result from gradients in the discharge of lower crustal material. Pressure driving the flow is tied to the thickness of the upper crust, ho, which is non-uniform due to thickening of crust to form the Tibetal Plateau.

I show their proposed system in Figure 4. The rate of change in thickness of the lower crustal channel is determined by the gradient in the discharge of lower crustal material down the channel. In other words,

$$\frac{\partial z}{\partial t} = \frac{\partial h}{\partial t} = -\frac{\partial Q}{\partial x} \tag{3.10}$$

where h is the thickness of the channel, Q the discharge of viscous material along it. Again, the discharge, Q, is the volume of material per unit time per unit width of channel [=]L^3/LT or L^2/T. Since the thickness of upper crust above the viscous channel, h_o, does not change with time, the rate of change of elevation of the surface, z, simply follows the rate of change of thickness in the channel. It is, however, important that the thickness of crust h_o does vary in space. It is the thickened crust beneath the Tibetan plateau that leads ultimately to its heating, which in turn leads to its reduction in viscosity and hence tendency to flow.

As in the rebound case, the discharge in the viscous channel is calculated from the integral of the velocity profile

$$Q = \int_0^h U(z)dz = -\frac{1}{12\mu}\frac{dp}{dx}h^3 \tag{3.11}$$

While at first the pressure gradient is set by the gradient in thickness of upper crust, gradients in lower crustal thickness will begin to play a role as it evolves:

$$\frac{dp}{dx} = \rho_c g \frac{d\left[h_o + h\right]}{dx} \tag{3.12}$$

Inserting this expression into the statement of conservation of volume in the channel results in

$$\frac{dh}{dt} = \frac{\rho_c g}{12}\frac{\partial\left[\frac{h^3}{\mu}\frac{\partial(h_o + h)}{\partial x}\right]}{\partial x} \tag{3.13}$$

For now let's assume that the viscosity of the lower crust is uniform, so that we may pull it out of the derivative. Taking the derivative of the remaining product results in two terms

$$\frac{dh}{dt} = \frac{\rho_c g}{12\mu}\left[h^3\left(\frac{\partial^2 h_o}{\partial x^2} + \frac{\partial^2 h}{\partial x^2}\right) + 3h^2\left\{\left(\frac{\partial h}{\partial x}\right)^2 + \frac{\partial h_o}{\partial x}\frac{\partial h}{\partial x}\right\}\right] \tag{3.14}$$

Taking the leading term in this expression, dominated by h^3, leaves us with the simpler

$$\frac{dh}{dt} = \frac{\rho_c g}{12}\frac{h^3}{\mu}\left[\left(\frac{\partial^2 h_o}{\partial x^2}\right) + \left(\frac{\partial^2 h}{\partial x^2}\right)\right] = \kappa\left[\frac{\partial^2 h_o}{\partial x^2} + \frac{\partial^2 h}{\partial x^2}\right] \tag{3.15}$$

The first term corresponds to a couplet of source-sink associated with the non-changing but non-uniform thickness distribution of upper crust. The second term then acts to smear out this constant source, leading to thickening of lower crust (and associated inflation of topography) on the lengthening ramp, and simultaneous deflation of the lower crust below the edge of the plateau. This is a diffusion equation, analogous to the one derived in our treatment of thermal problems except for this spatially distributed source term. I have made the analogy explicit by assigning an effective diffusivity, κ, to the collection of terms in front of the curvature. This constant reflects the efficiency with which changes in thickness occur. As expected, this efficiency is low when the viscosity

is high, and is greatly enhanced when the channel is thick; the collection h^3/μ is difficult to disentangle.

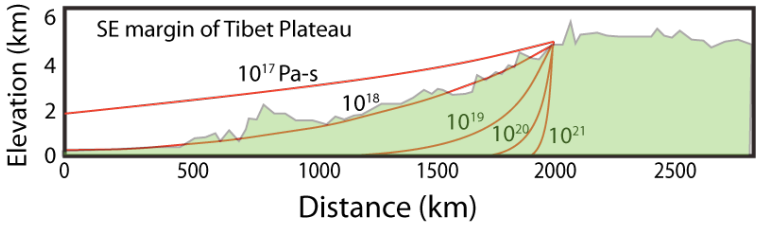


Figure 5. Topographic profile of ramp leading from Tibetan Plateau to the SE. Lines show solution for expected profile given assumed viscosity of the lower crust channel. (after Clark and Royden, 2000, Fig. 4a)

Clark and Royden present results from calculations (Figure 5) assuming a channel thickness h of 15 km. Given this choice of channel thickness, the best-fitting viscosity of the lower crustal material is 10^{18} Pa-s.

Gooshing of mantle across the continental edge

I should probably stop right here but I have to put out there another idea. We have talked about gooshing the mantle about by ice loads, and gooshing the lower crust by topographic loads. There is at least one other case worthy of our attention. When a huge ice sheet is constructed on a continent, that water comes from somewhere and that somewhere is the ocean. It gets there circuitously, delivered as snow by storms, but its ultimate source is evaporation from the ocean. So when the ice sheet volume is large the ocean volume is small. To give this a scale, at the Last Glacial Maximum some 20 thousand years ago, sealevel was drawn down about 120 m relative to today. Just when the land is pushed down by the giant plunger of an ice sheet or two in the northern hemisphere, forcing mantle to flow outward away from the load, the ocean basins of the world are being unloaded. The converse is also true: when the ice sheets dwindle, mantle rushes back in, while at the same time the ocean basins are being loaded up, and mantle should be pushed away from them. Consider then a continental margin well away from any ice load. The variations in the ocean load adjacent to the continental margin should drive a gooshing of the mantle back and forth across the continental margin. Let's marry the two examples we have discussed so far, the ice load and the topographic ooze, and sprinkle in a periodic variation in the system to construct a model of this situation. Our goal is to assess the amplitude of the effect on the topography of the edge of the continent, and the distance inland over which this signal declines.

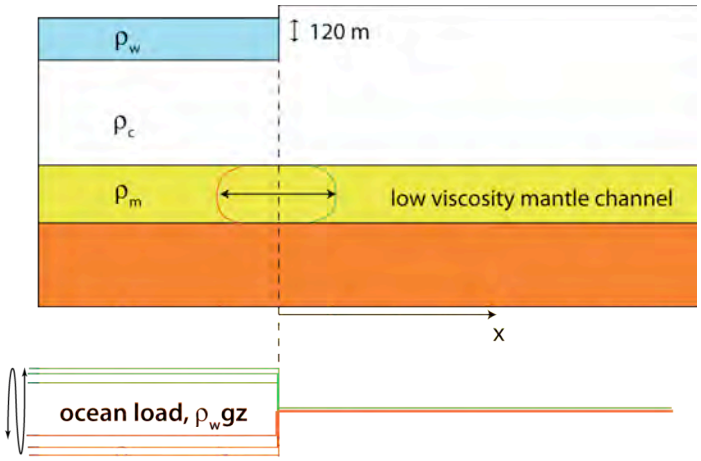


Figure 7. Schematic diagram illustrating the possible transport of mantle in a low viscosity channel in response to oscillating sea level. The ocean load varies with sea level, which in the Pleistocene has varied by 120 m between glacial lowstands and interglacial highstands with periods of 20-100 ka. At highstands (green), mantle is forced from beneath the ocean load toward the interior of continents (green velocity profile), while at lowstands (red), mantle will be pushed toward the ocean basins (red velocity profile). Our goal is to predict the resulting pattern of thickening of the channel, and consequent raising or lowering of the land surface.

Assume a viscous channel in the upper mantle, beneath a uniform crust. Consider the simplest case of a very steep continental margin, in which the full 120 m swing of sealevel does not cause significant lateral migration of a coastline (Figure 6). Now allow sealevel to vary sinusoidally with a fixed amplitude, Δh_{sea} and period, P:

$$h_{sea} = h_{sea} + \Delta h_{sea}(\sin(2\pi t / P) \tag{3.16}$$

This variation in sealevel translates into a variable pressure at the depth of the mantle channel, which in turn drives variation in the pressure gradient in the channel that causes transport of mantle. In addition, the thickness variation in the mantle then can drive flow as well, as the mean pressure in the thicker channel is greater than that in the thinner portion of the channel.

The conservation of fluid in the mantle channel is identical to that we wrote for the Tibetan lower crustal ooze (equation 10), as is the equation for lateral discharge of mantle. What differs is the pressure field driving flow, which now includes that of the dynamic oceanic load. The pressure is therefore

$$P = \rho_w g h_{sea} + \rho_c g h_c + \rho_m g(h/2) \tag{3.17}$$

When we acknowledge these contributions from all components of the load, the discharge becomes:

$$Q = -\frac{1}{12\mu}\frac{dp}{dx}h^3 = -\frac{h^3}{12\mu}\left[\frac{\rho_w g h_{sea} + \rho_c g h_c + \rho_m g(h/2)}{dx}\right] \tag{3.18}$$

Because the thickness of the crust does not vary in time, and it is largely uniform in space in this problem, I will drop this middle term. The statement for conservation of volume in the mantle channel then becomes:

$$\frac{dh}{dt} = \frac{\rho_c g}{12\mu}\left\{h^3\left[\left(\frac{\partial^2 h_o}{\partial x^2}\right) + \left(\frac{\partial^2 h}{\partial x^2}\right)\right] + 3h^2\left[\left(\frac{\partial h}{\partial x}\right)^2 + \frac{\partial h}{\partial x}\frac{\partial h_o}{\partial x}\right]\right\} \tag{3.19}$$

This formulation explicitly acknowledges that the thickness of the channel, h, varies, which in turn causes variations in the discharge through the strong nonlinear h^3 dependence.

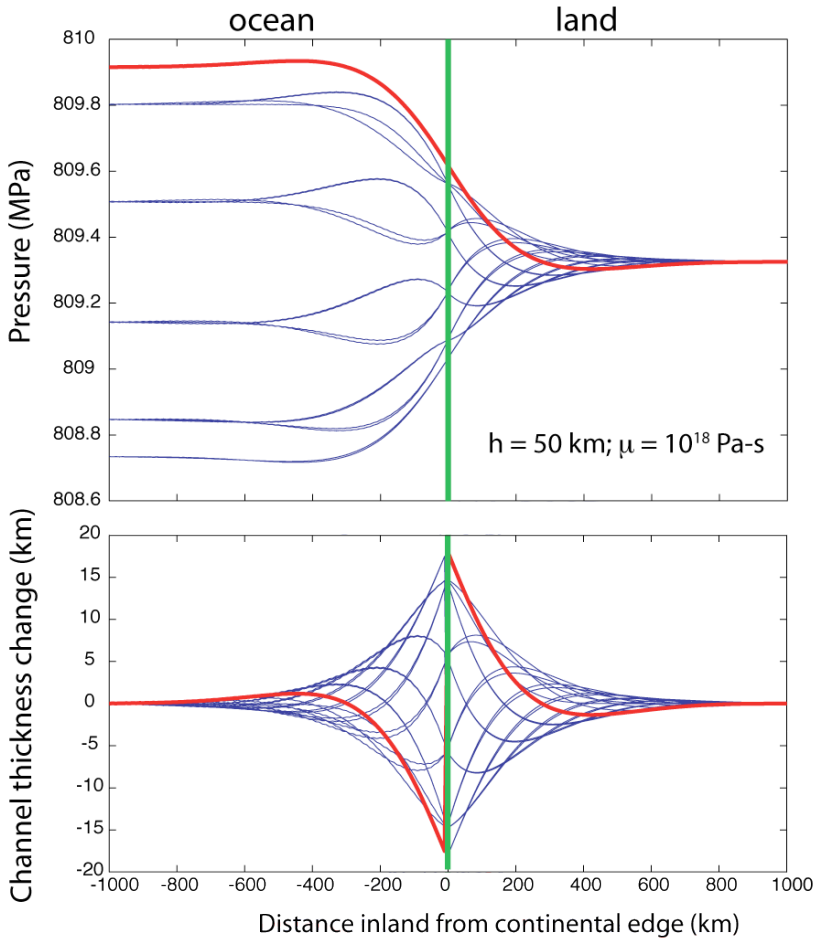


Figure 7. Elevation changes associated with periodic variations in sealevel. Top: Pressure field calculated in mid-mantle channel at 10 times within a sinusoidal variation in sealevel with period 20ka, full amplitude 120 m (two oscillations shown). Bottom: Mantle channel thickness change driven by gradients in discharge of viscous channel material. This should result in elevation changes in the overlying landscape that extends 100s of km from the continental margin. Last time stamp shown in red corresponds to sealevel highstand.

I have used a numerical model to explore the behavior of the system when forced with a 120 m oscillation in sealevel (Figure 7). You can see that the system displays a diffusive behavior. It displays both an exponential decay of amplitude with distance from the margin, and a phase lag, much like the solution for temperature within a half space when forced by oscillation of the surface temperature (e.g., Gold and Lachenbruch, 1973; Turcotte and Schubert, 2002). In the thermal case, the length scale that dictates both the decay rate and the time lag is set by the square root of the thermal diffusivity and the period of the oscillation:

$$L = \sqrt{\frac{\kappa P}{\pi}} \tag{3.20}$$

The same appears to be the case in this system. Numerical experiments show that the penetration of the effect into the continent indeed depends upon both the period of the oscillation and those variables that take the place of the diffusivity:

$$\kappa = \frac{\rho g h^3}{12\mu} \tag{3.21}$$

While this problem has yet to be fully exploited, I challenge the reader to look for hints that this gooshing of the mantle across the continental edge occurs. It seems to me that the geomorphic signal of this phenomenon will be best displayed where large rivers approach the coastline. Large rivers have small slopes. For example, the slope of the Amazon is around 1 cm/km, or 10^{-5}, while that of the Mississippi is perhaps a few times this ($2x10^{-5}$ on the delta itself; see Syvitski and Saito (2007)). The smaller the slope, the more likely it will be tweaked by the small tilts of the continental edge imposed by the movement of mantle across the margin. The maximum tilts in the case I have illustrated are about 15 m/300km, or $5x10^{-5}$; given the slopes of major rivers as they approach the coast, this tilt is worthy of discussion. Note also that a topographic sag comes and goes well inland of the margin – 500 km inland in the case I have illustrated – and that the sagging is out of phase with adverse tilt at the margin.

In working this problem (more or less as a teaser) I have not accounted for the rigidity of the lithosphere above the channel, which will serve to smooth any sharp gradients in the predicted thickness of the mantle channel. These occur at the continental margin. I also admit that the larger effect on a river draining the continent is the more obvious and dramatic oscillation of baselevel. During glacial times these major rivers seek to join an ocean (baselevel) that is 120 m below present, and should incise their margins as they seek that level. The greater hope to find the effect I have illustrated comes instead from its effects well away from the margin, and in any temporal lag of the signal. After all, sealevel reached its present highstand roughly 6000 years ago, while the high viscosity of the mantle channel should result in continued, ongoing warping of the margin.

In this section I have explored a few impacts of mantle physics on the overlying topography. These examples serve not only to illustrate another application of the principles of conservation, but to alert the geomorphologist that what happens deep in the Earth does indeed matter to its surface. These deep gooshings constitute the largest length scale processes to which the Earth's surface is subjected.

References

Andrews, J. T., 1968. Postglacial rebound in Arctic Canada: similarity and prediction of uplift curves. *Can. J. of Sci.* 5: 39

Clark, M. K., and Royden, L. H., 2000. Topographic ooze: Building the eastern margin of Tibet by lower crustal flow. *Geology* 28: 703–706.

Davies, G. F., 1999. *Dynamic Earth: Plates, Plumes and Mantle Convection*, Cambridge Press, 458 pp.

Peltier, W. R. and Andrews, J. T., 1976. Glacial-isostatic adjustment I. The forward problem. *Geophys. J. R. Astr. Soc.* 46: 605-646 .

Schoenbohm, L. M., Burchfiel, B. C., Liangzhong, C., 2006. Propagation of surface uplift, lower crustal flow, and Cenozoic tectonics of the southeast margin of the

Tibetan Plateau, *Geology* 34 (10): 813–816; doi: 10.1130/G22679.1

Shen, F., Royden, L. H., and Burchfiel, B. C., 2001, Large-scale crustal deformation of the Tibetan Plateau, *J. Geophys. Res.*, 106 (B4): 6793-6816.

Syvitski, J. P. M. and Saito, Y., 2007. Morphodynamics of deltas under the influence of humans. *Global and Planetary Change* 57: 261-282.

Walcott, R. I., 1972. Late Quaternary vertical movements in eastern North America: Quantitative evidence of glacio-isostatic rebound. *Rev. Geophys. Space Phys.* 10: 849-884.

4. The climate system

View of western hemisphere, with thin atmosphere visible on fringe of the Earth, and storm systems involved in the transport of heat within the atmosphere. Image from NASA (Visible Earth team) http://visibleearth.nasa.gov/.

We have seen that heat is moved about within rocks and soil near the surface of the earth largely by conduction. Any heat that arrives at the Earth's surface from below reflects very slow cooling of the Earth. In contrast, heat arrives at the Earth's surface by radiation from the Sun, at a rate (flux) that is greater by 5 orders of magnitude (roughly 1000 W/m^2 from the sun vs 0.04 W/m^2 from below). Once deposited on the earth's atmosphere, oceans and terrestrial surface through radiation, heat is moved about within the ocean and atmosphere by bulk motion of those fluids. In the first part of this section I address what controls the surface temperature of the Earth. In the second section I ask what pattern of poleward energy transport is required to allow the earth's climate system to remain steady.

What controls the surface temperature of the Earth?

Ice, clouds and oceans make the Earth seen from space largely blue and white. The Earth is unique in the solar system in that the water on or near the Earth's surface exists in solid, liquid, and vapor states. But why is Earth the "water planet"? Liquid water is stable in only a narrow range of pressure and temperature conditions. That the Earth's surface and atmosphere is found near this magic triple point of water requires

understanding what sets the surface temperature of the Earth. This can be cast first as a problem in conservation of energy, and then as a radiation problem.

Consider for now the Earth's surface and atmosphere as a simple box. This is our system, and we will quantify the energy that moves across the boundaries of this system. Later we will ask what drives energy about within the box, but that is later. It is safe then to make the statement of conservation of energy:

The rate of change of energy in the surface-atmosphere system = rate of energy delivered into it – rate of energy loss from it

In this section I will ask the simplest question of this system. I will dodge rates of change of the energy content in the system, and skip to the equilibrium or steady state situation. In this case, the left hand side of this word equation will be zero, and the rates of energy crossing the boundaries of the system must be equal: inputs must equal outputs.

Energy enters the box by two means: heat is gained by the system by both conduction of heat from the earth's interior, and radiation from the sun. Heat is lost from the system by radiation of the Earth's surface and atmosphere. I state this for completeness. In reality, the contribution from conductive cooling of the Earth (a heat flux of 41 milliwatts/m^2 is a rough global average) is trivial compared to that received from the Sun (as we will see this exceeds 1 kilowatt/m^2). Given this discrepancy of more than 5 orders of magnitude, I will ignore the conductive contribution, which converts the problem to a pure radiative balance. The incoming radiation from the very hot sun a long way away is balanced by the outgoing radiation of the much cooler Earth.

The inputs and outputs are illustrated in Figure 4.1. The radiative inputs are

$$E_{in} = Q_o(1-\alpha)\pi R_e^{\,2} \qquad (4.1)$$

where R_e is the radius of the Earth, Q_o is the radiative energy flux from the sun at the distance equal to the radius of the Earth's orbit, otherwise known as the solar constant, and α is the average albedo of the Earth. Note that the solar beam is intercepted only by an area equal to the area presented to the sun: that of a circle the radius of the Earth. The units of an energy flux, Q, are energy per unit area per unit time. As written in equation 1, the energy term therefore has units of energy per time – we should be calling it power. The term (1-α) represents the fact that not all energy arriving from the Sun is accepted into the atmosphere-surface system, but is instead rejected at the boundary; it is reflected away.

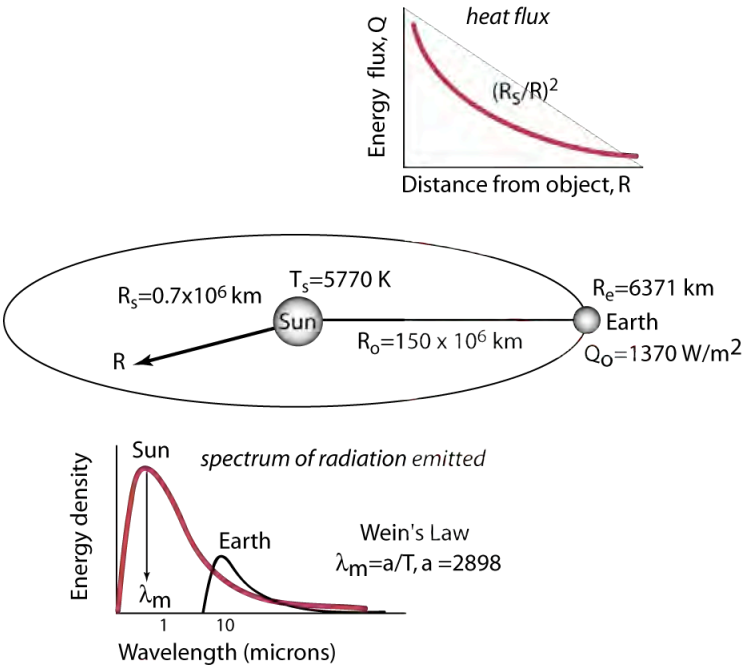


Figure 2.1 Elements of the radiation balance for the Earth. Middle: Earth shown in orbit around Sun, with radii of Sun, Earth and orbit, temperature of the Sun's surface, and the energy flux Q_o arriving at the outer edge of the Earth's atmosphere. Top: Inverse - square law for energy flux as a function of distance between the radiating source and the object. Bottom: Black body radiation spectra for both Sun and Earth, showing both the dependence of the peak wavelength on surface temperature, and of the total energy emitted (integral beneath the curves).

The radiative loss of energy to space takes on a similar form, modified only by the fact that the entire surface area of the Earth ($=4\pi R_e^2$) participates in the loss:

$$E_{out} = Q_e 4\pi R_e^2 \tag{4.2}$$

where Q_e is the radiative energy flux to space from Earth. With no further ado, we may now calculate this outgoing radiative flux:

$$Q_e = \frac{Q_o(1-\alpha)}{4} \tag{4.3}$$

The outgoing flux is tied directly to the incoming solar flux and to the albedo of the Earth. The 4 in the equation represents the difference between the area accepting energy from the Sun and that across which energy leaks back to space.

But how do we calculate the energy fluxes, and how do we get to temperature from an energy flux? It is after all the temperature of the Earth's surface that we wish to determine. Radiative flux from a black body obeys the following formula:

$$Q_s = \sigma T_s^4 \tag{4.4}$$

where σ is the Stefan-Boltzman constant ($5.67\text{x}10^{-8}$ watts-m^{-2}K^{-4}), and T is the surface temperature of the body emitting the radiation. Here I have written this general expression for the specific case of energy flux from the sun, Q_s, and the surface temperature of the sun, T_s. In a transparent (non-absorbing) medium, such as outer space, the same total energy must pass through spherical shells of increasing radius. The energy is conserved. This requires that

$$QA = Q_s A_s \tag{4.5}$$

where A is the surface area of the shell and Q is the heat flux through the shell at some distance centered around the Sun. At increasing distances from the Sun, the radiative energy flux Q must decline as the area of the sphere through which it is transmitted increases. Solving this for the energy flux at a specific distance from the radiating object (here the Sun, with radius R_s), we get

$$Q = Q_s \frac{A_s}{A} = Q_s \frac{\pi R_s^2}{\pi R^2} = Q_s \left[\frac{R_s}{R} \right]^2 \tag{4.6}$$

The energy flux falls off with the inverse of the square of the distance from the object (Figure 4.1). This is known as the inverse square law. If the arbitrary R in the above equation is replaced with R_o, the radius of the Earth's orbit around the Sun, then we can calculate the energy flux arriving at the top of the Earth's atmosphere responsible for warming up the atmosphere-surface system of the Earth. The distance R_o is roughly 150 million km (or 93 million miles, as you may have memorized in elementary school). The radius of the sun, R_s, is a little less than ¾ of a million km. We can now assemble the bits of the problem into one final equation. Equation 4.3 for the energy flux at the Earth's orbit may now be transformed to

$$\sigma T_e^4 = \sigma T_s^4 \left(\frac{R_s}{R_o} \right)^2 \left(\frac{1-\alpha}{4} \right) \tag{4.7}$$

Cancelling the σ and taking the fourth root results in the formula for the temperature of the atmosphere-surface system of the Earth:

$$T_e = T_s \left(\frac{1-\alpha}{4} \right)^{1/4} \left(\frac{R_s}{R_o} \right)^{1/2} \tag{4.8}$$

We must also estimate α, the albedo of the planet. The albedo is a measure of the reflectivity of the planet, 1 being perfectly reflective, 0 being perfectly absorbing. Only that portion of the sun's beam that is not reflected contributes to heating the atmosphere-surface system. The global average albedo of the Earth is roughly 0.35. (While outside the scope of this Little Book, I note that the early estimates of this number come from measurement of the brightness of the new Moon, which is illuminated not directly from the Sun but from the light reflected from the Earth.)

We can now see that the mean temperature of any planet is set by the temperature of the Sun, by the radius of its orbit around the Sun, and by the average albedo of the planet when seen from space. Before leaving this simple equation, inspect it. This is always a good practice, as you want to assure yourself that you have not made any serious errors in the math, and you want to milk the equation for as much insight into the system behavior as possible. You may wish to know how sensitive the temperature is to the variables contained in the equation. The temperature of the planet goes linearly with the temperature of the sun's surface; if the temperature of the Sun goes up 1%, so does the planetary surface temperature. The planetary surface temperature declines as the radius of the orbit increases; the dependence is as a square root rather than linear. The sensitivity of the planet surface temperature to the albedo is even smaller, going as the fourth root.

Let's now plug in numbers. The temperature of the sun is roughly 5770K. How do we know this? No one has been up there to measure it. We know it from measurement of the peak wavelength, λ_{max}, of the light emitted, and the relationship between this and the surface temperate, T, known as Wein's Law: $\lambda_{max} = a/T$. The constant $a = 2898$ if the temperature is in Kelvins, and the wavelength is in microns (10^{-6} m). As the temperature of the radiating body increases, the peak wavelength of the radiation emitted declines. Given these numbers, we calculate first that the energy flux at the top of the Earth's atmosphere is 1372 W/m^2. This is the solar constant, which you can see is more than five orders of magnitude greater than the conductive heat flux from the Earth's interior. Finally, we can also calculate that the mean temperature of the Earth's surface should be about 255K, or about –18°C. Clearly this is not the case. If it were, the Earth's surface would be frozen solid, and water would not exist in either liquid or vaporous forms. We can see from the plot of expected blackbody surface temperatures of the terrestrial planets (Figure 4.2) that except for Venus the actual and theoretical temperatures are not really too far off. But being as little off in the case of Earth has huge consequences. What have we done wrong in the calculation, or what is missing?

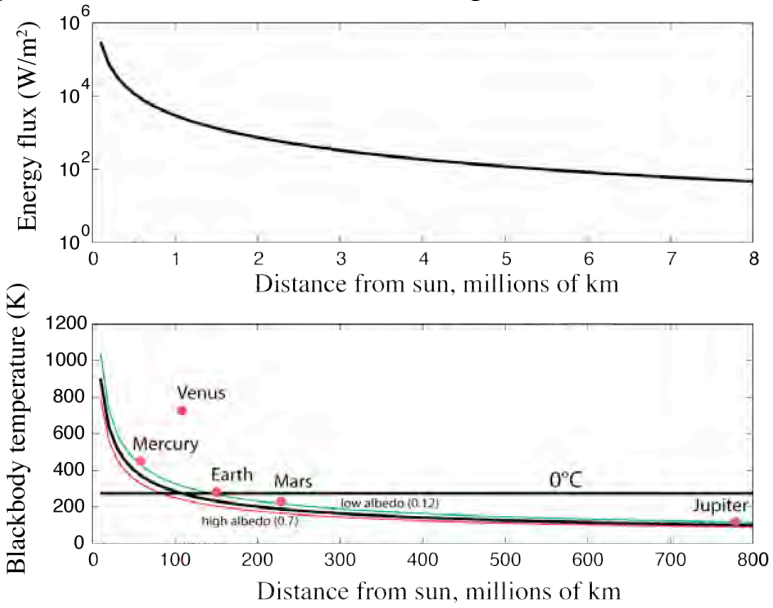


Figure 2.2 Calculated solar radiation arriving at the orbits of planets out to Jupiter (top plot, showing inverse square law) and the expected surface temperatures of these planets (bottom) based upon the assumed albedos (labeled). Dots correspond to measured blackbody surface temperatures. Intermediate albedo case (0.3) corresponds roughly to Earth's present measured albedo. That Earth lies above the expected temperature associated with its albedo indicates the role of the greenhouse gases in its atmosphere. Only Venus lies outside the expected envelope, owing to its extreme greenhouse gas atmosphere.

First of all, we must admit to the sloppiness of this statement. The surface temperature of the Earth is measured at the base of the system we have been considering. While this is indeed what most interests human and most other biological occupants of the Earth, it is not what outer space "sees". As measured from outer space, the radiative temperature or black body temperature of the earth would indeed be about 255K; it would see radiation coming from mid-atmospheric levels, which are much cooler than those on the surface of the Earth. In other words, we must begin to acknowledge that the "system" we are

assessing has a vertical dimension and that the temperature in it is not uniform. We will see in a subsequent section that this energy lacks lateral uniformity as well, and that this lack of uniformity drives the entire climate system.

Second, we have not taken into account the chemistry of the Earth's atmosphere, which contains gases, some of them in trace amounts, that are excitable by particular slices of the spectrum of radiation emitted by the Sun and the Earth. The excitation of the gases leads to absorption of some of the energy and alters the simple radiation balance we have just calculated. Many of the gases of concern are molecules consisting of two or more atoms bound by bonds of particular lengths and strengths. It is the vibrations and rotations of atoms in these molecules that are excited at specific wavelengths, which effectively consume some fraction of the energy in that wavelength band of the radiation. One can see from Figure 2.3 how gases like H_2O, CO_2, CH_4, O_2 and O_3 absorb energy in specific wavelength bands. Many of these wavelengths are around 10 microns, in what we call the infrared band. Note also that light of wavelengths around 0.5 microns, the visible spectrum, is not absorbed by these same gases. Radiation in most of the visible spectrum is not affected by the atmosphere.

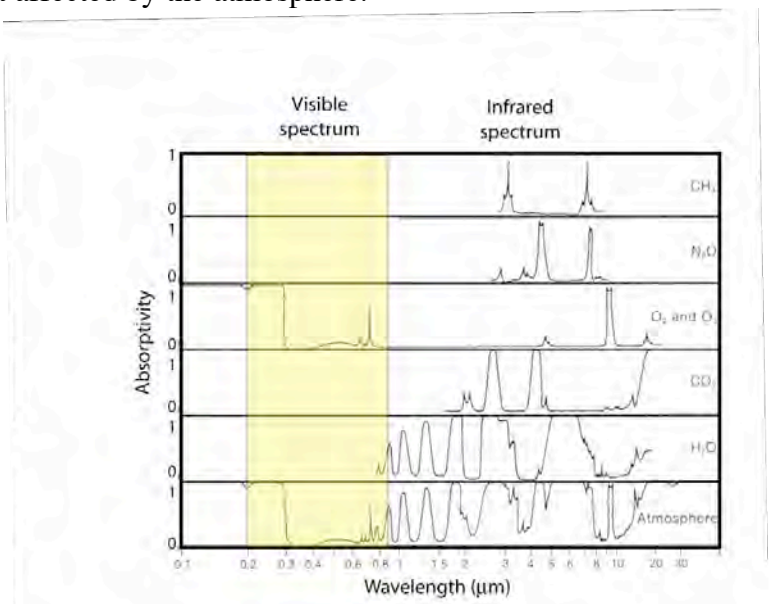


Figure 2.3 Absorption of radiation as a function of wavelength for various molecular components of earth's atmosphere and for the atmosphere as a whole. (from Fleagle and Businger, An Introduction to Atmospheric Physics, 1963, Academic Press.)

Now recall Wein's Law, which implies that a body with a temperature of around 0°C will radiate at about 10 microns, while a body with the Sun's temperature will radiate at about 0.5 microns. First, it is not a coincidence that wavelengths around 0.5 microns are called visible light – humans and most other animals have evolved to utilize this portion of the spectrum. Second, it is in the longer (infrared) wavelength band, that being emitted by an Earth with a surface temperature around 15°C, that is most efficiently absorbed by the molecules listed above. It is therefore this outgoing radiation that is intercepted by the atmosphere and used to excite atmospheric gases. The higher the concentration of these gases, the more of this wavelength radiation will be absorbed, and prevented from getting to outer space to balance the incoming radiation. Therefore, the Earth's surface will have

to be yet hotter to produce enough outgoing radiation to balance the incoming solar radiation. Hence the surface of the Earth is warmer than we calculated when we ignored the atmosphere. These so-called greenhouse gases are therefore very important for allowing the surface temperature of the Earth to be a b it above 0°C, rather than a bit below it. And it is presently the human alteration of the concentrations of these same gases that is forcing the global change of climate in the last hundred years.

Redistribution of heat by the climate system

We now embrace the non-uniform distribution of energy input to the atmosphere-surface system of the earth, and ask what pattern of poleward energy transport is required to allow the earth's climate system to remain steady. The basic problem is that much more heat arrives at low latitudes than at high latitudes due to the curvature of the earth. In the absence of any transport, this would heat up the equator and cool off the poles to temperatures that are locally in radiative balance. This is certainly not the case (the equator would be much hotter than it is, the poles much colder than they are). Again, we can start from a conservation equation, this time one for energy. The word picture is:

The rate of change of energy = incoming energy – outgoing energy

There are two flavors of energy transport, one by radiation (both incoming shortwave and outgoing longwave), and advection of heat in either the ocean or the atmosphere (Figure 1). If the annual mean temperature at any given location on the earth is to remain steady over long time scales, then there must be a balance between all of the inputs and outputs.

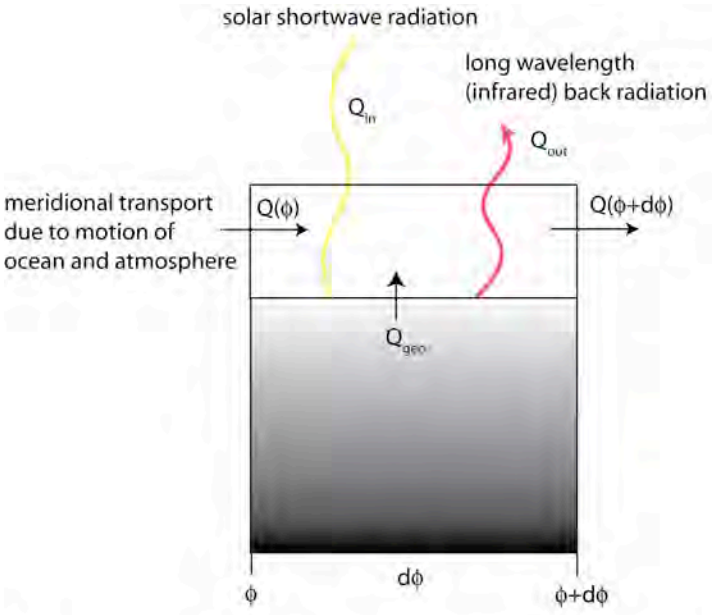


Figure x. Elements of the climate energy system involved in the energy balance at a given latitude. Combination of incoming shortwave solar radiation and outgoing longwave radiation sets net radiative input or output. Geothermal input is trivial compared to the other fluxes. If the temperature in the latitudinal band is to stay steady, any net radiation input or output must be balanced by gradients in meridional (N-S) transport accomplished by the oceans and atmosphere.

The sum of incoming shortwave solar radiation and outgoing longwave infrared radiation sets the pattern of net radiation on the earth. The incoming radiation, or more formally, the flux at a given latitude, depends upon the inclination of the earth's surface relative to the sun averaged over a year. The instantaneous solar radiation is set by the latitude and the tilt of the Earth's axis of rotation:

$$Q_{in} = (1-\alpha)Q_o \cos(\phi + \theta) \tag{4.9}$$

where Q_o is the solar constant, or roughly 1372 W/m^2, α the albedo of the Earth (about 0.35), and θ is the tilt of the earth's axis relative to the plane of the ecliptic. This must be summed over the year, meaning that the effective tilt rotates through from -23 to 23° over 365 days. Another element of reality is that the albedo of the Earth is not uniform. The albedo is affected by the differing albedos of clouds, land, and ice, none of which are uniformly distributed on the globe. A net positive radiation input occurs in low latitudes (Figure 4.4), and a net negative radiative loss occurs in high latitudes. If this were all that were happening, the poles would cool due to this continued net loss of heat by radiation, while the equatorial zone would warm. That the Earth's mean annual temperature is roughly steady (yes, the earth is warming, but it is warming slowly and due to other processes we are ignoring here) implies that the misbalance in radiative heat must be offset by other heat transport processes. This is indeed the role of (the duty of) the climate system: the oceans and atmosphere must accomplish the transport of heat by advection, the transport of heat by a fluid in motion (see a later chapter on Advection). Before exploring in detail the means by which the heat is transported, i.e. how these duties are partitioned between atmosphere and oceans, we can calculate just how much heat must be transported by the climate system for a heat balance to be assured.

The basic equation is one of conservation of heat. We will see this in several other contexts, but for now consider a system in which both radiative and advective transport occurs. The basic statement of conservation of heat is: the rate of change of heat at a point on the Earth is equal to the inputs of heat from all sides of that parcel and the losses of heat from that parcel. Put mathematically, this is:

$$\frac{dE}{dt} = q_{in} - q_{out} = Q_{geo}A(\phi) + Q_{rad\,net}A(\phi) + Q_{o+a}(\phi) - Q_{o+a}(\phi + d\phi) \tag{4.10}$$

where the 1st term on the right hand side represents the input from geothermal heat flux, the 2nd term represents the net radiative input or output from a latitudinal band with area A, and the other terms represent the meridional (N-S) poleward transport of heat by ocean and atmosphere across latitude ϕ and across $\phi+d\phi$. We can then evaluate the steady case in which $dE/dt = 0$. Setting the left hand side of the equation to zero reveals that the gradient of the oceanic and atmospheric heat flux must be equal to the net radiation input plus any geothermal input. As the latter is indeed trivial relative to the flux from the sun (tens of milliwatts/m^2, *vs* hundreds of watts/m^2), we will ignore it hereafter. The balance is illustrated in Figure 4.5 for examples in both equatorial and polar latitudinal bands.

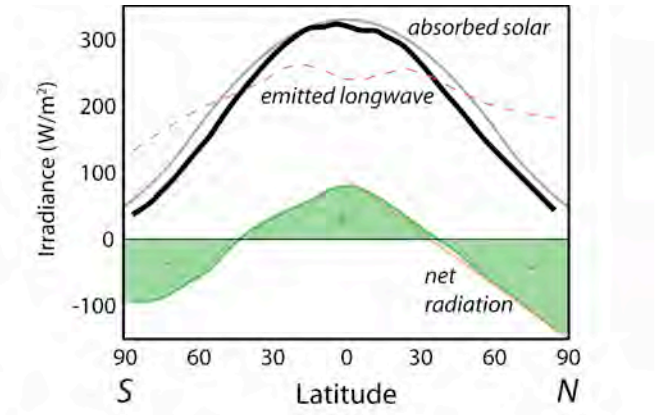
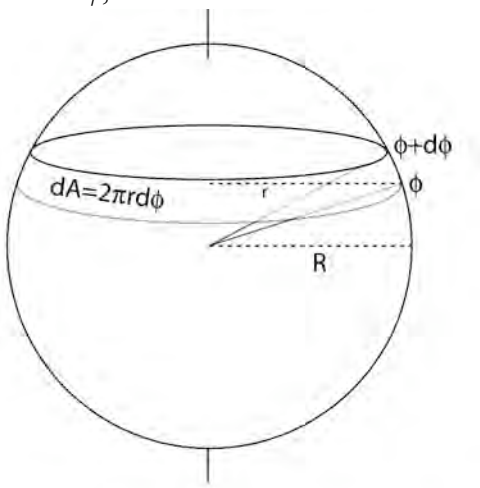


Figure x.x Absorbed, emitted and net radiation averaged in latitudinal bands (dark line). Also shown is a calculation of the average incoming solar radiation, as discussed in the text (smooth gray line). Note the net gain of heat from radiation in low latitudes, and the net loss at high latitudes. [after Hartmann, 1994, figure 2.12] Red line behind net radiation curve shows simplified pattern used in calculation of the poleward transport needed to balance the heat.

We can therefore visualize, before we actually perform the calculation, what the shape of the transport profile must look like: it must increase with latitude in the equatorial zone; it will reach its maximum where the net balance is zero, and will then decline toward the poles. Formally, the transport of heat required of the climate system (the oceans and atmosphere) is therefore the integral of the net radiation balance. The integration requires that we take into account the area in each latitudinal band, which varies with latitude (Figure 4.6). An increment of area in the integration, $dA=2\pi r d\phi$, involves the changing local radius of the Earth, $r = R\cos\phi$, where R is the true radius of the Earth.



Incremental area dA used in the integration of radiation budget on Earth. As the local radius to the axis of rotation, r, declines as latitude ϕ increases, the increment of area declines as well.

The integral is therefore

$$Q_{transport} = \int_0^{\phi} Q_{rad\,net}\, 2\pi R\cos\phi d\phi = 2\pi R\int_0^{\phi} Q_{rad\,net}\cos\phi d\phi \qquad (4.11)$$

This integration is shown in Figure 4.7. Indeed, it shows the features we expected. In addition, the measured transport by the oceans and the atmosphere is plotted. The most recent summaries of these transport terms suggest that the oceans dominate in transporting heat poleward at low latitudes, but become minor players at higher latitudes, where the atmosphere reigns.

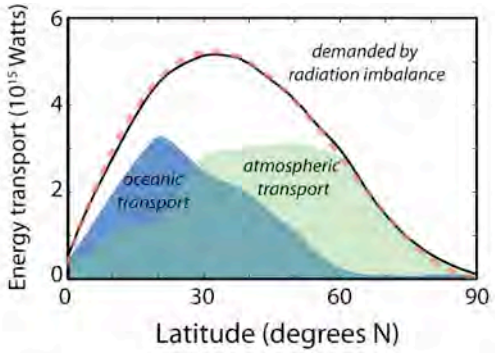


Figure x.x Transport of energy in petawatts demanded by the imbalance of net radiation, and the patterns of transport accomplished by ocean and atmosphere components of the climate system. [after Hartmann, 1994, Figure 2.14; adapted from Haar and Oort, 1973] Smooth red dashed curve is poleward transport calculated assuming a simple linear decline in net radiative balance depicted on Figure x.x.

I have shown only the pattern in the northern hemisphere. That in the southern hemisphere would be similar, although not quite as simple given the nonlinear distribution of the net radiation shown in Figure 4.5.

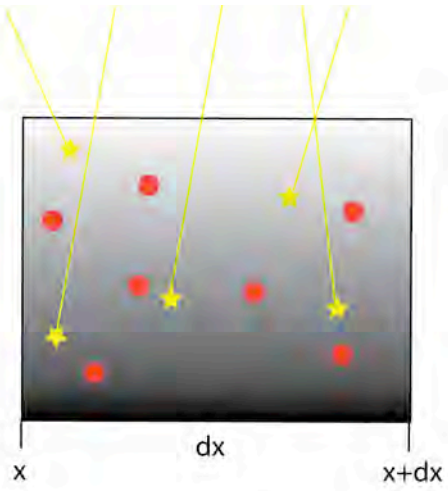
In these first two chapters we have explored heat transport and its relevance to broader Earth system issues. In both cases we made substantial progress by appeal to the conservation of heat, and then worried through the formulation of the specific transport processes involved. We have seen examples of heat transport by conduction, radiation and advection. In this last problem, dealing with the climate system, we were also forced to work in a spherical geometry, which altered the details of the integration of the conservation equation to arrive at that for the fluxes.

5. Dating using cosmogenic radionuclides

Sampling of bedrock in Sam Ford Fjord, Baffin Island, for cosmogenic ^{10}Be in order to quantify the timing of last glacial ice retreat from the fjord.

In geomorphology we often obtain rates of various surface processes by dating geomorphic features in the landscape. For example, the rate of river incision into bedrock can be obtained by dating bedrock (so-called strath) terraces at a measured height above the modern stream. In certain special settings, river incision can also be constrained by dating caves in the valley walls that were once occupied by the river as it sliced through the rock. Recently, dating of both of these features has been achieved by using the concentrations of cosmogenic radionuclides. These are born of cosmic ray interactions with atoms in minerals in surficial rock and soil, and they decay by radioactive fission. It is the concentration of these rare nuclides in soil and rock at the earth's surface that we use to deduce timing in the landscape. Given that nuclides are both produced and decay, we must take both processes into account in crafting an equation that governs the evolution of their concentration through time. The word picture for the system is (Figure 1):

Rate of change of number of atoms in the box = production of atoms in the box – decay of atoms in the box



Production and decay of cosmogenic radionuclides. Cosmic rays interact with atoms in surficial materials to produce new nuclides (stars). Radionuclides decay (red) with probability set by the decay constant. The rate of change of the concentration of radionuclides is therefore set by any mismatch between the birth and death rates.

If we know the volume of our box, the control volume, we can transform the number of atoms in the box to the concentration of atoms in the box, C. This can be cast as:

$$\frac{dC\rho dxdydz}{dt} = P\rho dxdydz - \lambda C\rho dxdydz \tag{5.1}$$

where C is the number of atoms per unit mass of quartz (atoms/gm qtz), ρ the density of the rock (kg/m^3), P the production rate of new nuclides per unit mass of quartz (atoms/gm qtz – yr), λ the decay constant for the nuclide (1/yr), and $dxdydz$ the volume of the box. This can be converted to an equation for the rate of change of nuclide concentration per unit mass of material (say atoms ^{10}Be per gram of quartz, something we measure) by dividing by the mass, $\rho dxdydz$, leaving

$$\frac{dC}{dt} = P - \lambda C \tag{5.2}$$

We can anticipate the shape of the solution by inspecting this equation. At early times, low concentrations of nuclides will result in only small contributions from the second term, and the rate of growth of the concentration should be steady, set by the production rate P. As concentration increases, however, the decay term will grow, forcing the rate of change of concentration to decline. Ultimately, the concentration should become high enough that the production and decay terms balance, and a steady concentration should then be achieved. This would occur when $C = P/\lambda$. Indeed, the solution to this equation shown in Figure 2 reveals just this behavior:

$$C = \frac{P}{\lambda}\left[1 - e^{-\lambda t}\right] \tag{5.3}$$

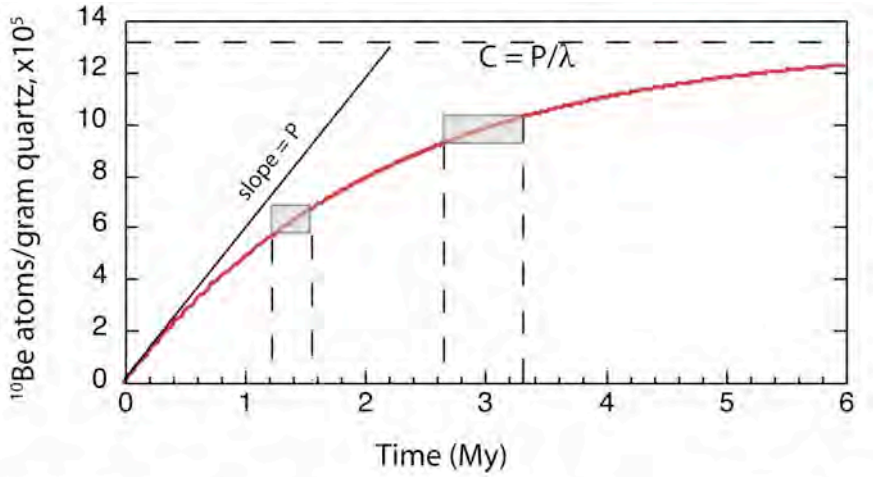


Figure x.x. Approach of ^{10}Be concentration toward secular equilibrium. Initial slope is set by the production rate. The time scale for the approach is set by the decay constant, the inverse of which, the "mean life", is 2.18 Ma. Uncertainty in measurement of [^{10}Be] (heights of gray boxes) leads to uncertainty in age (dashed lines) that increases with the age. At ages of several times the mean life, the concentration is no longer a good clock.

The early growth rate (slope on the plot) is set by the production rate, P. The asymptotic concentration is P/λ, at which point the system is said to be in "secular equilibrium". This plot also reveals the time scale over which the concentration of a radionuclide will be useful as a clock. Once the concentration gets so close to this asymptotic value (P/λ) that it is no longer changing significantly with time, the concentration can no longer be used to measure time. The real limiting time for the method is dictated both by the characteristic time of decay, (or the "mean life", $1/\lambda$), and by our ability to measure concentrations (see Figure 2, with error bars).

Examples

Now let's turn to a few examples of the use of this method. This will force us to acknowledge several important features of the system. I urge the interested student to explore this system further in extensive recent reviews by Granger and Muzikar (2001), by Cerling and Craig (1994) and by Phillips and Gosse). We are just scratching the surface here. First consider the history of radionuclide concentration a rock as it is exhumed. We will ultimately sample this rock and convert the measured concentration in the surface sample to deduce the rate at which it was exhumed. We again employ the master equation (2) for the rate of change of concentration but now take into account that the production rate P will evolve as the rock parcel moves closer to the surface. Measured production rate profiles show that the production rate falls off exponentially with distance into any material, the length scale that scales the decline being dependent on the density of the material:

$$P = P_o e^{-z/z^*} \tag{5.4}$$

where P_o is the production rate at the surface, and z^* is the length scale characterizing the rate of decline of production rate with depth. The length scale $z^* = \Lambda / \rho$, where Λ is the mass interaction scale, and characterizes the interaction rate of the cosmic rays with any

material [=] M/L^2, and ρ is the density of the material [=] M/L^3, here rock. If we assume that the exhumation rate is steady, at a rate e, then the conservation equation becomes

$$\frac{\partial C}{\partial t} = P_o e^{-(-\varepsilon t)/z^*} - \lambda C \tag{5.5}$$

where I have taken time, t, to be negative in the past, and 0 at the present. We can work a simple end-member case. Assuming that the exhumation rate is high enough that the newly created nuclides do not have much time to decay before the rock reaches the surface, we may ignore decay. The ordinary differential equation becomes

$$C = \int_{-\infty}^{0} P_o e^{-(-\varepsilon t)/z^*} dz \tag{5.6}$$

This may then be integrated to yield

$$C = \frac{P_o z^*}{\varepsilon} \tag{5.7}$$

Let us inspect the solution to assure that it behaves as we expect it. First, is it dimensionally correct? Yes. The concentration is the product of a production rate with a time scale. Here the relevant time scale is z^*/ε. As the exhumation rate increases, the rock spends less time in the production zone (within a few z^* of the surface), and the resulting concentration at the surface declines.

The fuller expression for the surface concentration, taking decay of nuclides into account, is

$$C = \frac{P_o}{(\varepsilon / z^*) - \lambda} \tag{5.8}$$

Inspection of this equation reveals that in the limiting case of very low λ (long half life), or very rapid erosion, high ε, it becomes equation 5.7. In other words, as long as decay is minor over the relevant time it takes to erode through the region in which production occurs, we are safe with equation 5.7. This method has been used to determine bedrock erosion rates at many sites around the world (e.g., see an early summary in Bierman, 1994). Examples of bedrock lowering rates in the alpine and desert areas of western North America reveal very slow rates of exhumation that are all only a few microns per year, or a few meters per million years.

Armed with this knowledge, now consider the two cases I mentioned at the outset in which CRNs are used to date river terraces and caves. In both cases we measure the concentration of radionuclides in sediment.

Depositional surfaces

More common than bedrock surfaces are sediment-capped surfaces, some simply mantling bedrock surfaces with several meters of sediment (marine terraces, fluvial terraces, pediments), others filling the landscape more deeply (alluvial fans, fill terraces, moraines). These too, if sampled well away from edges of the surface that might be experiencing significant modification by either erosion or deposition, can be dated using cosmogenic radionuclides. There is, however, a problem. All of the sediment that accumulated to form the deposit came from elsewhere, therefore spent some time within

a few meters of the earth surface, and therefore had the chance to accumulate cosmogenic radionuclides. This "inherited" component can be large compared to that obtained while sitting on or in the surface we wish to date. In addition, it will inevitably vary from one grain to another. Each has its own history of exposure, and hence will arrive on the surface with its own inheritance. This gives rise to considerable scatter if one dates single cobbles on a surface. How do we see through this problem of inheritance? The solution is expensive. A method has evolved in which one measures the concentrations of several samples (hence the expense) in a vertical profile into the surface, each sample being an amalgamation of equal mass from many clasts. The amalgamation process effectively averages out the inheritance, the stochastic component in the concentration, to which has been added a deterministic component that varies systematically with depth due to the decline in production with depth. The expected profile is a shifted exponential (Figure 3)

$$C = C_{in} + P_o T \, e^{-(z/z^*)} \tag{5.9}$$

The shift being is the mean inheritance of the sample, C_{in}, and the exponential portion of the profile represents the post-depositional accumulation of nuclides. Once the shift is constrained, which is best accomplished with one or more samples from several meters depth, the remaining exponential can be solved for the time since deposition, T.

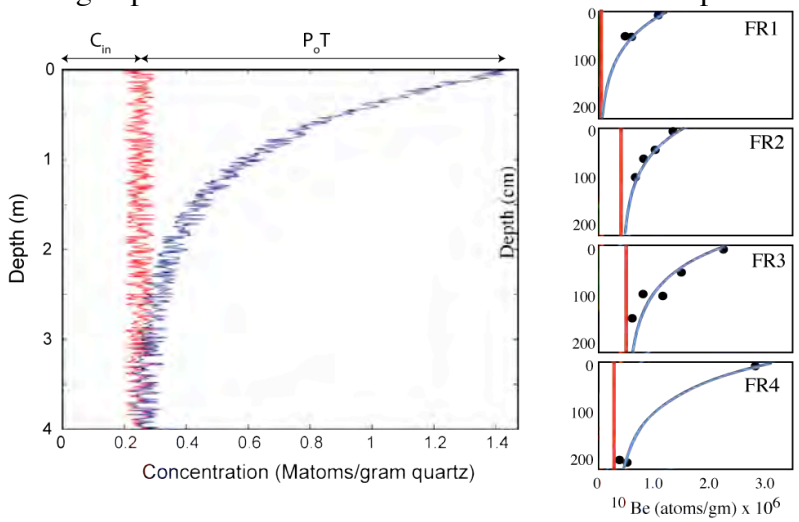


Figure x. Left: Numerical calculation of the expected shifted exponential of cosmogenic radionuclides in a deposit. The red profile reflects the inheritance of radionuclides, averaged over many grains. The blue profile is the total concentration, which includes both inheritance and post-depositional grow-in of nuclides. Scatter at any depth largely reflects variaton in the inheritance, as post-depositional production should be determined well by the depth into the deposit (i.e. it is deterministic). Right: Profiles of ^{10}Be from the Fremont River terraces of central Utah constructed from 3-4 samples, showing best fits of shifted exponential profiles. The ages of the terraces increase from FR1 through FR4, with increasing height above the river. The inheritance appears to vary from terrace to terrace. Adapted from Repka et al., EPSL.

This method has been used to date both marine terraces and fluvial terraces. In a fluvial example, the terraces that bound both the Wind River in Wyoming and the Fremont River in Utah show this shifted exponential well (Figure 3). As you might imagine, the inheritance itself can be used to constrain rates of exhumation in the landscape contributing sediment to the river. If we have gone through the expensive exercise of documenting the inheritance, we might as well use it!

Caves

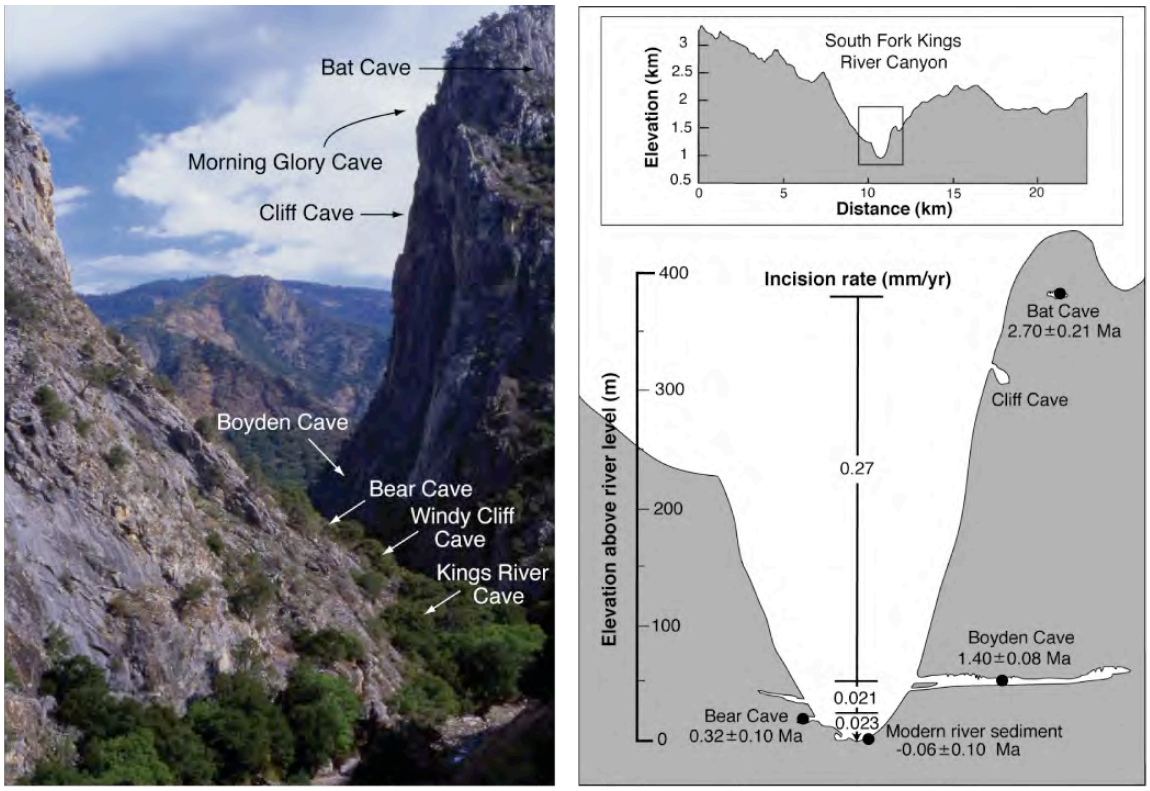
We can even date cave deposits using cosmogenic radionuclides. The method is based upon the fact that both ^{10}Be and ^{26}Al are produced in the same materials (quartz), and that ^{26}Al decays roughly twice as fast as ^{10}Be. If quartz-rich sediment is washed into the cave by the river responsible for the dissolution of the rock, and is then sequestered far enough underground to prevent further production of radionuclides within the sediment, then the differential decay of the nuclides results in decay in the ratio of their concentrations. This ratio can then be used as a clock. Mathematically, the ratio may be expressed as

$$R = \frac{N_{oBe} e^{-t/\tau_{Be}}}{N_{oAl} e^{-t/\tau_{Al}}} = R_o e^{-t/\tau_R} \tag{5.10}$$

where

$$\tau_R = \frac{\tau_{Be}\tau_{Al}}{\tau_{Be} - \tau_{Al}} \tag{5.11}$$

Here R_o is the initial ratio (most often likely the production ratio, 6.1), and the mean life of the ratio, τ_R, is 1.9 Ma. In principal, the ratio can therefore be used to date sediment as old as several of these time scales, or roughly 5 Ma. This is the essence of what has become known as the burial age method (Granger et al., 1997). As with most methods, it has several advantages, but a few drawbacks. One principal advantage is that the method is immune to any temporal variations in the production rate, as it is simply decay that is being measured. The necessary depth of burial is a couple tens of meters, deep enough to prevent production by both spallogenic and muogenic processes; most caves are at least this deep. The cave sediment must be quartz-rich, meaning that somewhere in the headwaters of the cave there must be a quartz-rich source of sediment. As it is only the inherited nuclides that we are counting with this method, the concentrations of cosmogenic radionuclides starts out small, and decays from there, halving every 2 Ma. Finally, we must know well the initial ratio of the nuclides, R_o. That we know this to only roughly 10% (6.1 being the presently accepted value of the ratio of production rates, R_o), the ratio can only be known to that level. This places a lower limit on the utility of the method. Enough decay must have occurred to lower the ratio by more than 10% in order to have any confidence in the age. Practically, this means that caves less than 300 ka cannot be dated reliably at the moment. Nonetheless, that we can date these voids beneath the earth's surface at all is remarkable, and opens up their use as a means of documenting the rates of incision of the streams responsible for them.



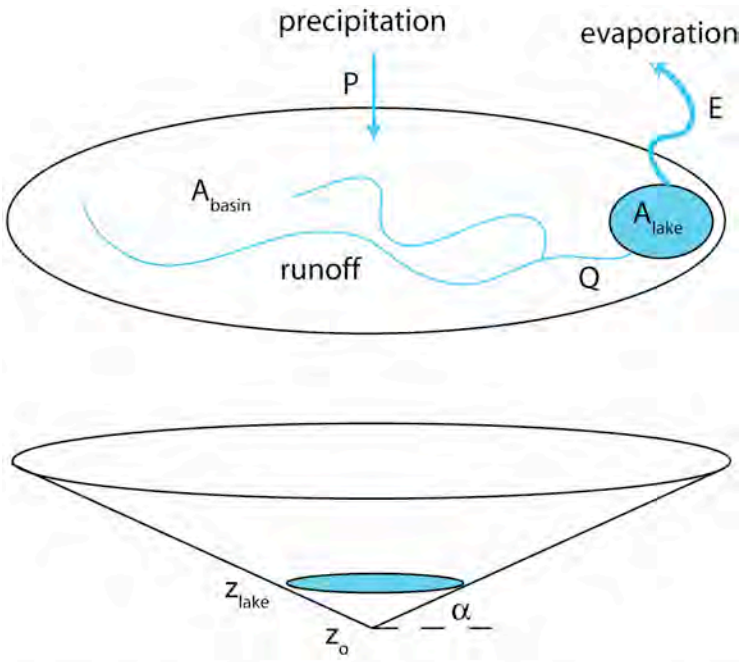
River incision history derived from dated caves in the Kings Canyon, western Sierra Nevada, California. Left: Photo of inner gorge of South Kings Canyon, looking upstream, with caves in valley wall labelled (courtesy of Greg Stock). Right: Broad view cross-canyon topographic profile (upper) and close-up (lower) with ages of caves derived from burial age method. It appears that rapid incision from 2.7 - 1.4 Ma was followed by much slower incision. After Stock et al., 2004&2005.

The examples shown are but the tip of the cosmogenic iceberg. Many other applications can be found in a growing literature on the topic. These include documentation of the basin-averaged rate of exhumation, dating of stratigraphy, and use as a tracer in mixing models of sediment in littoral settings.

6. Lakes

Small lake in Gore Range, central Colorado.

Lakes reflect geological accidents. Many lakes require that some obstacle pond up a river en route to the sea. This obstacle could be a moraine, a landslide, or a body of ice (ignoring for the moment both human and beaver activities). Such damming could reflect tectonic emplacement of a barrier by faulting, creating a closed basin with no outlet to the ocean. We would like to know what sets the level of a lake. We expect the lake level will reflect some aspect of the climate, making ancient lake levels proxies or records of climate in the past. But in order to be more quantitative, to turn an ancient lake shoreline elevation into a quantitative proxy of past climate, or to predict how changes in climate might affect future lake levels, we must be rigorous about what determines the lake level. The failure of natural dams that hold back lakes can also play an important role in the geomorphology of the rivers downstream by producing floods far larger than those that the weather can produce. In this section we will ask two questions: how climate sets lake level in a specific lake basin, and what sets the peak discharge from a lake when its natural dam breaks.



Lake water balance. At steady state, the inputs of water to the lake through direct precipitation onto the lake, and runoff from the landscape are balanced by evaporation from the lake surface.

Consider the lake basin illustrated in Figure 1. The lake with area A_{lake} occupies some fraction of the basin whose total area is A_{basin}. We begin with the relevant statement of conservation, here of lake volume.

Rate of change in storage of water = Rate of inputs of water – rate of outputs of water

This change in storage will be reflected in the change in lake level, which is our most easily measured quantity, and the most important quantity for those living on the edge of the lake. Let's look at each of the terms in this equation one at a time.

Rate of change of storage Rate of change of something means the partial derivative of that quantity. Here the quantity is the storage of water, by which we mean the volume of water in the lake, V. Therefore, the left hand side of the equation for conservation is $\frac{dV}{dt}$.

Inputs Water arrives at the lake either as direct precipitation onto the lake surface, or as runoff from the basin contributing to it. As each term in the expression must have units of volume per time, we write that the direct precipitation adds water to the lake at a rate of PA_{lake}. Runoff from the landscape we can write as the sum over all rivers discharging into the lake, Q, where Q is reported in m^3/second, or its equivalent over a year. Therefore, the inputs are $PA_{lake} + Q$.

Outputs A lake can lose water in several ways. These include leakage in groundwater, spill over the top of the dam out of the basin, and evaporation (again ignoring human

withdrawals). The major output from a closed basin is through evaporation, which can be written as EA_{lake}.

We now assemble these terms into our final expression for conservation of lake volume:

$$\frac{dV}{dt} = PA_{lake} + Q - EA_{lake} \tag{6.1}$$

Following the strategy of doing the simplest exercises first, we ask what is the level of the lake when it is in steady state. This allows us to ignore any rates of change, in which case the left hand side of the equation vanishes. At steady state, the inputs must equal the outputs, meaning that the sum of the two must be zero. It sounds quite simple. But imagine that I look up values of P and E from the climatology of a region, and Q from a water resources report, or USGS web site. I can isolate A_{lake} on one side of the equation, and solve for it. But I am not interested in the area of the lake. Instead I wish to know the level of the lake, z_{lake}. I need more information. I must know how to translate lake area into lake level. This translation comes from the topography of the basin. More specifically, it requires we know the "hypsometry" of the basin, its area as a function of elevation, $A(z)$.

A steady conical lake

Imagine first a geometrically simple basin in which the topography is an inverted cone with a spreading angle α. The area of the base of the cone is that of a circle whose radius is $R=z/\tan(\alpha)$. The area of the lake is then related to the lake level through

$$A_{lake} = \pi R^2 = \pi(z_{lake} / \tan\alpha)^2 \tag{6.2}$$

We may now isolate z_{lake} on the left hand side and solve for the lake level as a function of the local climatic variables, P and E. In order to simplify the exercise by one more step, consider that we can cast the expected discharge of a river as the product of the effective precipitation with the basin area. Here the relevant basin area contributing runoff would be that not covered by the lake, or A_{basin}-A_{lake}. The resulting formula for inputs would then be

$$PA_{lake} + P_{eff}(A_{ba\sin} - A_{lake}) = PA_{lake} + \beta P(A_{ba\sin} - A_{lake}) \tag{6.3}$$

where β is a runoff coefficient, the fraction of precipitation that participates in runoff. The full water balance for the basin is now

$$EA_{lake} = PA_{lake} + \beta P(A_{ba\sin} - A_{lake}) \tag{6.4}$$

Using the approximation that the lake area is small relative to the basin area (this is the case in most basins), this reduces to

$$EA_{lake} = PA_{lake} + \beta PA_{ba\sin} \tag{6.5}$$

Solving for the area of the steady state lake, this becomes:

$$A_{lake} = \frac{\beta P}{E - P} A_{ba\sin} \tag{6.6}$$

Note that before we proceed, the expected lake area is indeed small relative ot the basin. In the arid western North America, for example, a common annual evaporation might be several m (call it 2 m), while the annual precipitation might be a few tens of cm (call it 20 cm). Even if the efficiency of converting precipitation to runoff is 100% (β=1.0), the

lake area becomes 0.2/(2-0.2) = 0.11 of the basin area. Finally, converting the lake area to lake level through the hypsometry, we have a formula worth working with:

$$\pi(z_{lake} / \tan\alpha)^2 = \frac{\beta P}{E - P} A_{ba\sin} \tag{6.7}$$

Solving for the lake level, this becomes:

$$z_{lake} = \sqrt{\frac{(\tan\alpha)^2 \beta P}{\pi(E - P)} A_{ba\sin}} \tag{6.8}$$

Given all these steps, it is worth checking the solution to make sure it makes sense. First, the final formula had better have the correct dimensions – here, a length scale. The right hand side is indeed a length scale, because π, α and β are dimensionless, P and E have the same units (hence their dimensions cancel out), and we are therefore taking the square root of an area. Does it have the right sign (we expect z_{lake}>0)? As E-P is always > 0, and all other variables are positive, we are ok there as well. Now does it behave as we expect it to as the basin changes? As the basin area increases, the contributing area increases and we ought to have more water delivered to the lake. As the basin slope, α, increases, the basin becomes a narrower cone, and the same lake area is not achieved until the lake is deeper. So this makes sense as well. It appears then that for a given basin the lake level goes as

$$z_{lake} \sim \sqrt{\frac{\beta P}{(E - P)}} \tag{6.9}$$

Another way to formalize this relationship is to non-dimensionalize the equation, isolating those variables that control the basin geometry on the left and those that control the climate on the right:

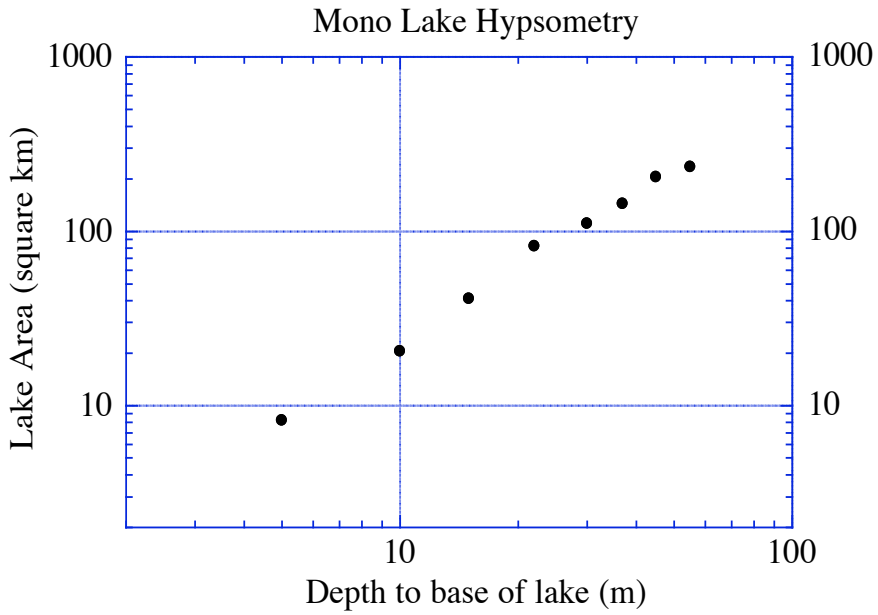
$$\frac{z_{lake}}{\tan\alpha \sqrt{A_{ba\sin} / \pi}} = \sqrt{\frac{\beta P}{(E - P)}} \tag{6.10}$$

Real basin geometries

Of course, real basins are not simple cones. Digital topographic maps, however, ease the construction of hypsometries, so that our assumption of a conical lake may be relaxed. Figure 2 displays the hypsometric curve for Mono Lake basin in eastern California. Shown in log-log space, the *A(z)* data reveal a straight line whose slope is about 1.4. This means that the area goes as the 1.4 power of the lake level (as opposed to the 2nd power as in our conical lake). In other words,

$$A = a(z - z_{base})^{1.4} \tag{6.11}$$

where z_{base} is the elevation of the base of the lake basin, and a is a constant we can evaluate by the intercept on the plot. The same analysis can be performed on this real basin. One can still derive the relationship between lake level and climate variables using the *A(z)* function in equation 13 in place of equation 2.



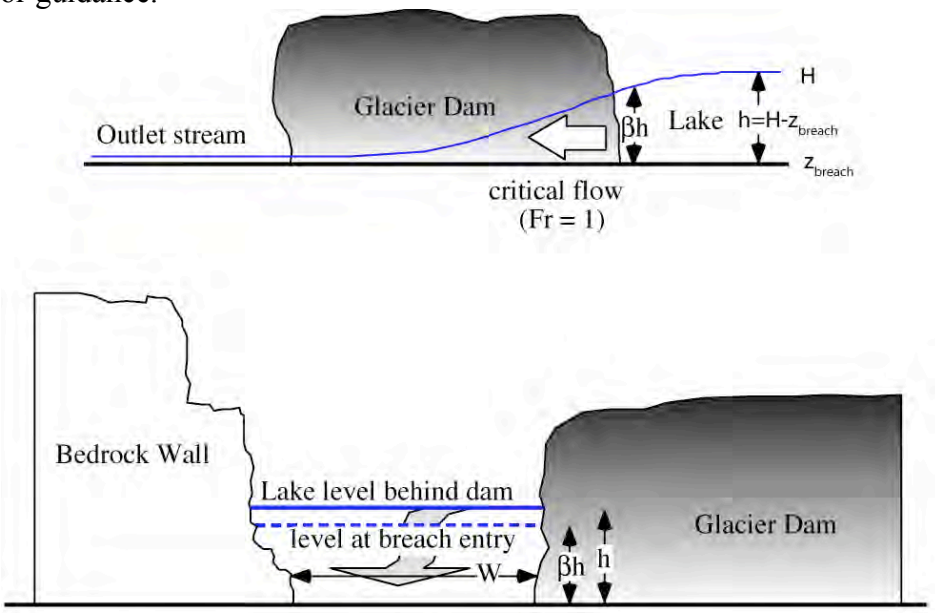
In addition, one might want to drive the components of the climate with real time series of climate, or projected time series of future climate. Here one would want to drive evaporation with more fundamental climate variables such as temperature, wind and humidity. For examples of this sort of modeling effort, see Hostetler and others (1994), or Menking et al. (2004).

Big floods

The largest floods by far are those generated by the failure of dams. In non-dammed landscapes water moves relatively rapidly off the landscape. It passes at cm/second over hillslopes (overland flow) or more slowly through the shallow groundwater system to the edge of a stream, and then down the tributary network to the ocean. Typical rates of flow in the fluvial system are 1 m/s. Even in very long river systems (the longest reach 6000 km, or 6×10^6 m) this speed translates into a travel time of 0.2 years [here I used a very handy shorthand that there are π times 10^7 seconds, or about 30 million seconds]). This means that outside of the deeper groundwater systems, surface water spends less far than a year moving across the landscape. In contrast, dams serve to store water on the landscape for years to decades, even centuries. Their failure releases this water over a very short period of time. These events can punctuate the riverine flood record, and wreak such havoc on the river that it retains the record of these largest floods. The Snake River corridor in Idaho faithfully records the one-time passage of the great Bonneville Flood released from Lake Bonneville in the great Basin about 14500 years ago. This flood was first proposed by Gilbert (1890). Its effects on the Snake River corridor were described by Malde (1968), and more recently addressed by O'Connor (1993). The channeled scablands of Columbia record the repeated floods from failure of glacially dammed Lake Missoula during the last glacial maximum (e.g., Baker (1978); Waitt (1985). While the alluvial dam of lake Bonneville could not rebuild, the ice dam of the

Okanogan Lobe of the Cordilleran ice sheet was capable of healing. These floods, called the Spokane Floods, occurred about every 60 years!

Here our task is to estimate how large the peak discharge might be from failure of such dams. I have turned to papers by Walder and Costa (1996) and Walder and O'Connor (1997) for guidance.



Sketch of breach geometry in a glacially dammed lake. Lake level above the base of the breach will drop as the water drains through the breach into the outlet stream. The worst case scenario (largest peak discharge) corresponds to instantaneous insertion of the full breach of width W and level z_{breach} (after Walder and Costa, 1996).

That the release of water is so rapid compared to the inputs allows us to simplify the water balance equation by ignoring all inputs and all losses other than discharge through the breach:

$$\frac{dV}{dt} = -Q_{out} \tag{6.12}$$

Climate becomes irrelevant, and we focus on what sets the discharge out of the lake, through the breach in the dam (Figure 3). This outlet discharge will depend on both the area of the breach, and the speed of the flow. In other words, $Q_{out} = WHU$, where W is the width of the breach, H the height of the water surface above the base of the breach, and U the mean speed of the water. We have a few pieces to put in place before we can proceed. We need to specify the initial breach size and perhaps its evolution. We need an equation for what sets the average speed of flow through the breach. And on the left hand side of the equation we need to relate the lake volume to the lake level. The worst-case scenario would be that the breach is instantaneously as large as it will get. It fails so rapidly that the detailed evolution of the breach is irrelevant. This results in the worst case because the cross sectional area of the water flowing through the breach is its maximum value. If we are trying to estimate the largest flood a lake could produce, which one might want to know if living in a town on the floodplain downstream, this is

the calculation to make. Here we make use of an observation that the flow through a notch goes "critical" somewhere within its passage through the notch. By critical here I mean that its Froude number goes to 1. The Froude number is one of those dimensionless numbers in fluid mechanics, and is defined by $Fr = \frac{\bar{U}}{\sqrt{gH}}$, where H is the flow depth and g the acceleration due to gravity. This number reflects the relative importance of inertia and gravity in the problem, and discriminates between tranquil flow (Fr<1) and supercritical or shooting flow (Fr>1). In the breach, then, we can say that

$$\bar{U} = \sqrt{g\beta(H - z_{breach})} \tag{6.13}$$

where β is a factor between 0 and 1 reflecting this "somewhere" factor. It is taken to be about 2/3 by Walder. The water balance equation now reads

$$\frac{dV}{dt} = -W\beta(H - z_{breach})\sqrt{g\beta(H - z_{breach})} = -Wg^{1/2}\beta^{3/2}(H - z_{breach})^{3/2} \tag{6.14}$$

The rate at which water leaves the lake is set by the 3/2 power of the flow depth. Now we need to modify the left hand side. We need a relationship between lake volume and lake level. In the simple conical lake we saw that this relationship was a power law, and in Mono Basin we again saw a different power law. For convenience, then, let's simply assert that lake volumes will depend upon the lake level taken to a power, p, that we can determine from maps or DEMs of real basins. The simplest expression is therefore

$$V = \alpha(H - z_{breach})^p \tag{6.15}$$

If we want on the left hand side of equation 14 an expression for rate of change of lake level, rather than rate of change of lake volume we must use the chain rule:

$$\frac{dV}{dt} = \frac{dV}{dh}\frac{dh}{dt} = \alpha p h^{p-1}\frac{dh}{dt} = \alpha p(H - z_{breach})^{p-1}\frac{d(H - z_{breach})}{dt} \tag{6.16}$$

where I have used the flow depth h=H-z_{breach}. Inserting this in equation 14, and dividing by the factor in front of the derivative, we arrive at a final equation for the evolution of the flow depth:

$$\frac{\partial(H - z_{breach})}{\partial t} = -\frac{Wg^{1/2}\beta^{3/2}(H - z_{breach})^{3/2}}{\alpha p(H - z_{breach})^{p-1}} = -\frac{Wg^{1/2}\beta^{3/2}}{\alpha p}(H - z_{breach})^{3/2-p+1} \tag{6.17}$$

This is still a little intimidating. Using instead our definition of flow depth above the breach, h, and recognizing that all variables besides this may be collected into a single constant, this becomes more simply

$$\frac{dh}{dt} = -\gamma h^{3/2-p+1} \tag{6.18}$$

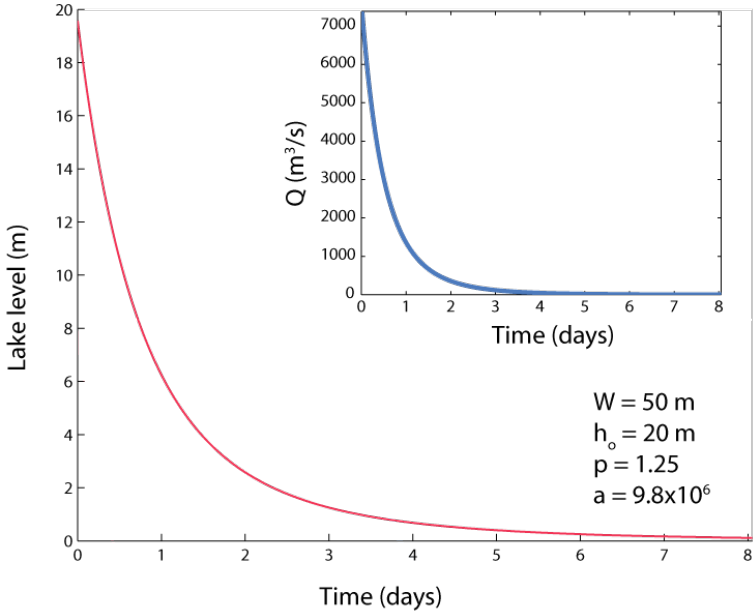
Let's dissect this a bit. I usually turn to the simplest possible case. Here this would where p=3/2. The power on h then becomes 1, the equation becomes the 1st order linear ODE that we encounter in the first pages of a textbook on ODEs, and the solution of the equation is an exponential

$$h = h_o e^{-t/\tau} \tag{6.19}$$

where h_o is the initial height of the lake above the breach, and τ is a characteristic timescale for the decay of lake level (and hence discharge of water from the lake). Noting that in this case, the constant γ has units of 1/time, the characteristic time scale τ=1/γ, or

$$\tau = \frac{\alpha(3/2)}{Wg^{1/2}\beta^{3/2}} \tag{6.20}$$

Does this make sense? The main variables here are W, the width of the breach, and α, which scales the lake volume. We expect that a wider breach will release water more rapidly, meaning that the flood would decline more rapidly. This is indeed the case; the time scale τ is shortened as W increases. On the other hand, the larger α is, the greater the lake volume for a given lake depth. It will therefore take longer for the lake to draw down, all other things being constant. It therefore makes sense that α is in the numerator.



Lake level history and associated discharge through a breach with specified dimensions W and ho, calculated using a numerical code. Note that while the lake level is roughly exponential (not quite exponential because p does not equal 1.5), the discharge drops even more rapidly.

There is no reason that real lake basins should have this magic geometry in which p=3/2. I present in Figure 4 an example of numerical integration of these equations for discharge and lake level using p = 1.25. The equation governing evolution of the lake level will therefore not reduce to the simple linear form whose solution we know to be an exponential. But most lake basins display p>1, and we don't expect wildly different behaviors of the resulting lake level evolution equation. Before moving on, we should make sure we understand why the lake level behaves at least roughly as an exponential. Given the assumptions we have made (most crucially, instantaneous insertion of the full breach), we expect that the maximum discharge will occur immediately after the breach opens. It is then that the cross sectional area of the flow passing through the breach is greatest, and it is then that the flow speed should be greatest. As the lake level drops, both the cross section of the flow and the mean speed drop. And they do so consistently until there is no more water to extract.

I note in closing this section that we are in a position to estimate the peak discharge. As I have just said, this occurs when $h = h_o$. Hence

$$Q_{peak} = Wg^{1/2}\beta^{3/2}h_o^{3/2} \quad (6.21)$$

For example, if a dam fails, whose height is 50 m, and the breach widens rapidly to 100 m, the peak discharge would be $Q_{peak} = 100*10^{1/2}*(2/3)^{3/2}*50^{3/2}$, or 60,300 m^3/s. Note that because the thickness of water above the breach base falls off roughly exponentially, the discharge will decline even more rapidly than an exponential, as is seen in Figure 4. While these equations serve as a first order predictor of time scales and discharges, I note that a far more comprehensive treatment is available in Walder and Costa (1996) for ice dams and in Walder and O'Connor (1997) for earthen dams. Because these authors treat the evolution of a breach, they can relax the assumption of the instantaneous insertion of the breach, which in turn reduces the magnitude of the peak discharge. The longer it takes to reach the maximum width and depth of the breach, the more water will have been pulled out of the lake before the leak is most efficient.

References

Baker, V. R., 1978, Paleohydraulics and hydrodynamics of scabland floods, *in* Baker, V. R., and Nummedal, D., eds., *The Channeled Scabland: A guide to the geomorphology of the Columbia Basin, Washington.* Comparative planetary field conference, June 5-8, 1978. Planetary Geology Program, Office of Space Science, National Aeronautics and Space Administration, p. 59-79.

Gilbert, G. K., 1890, Lake Bonneville. *U.S. Geological Survey Monograph 1*, 438 p.

Hostetler, S. W., F. Giorgi, G. T. Bates and P. J. Bartlein, 1994, The role of lake-atmosphere feedbacks in sustaining paleolakes Bonneville and Lahontan 18,000 years ago. *Science* 263:665-668

Menking, K. M., Anderson, R. Y., Shafike, N. G., Syed, K. H., and Allen, B. D., 2004, Wetter of colder during the Last Glacial Maximum? Revisiting the pluvial lake question in southwestern North America. *Quaternary Research* 62: 280-288.

Malde, H. E., 1968, The catastrophic late Pleistocene Bonneville Flood in the Snake River Plain, Idaho, *U.S. Geological Survey Professional Paper 596*, 52 p.

Menking, K.M., Syed, K.H., Anderson, R.Y., Shafike, N.G., and Arnold, J.G., 2003, Model estimates of runoff in the closed, semiarid Estancia basin, central New Mexico, USA. *Hydrological Sciences Journal*, v. 48, p. 953-970.

O'Connor, J. E., 1993, Hydrology, hydraulics, and geomorphology of the Bonneville flood. *Geological Society of America Special Paper 274*, 83 p.

Waitt, R. B., 1985, Case for periodic, colossal jokulhlaups from Pleistocene glacial Lake Missoula *Geological Society of America Bulletin* 96(10): 1271-1286.

Walder, J. S. and Costa, J. E., 1996, Outburst floods from glacier-dammed lakes: the effect of mode of lake drainage on flood magnitude. *Earth Surface Processes and Landforms*, 21 (8), 701-723.

Walder, J. S. and O'Connor, J. E., 1997, Methods for predicting peak discharge of floods caused by failure of natural and constructed earthen dams. *Water Resources Research*, 33 (10), 2337- 2348.

7. Continuity or the Conservation of Mass

Variegated Glacier, Alaska, immediately after its 1982-1983 surge. The highly crevassed surface is characteristic of a recent surge, as each parcel of ice near the surface has experienced high lateral and longitudinal strain histories.

One of the most general and most powerful statements one can make in the earth sciences is that mass must be conserved. In this section we will formalize this statement to arrive at a compact form that we can then manipulate for special cases. The case we will focus on is that in which the material is incompressible, as this is indeed the case in most geological and certainly most geomorphic problems. As a point of terminology, when you hear someone say "by appeal to continuity", they typically mean that "when mass is conserved".

Consider a fixed volume with sides *dx*, *dy* and *dz*. We call this the control volume in the problem, and can dictate that it not change in size. The word statement for conservation of mass in this control volume is simply:

The rate of change of mass within a control volume = the sum of the masses leaving or entering through its edges

$$\frac{\partial(\rho dxdydz)}{\partial t} = inputs - outputs \qquad (7.1)$$

The left hand side represents the rate of change of mass within the control volume of fixed volume *dxdydz* (Figure 1). Note that we have not allowed any sources or sinks of mass within the volume; it can neither be created nor destroyed.

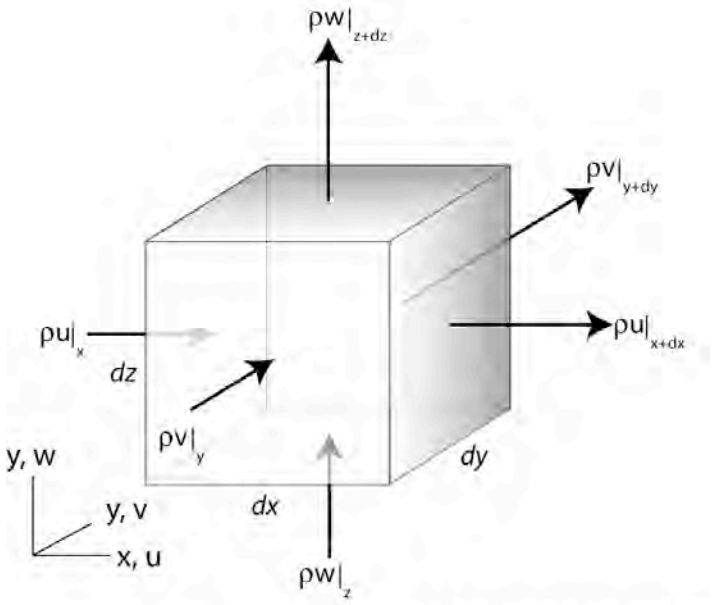


Figure 1. Conservation of mass in a control volume dxdydz. Rate of change of mass in the volume is dictated by rates of arrival and loss of mas across each of the walls. Mass fluxes (arrows) are depicted here in a pattern in which the gradients in mass flux in each direction are positive (more leaves the box on the "downstream" side than arrives on the "upstream" side). Coordinate system depicted shows u, v and w velocity components in the x, y and z directions, respectively.

We now need expressions for the inputs or outputs of mass through the walls of the control volume. These will be the products of mass fluxes with the areas of the walls. For example, the transport of mass across the left wall is the mass flux ρu evaluated at that wall, i.e. at the position x. Let's make sure we agree that the product ρu is a mass flux. The units of this product are $M/L^3 * L/T = M/L^2T$. According to the definition I am using throughout this text, this is indeed a mass flux: a mass per unit area per unit time. To determine the total mass transported across the wall per unit time is then the product of this mass flux with the area of the wall. Taking each term in succession, representing total transport of mass across each wall, we now have

$$\frac{\partial(\rho dxdydz)}{\partial t} = (\rho u)\big|_x dydz - (\rho u)\big|_{x+dx} dydz + (\rho v)\big|_y dxdz - (\rho v)\big|_{y+dy} dxdz + (\rho w)\big|_z dxdy - (\rho w)\big|_{z+dz} dxdy \qquad (7.2)$$

This is a very general statement for conservation of mass, or continuity. We may divide through by the volume of the parcel to simplify this, collecting terms on the right hand side so that they are recognizable as derivatives:

$$\frac{\partial \rho}{\partial t} = -\left[\frac{(\rho u)\big|_{x+dx} + (\rho u)\big|_{x}}{dx}\right] - \left[\frac{(\rho v)\big|_{y+dy} + (\rho v)\big|_{y}}{dy}\right] - \left[\frac{(\rho w)\big|_{z+dz} + (\rho w)\big|_{z}}{dz}\right] \tag{7.3}$$

Acknowledging that each of the terms in square brackets is a derivative (by definition, as the element shrinks, in the limit as *dx, dy, dz*->0), this simplifies to

$$\frac{\partial \rho}{\partial t} = -\frac{\partial(\rho u)}{dx} - \frac{\partial(\rho v)}{dy} - \frac{\partial(\rho w)}{dz} \tag{7.4}$$

Each of the terms on the right side represents accumulation or loss of mass due to gradients in transport in a particular direction. Make sure you understand why the minus signs are there. Consider the first term as an example. The derivative is positive when more mass leaves across the right hand wall than arrives through the left hand wall. This should reduce the mass in the box, and the negative sign assures that this is the case: positive gradients lead to negative rate of change of mass in the box.

We have assumed nothing at all about the substance being conserved. It could for all we know be the mass of ice in a glacier or the mass of fish over a shoal or the mass of peas on your plate. What is the next step? The next step is to recognize that the quantities on the right hand side are products of density and speed, both of which, in general, are variables. We must therefore use the product rule to expand each of these derivatives. While this action will double the number of terms on the right, it will allow us to recollect them in two groups with different meanings.

$$\frac{\partial \rho}{\partial t} = -\rho\frac{\partial u}{dx} - u\frac{\partial \rho}{dx} - \rho\frac{\partial v}{dy} - v\frac{\partial \rho}{dy} - \rho\frac{\partial w}{dz} - w\frac{\partial \rho}{dz} \tag{7.5}$$

Moving them to the left hand side those terms with density derivatives, and factoring out the density from the remaining terms on the right leaves

$$\frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{dx} + v\frac{\partial \rho}{dy} + w\frac{\partial \rho}{dz} = -\rho\left[\frac{\partial u}{dx} + \frac{\partial v}{dy} + \frac{\partial w}{dz}\right] \tag{7.6}$$

We can now recognize the left hand side as the total derivative (or substantial derivative, or derivative following the material; see the chapter on substantial derivative). When the derivative of some quantity following the fluid is zero, it means that that quantity is not changing as the material moves. In this case, the quantity is density; when it does not change while following the parcel means that the material is incompressible. In other words,

$$\frac{D\rho}{Dt} = 0 \tag{7.7}$$

is shorthand for an incompressible material. Under such conditions, then, the left hand side goes to zero, we can divide by (-ρ), and the continuity equation reduces to

$$\frac{\partial u}{dx} + \frac{\partial v}{dy} + \frac{\partial w}{dz} = 0 \tag{7.8}$$

The shorthand for this equation, in vector notation, is

$$\nabla \cdot U = 0 \tag{7.9}$$

where both the del operator and the velocity U are vectors. Whether in the component notation or vector notation, this incompressible form of the continuity equation is very commonly the first equation in a paper on fluid mechanics. Because it is so common, it

is worth absorbing deeply. Each of the terms in equation 1.8 is a strain rate. They are equally well called gradients in velocity or strain rates. Their units are (L/T)/L, or T^{-1}. Since strain is dimensionless, a change in length divided by an original length, this is a rate of strain. If the material is incompressible, then a positive strain rate in one dimension must be compensated for by a strain rate in one or the other or both other dimensions. Say there is a positive gradient in speed in the x direction, meaning that $\partial u / \partial x > 0$. This means that the material is stretching in the x direction. If the material is incompressible, it must simultaneously be thinning in at least one of the other dimensions. This thinning, in which particles of material are coming closer together, requires negative velocity gradients.

The Variegated Glacier surge

We can straight-forwardly apply this concept to explain the pattern of thickening associated with the propagation of a glacier surge down the Variegated Glacier in Alaska in 1982-1983 (Kamb et al., 1985). While most glaciers are well-behaved, and dutifully transport the excess ice arriving in the accumulation zone toward the ablation zone year after year, some small fraction of glaciers are not. These ill-behaved glaciers are called surging glaciers. They wait for many years to decades before performing this principal task of a glacier, and then do so in dramatic fashion. Glacier surges require rapid motion of the glacier, which is accomplished by rapid basal motion rather than rapid internal deformation. The surge is typically initiated in the upper reaches of the glacier, and propagates downvalley, bringing with it huge amounts of ice. The pattern of strain experienced by a parcel of ice as it is transported during a surge results in a highly crevassed surface on which one should not even imagine traveling.

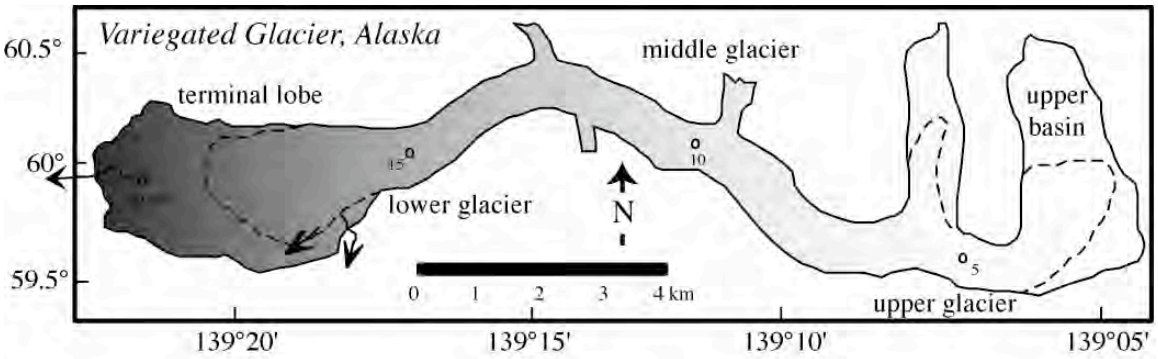


Figure 2. Map of the Variegated Glacier, Alaska, showing area of glacier involved in the 1982-83 surge (dashed lines), a few of the km stake locations, and the locations of the major outlet streams near the terminus. (after Kamb et al., Science 227, 1985)

Seen at a point, say through the lens of a fixed camera pointed at the ablation zone, the glacier is at first relatively thin because so little ice has been delivered to the ablation zone for years, then thickens rapidly as the bulge of the surge arrives. Our task is to understand why this thickening occurs, and at what rate it should occur if we know something about the down-valley pattern of horizontal motion. In other words, for a measured pattern of *u(x)*, what is the expected pattern of thickening, *dH/dt(x)*? The

pattern of ice speeds, as documented from repeated surveys of stakes on the glacier surface, is shown in Figure 3.

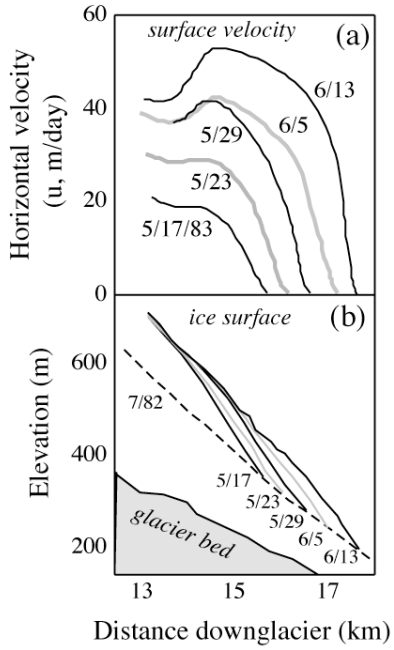


Figure 3. a) Ice surface velocity profile, u(x), and b) ice surface topography, z(x), in a 3 km reach of the lower glacier, during a one-month period of the 1983 surge of the Variegated Glacier, Alaska. (after Kamb et al., 1985, Science 227, Figure 4)

The propagation of the surge front into the reach of glacier between km 12 and 18 is clear, as ice speed at a point increases from <1 m/d to > 40 m/d. The spatial pattern of horizontal speed at any time shows a very strong negative gradient of order -20 m/d per km. The pre-surge ice thickness in this reach of the glacier is roughly 150 m (dashed line in (b)). Let's see if we can get in the ballpark with an estimate of the pattern of thickening using our concept of continuity. Our goal is an expression for the expected pattern of vertical velocities at the ice surface, i.e. for *w(H)*, where *H* is the thickness of the ice. From equation 1.8, we have

$$\frac{\partial w}{dz} = -\frac{\partial v}{dy} - \frac{\partial u}{dx} \tag{7.10}$$

The rate of change of vertical velocity, *w*, with height above the bed, *z*, must equal the sum of the velocity gradients in *x* and *y* speeds. But the glacier is well-confined within its walls, and not much of the crunch associated with the huge gradient in horizontal down-valley speed is taken up by gradients in the lateral direction. So to first order we may ignore *dv/dy*, resulting in

$$\frac{\partial w}{dz} = -\frac{\partial u}{dx} \tag{7.11}$$

We need to integrate this equation in the vertical dimension in order to estimate *w* at a given height above the bed, *z*. In doing so we can make use of another simplification: as the horizontal speeds of a surge are entirely accomplished by basal motion or sliding of the glacier against its bed, they are independent of height above the bed. The glacier is

moving like a plug; $u(z) = U_{slide} = U_{surface}$. Therefore both u and du/dx are independent of height above the bed. This means that

$$w(z) = \int_0^z \frac{\partial w}{dz} dz = -\int_0^z \frac{\partial u}{dx} dz = -\frac{\partial u}{dx} \int_0^z dz = -\frac{\partial u}{dx} z \qquad (7.12)$$

We wish to know the vertical speed at the ice surface, i.e. at $z=H$. The final expression for the expected vertical (uplift) speed is

$$w(H) = -\frac{\partial u}{dx} H \qquad (7.13)$$

The uplift speed or thickening rate of the glacier should be the product of the ice thickness with the local gradient in the horizontal ice speed. The greater the rate at which the column of ice is being squeezed by the horizontal strain rate, and the greater the original height of the column being squeezed, the greater the thickening rate. We can see qualitatively from Figure 3 that this is the case during the Variegated Glacier surge. But let's get quantitative. On May 17th the gradient at km 15, in the middle of the strong speed gradient, is about -20m/d-km (or, stated another way, a horizontal strain rate of 0.02 d^{-1}), and the pre-surge ice thickness is 175 m. From equation 1.13 we predict a vertical speed of the ice surface of about 3.5 m/d. In the 6 days between May 17th and 23rd, the ice thickened by about 28 m at km 15. Our calculated total uplift is 6 d*3.5m/d = 21 m. We are indeed in the ballpark.

Now inspect the lead photograph to this section. Seen immediately after the surge of 1982-83, the surface of the Variegated is diced into a city of ziggurats. These intersecting sets of crevasses, some of them many tens of m deep, reflect the strain history to which the surface has been subjected. The first set was longitudinal: it went up and down the glacier. The insertion of these crevasses occurred at a time when the glacier was thickening. It thickened most where the original thickness was greatest: in the centerline. This led to steepening and stretching of the ice surface in the lateral dimension. At the surface, the strain rates must have exceeded the strength of the ice, and it broke elastically to create this crevasse set. At this point the ice is thick, and broken into longitudinal slices. The second set of crevasses occurred after the surge front passed. Note in Figure 3a that the spatial pattern of horizontal speed at any time reaches a peak and then falls off upglacier. Upglacier of the peak in surge speed, a positive gradient exists in the glacier speed. For example, on 6/13, the speed increases downglacier from 41-53 m/d over about 1 km. The +0.012 d^{-1} positive horizontal strain rate would lead to stretching in the downglacier direction. Indeed, the second set of crevasses is normal to this stretching direction; the ice at the surface again fractured when subjected to such great rates of horizontal strain.

References

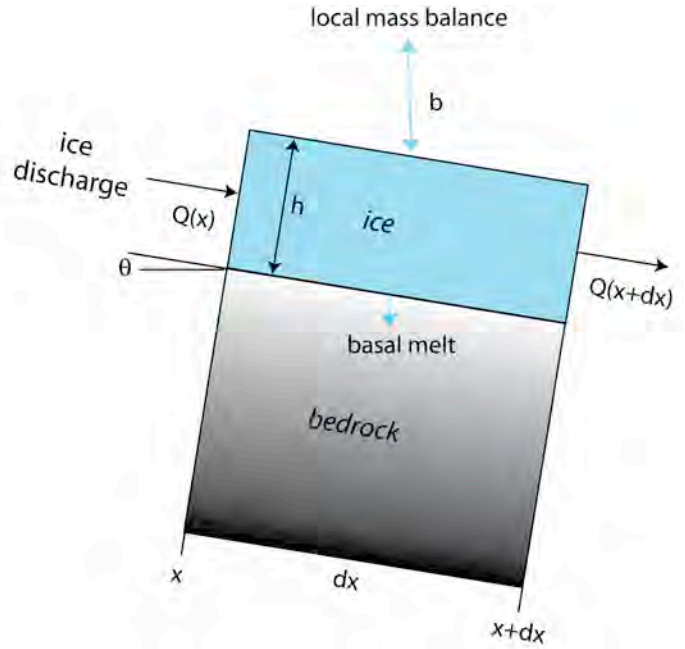
Kamb, B., C. F. Raymond, W. D. Harrison, H. Englehardt, K. A. Echelmeyer, N. F. Humphrey, M. M. Brugman and W. T. Pfeffer, 1985, Glacier surge mechanism: 1982-83 surge of Variegated Glacier, Alaska. *Science* 227: 469-479.

Sharp, M., W. Lawson, and R. S. Anderson, 1988, Tectonic processes in a surge-type glacier -- an analogue for the emplacement of thrust sheets by gravity tectonics. *Journal of Structural Geology* 10: 499-515 .

8. Glaciers

Palisades Glaciers, eastern Sierras, California, shown in September 1990. Little Ice Age moraines show ice extent roughly 100 years ago, and ELA separates accumulation above from ablation zone below.

Glaciers grace the tips of our highest mountain ranges. They are responsible for erosional modification of the mountains. And as they are sensitive to climate, they are in some sense the canaries in the coal mine of global climate change. We would like to be able to predict the position of the terminus of a glacier, and just how this reflects the climate. Once again, the problem comes down to a balance, this time of mass of ice. One may find in the bible of the glaciologists, Paterson's *The Physics of Glaciers*, now in its 3rd edition, at least one chapter on mass balance alone.



Glacier ice balance. Meteorology dictates the local mass balance, which is positive (downward arrow) above the ELA, and negative (upward arrow) below it. Ice discharge is accomplished by both basal sliding and internal deformation of the ice.

The word picture describing the volume balance set up in Figure 1 is:

Rate of change of ice volume = rate of ice input – rate of ice output

One may formalize the illustration of the mass balance shown in Figure 1 with the following equation:

$$\frac{\partial h}{\partial t} = b - \frac{1}{W}\frac{\partial Q}{\partial x} \tag{8.1}$$

where W is the local width of the glacier. Let's talk about each of the terms. Mass can be lost or gained through all edges of the block we have depicted (the top, the base, and the up- and down-ice sides). Here the b represents the "local mass balance" on the glacier surface, the mass lost or gained over an annual cycle. It is usually expressed as a change in height of the ice surface, i.e., meters of ice equivalent. This quantity is positive where there is a net gain of ice mass over an annual cycle, and negative where there is a net loss. Where the mass balance is zero defines the "equilibrium line altitude", or the ELA, of the glacier (see Figure 2). The mass balance reflects all of the meteorological forcing of the glacier, both the snow added over the course of the year, and the losses dealt by the combined blows of ablation (melt) and sublimation. Where the annual mass balance is positive, it has snowed more than it melts in a year, and vice versa. To first order, because it snows more at higher altitudes, and melts more at lower altitudes, the mass balance always has a positive gradient with elevation.

One may measure the health of a glacier by the total mass balance, reflecting whether in a given year there has been a net loss or gain of ice from the entire glacier. This is simply

the spatial integral of the product of the local mass balance with the hypsometry of the valley:

$$B = \int_0^{z_{max}} W(z)b(z)dz \tag{8.2}$$

This exercise is carried out on at least a dozen Norwegian glaciers on an annual basis. The Norwegians are interested in the health of their glaciers in large part because a significant portion of their electrical power comes from subglacially tapped hydropower sources.

It is a common misconception that a considerable amount of melting takes place at the base of a glacier, because after all the earth is hot. Note the scales on the mass balance profiles. In places, many meters can be lost by melting associated with solar radiation. The heat flux through the earth's crust is about 41 mW/m^2 (defined as one heat flow unit). This is sufficient to melt about 5 cm of ice per year, a trivial amount when contrasted with the high heat fluxes powered in one or another way by the sun. As far as the mass balance of a glacier is concerned, then, there is little melt at the base of the glacier.

If nothing else were happening but the local mass gain or loss from the ice surface, a new wedge of snow would accumulate, which would be tapered off by melt to a tip at the ELA (or snowline) each year. Each successive wedge would thicken the entire wedge of snow above the snowline, and would increase the slope everywhere. But something else *must* happen, because we find glaciers poking their snouts well below the ELA, below the snowline. How does this happen? Ice is in motion. This is an essential ingredient in the definition of a glacier. Otherwise we are dealing with a snowfield. The Q terms in the mass balance expression reflect the fact that ice can move downhill, powered by its own weight. Ice has two technologies for moving, one by basal sliding, in which the entire glacier moves at a rate dictated by the slip at the bed, the other by internal deformation, like any other fluid (see Figure 7). We will return to a more detailed treatment of these processes in a bit. Know for now that the ice discharge per unit width of glacier, Q, may be defined as the product of the mean velocity of the ice column, <u>, with the thickness of the glacier, h.

Given only this knowledge, we can construct for ourselves a model of a glacier in steady state, one in which none of the variables of concern in the mass balance expression are changing with time. Setting the left hand side (lhs) to zero, then, we see that there must be a balance between the local mass balance of ice dictated by the meteorological forcing (the climate) and the local gradient in the ice discharge.

$$Q(x) = \int_0^x W(x)b(z(x))\,dx \tag{8.3}$$

Here we have taken $x = 0$ at the up-valley end of the glacier. Ignoring for the moment the width function $W(x)$, reflecting the geometry (or really the hypsometry) of the valley, we see that for small x, high up in the valley, since the local mass balance is positive there the ice discharge must increase with distance downvalley; conversely, it must decrease for distances below that associated with the ELA. The ice discharge must therefore go through a maximum at the ELA. In addition, since the discharge Q must increase at least

as fast as the ice thickness (Q has H in the expression already, and if <u> increases with H it will be even more sensitive to H), we might expect that the ice thickness is also a maximum at the ELA.

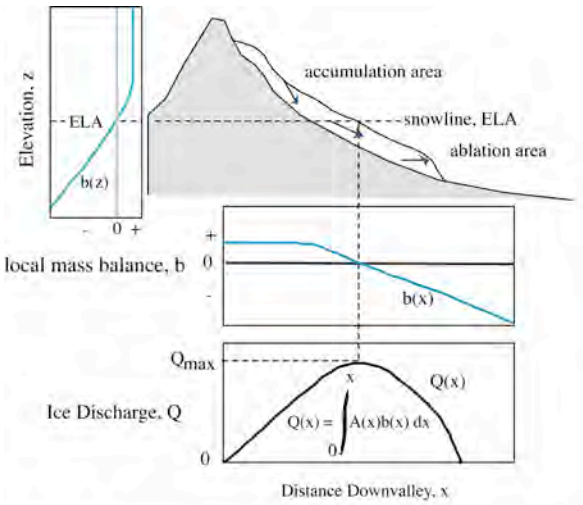


Figure 2. Schematic diagrams of a glacier (white) in a mountainous topography (gray) showing accumulation and ablation areas on either side of the equilibrium line altitude (ELA). Mapped into the vertical, z (left diagram), the net mass balance profile, b(z), is negative at elevations below the ELA and positive above it. We also show the net balance mapped onto the valley- parallel axis, x (follow dashed line downward), generating the net balance profile b(x). At steady state the ice discharge of the glacier must reflect the integral of this net balance profile (bottom diagram). The maximum discharge should occur at roughly the down-valley position of the ELA. Where the discharge goes again to zero determines the terminus position.

Many mass balance profiles show an essentially linear increase in the local mass balance, b, with elevation, z, defined by the gradient of balance with elevation, γ, and the elevation at which the balance crosses zero, the equilibrium line altitude, denoted here by z_{ela}:

$$b = \gamma(z - z_{ela}) \tag{8.4}$$

Conservation of ice in a representative valley-spanning column of width W requires that the rate of change of ice cross-section, HW, include meteorologically driven loss or gain of ice through the top of the column, bW, and the divergence of ice flux through the up-glacier and down-glacier sides of the column. In 1D, this becomes

$$\frac{\partial(HW)}{\partial t} = bW - \frac{\partial Q}{\partial x} \tag{8.5}$$

For a steady state glacier the left hand side equals zero, and the divergence of ice flux must equal the volume of ice gained or lost through the top of the column:

$$\frac{dQ}{dx} = bW \tag{8.6}$$

The steady-state pattern of ice discharge, $Q(x)$ becomes:

$$Q(x) = \int_0^x Wb\,dx \tag{8.7}$$

The location of the glacial terminus lies at the x-position where the ice discharge down-valley of the ELA vanishes. Using analytical and numerical solutions of this equation, we explore the effects of the climate through $b(x)$ and of valley geometry through $W(x)$.

To illustrate this, we show in Figure x a simulation of the evolution of a small alpine glacier in its valley, starting with no ice and evolving to steady state. We impose a simple capped linear mass balance profile, and hold it steady throughout the model run. Similar modeling exercises have been used recently to explore the sensitivity of alpine glaciers to climate changes in the past and in the future (e.g., Oerlemans, 2005).

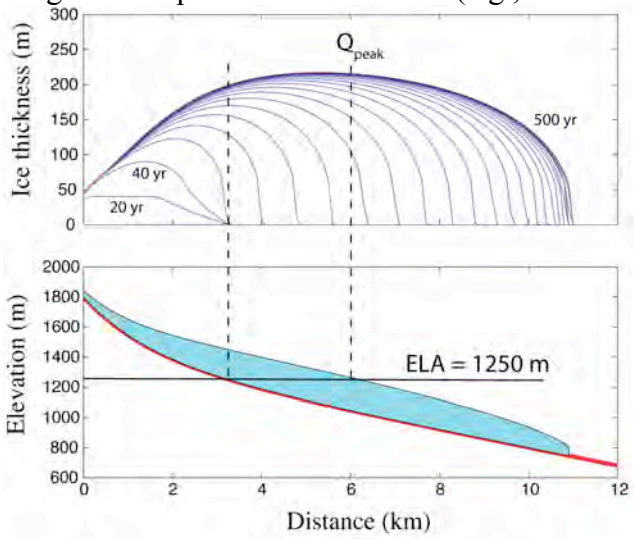


Figure x. 1-D glacier model in a steady climate with ELA=1250 m. 500 year simulation with snapshots every 20 years. top: Ice thickness evolution. First 40 years shows no ice below the intersection of ELA with valley floor (left dashed line). Final 100 years is nearly steady, showing only slight advance and steady thickness throughout. Peak ice discharge corresponds to peak ice thickness, and to intersection of ice surface with ELA (right dashed line). Note this is considerably below the ELA intersection with the bedrock floor, demonstrating a positive ELA feedback. Bottom: glacier shown at end of 500 year simulation.

This exercise also yields another interesting result. In steady state, we find that within the accumulation area, the ice discharge must be increasing down-valley in order to accommodate the new snow (ultimately ice) arriving on its top. Conversely, the ice discharge must be decreasing with down-valley distance in the ablation region. This has several important glaciological and glacial geological consequences. First, the vertical component of the trajectories of the ice parcels must be downward in the accumulation and upward in the ablation zone, as shown in all elementary figures of glaciers, including Figure x. As a corollary, debris embedded in the ice is taken toward the bed in the accumulation zone and away from it in the ablation zone. Glaciers tend to have concave up-valley contours above the ELA, and convex contours below. Debris therefore moves away from the valley walls in the accumulation zone and toward them in the ablation zone. This is reflected in the fact that lateral moraines begin at roughly the ELA. This observation is useful if one is trying to reconstruct past positions of glaciers in a valley, or more particularly to locate the past position of the ELA. As the ELA is often taken as a proxy for the 0°C isotherm, it is a strong measure of climate, and hence a strong target for paleoclimate studies.

This straight-forward exercise should serve as a motivation for understanding the mechanics of ice motion. This is at the core of all such simulations. It is what separates

one type of glacier from another. And whether a glacier can slide on its bed or not dictates whether it can erode the bed or not -- and hence whether the glacier can be an effective means of modifying the landscape.

9. Weathering alteration of deposits

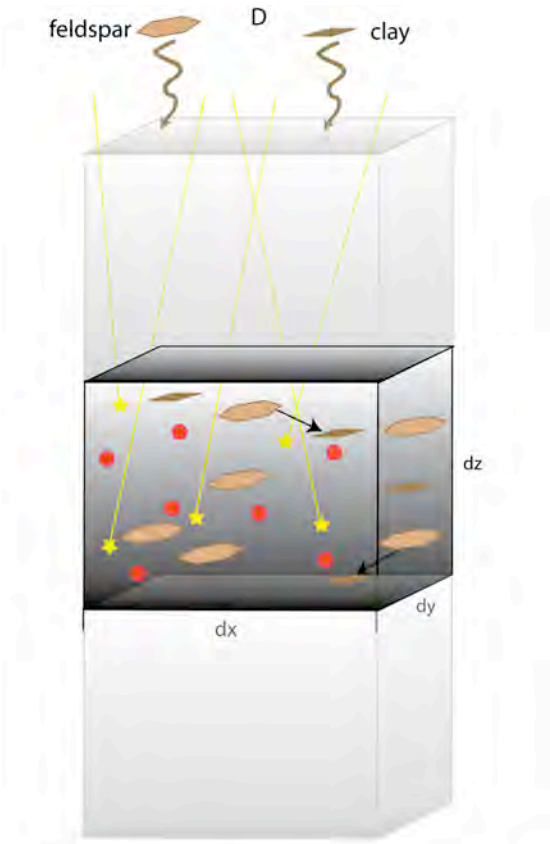
Consider the case of an accumulating deposit, in which proximity to the surface promotes chemical alteration of the accumulating sediment. This ought to produce a signature of the sediment having spent time near the surface, and one might expect that the signal would be stronger the more time the material spends in this zone. The depositional setting might be eolian, say at a site of loess accumulation, or it might be slow accumulation in a floodplain by occasional fluvial overbank deposition. We can again formalize the problem by employing the principle of conservation. One of the chief differences between this case and the others we have dealt with is that we will follow the parcel of sediment as it gets buried. The reference frame, in other words, follows the parcel, rather than being fixed in space. The word picture might be:

Rate of change of weathering product in the parcel = rate of addition of weathering product in the sediment + rate of production of weathering product in the parcel – rate decay of weathering product in the parcel

This statement allows for the real possibility that the sediment being added to the deposit has some original concentration of the mineral of concern (Figure 1). To be specific, let us say that our weathering product is a clay, although the set-up should be generic enough to accommodate the evolution of magnetically susceptible grains, the concentration of cosmically produced radionuclides, and so on. Call the concentration of clay in the incoming sediment C_o, while that in the deposit is C. We must acknowledge that the deposit is thickening, let us say at a rate $\dot{D}$. The translation of the word picture then becomes

$$\frac{D(Cdxdydz)}{Dt} = Pdxdydz - Rdxdydz \tag{9.1}$$

where here I have used P to represent production rate of new clay, and R the removal rate of clay, taken per unit volume of sediment. Note that since we are following a parcel of sediment once it is deposited, there are no terms reflecting loss or gain of sediment across the edges of the box. I have used the D/Dt notation to acknowledge that we are following a parcel of material; this is called the total derivative, or substantial derivative (see fuller discussion of advection in the Advection chapter). We will argue about the more general case in a moment, but let's address the easiest first. The addition of sediment in the top of the box occurs only when it is at the surface, a fact that we will take care of in the boundary conditions to the problem.



Production (stars) and decay (dots) of cosmogenic radionuclides. The rate of change of the concentration of radionuclides is therefore set by any mismatch between the birth and death rates. Also shown is deposition of clays and feldspars at this site, either by eolian or fluvial processes, leading to accumulation at a rate D, and production of new clays by weathering of feldspars.

A simple case: radionuclide concentration profiles

Here we ask what is the expected profile of concentration of a radionuclide, say ^{10}Be. The production rate will be due to cosmic ray bombardment, which declines with depth at a rate that is dictated solely by the density of the material through which the particles are passing. The production of radionuclides does in fact follow an exponential decay with depth, the length scale z^* being about 0.8 m in typical soils. In this case we must also acknowledge the potential for decay. This will come in on the negative side of the ledger, represented by R in the above equation. The concentration we seek is the number of atoms of ^{10}Be per gram of quartz. So the concentration of quartz, which might vary with depth, is taken out of the problem - it only becomes an operational problem, as one needs a certain number of nuclides (say a few million) to be able to measure the concentration of ^{10}Be. The concentration of quartz therefore becomes important in setting the size of sample one collects in the field. The equation for production becomes

$$P = P_o e^{-z/z^*} \tag{9.2}$$

while that for decay is simply

$$R = \lambda C \tag{9.3}$$

Assembling these into a single governing equation

$$\frac{DC}{Dt} = P_o e^{-z/z^*} - \lambda C \tag{9.4}$$

Let us first inspect the equation. In the simplest case let us ignore decay (or, equivalently, use a stable nuclide). The 2nd term becomes unimportant, and we simply watch the accumulation of nuclides as the parcel is slowly buried at a rate D.

$$C = C_o + \int_o^T P_o e^{-Dt/z^*} dt \tag{9.5}$$

Here the 1st term represents the concentration of nuclides in the incoming quartz (in other words, "inheritance" of nuclides), while the 2nd term represents the newly produced nuclides. In this case we know how to relate time to depth through the steady deposition rate, giving us

$$C(z) = C_o + \int_o^{z/D} P_o e^{-Dt/z^*} dt = C_o - P_o z^*/D\left[e^{-Dt/z^*}\right]_0^{z/D} \tag{9.6}$$

or

$$C = C_o + P_o(z^*/D)[1 - e^{-z/z^*}] \tag{9.7}$$

Note that the concentration asymptotically approaches a maximum that is dictated by the time the parcel spends within the production zone, which in turn is set by the combination z^*/D. The faster the deposition rate, the more rapidly the parcel transits the production zone (Fig. x), and the lower the asymptotic concentration.

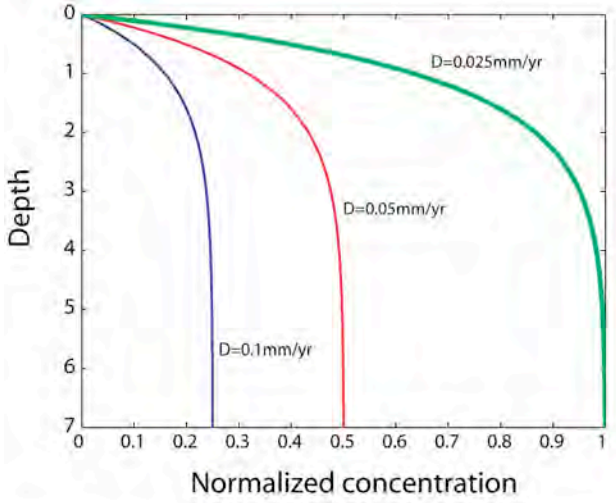


Figure x. Concentration profiles in a steadily depositing system for three deposition rates differing by 2 and 4-fold. Depths normalized to production length scale, z*, and concentration profiles normalized by the asymptotic concentration in the low deposition rate case.

If we allow decay of radionuclides, we can imagine that the profiles will droop back toward zero concentration at a rate dictated by the decay timescale. Here I have constructed a simple numerical model in which I organize the code to track both the production and the decay in each parcel of sediment. In other words, I employ exactly these conservation principles to organize the code. The top of the profile looks the same, appearing to approach an asymptotic value (Figure x), but this time the decay takes over once the parcel is deep enough for decay to outpace new production. This will be the case forevermore, as the parcel gets deeper and deeper, further reducing the production

rate. The profile eventually settles into an exponential decline as we expect from radioactive decay, this time depth replacing time through the relation $z=Dt$.

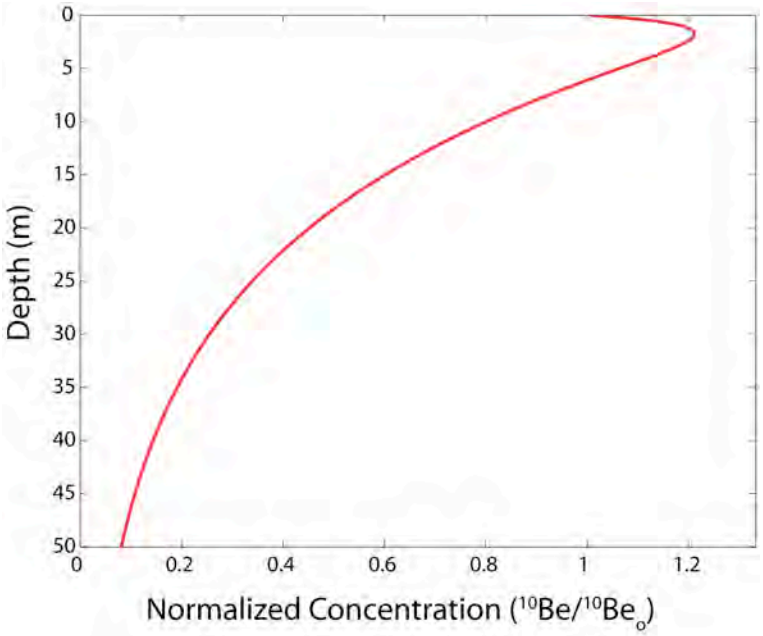


Figure x. Hypothetical profile of a radionuclide concentration expected in a steadily accumulating loess stratigraphy. Near the surface, the increase in concentration represents grow-in of new nuclides. Below the peak concentration, decay exceeds new production, and the concentration declines toward zero.

The case we have just presented is analogous to one discussed in Anderson and Hallet (1996), in which the system was a loess deposit in which the concentration being calculated was that of magnetically susceptible grains (e.g. magnetite). There, the magnetic grains could arrive on the airstream from the wind, setting the inheritance, and they could be produced in the accumulating soil (i.e. pedogenic). The measurement of magnetic susceptibility of a soil is then the easily measured proxy for the concentration of such minerals. We treated this steady state case as an end-member case, and went on to discuss the roles of variation in deposition rate (D), of pedogenic production rate (P_o), and of collapse of the soil, all of which should vary with climate.

Weathering in soils

In the cosmogenic problem worked above we did not allow the minerals in the soil to evolve. But of course it is exactly this evolution that determines the texture, structure and the mineralogy of a soil. In particular, the clay content of a soil is all-important. While some of these clays may be original clays in the deposit, much of the clay is derived from in situ weathering. Here I approach the more general case in which clays are produced from weathering of feldspars. Let us assume that the clay can no longer change once it is produced: it is an end-member clay. In this case, the removal rate R is zero (we will not allow it to elutriate either, or leak downward out of our parcel). The complexity, and the interesting twist to the problem, comes in the fact that the production rate of new clay ought to depend upon the concentration of feldspar in the parcel, C_f. This requires that we track the concentrations of both minerals. Certainly we would all agree that if the entire

deposit were quartz, say, there would be no production of clay. The resulting equations for rate of change of mass of clay and of feldspar in the parcel are:

$$\text{Clay: } \frac{D\left(\rho_b C_c dxdydz\right)}{Dt} = (1-f)R\gamma C_f dxdydz \tag{9.8}$$

$$\text{Feldspar: } \frac{D\left(\rho_b C_f dxdydz\right)}{Dt} = -R\gamma C_f dxdydz \tag{9.9}$$

I have written the equations such that:

- C_f, and C_c are concentrations of feldspar and clay, respectively, in mass per unit mass [M/M]

$\cdot \rho_b$ corresponds to bulk density of regolith [M/L^3]

- R is the rate of loss of feldspar, in mass per unit area per time [M/L^{2}T]

$\cdot \gamma$ is the reactive surface area of mineral per unit volume [L^2/L^3]

and

- f is the fraction of the feldspar loss that goes directly to solution, while (*1-f*) goes to the solid phase, clay. In other words, we acknowledge that the system is not perfectly conservative in that we allow solute leakage from the system.

To simplify a little, we may divide both sides by the volume of the parcel, *dxdydz*. All that remains is to be more specific about the production rate, R. For the moment, I will dodge the detailed chemistry of the problem, and posit that the reaction involved is such that the production rate falls off with distance from the surface. Here I can lean on the fact that the solution accomplishing the weathering is furthest from equilibrium where the water is most dilute, which is nearest the surface. In addition, this is where the temperatures are greatest in the deposit, within the thermal boundary layer associated with both annual and daily fluctuations of temperature. For simplicity, I therefore propose that we take the production rate of clays per unit surface area of feldspar to go something like:

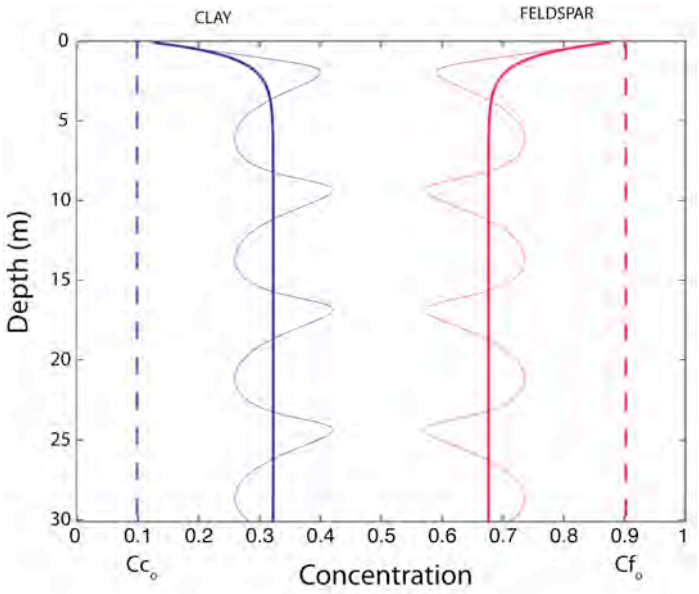
$$R = R_o e^{-z/L} \tag{9.10}$$

where R_o is the reaction rate at the surface (z=0), and L is a length scale over which the reaction rate falls off by 1/e. We can argue later about what might set this length scale, L, but I anticipate that it ought to involve climatic features such as mean annual temperature and total precipitation. Re-writing the evolution equations now yields:

$$\text{Clay: } \frac{D\left(\rho_b C_c\right)}{Dt} = (1-f)R_o e^{-z/L}\gamma C_f \tag{9.11}$$

$$\text{Feldspar: } \frac{D\left(\rho_b C_f\right)}{Dt} = -R_o e^{-z/L}\gamma C_f \tag{9.12}$$

I have again embedded this pair of equations in a numerical model. One can see the expected symbiotic relationship, with clay increasing with depth, and feldspar declining with depth in the steady deposition case (Figure x). I show as well the result when the deposition rate is allowed to vary with time, with an amplitude of one half of the mean rate. In the case shown, the incoming sediment is 90% feldspar and 10% clay. Times of low deposition rate result in more residence time in the soil "reactor zone" and result in commensurately higher degrees of weathering of feldspar to clay.



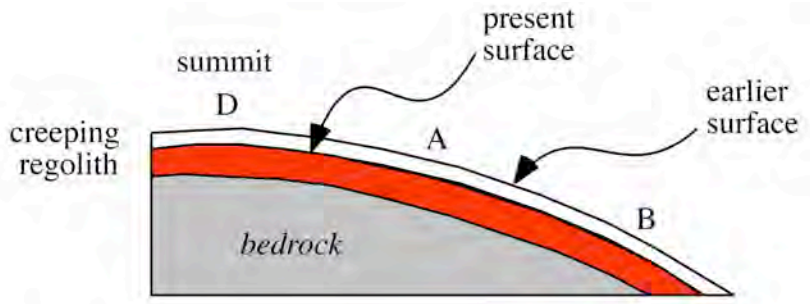
Profiles of clay and feldspar in an accumulating loess sequence in the face of both steady and oscillating deposition rates. Concentration of clay in the incoming dust is 10%., the remainder being feldspar in this hypothetical case (dashed lines). Steady case results in profiles of clay that increases asymptotically with depth. Oscillating deposition results in variable residence times in the near-surface weathering reactor; slow deposition translates to high residence times and more extensive conversion of feldspar to clay.

One might imagine a stack of paleosols in a floodplain deposit, or in a loess sequence, recording a signal of variation in climate.

10. Hillslopes

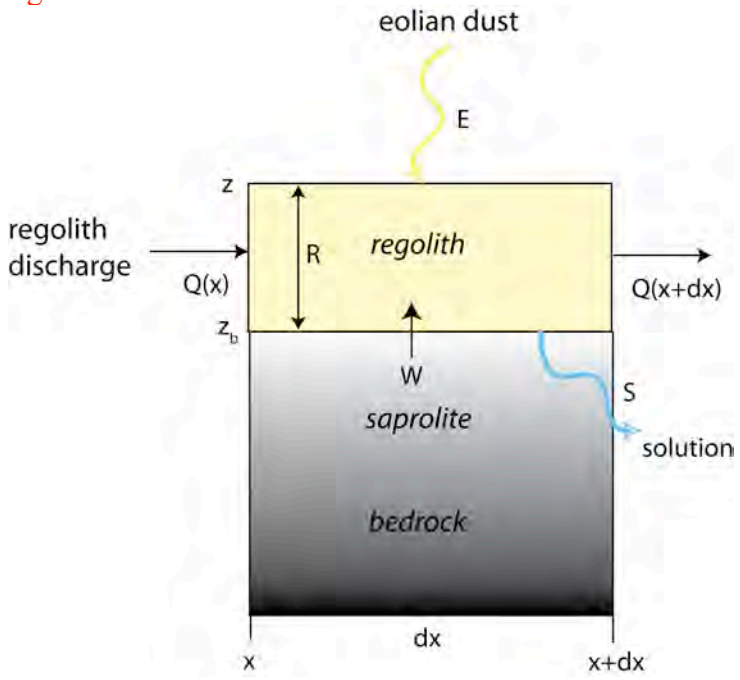
Badland slopes and buttes near Goblin State Park, Utah. Layered shale bedrock shows through thin mantle of regolith.

Hills comprise the majority of landscapes. In fact, on the Moon there is nothing else: no channels, just hills. Consider now the shape of a hillcrest. Whether on the Moon or on Earth, most hilltops are convex up. They are rounded, not sharp. G. K. Gilbert recognized this generality in the early 1900s, and crafted a wonderful word picture of the problem, accompanied by the classic diagram I have reproduced in Figure 1 (Gilbert, 1909). I present here both his word picture, and the generalization of it into a mathematical statement.



Schematic cross-section of a steady state convex hilltop. Summit at D and two equally spaced points A and B are shown. Surface of the hill is shown at two times, early (dashed) and present. The entire hilltop has been lowered by the same amount. Bedrock weathers to mobile regolith at the bedrock-regolith interface. Regolith produced between the summit, D, and the point A must be transported past A by creep, while all regolith produced between D and B must be transported past B. If the transport process is slope-dependent, the slope at B must be twice that at A. This is the essence of convex hilltops. (after Gilbert, 1909, Figure 1)

Let us recast the problem in its most general form. We start by defining the relevant conservation equation. Consider the mass of regolith within a box on the hillslope, shown in 1D in Figure 2.



The word statement representing conservation of regolith is:

rate of change of regolith mass [M/T]	=	rate of regolith inputs [M/T]	-	rate of regolith outputs [M/T]

Treating both sides of the box, this statement may be translated into mathematics:

$$\frac{\partial\left(\rho_b R dxdy\right)}{\partial t} = \rho_r \dot{W} dxdy + Q_x(x)dy - Q_x(x+dx)dy + Q_y(y)dx - Q_y(y+dy)dx + Edxdy - Sdxdy \quad (10.1)$$

where R is regolith thickness, $\dot{W}$ is the weathering rate [L/T], or the rate of production of regolith, ρ_r is the density of rock, ρ_b is the bulk density of the regolith, E is the rate of addition of mass through the top of the box by eolian processes, and S is the rate of loss of mass through solutes. For our purposes, let us assume that E and S are negligible, although we fully recognize that in some situations this is not the case [see for example Mudd and Furbish, 2006)]. In addition, if we align the x direction of our box with local slope, and assume that the slope has no curvature across-slope (more formally, cylindrically symmetrical), it is safe to ignore fluxes across the sides of the box (Q_y) and equation 9.1 becomes:

$$\frac{\partial\left(\rho_b R dxdy\right)}{\partial t} = \rho_r \dot{W} dxdy + Q_x(x)dy - Q_x(x+dx)dy \quad (10.2)$$

Noting that we may hold the length of the sides of the box to be unchanging with time, we may remove them from the derivative on the left hand side and divide through by $dxdy$. If we assume that the bulk density, ρ_b is also unchanging, we may extract it from the derivative as well, and the left hand side becomes simply the rate of change of regolith thickness with time. On the right hand side the dy vanishes entirely. The equation can then be arranged to become:

$$\frac{\partial R}{\partial t} = \frac{1}{\rho_b}\left[\rho_r \dot{W} - \frac{Q_x(x+dx) - Q_x(x)}{dx}\right] \quad (10.3)$$

From our introduction to calculus, we recognize the term in brackets on the right hand side is the spatial derivative of the regolith discharge, $\partial Q_x / \partial x$.

$$\frac{\partial R}{\partial t} = \left(\frac{\rho_r}{\rho_b}\right)\dot{W} - \frac{1}{\rho_b}\frac{\partial Q_x}{\partial x} \quad (10.4)$$

We have, in equation 9.4, a general statement of conservation of mass on a hillslope in one dimension. It is a partial differential equation, in that it contains derivatives of both space and time. The first term on the right hand side represents the source of regolith, modified by a bulking factor to accommodate how much the regolith is puffed up due to the insertion of porosity in the transformation from bedrock to regolith. The second term is formally the divergence of the regolith discharge, which simply represents the difference in the mass moving across the left of the box *vs* the right of the box. The steady state case that Gilbert addressed is captured by taking the time derivative to be zero. In this case,

$$\frac{dQ_x}{dx} = \rho_r \dot{W} \quad (10.5)$$

In words, equation 9.5 requires that the local production rate of regolith by weathering of the bedrock beneath (the right hand side) must exactly be balanced by the spatial change in the discharge of regolith with distance down slope (the left hand side). Otherwise, regolith will pile up in the box, or be slowly drained from the box. In Gilbert's word picture, the right hand side is one of the vertical arrows in the figure, which must be

accommodated by the downslope increase in discharge. Note as well that because the time derivative is taken to be zero, we have transformed the problem into an ordinary differential equation (ODE). In general, these are easier to solve than partial differential equations (PDEs).

This equation can of course be integrated once we know how the regolith production rate is distributed downslope, in x. In the Gilbert steady state form, this transformation rate must be uniform in x, so that the entire regolith-bedrock interface comes down at the same rate. In this steady form case, then,

$$Q_x = \rho_r \dot{W} x + c_1 = \rho_r \dot{W} x \qquad (10.6)$$

where we have taken the constant of integration, c_1, to be zero by asserting that at the top of the hill, at x=0, there is no discharge. The steady hillcrest requires that the regolith discharge linearly increase with distance downslope. Another way of putting it is that at each position on the hill, the discharge of regolith equals all of the regolith that is produced uphill of that spot. Since it was arbitrary in Figure 2 how long we made the first box, this statement reflects the need to transport out all regolith that was produced within the box.

That is about as far as we can go in the problem without getting more specific about just how the regolith moves, what rule governs how regolith transport is accomplished. This is an analogous place in the problem to that we achieved after casting the conservation of heat problem. All we had at that point was a statement that in which the steady case demanded that the heat transport took on a particular spatial pattern. We could not get all the way to a measurable quantity like temperature until we became more specific about how the heat was transported. In the problem presented in the 1st chapter, recall that we called upon conduction of heat as the specific transport process, and then could appeal to Fourier's Law to connect heat transport to temperature, in particular to temperature gradients. This is called closing the problem.

In order to make progress on the shape of a hillslope, we must now close the conservation of regolith equation we have just developed by appeal to a specific transport process. More concretely, we must address what dictates the regolith discharge, Q. A very general regolith flux or discharge rule might look something like:

$$Q = -kx^m \frac{dz}{dx}^n \qquad (10.7)$$

where m and n are constants setting the relative importance of distance from the divide, x, and of slope, dz/dx, respectively, and k sets the efficiency of the process. That the topographic slope is involved is not a surprise, as it is gravity that drives most hillslope and for that matter most geomorphic processes. In the diffusive case, m=0 and n=1, in other words the regolith discharge depends solely, and linearly, on the local slope, dz/dx, and the distance from the divide is irrelevant. In this case, our general statement for conservation of regolith, equation 10.4, becomes

$$\frac{\partial R}{\partial t} = \frac{1}{\rho_b}\left[\rho_r \dot{W} - \frac{\partial(-k(\partial z / \partial x))}{\partial x}\right] \qquad (10.8)$$

If we make the further simplifying assumption that the hillslope efficiency, k, is spatially uniform, i.e., does not depend on x, then we can remove the k from the derivative to obtain:

$$\frac{\partial R}{\partial t} = \left(\frac{\rho_r}{\rho_b} \right) \dot{W} + \kappa \frac{\partial^2 z}{\partial x^2} \tag{10.9}$$

This is classified as a diffusion equation with a source term. We have bundled $\kappa = k/\rho_b$, which has been called the landscape diffusivity in analogy with the thermal diffusivity, $\kappa = k/(\rho c)$. The coupling of this to the bedrock topography comes through

$$\frac{\partial z_b}{\partial t} = -\dot{W} \tag{10.10}$$

This acknowledges that as the bedrock is transformed into regolith, the bedrock interface, z_b, is lowered. The topography itself is simply

$$z = z_b + R \tag{10.11}$$

In general, these diffusion equations can be solved analytically only under very simple conditions. Because the diffusion equation is so common, solutions for many particular conditions are available. One useful compendium is Carslaw and Jaeger's Conduction of Heat in Solids, which has much broader application than its title implies. These solutions were appealed to in the pioneering work of Culling (e.g., Culling 1960, 1965).

Returning to the Gilbert steady state case, we require the solution for the case in which the left hand side is set to zero. This transforms the equation into an ordinary differential equation, one in which the dependent variable, z, depends upon only one variable, x. If the regolith production rate is uniform (not a function of x), and the slope efficiency k is also uniform, then:

$$\frac{d^2 z}{dx^2} = -\frac{\rho_r}{\rho_b} \frac{\dot{W}}{\kappa} \tag{10.12}$$

This says that the curvature, the left hand side, is a constant. Note that the curvature is negative, meaning the topography is convex up. We wish to know not the curvature but the topography itself, $z(x)$. This will require two integrations, each requiring evaluation of a constant of integration, for which we appeal to known boundary conditions.

Equation 10.12 can be integrated once to obtain the slope as a function of position:

$$\frac{dz}{dx} = -\left[\frac{\rho_r \dot{W}}{\kappa \rho_b} \right] x + c_1 \tag{10.13}$$

As we argued above, the constant of integration, c_1, is 0 because at x=0 (the hillcrest) the slope is zero. A second integration yields the topographic profile across the hilltop, $z(x)$:

$$z = -\left[\frac{\rho_r \dot{W}}{2\kappa \rho_b} \right] x^2 + c_2 \tag{10.14}$$

where c_2 is another constant of integration. Note that c_2 is an elevation. We may choose it by dictating the elevation at some location. One choice, the easiest, is the hillcrest, where x=0, where we can dictate that the elevation = z_{max}. Then $c_2 = z_{max}$, and the equation becomes

$$z = z_{\max} - \left[\frac{\rho_r \dot{W}}{2\kappa\rho_b}\right] x^2 \qquad (10.15)$$

This describes an inverted parabola, a simple geometrical shape with the characteristic that the slope is zero at the crest and linearly increases with distance from the crest. This is the essence of a steady state diffusive landscape: hilltops are convex, and in fact are parabolic.

Alternatively, we could specify the elevation of the edges of the hillslope, at the channel that bounds the hillslope, at $x = L$. Asserting that the channel is incising at a steady rate e, the boundary condition becomes

$$z(L) = z_o - \dot{e}t \qquad (10.16)$$

and we use this to solve for the constant of integration, c_2:

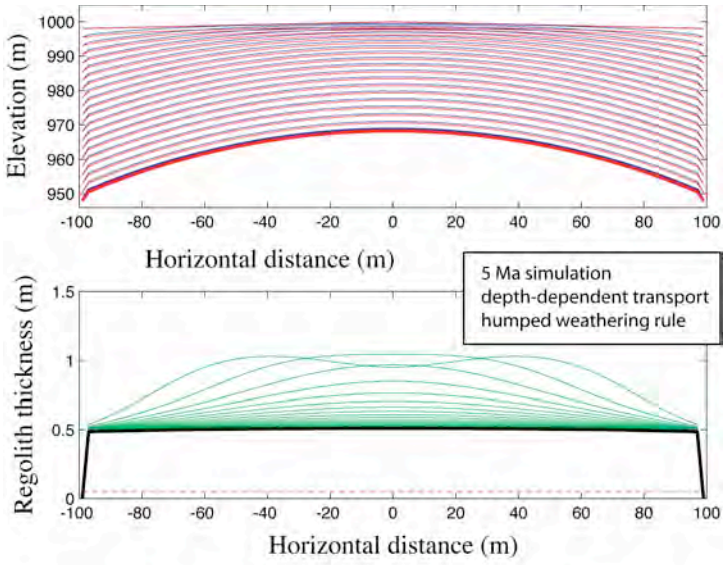
$$c_2 = z_o - \dot{e}t + \left[\frac{\rho_r \dot{W}}{2\kappa\rho_b}\right] L^2 \qquad (10.17)$$

When inserted into equation 10.14, we see that the topography can be described with

$$z = z_o - \dot{e}t + \left[\frac{\rho_r \dot{W}}{2k}\right]\left(L^2 - x^2\right) \qquad (10.18)$$

where we have employed $k=\kappa\rho_b$. The first two terms represent the channel incision (the boundary conditions for the hillslopes), while the third term represents the parabolic shape of the interfluve. The expression in square brackets is the local relief of the landscape, the difference between *z(0)* and *z(L)*. Let us see if this makes any sense. Recall that the regolith discharge must increase linearly with distance from the crest. Inspection of equation 10.18 reveals that for a given hillslope length L and channel incision rate e, the relief goes up as the rate of regolith production goes up, and the relief goes down as the landscape efficiency increases. For a given k, a higher rate of regolith production requires greater slopes everywhere, which in turn requires higher local relief. On the other hand, for a given W, the more efficient the hillslope transport is, the lower the slopes required for the required transport, and hence the lower the relief.

One can also model the evolution of hillslopes using a numerical model. In the simulation shown in Figure 3, I depict the evolution of a single interfluve over 5 million years. Embedded in the code are rules for the generation of regolith from bedrock, W, and for the transportation of regolith once it is generated, Q. While these rules are more fully discussed in Anderson (2002), suffice it to say here that the transport rule is one in which Q increases linearly with slope. As this should result in a parabolic steady state profile, we start out with that expectation. You can see that this is indeed the shape the hilltop finally takes on, but here we can see just how long it takes to get to that shape (the "transient response", and the kinds of regolith thickness profiles that might occur during this transition to the steady form.



Hillslope evolution in the face of regolith production and diffusive downslope transport processes. Top: evolution of the topographic profile from flat initial condition to steady-state parabolic shape. red=regolith-bedrock interface; blue: regolith-atmosphere boundary. Bottom: evolution of regolith thickness. Initial thickness 0.05 m thick (red dashed) evolves to steady uniform thickness of 0.5 m (black), with intermediate times in which regolith is greater in thickness (green). (rules described more fully in Anderson, 2000)

Although we will not progress deeply down this line, we can also expect certain behavior of the system in transient cases. Inspection of the original diffusion equation reveals that places of any high initial curvature (another name for sharp corners) will be smoothed most rapidly. This is another case in which nature is shown to abhor sharp corners. Sharp corners in topography, in temperature, in concentration of ink in water… all of these are smoothed by diffusive processes.

Clearly, the next steps involve opening up the black box that is Q. What sets the efficiency of hillslopes in passing regolith downslope? How does climate come into the act? How does the nature of the regolith, its grainsize distribution, its cohesion, etc., play into the efficiency of the transport process. While this is beyond the scope of this little book, we point you toward recent work on these problems. For rainsplash, see Furbish et al. (2007). For biological agents, see Gabet et al. (2003) and see Furbish et al. (in prep.). For frost heave, see Anderson (2002), as used in the simulation shown. For the general case of nonlinear creep, see Roering et al. (2001). For discussion of influence of dissolution on hillslope form, see Mudd and Furbish (2006).

11. Groundwater

The least visible portion of the water cycle involves water traveling through the subsurface. This transport and storage system is fed by infiltration from the surface, and delivers water to streams and oceans. As the speed of water through the pores of soil and rock is very small, and the volume of water involved is large, the time water spends in the subsurface can be quite long. The slow speeds serve to establish a long timescale filter to the inputs, allowing the outputs from the subsurface water system to be lagged significantly relative to the inputs. This lag is important in that the return of groundwater to the surface at the edges of our streams supports their baseflow long after the snowmelt or rainfall inputs have ended.

Geomorphologists are also interested in groundwater as a common trigger for landslides. The water table height, and the dynamic pressures associated with flow of water in the subsurface can support a fraction of the weight of the overlying sediment or rock, reducing the effective normal stress, promoting its failure. Large landslides can occur well after water is delivered to the surface as either rain or snowmelt. The lag reflects the long timescale of water movement within the subsurface – the deeper the path, the longer the timescale.

Water from either rain or snowmelt must first move through the unsaturated zone near the surface before it reaches the saturated zone. The unsaturated zone is called the vadose zone, the saturated zone is the phreatic zone or groundwater, and the interface between the two is the groundwater table. Movement of water in the vadose zone is very complicated, as the efficiency with which water moves is highly dependent upon the water content of the soil – i.e. the degree to which the pores are filled. This is comparable to heat problem in which the thermal conductivity depends upon the temperature. This was codified early in the 1900s first by Gardner (1919) and then by Richards (1931), in what is now known as the Richards equation (see review in Raats and van Genuchten, 2006). We will dodge this zone entirely and focus on flow within the saturated zone, fed by infiltration rates through the vadose zone that we will simply prescribe. We will see that this is complicated enough, and serves the present purpose of demonstrating the relevance of the conservation principle in setting up the problem.

As usual, we approach the problem by combining a statement for conservation with one for flux. Consider an element of the subsurface with a characteristic porosity and hydraulic permeability. As we are limiting the discussion to the groundwater system, which by definition is saturated, we may assume that all pores are filled with water. We seek a mathematical statement that captures this word statement:

The rate of change of water in the element = the rate of inputs of water – the outputs of water + rate of local gain or loss of water

The last term reflects the possibility that water can be either captured or freed from its fluid form by mineral reactions. For example, in accretionary wedges at subduction zones, where water pressures are all-important in setting the material strength of the wedge, transformation of smectite to illite represents such a source of water (e.g. Saffer and Bekins, 1998, 1999). We will ignore this complexity in the treatment below. So also, we will ignore the reality that the porosity and hence permeability of a rock or soil can be dynamic due to compaction and/or to cementation; we will assume that porosity and permeability are constant.

We are looking for an equation that describes the evolution of the groundwater table in both space and time. The control volume is a vertical column extending down from the surface with horizontal dimensions *dx* and *dy*, within which the water table exists at a height *h* above some arbitrary datum (Figure 1).

We seek the simplest statements. The conservation statement then becomes:

$$\frac{\partial (Shdxdy)}{\partial t} = q_x\big|_x - q_x\big|_{x+dx} + q_y\big|_y - q_y\big|_{y+dy} + Idxdy \qquad (11.1)$$

where *S* is the storativity of the aquifer. Also called the effective porosity, this is fraction of the sediment or rock that can be occupied by water that can freely come and go, the remainder being held firmly at mineral interfaces; it is smaller than the porosity. The lefthand side represents the rate of change of volume of water within the groundwater in the control volume, where h is the top of the saturated zone. Dividing both sides by the planview area of the element, *dxdy*, and assuming that the storativity does not change in time, in the limit as *dx* and *dy* approach zero, we have:

$$\frac{\partial h}{\partial t} = \frac{1}{S}\left[I - \frac{\partial q_x}{\partial x} - \frac{\partial q_y}{\partial y} \right] \qquad (11.2)$$

This is a general evolution equation for the height of the groundwater table. It should look familiar, as it contains terms for a source, and for gradients in the fluxes. Without further analysis, we can already anticipate behavior of the system in some simple situations. For example, consider the 1-D case, for groundwater flow only in the *x* direction (all derivatives in *y* therefore vanish), when fed by a uniform input of water by infiltration from the vadose zone above. In other words, let *I* be uniform at I_o. Once the system has become steady (all time derivates vanish), the groundwater discharge in the *x*-direction must be

$$q_x = I_o x \qquad (11.3)$$

Groundwater discharge should increase linearly with distance form the groundwater drainage divide. The reason for this behavior in the steady state directly mirrors other cases we have seen, for example the linear increase in either overland flow of water, or of regolith discharge on a hillslope. The discharge must increase downslope from the divide in order to transport the added increment of water or regolith.

To proceed further toward a prediction of the distribution of heights of the water table, we must assume some transport rule for the water. At this point in the development, we have not specified the physics of transport, but have simply exercised the principle of

conservation. We need something analogous to Fourier's law to "close" the equation, to transform it from one in fluxes to one in water table heights. Here we appeal to Darcy's Law, in which the transport in a given direction (here x) is given by

$$q_x = -K\frac{\partial H}{\partial x} \tag{11.4}$$

where H is the total head, and K is the hydraulic conductivity of the material. Let's inspect each of these components. The conductivity must depend not only upon the permeability of the material, κ, but upon the viscosity of the fluid being pushed through it, μ:

$$K = \frac{\kappa}{\mu} \tag{11.5}$$

In our problems this fluid will be water, and any variability in its viscosity will be due to the temperature dependence of viscosity. The permeability is set by the sizes of the pores and their degree of connectivity; it varies by many orders of magnitude. Many models exist of permeability in porous media, most of them consisting of flow through pipes of varying size, set by the effective pore diameters (see for examples the treatments in Furbish (1997) and in Turcotte and Schubert (2002)). The total head is defined as the head associated with elevation of the parcel of water (the elevation head) and that associated with the pressure (the pressure head).

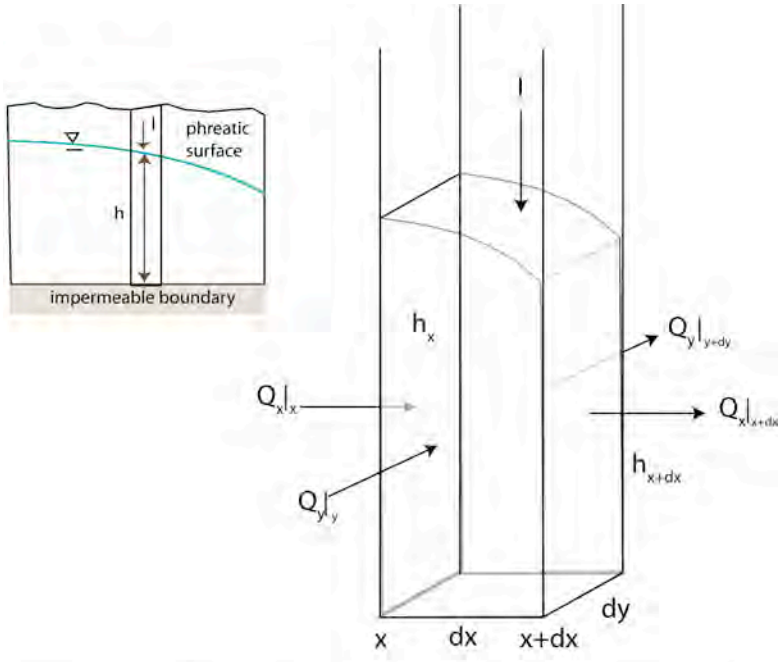


Figure 1. Schematic diagram for a water balance in a phreatic aquifer with source from above by infiltration, I, and lateral discharge of water through the sides of the element. Mismatches of inputs and outputs will drive change in the storage of water in the element, reflected in the height of the water table, h.

In the case of a phreatic aquifer, as illustrated in Figure 1, the discharge Q through the side of the control volume is the product of this Darcian flux (q, with units of volume/area/time) with the height of the water table, h:

$$Q_x = -kh\frac{\partial h}{\partial x} \tag{11.6}$$

Here the combination *kh* is often called the transitivity of the aquifer. Combining this with the equation for continuity of groundwater (equation 2), we obtain

$$S\frac{\partial h}{\partial t} = \frac{\partial}{\partial x}\left(kh\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(kh\frac{\partial h}{\partial y}\right) + I \tag{11.7}$$

This had better make sense in the simplest cases. For example, if the water table were flat and there were no flow, all terms on the right hand side except the source term, *I*, disappear, and we are left with

$$\frac{\partial h}{\partial t} = \frac{I}{S} \tag{11.8}$$

The rate of water table rise depends linearly on the rate of infiltration and inversely as the storativity. For low storativity (effective porosity), the rate of water table rise is magnified, simply because there aren't as many pores into which to stick the water. This makes sense.

The Dupuit case

Now consider a steady case in which infiltration is spatially uniform, and flow is in one direction, call it *x*. If the lower boundary is in fact impermeable, and the length scale of the hillslope is much greater than the height of the water table, the flow in the majority of the aquifer will be roughly horizontal. Under these conditions, dubbed the Dupuit approximations after Jules Dupuit (1804-1866, a French economist and civil engineer responsible for supervision of construction of the Paris sewer system), the equation above reduces to

$$0 = \frac{d}{dx}\left(kh\frac{dh}{dx}\right) + I_o \tag{11.9}$$

This is at least now an ODE rather than a PDE. We have made progress. Further assuming that the hydraulic conductivity, *k*, is uniform, we may remove it from the derivative, yielding

$$\frac{d}{dx}\left(h\frac{dh}{dx}\right) = \frac{-I_o}{k} \tag{11.10}$$

This looks messy, as it is nonlinear (the product of a variable and its derivative), but we may employ a trick. Recognizing that $d(h^2)/dx = 2h\, dh/dx$, this ordinary differential equation becomes

$$\frac{d^2\left(h^2\right)}{dx^2} = \frac{-2I_o}{k} \tag{11.11}$$

This is a second order ODE for h^2. We may now proceed to integrate both sides of the equation twice in order to solve for h^2, and then take the square root:

$$\frac{d\left(h^2\right)}{dx} = \frac{-2I_o}{k}x + c_1 \tag{11.12}$$

where c_1 is a constant of integration. We will evaluate this after a second integration. This yields

$$h^2 = \frac{-I_o}{k}x^2 + c_1 x + c_2 \tag{11.13}$$

We now impose boundary conditions, knowledge of the system at its boundaries. Here we will assert that the groundwater table intersect the top of the streams that bound the interfluve, here taken to be h_o at $x=0$ and h_L at $x=L$. In general these could be at different elevations. The condition at $x=0$ requires that $c_2=h_o^2$. The condition at x=L then results in

$$c_2 = \frac{\left(h_L^2 - h_o^2\right)}{L} + \frac{I_o}{k}L \tag{11.14}$$

The final formula for the shape of the water table is therefore

$$h^2 = h_o^2 + \frac{I_o}{k}\left(Lx - x^2\right) + \left(\frac{x}{L}\right)\left(h_L^2 - h_o^2\right)$$

$$or \tag{11.15}$$

$$h = \sqrt{h_o^2 + \frac{I_o}{k}\left(Lx - x^2\right) + \left(\frac{x}{L}\right)\left(h_L^2 - h_o^2\right)}$$

This is an equation for an ellipse, and is depicted in Figure 2 for several ratios of I_o/k. The height of the water table is scaled by the ratio I_o/k. The stronger the source term, I_o, the higher the arch of the water table, and the more efficient the flow of water within the groundwater system, scaled by k, the lower the water table arch.

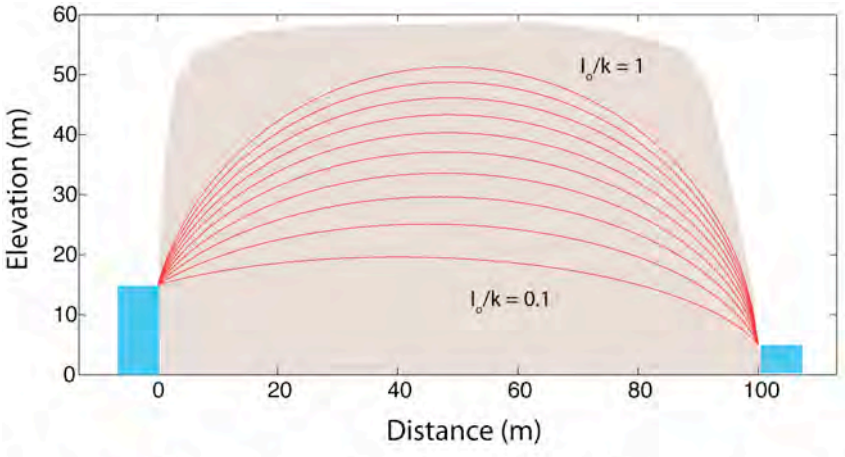


Figure 2. Calculated phreatic surfaces (groundwater tables) for ten ratios of I_o/k, in the subsurfcae of an interfluve between two bounding channels with differing stages, 15m and 5 m. Each water table describes an ellipse. The greater the ratio of the infiltration rate to the hydraulic conductivity, the greater the arch in the groundwater table..

I note that this form is perfectly analogous to the expected shape of the topography in the face of a source strength (weathering rate) and efficiency of regolith transport, as

discussed in the hillslope section. In that system we expect parabolas; in this system we expect ellipses.

Obviously, pumping from a well into a water table will result in a transient lowering of the phreatic surface while the pump is on, and a rebound of the cone of depression after the pump is turned off. Transient cases of more geomorphic interest include the response of the groundwater table to oscillations in the stage of the adjacent river. We can anticipate that as the river stage rises, the water table may tip back toward the interfluve, while as it drops, the groundwater table gradient will increase toward the river. The former represents recharge of the groundwater table, the latter leakage from it.

References

Anderson, M., 2007, Introducing groundwater physics. *Physics Today* May 2007.

Bear, J. 1988, *Dynamics of Fluids in Porous Media*. New York: Dover Press, 784 pp.

Bear, J. 1979, *Hydraulics of Groundwater*. Dover Press. 569 pp.

Freeze, R. A. and J. A. Cherry, 1979, Groundwater, Prentice-Hall, Englewood Cliffs, 604pp.

Gardner, W., 1919, The movement of moisture in soil by capillarity, *Soil Science* 7: 313-317.

Furbish, J. D., 1997, *Fluid Physics in Geology*, Oxford Press: New York, 476 pp. (see ch 9, 13)

Hornberger, G. M., J. P. Raffensperger, P. L. Wiberg, and K. N. Eshleman, 1998, *Elements of Physical Hydrology*, The Johns Hopkins University Press, 312 pp.

Middleton, G. V. and P. R. Wilcock, 1994, *Mechanics in the Earth and Environmental Sciences*, Cambridge University Press.

Raats, P. A. C. and M. T. van Genuchten, 2006, Milestones in soil physics, *Soil Science* 171:S21–S28.

Richards, L. A., 1931, Capillary conduction of liquids through porous mediums. *Physics* 1:318–333.

Saffer, D. M. and B. A. Bekins, 1998, Episodic fluid flow in the Nankai Acceretionary Complex: Timescale, geochemistry, flow rates, and fluid budget, *J. Geophys. Res.*, 103 (B12): 30,351-30,370

Saffer, D. M., and B. A. Bekins, 1999, Fluid budgets at convergent plate margins: Implications for the extent and duration of fault zone dilation, *Geology* 27 (12): 1095-1098.

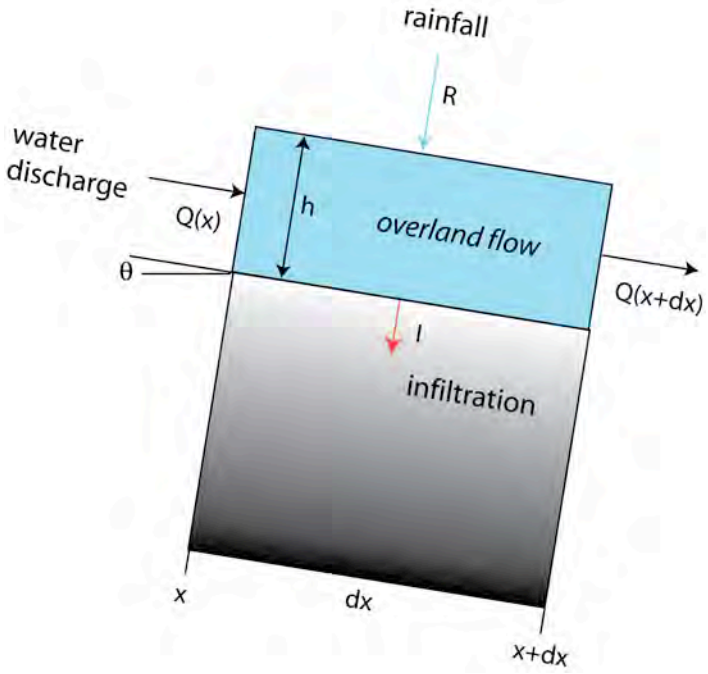
Turcotte, D. L. and G. Schubert, 2002, *Geodynamics*, 2nd edition, Cambridge University Press.

12. Overland flow of water

Sunset supercell thunderstorm above South Park, Colorado.

The movement of water across and within the landscape accomplishes most of the geomorphic work in molding the landscape. Now that we have set up the balance for regolith conservation, and have determined that it is the spatial pattern of regolith discharge that "counts" in thickening or thinning the regolith, one might imagine that the next step in addressing the evolution of landscapes would entail treating how sediment or regolith is transported by motion of water across it.

Under certain circumstances, water flows across the landscape as a sheet. In particular when rainfall rates are very much higher than infiltration rates, and the surface is not deeply crinkled into little gullies and rills, this process will occur. As it was first studied by Robert Horton in the 1930s and 1940s (e.g., Horton, 1945), this runoff process has become known as Horton overland flow. The sediment transport accomplished by overland flow is called sheetwash. Once again, we can formalize this problem by setting up a conservation of something in a control volume. Here the "something" is water atop the land surface.



Consider a simple case of a planar slope at an angle θ onto which rain is falling at a rate R (mm/hr, say) (Figure 1). Rainwater is infiltrating at a rate I. The box is dx long along the slope, and again I ask you to imagine that the box in 2D has a length into the page of dy. In order to assess what the spatial pattern of sediment transport is, we must know the pattern of shear stress on the land, which in turn requires that we know the thickness of the overland flow, h, as a function of distance from the divide, x. The word picture that captures this sketch is

Rate of change of volume of water in the box = rate of gain of water in the box – rate of loss of water from the box

Note that since the density of the material in transport does not evolve in any way, we can choose to treat its volume, rather than its mass. The volume balance may be written

$$\frac{\partial(hdxdy)}{\partial t} = Rdxdy - Idxdy + Q_x(x)dy - Q_x(x+dx)dy \qquad (12.1)$$

where here the discharge is that of water on the slope. Dividing by the area of the box, $dxdy$, we obtain the evolution equation for the water thickness:

$$\frac{\partial h}{\partial t} = (R-I) - \frac{dQ_x}{dx} \qquad (12.2)$$

By now this should look pretty familiar. The rate of change is due to two processes: sources and sinks (here R-I), and divergence of the flux or discharge of the medium. We see immediately that it is the excess rainfall (R-I) that counts in setting the effective source of water on the hillslope. We now ask what must the discharge of water look like in steady state, after which we can ask what the pattern of thickness must be to achieve this.

Downpour in Blue Hills badlands near Caineville, Utah. Raingage in the foreground records rainfall intensity, *R*, while PVC supports center sonic rangers over channel to record stage of resulting flashfloods.

At steady state, the left hand side can be set to zero, leaving us with the equation for the gradient in water discharge:

$$\frac{dQ_x}{dx} = (R - I) \tag{12.3}$$

As the rainfall rate and the infiltration rate are assumed to be uniform in space and unchanging in time in the steady case, this may be simply integrated:

$$Q = (R - I)x \tag{12.4}$$

As in the case of regolith on a hillslope, the steady discharge of water linearly increases with distance from the divide, at a rate dictated by the source strength. This is such a common finding in working these problems that it deserves a pause to think about its implications. The discharge *must* increase if continuity is to be obeyed. The system must figure out how to accomplish this. In general, the discharge (per unit contour length) is the product of the thickness of the material and its mean speed. The system then has two choices. Either the material (regolith, water) must thicken, or it must speed up, or some combination of the two.

In order to proceed, we need to relate the water discharge to the water thickness. In this simple case, the slope is uniform, and the only adjustable variable is flow thickness, h. The water discharge per unit contour width, Q, is equivalent to the product of the mean speed of the water and the flow depth: $Q=Uh$. We now need a formula for the mean speed. While there are several well-known formulas, we choose here the Darcy-Weisbach formula, in which

$$U = \frac{1}{f}\sqrt{ghS} \tag{12.5}$$

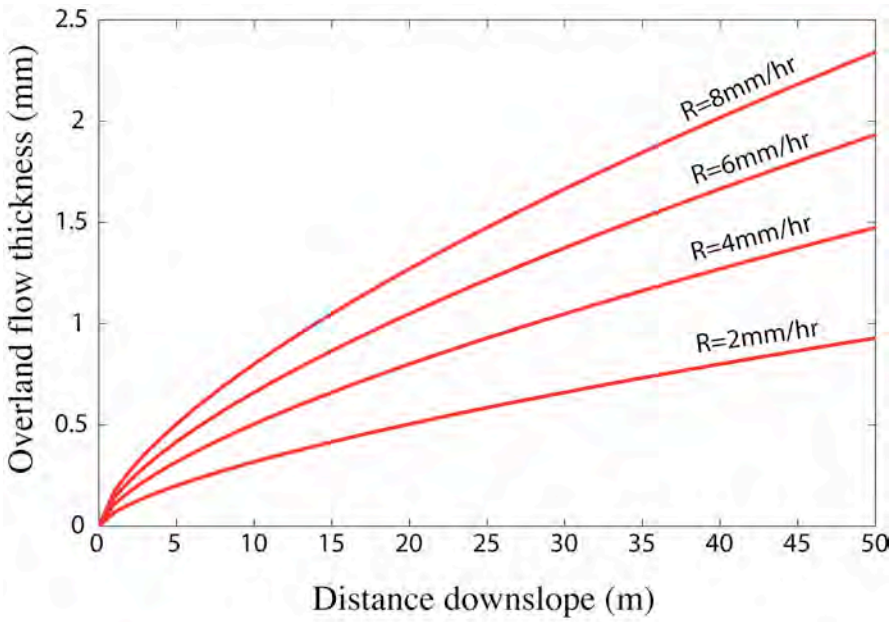
where f is the Darcy-Weisbach friction factor. The specific discharge of water is then

$$Q = Uh = \frac{1}{f}(gS)^{1/2} h^{3/2} \tag{12.6}$$

Equating equations 11.6 and 11.4, we can now solve for the dependence of the flow thickness on position, *h(x)*:

$$h = \left[\frac{f(R-I)}{(gS_o)}\right]^{2/3} x^{2/3} \tag{12.7}$$

The flow thickens with distance from the divide at a rate that is slightly less than linear, as can be seen by the 2/3 power of x (Figure 3). The constant out in front sets the thickness scale. Again, it is worth inspecting the solution to see if we can make sense of what is in the denominator and what in the numerator. It makes sense that the flow ought to be thicker at a given distance x if the excess rainfall is greater. If the slope is steeper, it also makes sense that the flow requires less thickness to accomplish the needed discharge, so that as S_o increases, h declines. The slope S_o should therefore be in the denominator, as it is.



Thickness of overland flow for four different rainfall intensities. Thickness increases as 2/3 power of distance downslope, and increases with rainfall intensity. All values of infiltration capacity, slope (constant slope in all cases), and roughness factor are the same for all cases.

While this process of overland flow, and the sediment transport that it can accomplish, called sheetwash, are now thought to occur on a small fraction of the Earth's surface, it is nonetheless instructive to have a solution for the end-member case against which to view all other processes. The analysis still pertains to parking lots, which cover a growing fraction of the Earth's surface.

13. Settling and Unsettling Problems

The motion of objects through fluids is important in many corners of earth sciences. Raindrops and snowflakes and sand and dust all interact with air, mineral grains move in the water column. Mantle plumes move through the mantle, blobs of magma move through the crust, and bubbles of gas rise through the magma. Animals move through the ocean, and their fecal pellets do too. And many sports involve moving objects through fluids: cricket, baseball, soccer, tennis and even badminton. In track and swimming events it is not a ball but the athlete moving through the fluid. These problems are connected through the physics of motion of an object through a fluid. In this section we will set up the problem in its simplest form, by treating the simple vertical motion of these objects. In other words, we will work a 1D problem. But the physics we set up is relevant and perfectly transportable into the more generic case of motion in any direction, of any object in any fluid. We will find that there are very specific attributes of the fluid and of the object being transported through it that dictate the speed of the object.

In contrast to the other problems we have set up in this book, we begin here by conserving something that we cannot see or touch: momentum. You are no doubt familiar with Newton's 2nd law, the shorthand for which we have all memorized as $F=ma$, where F is a force, m the mass of the object and a its acceleration. The more robust and more useful way to cast this is by turning it around to become $ma = F$, which

is of course legal, and then recalling that acceleration is just the rate of change of velocity. The left hand side then becomes *m(dv/dt)*. But as long as the mass is not changing we can place it inside the derivative, and then recognize that *mv* is momentum. Given this, we can then recast Newton's second law as:

The rate of change of momentum of an object = the sum of the forces acting upon it

$$\frac{\partial(ma_i)}{\partial t} = \sum F_i \tag{13.1}$$

where *i* represents the direction. We need expressions for the mass of the object and for the forces acting on it. In the case of a spherical object, and many of the objects listed above can be assumed to be spherical at some level, the mass is then the volume of a sphere times its density.

$$m = \rho V = \rho_p \pi \frac{D^3}{6} \tag{13.2}$$

where the subscript *p* indicates a property of the particle. The forces acting on the sphere come in two flavors: one associated with the pull of gravity, the other with any motion of the object through the fluid. The first we call the weight, or more formally the buoyant weight of the object; the latter we call the drag force.

Many of the problems mentioned in the introduction involve only vertical motion of the particle or object. These are settling problems, and let's take the settling of a raindrop in the atmosphere as an illustrative case. Our goal will be to estimate the speed at which a raindrop falls within the atmosphere. We can imagine that this depends upon the size of the drop. The force balance on a raindrop is illustrated in Figure 1.2. The vertical coordinate is *z*, taken to be positive upward, and, as is common in fluid mechanics, we call the vertical velocity *w*. The only acceleration in this case is in the vertical, meaning that we can write the acceleration of the drop to be the rate of change of the vertical speed, or *dw/dt*. The weight acts downward, in the negative *z* direction, and may be written as $F_W = mg$. But the *buoyant* weight of an object results from its immersion in a fluid of finite density. The buoyant weight is therefore

$$F_b = m_p g - m_f g = \left(\rho_p - \rho_f\right)\pi \frac{D^3}{6} g \tag{13.3}$$

This is the object's weight less that of the fluid it displaces. Note that if the density of the object and the fluid are the same, the buoyant force vanishes, and the object should be perfectly happy to sit still in the fluid; it is neutrally buoyant.

The opposing force acting to retard the fall, the drag force, is a little more complicated, as it depends strongly upon the speed of the object relative to the fluid. It is this feedback that allows there to be a finite, specific settling speed, or terminal velocity, or fall velocity, for an object. A general formula for the drag force is:

$$F_{drag} = C_d \left[\frac{1}{2}\rho_f w^2\right]\left(\frac{\pi D^2}{4}\right) \tag{13.4}$$

where ρ_f is density of the fluid through which the particle is moving (air in this case), and C_d is a non-dimensional parameter called the *drag coefficient*. While this looks complicated and might be difficult to commit to memory, it really has just three pieces. The piece I have put in [square] brackets is a dynamic pressure, which as all stresses do has units of a force per unit area. It represents the pressure difference between the upstream and downstream faces of the object. We see terms that look like this whenever we see dynamic pressures, those associated with motion of a fluid, represented in the equations for fluid motion. This is the pressure that is necessary to bring the fluid to rest at the edge of the object (which must be the case, as the fluid does not flow *through* the object, but around it).

As an aside, it is this relationship between pressure and local flow speed that is used to assess the airspeed of an airplane in what we call pitot tubes (named after Henri Pitot (1695-1771) a French hydraulic engineer who invented them in 1732 while exploring the flow speeds in the river Seine). These are the little L-shaped hook-like probes attached to the nose that you can see whenever you board a plane. The dynamic pressure, $P_{dynamic}$, is measured, and local airspeed is deduced from the formula $U = \sqrt{2(P_{dynamic} - P_{static}) / \rho_a}$. Here the static pressure is measured by air inlets along the side of the tube, while the dynamic pressure is measured at the stagnation point at the forward tip of the tube. The difference in pressures is measured using a differential pressure gauge. This yields the relative speed of the plane and the air. Simultaneous measurement using GPS documents the plane's motion relative to the ground. The difference between the speeds is due to the wind speed of the air through which the plane is pushing, something the pilots like to know.

The second piece, in (round) brackets, is the surface area of the object presented to the flow. In the case of a spherical object, this is a circle with diameter D. The drag force, then, is the product of a pressure (force per unit area) and an area. What remains, the third piece, is a dimensionless efficiency factor, the drag coefficient. It turns out that the drag coefficient is also, under some circumstances, dependent upon the speed of the object. The value of the drag coefficient depends upon whether the flow around the object (here our raindrop) is laminar or turbulent. In fluid mechanics, we rely upon a quantity called the Reynolds number, *Re*, to determine whether the flow field is laminar or turbulent. It honors the contributions of Osborne Reynolds (1842-1912), a 19th century British professor of engineering specializing in fluid mechanics. This is only one of the many non-dimensional numbers encountered in fluid mechanics problems, all of them named for famous fluid mechanicians. The Reynolds number represents the relative importance of inertial forces and viscous forces in the problem. The formula for the Reynolds number is

$$\mathrm{Re} = \frac{wD\rho_f}{\mu} = \frac{wD}{\upsilon} \tag{13.5}$$

where μ is the dynamic viscosity of the fluid, and ν the kinematic viscosity = μ/ρ_a. In this generic formula, the D represents a "relevant length scale" in the problem, while w is a "relevant speed" in the problem. Here the relevant length scale is the diameter of the object, while the relevant speed is the relative speed of the object and the fluid. Given

that we are assuming in the settling problem that the fluid is at rest, this speed becomes simply the vertical speed of the object. In channel flows, in contrast, the relevant length scale is taken to be the depth of the fluid in the channel, and the relevant speed is taken to be the mean speed of the fluid. But note that there is a bit of arbitrariness in these choices. Why for example choose the diameter and not the radius of the object, or the mean speed as opposed to the maximum or surface speed? This arbitrariness is solved by appeal to tradition; while it is an important tenet of science to be skeptical of past work, we must also respect our scientific predecessors. In this case we acknowledge that scientists studying the interaction of spheres with a fluid have settled upon the use of diameter and not radius in the construction of the Reynolds number for such problems, and we faithfully follow their convention.

The dependence of the drag coefficient upon the Reynolds number is shown in Figure 1. Ignoring the little dip out there at very high *Re* (a feature dubbed the drag crisis, at *Re* of 10^5 or so), there are essentially two asymptotic expressions, one at very high *Re*, the other at low Re. At high Re, the drag coefficient becomes a constant, at about 0.4. And at low *Re*, the drag coefficient is inversely dependent upon Re, following the relationship $C_d = 24/Re$. Note that the low Re asymptote displays a negative slope of -1: one factor of ten (or 1 log unit) over for one factor of ten (or 1 log unit) down. As straight lines on log-log plots represent power laws, with the slope indicating the power, we see that $Cd \sim Re^{-1}$, or $1/Re$. The factor of 24 merely sets the vertical position of the line. A piecewise function that spans the transition region between low (<1) and high (>100) Reynolds numbers is nicely tabulated by Morsi and Alexander (1972), and was used to plot the remainder of Figure 1.

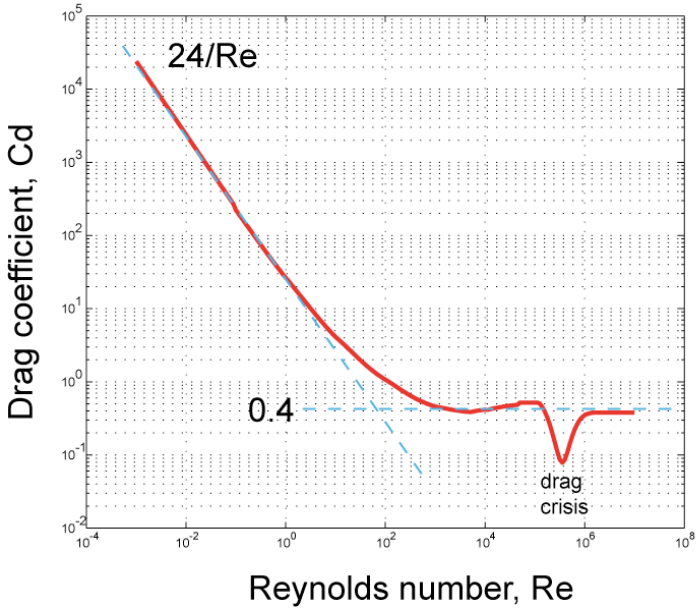


Figure x. Dependence of drag coefficient on Reynolds number. Low and high Re end-member cases are shown in dotted lines, with the transition region between. Here the dependence is taken from the piecewise function tabulated by Morsi and Alexander (1972).

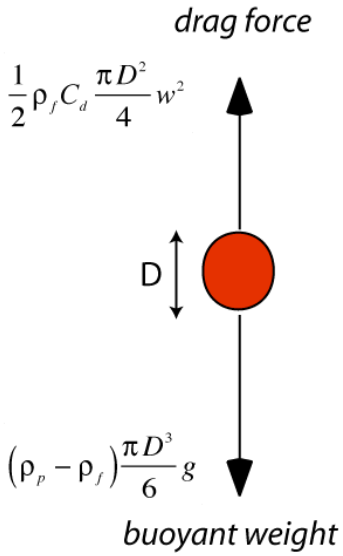


Figure Forces on a spherical particle falling in a still fluid. At the settling velocity, the drag force and the buoyant weight are exactly balanced. The drag coefficient, Cd, is a function of the Reynolds number, Re.

Now we have all the pieces to the problem. The force balance now reads:

$$m\frac{dw}{dt} = \rho_p \frac{\pi D^3}{6}\frac{dw}{dt} = \sum F = (\rho_p - \rho_f)\frac{\pi D^3}{6} g - \frac{1}{2}\rho_f C_d \frac{\pi D^2}{4} w^2 \qquad (13.6)$$

We can now divide through by the mass of the object to isolate an equation for the acceleration of the object in the vertical:

$$\frac{dw}{dt} = \frac{(\rho_p - \rho_f)}{\rho_p} g - \frac{3}{4}\frac{\rho_f}{\rho_p} C_d \frac{1}{D} w^2 \qquad (13.7)$$

To determine the settling velocity, which by definition occurs when the buoyant weight (represented by 1st term on the right hand side) is exactly balanced by the drag force (represented by second term on the right hand side), we simply set the left hand side of equation 13.7 to zero, plug in the relevant formula for the drag coefficient, and solve the equation. This results in two expressions for settling speed, one for low and the other for high Reynolds numbers. For low Re, the full equation becomes:

$$\frac{dw}{dt} = \frac{(\rho_p - \rho_f)}{\rho_p} g - \frac{3}{4}\frac{\rho_f}{\rho_p}\frac{24\mu}{\rho_f w D}\frac{1}{D} w^2 = \frac{(\rho_p - \rho_f)}{\rho_p} g - \frac{18\mu}{\rho_p D^2} w \qquad (13.8)$$

Setting the left hand side to zero to obtain the settling speed results in

$$\text{low Re}: w_{sett} = \frac{gD^2(\rho_p - \rho_f)}{18\mu} \qquad (13.9)$$

You may have heard of this as Stokes' Law, named for George Gabriel Stokes (1819-1903) the Cambridge professor of mathematics who derived the solution for dilute

suspensions of grains in viscous fluids. This is the same Stokes for whom the Navier-Stokes equation is named, the governing equation for fluid motion.

For high Re, the full acceleration equation becomes:

$$\frac{dw}{dt} = \frac{(\rho_p - \rho_f)}{\rho_p} g - \frac{3}{4}\frac{\rho_a}{\rho_p} 0.4 \frac{1}{D} w^2 = \frac{(\rho_p - \rho_f)}{\rho_p} g - \frac{3}{10}\frac{\rho_f}{\rho_p}\frac{1}{D} w^2 \qquad (13.10)$$

And the solution for a settling speed is:

$$\text{high Re}: w_{sett} = \sqrt{\frac{gD(\rho_p - \rho_f)}{0.3\rho_f}} \qquad (13.11)$$

But there is a problem. How do you know which formula to use? Here is one strategy. If you are doing such a calculation, you can and should check to make sure you have chosen the correct formulation of the drag coefficient, by assessing the Re associated with the calculated settling velocity. If you used the low *Re* formulation, and find *Re*>>1 for the calculated settling velocity, or vice versa, you used the wrong formulation.

In Figure 2 we show the fall velocities for objects in both water and air, calculated using a numerical integration of the equation of motion of the particles, using the full tabulation of $C_d(Re)$ given by Morsi and Alexander (1972).

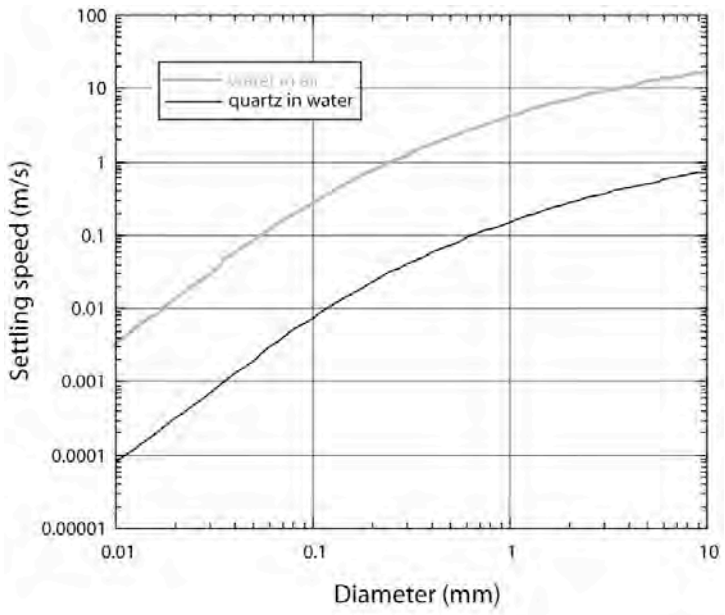


Figure 9.5 Settling velocity as a function of particle size for two cases: water droplets in air, and quartz grains in water (a) in linear-linear space, and (b) in log-log space. Although the settling velocity increases monotonically with particle size, note the steep increase (as D^2) for small particle sizes, and less dramatic increase (as $D^{1/2}$) for large particle sizes.

Now inspect this plot of settling speed *vs* diameter in log-log space. It has two straight limbs connected by a smooth curve. The straight limbs correspond to the two end-member cases of low and high Re. The low Re solution predicts that settling speed goes

as D^2, while the high Re case predicts that it goes as $D^{1/2}$. The slope of the plot for small diameters (hence small Re) indeed shows a slope of 2: two factors of 10 up for one factor of 10 over. And the slope for high diameters shows a slope of ½, also as expected.

As a part of his Ph.D. dissertation, Bill Dietrich assembled all the settling data for natural grains at that time, and produced a curve fit in which all this data was collapsed (Dietrich, 1982). He presents a fit for spherical particles using a 4th order polynomial:

$$\log(W^*) = -3.76715 + 1.92944(\log(D^*) - 0.09815(\log(D^*)^2 \\ -0.00575(\log(D^*)^3 + 0.00056(\log(D^*)^4 \quad (13.12)$$

where the settling speed has been non-dimensionalized using

$$W^* = \frac{\rho_f w_{sett}^{\ 3}}{\left(\rho_p - \rho_f\right) g \upsilon} \quad (13.13)$$

and the grain diameter is non-dimensionalized using

$$D^* = \frac{\left(\rho_p - \rho_f\right) g D^3}{\rho_f \upsilon^2} \quad (13.14)$$

He went on to show how the effects of grain shape may also be incorporated. This conveniently provides a single formula that may be used if you need to calculate a settling speed without the trauma of a numerical integration. I note in conclusion that given the kinematic viscosities of water and air, the Reynolds numbers for sand-sized grains will be in the transition region. This means that neither the high nor low Reynolds number formulae for settling speed will work. These formulae will both overestimate the settling speed, while Dietrich's formula captures this zone faithfully.

These results have several implications for earth science problems.

Raindrop transport

The efficiency of ejection of grains from a bare soil surface by raindrops depends upon the kinetic energy of the drop, ½ mw^2. The cubic dependence of mass, and hence volume on diameter suggests that KE will at least as the 3rd power of drop diameter. For small drops, the dependence of settling speed on D^2 adds another factor of D^4, suggesting KE~D^7, while for larger drops at the high Re end of the spectrum, the added dependence on speed is $(D^{1/2})^2$, suggesting KE~D^4. The largest drops will therefore do the lion's share of the work in moving sediment about. This highlights the importance of storms capable of producing the largest raindrops. These are the storms with tall convecting towers in which drops have sufficient time to grow to large sizes by coalescence with other drops.

Saltation vs suspension and the role of response time

The trajectories of grains in either air or water are dictated by how efficiently the particles respond to the turbulence in the fluid. Those that respond rapidly will become "suspended" in the flow, while those that respond sluggishly will either slide or hop short distances in what we call saltation trajectories. Formally, the difference results from differing response times of the particles. One estimator of the response times can be seen

from equation 12.8. To be dimensionally consistent with the other terms in the equation, the collection of constants in front of the last term on the right hand side must have units of 1/time. Its inverse is therefore a time scale,

$$T = \frac{18\rho_p D^2}{\mu} \tag{13.15}$$

What does this time scale represent? Let's see by solving the full equation 1.8 for the temporal evolution of the particle speed. The equation may be reduced to the simpler-looking

$$\frac{dw}{dt} = a - bw \tag{13.16}$$

This is a first order linear ordinary differential equation one finds in the first couple chapters of a book on ODEs. It looks a lot like the equation we solved for the evolution of cosmogenic radionuclide concentrations in Chapter 5. The *a* term represents a source, and the *-bw* term represents decay, or a feedback in the system. The solution to the equation is:

$$w = \frac{a}{b}(1 - e^{-bt}) = w_{sett}(1 - e^{-t/T}) \tag{13.17}$$

The ratio *a*/*b* is the settling speed we derived from setting the left hand side to zero to arrive at equation 1.9. Now we can see what this response time represents. If a particle is dropped from rest at time = 0, at a specific time *T* later, the particle will have achieved $(1\text{-}e^{-1})$ or roughly 2/3 of its final or terminal speed (Figure 3). By several times *T*, it will have reached its terminal or settling speed.

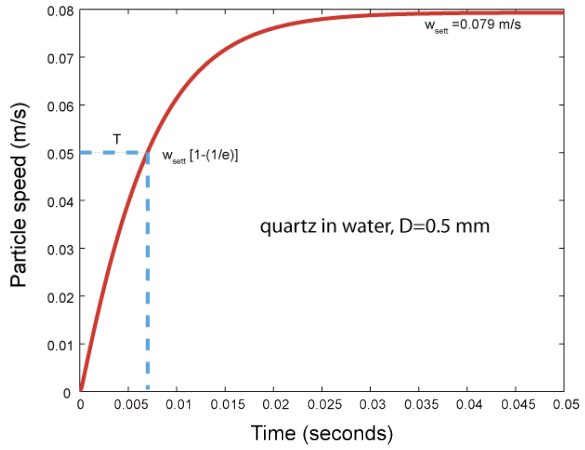


Figure x. Example of approach to settling speed, with characteristic response time T shown. Here the fluid is water (density = 1000 kg/m^3), and the spherical grain is quartz (density = 2650 kg/m^3). Within a few times T, the particle has achieved its settling speed.

Let's get back to our saltation problem (Have you heard Dylan Thomas reading his wonderful poem, "A Child's Christmas in Wales"? After a rousing monologue by the narrator, including descriptions of throwing snowballs at cats lined up in a row on the wall, the child whispers "Let's get back to the presents…".) A large particle (high *D*) with a long response time is less likely to respond to the turbulent fluctuations in air or water speed, and will carry out a smooth trajectory, while smaller particles with short response

times will perform more wiggly trajectories as they are embedded more deeply in the turbulence of the flow. The cases shown in Figure 4 illustrate how dramatic these differences can be. In each simulation, ten trajectories are shown, each with identical launch speed and angle. Each experiences turbulence in the flow near the bed that has the proper statistical description. In the saltation case, the trajectories are effectively deterministic: one can predict quite well the length of the trajectory and its speed and angle of impact knowing its initial conditions. This is not the case as smaller grains are launched. The process becomes stochastic as it reflects more and more strongly the stochastic nature of the turbulence, and we lose the ability to predict the outcome of a trajectory given only its initial conditions.

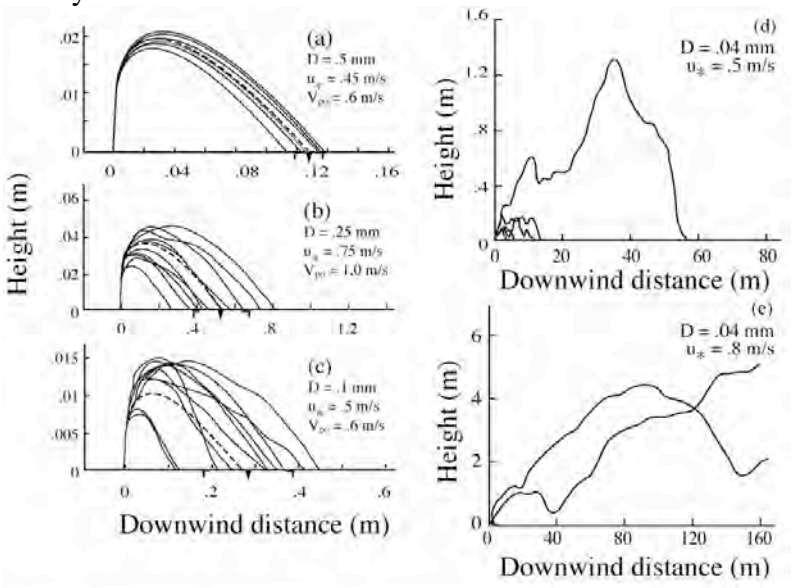


Figure 3.x Examples of calculated quartz grain trajectories in air. All ten trajectories shown in each panel are launched with the same angle and speed. Large grains traverse smooth saltation trajectories that do not vary much between realizations. As the grain size diminishes, or as the intensity of the turbulence increases (scaled here by the shear velocity, u*), the trajectories become more wiggly. Note also the change in horizontal scale between cases; some of the suspension trajectories cover many tens of meters. The process becomes more stochastic as grains take more random paths characteristic of suspension. (after Anderson, 1987 Journal of Geology)

Fecal pellets

Given the strong dependence of settling speed on particle diameter, imagine how much longer it would take organic matter to settle to the ocean floor before the 'invention' or evolution of fecal pellets – the packaging of poop. It has been suggested (Logan et al., 1995) that the evolution of fecal pellets allowed much more efficient delivery of organic material to the ocean floor, which in turn changed radically the cycling of elements in the ocean. The essence is that in order to reach the ocean floor any excreted organic material must be able to survive its transport through the oxygenated top layer of the ocean, and it is more likely to do so if packaged as pellets with high settling speed.

Explosivity of magmas

One of the main determinants of volcanic eruption style is how much gas the magma contains upon eruption. As the magma approaches the surface during emplacement

within the volcanic edifice, the pressure drops, allowing dissolved gases to be released. These gases then amalgamate into bubbles that are significantly less dense than the surrounding magma, and should rise toward the surface just as the air bubble rises within the honeybear on the breakfast table. This is settling upside down (shall we call it unsettling?). The moving object is the bubble and the fluid is the magma. The buoyant force is directed upwards, and the resistance due to drag, which always acts in the direction opposite to the relative motion of the fluid and object, is therefore directed downwards. We can therefore use the same formula to predict the rate of rise of a bubble. In particular, in the case of magmas, with viscosities of order 10^4-10^{16} Pa-s, the Reynolds numbers should be very small and we can use the low *Re* formula. As the rise speed has viscosity in the denominator, highly silicic magmas impede the rise of bubbles, and the gas will therefore be retained in the magma. It is these gas-charged magmas that upon eruption are driven to supersonic speeds in the momentum jet of a plinian eruption, propelled by the explosive expansion of the gas bubbles upon depressurization (see work of Steven Sparks (1997), and the extensive publications of George Walker and Lionel Wilson on explosive volcanism). The final expansion of the bubbles disaggregates the magma into very small shards, the old bubble walls, which then lose heat efficiently to the entrained air, allowing transformation of the momentum jet to a buoyant plume that can reach more than 10 km into the atmosphere.

Mantle plumes

At a larger scale, we can also think of rising mantle plumes using the same basic physics. The tops of these plumes are predicted to have a plume head that is roughly spherical, at least until they interact strongly with the rigid lithospheric lid (see Geoff Davies, 1999 book and simulations on his web site). Again, we can use the formula for the low Re case. Let's estimate how long it would take for a plume to rise to the base of the lithosphere from the core mantle boundary, roughly 3000 km below the surface. Taking the viscosity to be 10^{22} Pa-s, the diameter of the plume to be 200 km, and the density difference between the melt and the surrounding mantle to be 200 kg/m^3, the rise speed is 24 mm/yr, and the time to traverse the mantle is about 100 million years. These are long time scale earth events.

References

Anderson, R.S., 1987, Eolian sediment transport as a stochastic process: The effects of a fluctuating wind on particle trajectories. *Journal of Geology* 95: 497-512.

Davies, G. F., 1999, *Dynamic Earth: Plates, Plumes and Mantle Convection*, Cambridge Press, 458 pp.

Dietrich, W. E., 1982, Settling velocity of natural particles: *Water Resources Research*, 18(6): 1615-1626.

Edwards, B.F. *et al* 2001 Dynamics of falling raindrops, *Eur. J. Phys.* 22 113-118 doi:10.1088/0143-0807/22/2/302

Logan, G.A., Hayes, J.M., Heishma, G.B., and Summons, R.E., 1995. Terminal Proterozoic reorganization of biogeochemical cycles, *Nature* 376: 53-56.

Manga, M. and H.A. Stone (1994) Interactions between bubbles in magmas and lavas: Effects of deformation, *Journal of Volcanology and Geothermal Research*, vol. 63, 269-281.

Manga, M. (1997) Interactions between mantle diapirs, *Geophysical Research Letters*, vol. 24, 1871-1974.

Morsi, S.A. and A. J. Alexander, 1972, An investigation of particle trajectories in two-phase flow systems, *J. Fluid Mech.*, 55, 193-208.

Sparks, R.S.J., Bursik, M.I., Carey, S.N., Gilbert, J.S., Glaze, L.S., Sigurdsson, H., Woods, A.W. (1997) *Volcanic Plumes*, John Wiley & Sons Ltd, Chichester, UK, 574pp.

Syvitski, J.P.M., Murray, J.W., 1981. Particle interaction in fjord suspended sediment. Marine Geology 39 (3–4), 215–242.

Syvitsky, J.P.M. (ed.), 1991. Principles, Methods, and Application of Particle Size Analysis. University of Cambridge Press

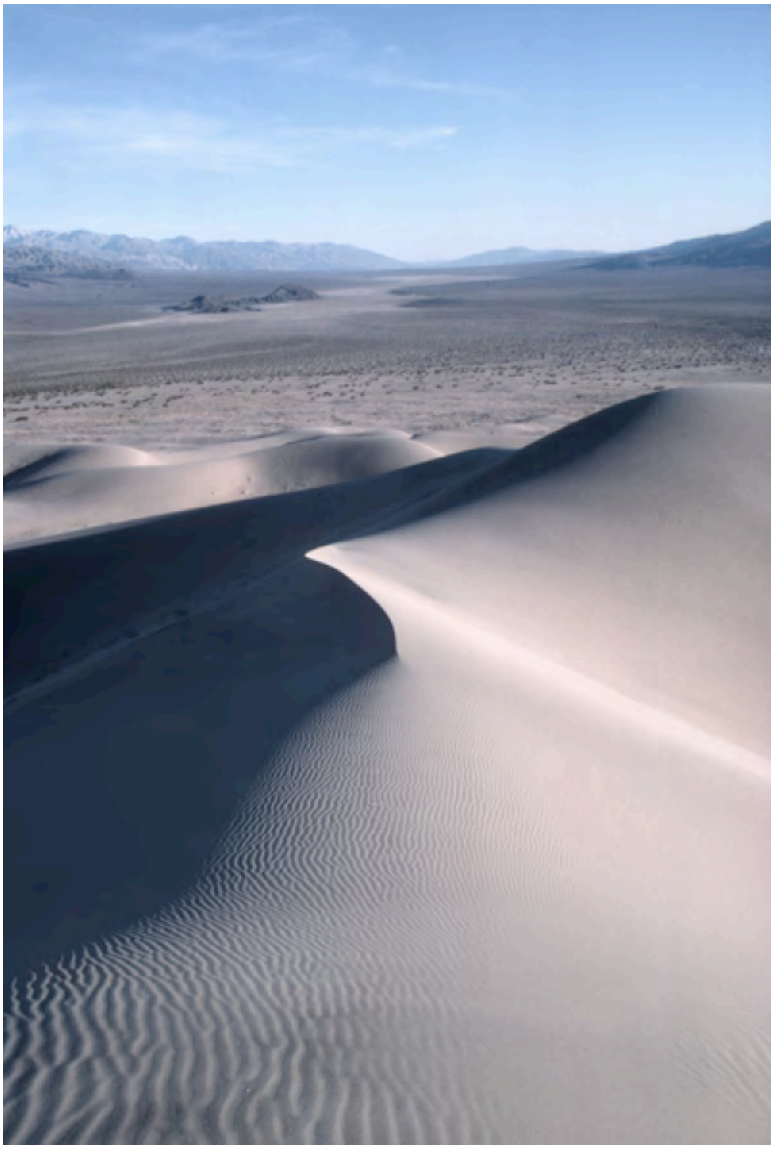
Van Rijn, L. C., 1993, Principles of sediment transport in rivers, estuaries, and coastal seas, Section 3.2.5: Particle fall velocities. Aqua Publications.

Vogel, S., *Life in Moving Fluids*.

Walker, G. P. L. 1983: Ignimbrite types and ignimbrite problems. *Journal of Volcanology and Geothermal Research 17*: 65–88.

Wilson, C. J. N.; Walker, G. P. L. 1985: The Taupo eruption, New Zealand. I. General aspects. *Philosophical Transactions of the Royal Society of London A314*: 199–228.

14. Sediment transport and bedforms



Arm of star dunes, northern Panamint Valley, California. Ripples in foreground record most recent winds from right to left.

The transport of sediment by a fluid, either water or air, inevitably results in the formation of waves or ridges on the bed, called bedforms. A flat bed is unstable to perturbations when subjected to an overlying fluid in motion. The backdrop for the study of these interesting examples of self-organization in geomorphology is conservation of sediment volume. The situation we choose to treat is one in which there is unidirectional flow of the fluid over the granular substrate, which in turn entrains grains from the substrate. Those grains traveling short hops before re-encountering the bed are called bedload and travel by saltation trajectories, while those that travel higher, more wiggly trajectories affected by the turbulence of the flow comprise the suspended load. The

control volume in this case is a box with a footprint *dx dy* (shown in Figure 1 in 1-D) and a top to which grains can either be gained or removed. The remainder of the bed does not participate, and merely serves to store grains that could be removed if significant erosion occurs. The quantity we wish to conserve is the volume of grains above some arbitrary level in the bed, the height of the bed being z_b. In other words, in this case the elevation of the top of the box is the variable in our problem.

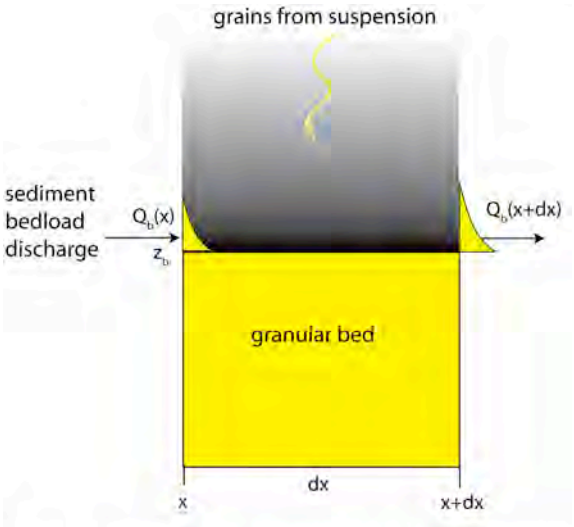


Figure x. Conservation of sediment in the face of sediment transport. This continuum formulation of sediment conservation leads to the Exner equation that dictates the dependence of erosion and deposition on the pattern of sediment transport.

The word picture for the system is

Rate of change of elevation of the sediment bed = rate of transport into left side of the column – rate of loss of sediment from the right side +/- rate of deposition or entrainment of suspended grains from the bed

Ignoring for the moment the suspended sediment, this may be cast formally as a continuity equation for the sediment

$$\frac{\partial z_b}{\partial t} = -\frac{1}{\rho_b}\frac{\partial Q_b}{\partial x} \qquad (14.1)$$

where Q_b is the mass transport rate of bedload sediment. This is commonly referred to as the Exner equation (Exner, 1920), representing conservation of sediment in a mobile bed. As usual, this forces us to address the spatial pattern of a flux, here the flux of bedload. Before entangling ourselves in the details of sediment transport mechanics, it should already be clear that erosion of the bed, where $dh/dt<0$, occurs where $dQ_b/dx>0$, in other words at sites of a down-flow increase in sediment transport. The opposite is also true: deposition occurs where $dQ_b/dx<0$. Put to words, sand accumulates in spots where more sand arrives from upstream than is lost downstream, and vice versa.

Now consider a bedform, as depicted in Figure 2. This could be either a dune or a ripple, and could be either eolian or fluvial.

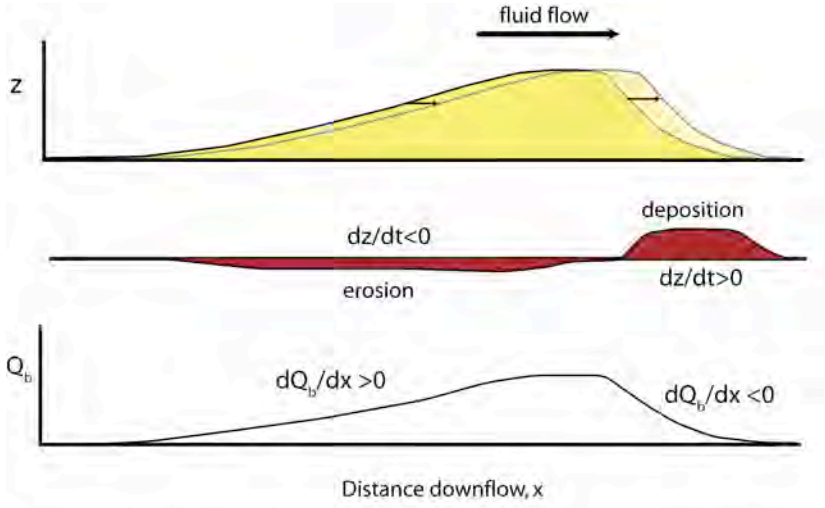


Figure x. Top: Migration of a bedform in the downflow direction, preserving its shape. Middle: the associated pattern of erosion on the stoss or upflow slope of the bedform, and deposition on the lee slope. Bottom: the inferred pattern of sediment transport, increasing on the stoss face, and decreasing on the lee.

If this form is to translate downstream with no net change in shape, we can already deduce what the pattern of sediment transport across the form must be. It must increase monotonically up the stoss face, reach a peak at the crest of the form, and then decline, as shown in the bottom panel of the figure.

15. Fluvial sediment balance

The problem of sediment transport in a fluvial system, viewed at a scale larger than that involved in bedforms, requires that we acknowledge the sources of sediment from hillslopes on the one hand, and from tributaries on the other. The former represent a continuous source, while the latter are a point source. The word picture of the system is

Rate of change of mass of sand in the alluvial reach = rate of sediment discharge into the upstream end of the reach – rate of loss of sediment from the downstream end of the reach + rate of inputs of sand from hillslopes on either side

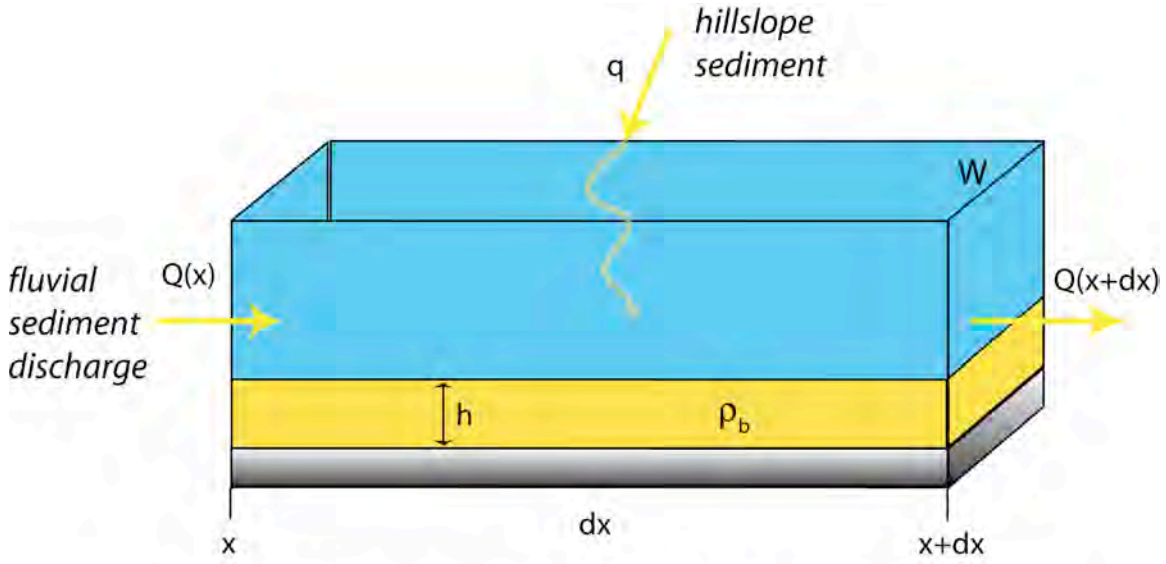


Figure x. Conservation of alluvial sediment in a fluvial reach of length dx and width W. The mass of sediment, and hence its mean thickness in the reach, will depend upon the gradients in downstream bedload sediment transport, and upon the strength of the local sources of bedload size sediment from the two adjacent hillslopes, 2*q.

Translating this word picture to math, the balance may be written:

$$\frac{\partial(\rho_b H W dx)}{\partial t} = Q_x - Q_{x+dx} + 2q_{hillslope}dx \qquad (15.1)$$

where the term in brackets on the left hand side is the mass of sediment in a reach of length dx, channel width W, alluvial sediment thickness H, and bulk density ρ_b. The last term on the right represents the inputs of sediment from the hillslopes, the 2 reminding us that it comes form both sides. I have ignored contributions from tributaries for the moment; they can be taken into account at tributary junctions. Dividing by the width of the channel, which we do not expect to change on short timescales, by the bulk density of sediment, and by the length of the reach, dx, this becomes

$$\frac{\partial H}{\partial t} = -\frac{1}{W\rho_b}\frac{\partial Q_x}{\partial x} + \frac{2}{W\rho_b}q_{hillslope} \qquad (15.2)$$

It looks similar to the Exner equation except that there is now a source term associated with sediment delivered from the hillslopes. A more apt comparison is therefore with the hillslope regolith problem, and any other in which a source is expected. The steady state problem may now be addressed, by setting the left hand side of the equation to zero, and integrating with respect to x. In the simple case in which the contributions from the adjacent hillslopes are spatially uniform, this leave us

$$Q = \frac{2}{W\rho_b} q_{hillslope} x \tag{15.3}$$

In this simplest case, the alluvial discharge must increase linearly with distance downstream, x. As the hillslope flux will depend upon the erosion rate of the hillslopes, the rate of increase in sediment discharge will depend upon the erosion rate of the landscape. This is perfectly analogous to the expected pattern of regolith discharge on a steady hillslope, and as we will see, to the expected pattern of coastal sediment transport in the face of a uniformly receding seacliff.

We can extend this analysis to a situation that is a little more complex and therefore a little more real. Let us imagine that the drainage basin is underlain by two lithologies that differ markedly in their erodibility and in their susceptibility to subsequent breakdown during fluvial transport (Figure 2).

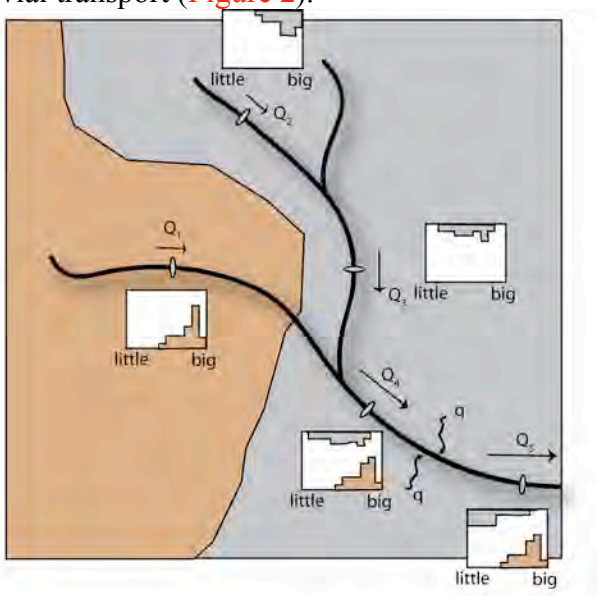


Figure x. Schematic of downstream fining of sediment from two tributaries with distinct lithologies. If the system is steady, total mass balance of sediment requires downstream increase in sediment discharge, Q, reflecting sediment contributions from hillslopes, q, and tributaries. The grainsize distribution of each lithology in the sediment mixture will vary with distance as well, reflecting the relative resistance of sediment of each lithology to comminution.

We may craft a sediment budget (balance) for not only the total sediment in transport, as given by the above equation, but for each lithology. This is straight-forward. From such a budget one could deduce the relative contributions from each portion of the catchment, and hence the relative erosion rates of basins underlain by each of the two lithologies. One can go even further, however, and balance each grainsize of each lithology. In this

case, the problem begins to take on elements of the demography problem of the biologists. Grains can erode, comminute, to become smaller (much like an animal can age), but cannot recombine to become larger (much as, sadly, animals cannot get younger). It is a one-way process. That this process takes place more rapidly in some lithologies than in others has long been known. Experiments designed to document the erodibility of clasts have been performed for many decades, with more and more realistic methods.

You can see from the histograms of grainsizes shown that the brown lithology breaks down much more slowly than the gray. Its grainsize distribution shifts much less with downstream distance. This means that by the last sample site the large grainsize classes are populated by the brown lithology almost exclusively. Think quartzite. Attal and Lave (2006) have nicely shown that this effect dominates the spatial pattern of grainsizes in the Marsyandi River, draining the Annapurna region of the Himalayas.

16. Coastal littoral sand budget

The same principles may be applied to problems involving coastal sediment. A littoral (from Latin *litus* for shore) cell is defined as being a self-contained unit of coastline within which the sand sources and sinks are contained (Figure 1). The top or up-coast boundary of a cell is defined as a line across the coast across which sand is not transported. The bottom of a cell is commonly a site where sand is lost from the littoral system down a submarine canyon. In other words, again sand does not cross the line normal to the coastline, but its instead lost from the coastal system. Sand moves along the beach or littoral system by longshore drift, being pushed by waves one little hop at a time. We may formalize the sediment balance in a portion of the cell just as on a segment of hillslope. We must consider both the sources and sinks of sediment, and the rate at which it is passed from one point to another within the system (the fluxes).

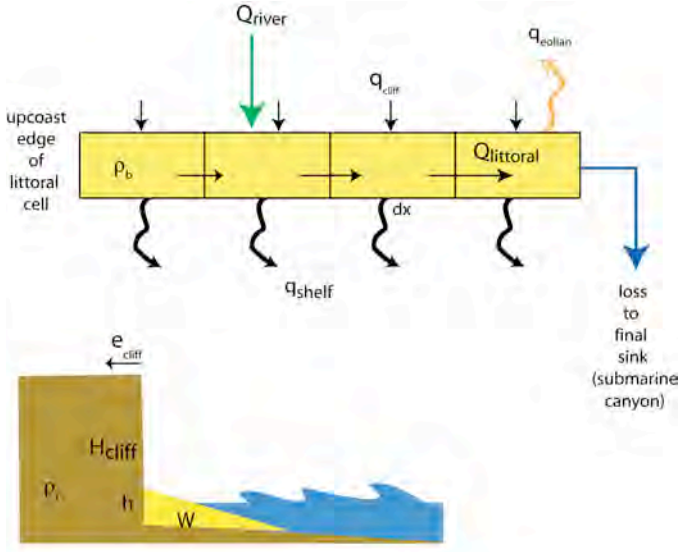


Figure x. Schematic diagram of sediment budget within a littoral cell, seen in planview (top) and in cross-section showing beach wedge (bottom). Sources of sediment include contributions from cliff retreat and rivers, while sinks include losses to the shelf in suspension, and loss to submarine canyons. In the case illustrated, the submarine canyon dictates the downcoast edge of the littoral cell. At steady state the volume of sand in the beach in each portion of the cell will remain the same, meaning that inputs and outputs of sand must balance. While not shown here, one would also cast this problem as a balance of each of many grain sizes, or of particular minerals.

As usual, the art lies in choosing what to balance or conserve: we will balance the volume of sand-sized sediment, that large enough to be retained within the beach or littoral system. Grains finer than a particular "cut off diameter", which may depend upon the wave energy available to cause significant suspension, will be lost to the shelf and will no longer count in the littoral sediment balance. For simplicity of discussion, let us call the grains large enough to be retained "sand", although in for example the Santa Cruz coastline of California, the cutoff diameter is about 0.2 mm. The sources of sediment include both rivers and sand from cliff retreat. Sinks include any loss to the shelf (we

could call this leakage), or to eolian dunes onshore, and the ultimate sink is submarine canyons that intercept the littoral drift.

In words, we may write the balance as

Rate of change of mass of sand in the beach wedge = rate of inputs of sand to the beach wedge – rate of loss of sand from the beach wedge

Translation into mathematics yields:

$$\frac{\partial M}{\partial t} = Q_{littoral}\big|_x - Q_{littoral}\big|_{x+dx} + q_{cliff}dx - q_{shelf}dx - q_{eolian}dx + Q_{river} \quad (16.1)$$

where M is the mass of sand in a reach of coastline of length dx, Q represents mass discharge (M/T), and q is specific discharge (M/LT). The mass of sand in an element of beach (Figure 1) is $M = \rho_b A dx = \rho_b W\bar{h}\, dx$, where A is the cross-sectional area of beach sand and ρ_b is its bulk density. This serves to organize our thoughts about what we must know about a specific field example. A river contributes sand to the littoral cell at a point, at a rate that reflects the basin area, A_b, the mean erosion rate in the basin, $\dot{e}_b$, and the fraction of the products that are sand, f_r.:

$$Q_r = \rho_r f_r \dot{e}_b A_b \quad (16.2)$$

where ρ_r is the density of rock. Cliffs contribute sand at a rate dictated by the cliff recession rate, the height of the cliff, and the proportion of that cliff material that is above the grain size threshold (Figure 1).

$$q_{cliff} = \dot{e}_{cliff} \rho_r f_{cliff} H_{cliff} \quad (16.3)$$

All of these terms can now be inserted into a final equation for the sediment budget in an element of the littoral cell. Doing so, and dividing by both the distance dx and the bulk density of littoral sand, ρ_{litt}, yields:

$$\frac{\partial A}{\partial t} = -\frac{1}{\rho_b}\frac{\partial Q_{littoral}}{\partial x} + \varepsilon_{cliff}\left(f_{cliff} H_{cliff} \frac{\rho_r}{\rho_b}\right) + \frac{\rho_{riv}}{\rho_b}\left[\sum_i f_{riv,i}\dot{e}_{b,i}A_{b,i}\delta(x - x_i)\right] - \frac{\rho_b}{\rho b} q_{shelf} \quad (16.4)$$

where we sum the river inputs over i rivers. Note that this is an evolution equation for the cross-sectional area of sand in a segment of beach, A, and that I have ignored for the moment the potential loss to eolian dunes. The equation is complicated, but explicitly accounts for all of the processes involved, and the relevant erosion rates of cliffs and river basins. The 1st term on the right corresponds to our now familiar gradient of flux, here the gradient of longshore drift of sand. The 2nd term represents the contribution from cliff retreat, the 3rd from rivers, and the 4th corresponds to losses of sand to the shelf.

Let us ask the question of what the pattern of sediment must be in order for the system to be steady. As in our treatments of hillslopes and glaciers, for example, if we assume that the system is steady, we may ignore the time-varying term on the left hand side, and integrate the remaining expression to obtain the expected spatial pattern of beach sand discharge. Between point inputs of sand from the rivers, the discharge of sand down-cell must simply be the integral of the contributions for the cliff back-wearing and losses to the shelf. If the back-wearing and shelf loss are uniform over this interval, this leads to linear increase in discharge:

$$Q = (\dot{e}_{cliff} f_{cliff} H_{cliff} - q_{shelf})x \qquad (16.5)$$

At the rivers, the discharge must experience a jump or step comparable to the fluvial input, Q_r. The resulting steady state pattern of downcell discharge is a linear increase punctuated by steps at the rivers (Figure 2). This kind of pattern has been observed in the Santa Cruz coastline, as documented by Perg et al. (2003) who combined a mixing model and concentrations of cosmogenic radionuclides in the littoral sand and its sources to document the rates involved.

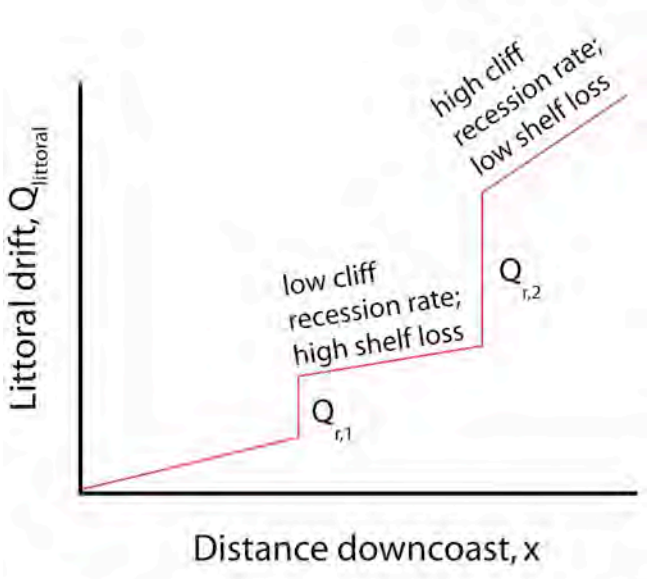


Figure x. Expected down-cell pattern of the steady state sand discharge. In this hypothetical case, two rivers contribute point-sources of sediment, the 2nd stronger than the 1st. The sand contributed from cliff recession is likewise stronger in the 2nd segment than in the 1st.

While to our knowledge this has never been done, one *could* write a similar balance for each of several grain sizes, and for each of several minerals. This would allow a formal way to address the issues of grain size and mineralogical evolution in a littoral system. As these characteristics of the beach are commonly and easily documented, they would provide strong constraints on models of provenance, and strong tests of our understanding of grain size evolution in coastal systems. It would become clear from writing the necessary equations that one must have terms for the rate of change of grainsize due to comminution during longshore transport. A large grainsize therefore becomes a source for smaller grainsizes. On the other hand, the opposite cannot happen, as small grains cannot reverse the process and become larger. One might imagine that the rate of comminution will depend upon mineralogy, with more susceptible grains becoming smaller more rapidly.

17. Momentum

In the other examples in this book, we have conserved quantities that can be seen or at least imagined. Ice, water, sediment, regolith -- all of these are tangible quantities. In this example we add momentum to our list of quantities to be conserved. Like conservation of energy, conservation of momentum is one of the great organizing principles in physics. Isaac Newton stated in his 2nd law that:

Rate of change of momentum of an object = sum of the forces acting upon it

As in this case we are considering the momentum of an object, there is no flow of momentum across the edges of the box. This simplification must be abandoned when we address the momentum of an arbitrary blob of fluid. While we have stamped in our memories from highschool physics the simple formula $F=ma$, where m is the mass of the object, a its acceleration and F the force causing its acceleration, the more general form of this is as written in the box. Our first task is to show that these are equivalent statements. Recall that acceleration is the rate of change of velocity: $a = \partial v / \partial t$, and that momentum is the product of an object's mass with its velocity, mv. Since in this case the mass of the object is unchanging in time, we may rewrite the right hand side of $F=ma$:

$$F = ma = m\frac{\partial v}{\partial t} = \frac{\partial mv}{\partial t} \tag{17.1}$$

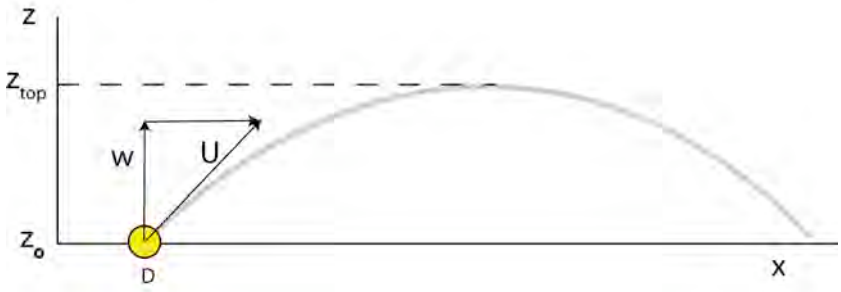
Turning this around and acknowledging that it is really the sum of all the forces acting on the object, this becomes

$$\frac{\partial mv}{\partial t} = \sum F \tag{17.2}$$

or the rate of change of momentum equals the sum of the forces acting on the object. In Newton's honor, the unit of force in the mks system is called the Newton = 1 kg m s^{-2}.

I will first use this equation to work a problem in sediment transport before introducing the more general statement relevant to fluids.

Imagine the sandy upwind side of a desert dune. The wind is blowing hard enough to entrain grains from the sand surface and accelerate them downwind. The sand on the surface is said to be saltating, a word derived from Latin *saltare* meaning to jump or to leap. The physics of this process is now known to include not only the aerodynamic drag that accelerates the grains downwind, but the energetic impact of grains with the surface. The impact of a sand grain that has been accelerated by the wind results in a "granular splash" in which numerous grains are ejected from the sand surface from nearby the impact site. We can employ Newton's 2nd law to predict the height to which these ejected grains will fly as they traverse their downwind-canted trajectories.



Trajectory of a grain of diameter D ejected from the sandy surface at a speed U, with vertical component w. Ignoring air drag, one can calculate both the hop time and the hop height, z_{top}.

Because the density of sand is roughly 2000 times that of the air into which it is ejected (ρ_a=1.22 kg/m^3; ρ_g=2700 kg/m^3), the height to which the grain is ejected can be estimated well by assuming that the drag force exerted by the air in the vertical direction is negligible. This is not the case in water, in which the density contrast is only about 2-fold; this is in fact a chief difference between sediment transport in air and water. The only force remaining in the problem is therefore the weight of the sand grain, F_w, or the force due to the acceleration of gravity, $F_w = mg$. Equating this with the rate of change of the grain's momentum in the vertical direction yields:

$$\frac{d(mw)}{dt} = -mg \tag{17.3}$$

where w is the vertical component of the grain's velocity (Figure 1). The minus sign on the right hand side acknowledges that while the vertical speed is taken positive upward, the acceleration due to gravity, g, is directed downward. Given that the mass of the grain is not changing, we can remove the mass from the partial derivative on the left hand side, and divide through by it. This results in the simple formula for the evolution of the vertical speed:

$$\frac{dw}{dt} = -g \tag{17.4}$$

We must now integrate this equation to obtain the history of vertical speed:

$$w = -gt + c_1 \tag{17.5}$$

where c_1 is a constant of integration. We can see that this constant must have units of a speed. Taking t=0, we can see that the constant is equal to the speed at time = 0, or the initial vertical speed, w_o.

$$w = w_o - gt \tag{17.6}$$

This vertical ejection speed is set by the physics of the granular impact process. Armed with this simple equation, we can now evaluate two properties of the grain trajectory: how long it is airborne, and how high it will hop. At the top of the hop, its vertical speed should go to zero. Setting the left hand side to zero suggests that the time to the top of the hop is simply $t_{top} = w_o/g$. Given that there are no other forces acting, the time to return to the sand surface should be the same, suggesting that the time to accomplish the entire hop is

$$t_{hop} = 2\frac{w_o}{g} \tag{17.7}$$

Typical ejection speeds are fractions of 1 m/s, while a few rare grains are ejected from the bed with several m/s speeds. For example, for w_o=0.1 m/s, the hop duration is 2*(0.1)/9.81 ~ 0.02 s.

The height to which the grain hops requires that we know its vertical position, not its vertical speed. Recognizing that the vertical speed is the rate of change of vertical position, or $w=dz/dt$, we can calculate the history of its vertical position by integrating its vertical speed:

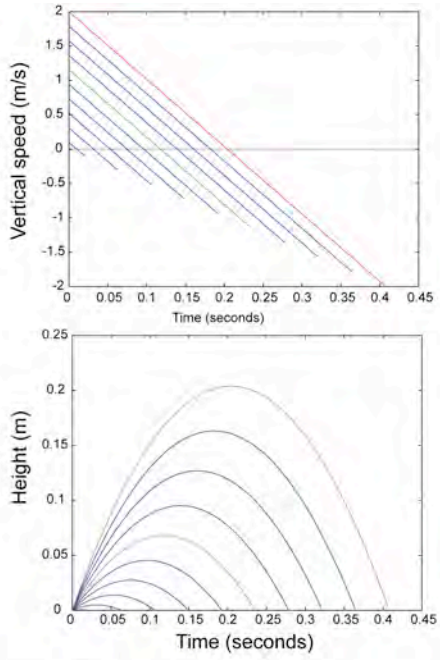
$$z = \int w dt = \int (w_o - gt) dt \tag{17.8}$$

Breaking this into two integrals, and integrating, this becomes

$$z = w_o t - \frac{1}{2} g t^2 + z_o \tag{17.9}$$

where for the 2nd constant of integration I have substituted the initial vertical position of the particle, z_o. This is the equation for a parabola (Figure 2). We can now ask what the particle height it at the time of the top of the hop, $t=t_{top}$:

$$z_{top} = z_o + \frac{w_o^2}{g} - \frac{1}{2} g \left(\frac{w_o}{g} \right)^2 = z_o + \frac{w_o^2}{2g} \tag{17.10}$$



Vertical speed (top) and vertical position (bottom) as functions of time in the case of ballistic (no drag) conditions. Hop time goes linearly with ejection speed, while hop height goes as the square of the ejection speed.

Relative to its original position, the particle reaches a height of $w_o^2/2g$. For our particle ejected at 0.1 m/s, this results in a hop height of 0.5 mm, or a few grain diameters for a typical fine sand particle of 0.1 mm diameter. Most of the grains in saltation, then, are traveling with hops that take them only a few diameters away from the sand surface. Those few grains ejected with speeds of 1 to 5 m/s, on the other hand, will hop to heights of 5 – 100 cm. These rare grains are the ones that will accumulate in the cuffs and the

pockets of your pants while you stand there admiring the beauty of the sand ripples as they translate downwind.

18. Advection and the substantial derivative

Although this phenomenon was introduced earlier, in our discussion of the climate, it is worth devoting some space to clarify it and to give you a couple more examples. I find that many students find the concept of advection confusing. We have already talked about how heat can conduct in a solid, moving from hot to cold. And we have talked about radiation, transporting heat through transparent media. There is a third mechanism, called advection, that plays a major role in transporting heat in moving media, like water and air. Advection is the main means of transport of momentum and heat (and many other quantities) in the atmosphere and in rivers. Advection of heat also provides an important analog for transport of any quantity in a moving medium.

When the medium is in motion, there is a possibility that the concentration of heat at a point, or the temperature, can change due to that motion. Whether this happens or not depends upon whether there is a gradient of the temperature as well, and in particular a gradient of temperature in the direction of motion. I will define advection as follows:

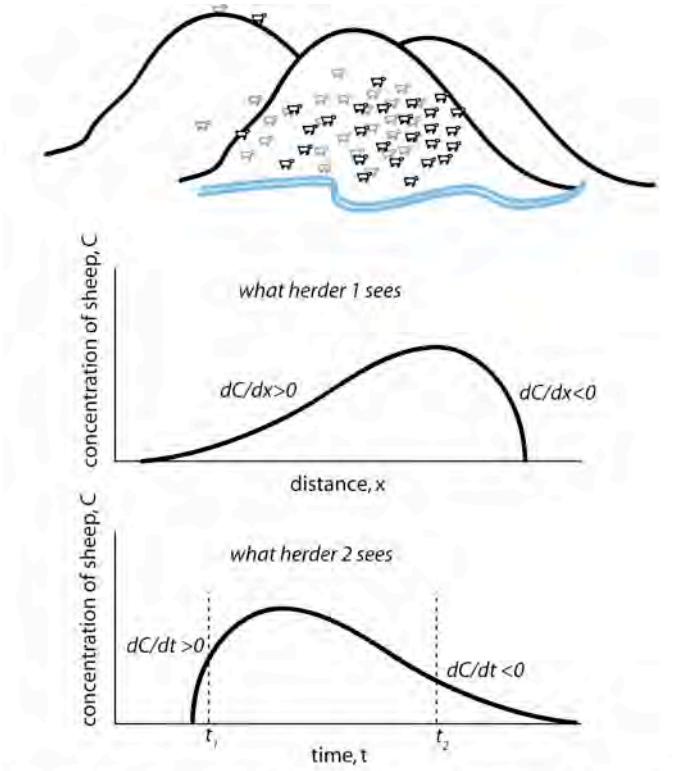
Advection The rate of change of some quantity due to material motion in the direction in which there is a gradient in the concentration of that quantity.

Advection can be very efficient because the transport is accomplished by the bulk motion of the fluid rather than by the exchange between individual atoms or molecules. The quantity is embedded in the fluid and moves with it. The process does not depend upon the passage of heat (or any other quantity) through random molecular motion, as in conduction of heat, but involves the bulk motion of the fluid.

I use this discussion of advection as an excuse to introduce two other topics. The role of advection can be illuminated by comparing the same phenomenon from two points of view. This in turn will lead to the introduction of the substantial derivative. But let us begin.

Consider again our herd of sheep, introduced at the beginning of the book. The herd of sheep is being watched by two shepherds. Shepherd 1, call him Elmer, is riding his horse along with the herd, while the other, shepherd 2, call him Clovis, is taking a break on a hillside across the stream, and is watching as the herd moves left to right across the opposite hillside (Figure 1). Sheep always move left to right. Let us say that the quantity of concern is the concentration of sheep, the number of sheep per unit area, say a hectare (100m x 100m). The sheep have arranged themselves into a band with a sharply defined leading edge, defined by a couple bellwethers and the sheep who follow closely on their heels, and a more diffuse strung-out back end with the slowpokes, the injured, and the renegades, who if it were not for the trusty sheepdog would be lost entirely. What does this look like to Elmer, the shepherd traveling with the band? It looks like a steady pattern with not much change as he slowly rides along with the band. Looking ahead he sees the strongly defined front, and behind he sees the diffuse trailing edge to the band. It looks the same another half an hour later, on a different hillslide, the entire herd having

moved hundreds of meters. But what does this look like to Clovis, the shepherd taking a break? Looking across the creek, he sees anything but a steady pattern. Staring at the same segment of hillside he has in his field of view, he sees first the arrival of the sharply defined front end of the herd, then the high concentration of the middle of the band, with Elmer at the edge of it, and finally he sees the diffuse tail go by, after which the concentration of sheep on this hillside returns to zero. From his perspective there has been a temporal pattern of sheep concentration, while for his buddy moving with the sheep the pattern was steady, not changing in time. What is seen depends upon the point of view. These points of view are the end-member cases, one still, the other traveling with the medium. These have names: that traveling with the medium is the Lagrangian point of view, while that pinned down at a point is the Eulerian point of view. (Joseph Louis Lagrange (1736-1813) was an Italian-born mathematician and astronomer, teacher of Fourier and Poisson; Leonhard Euler (1707-1783) was a prolific Swiss-born mathematician and physicist who spent most of his life in Germany and Russia).



Lagrangian and Eulerian points of view and advection. Top: pattern of sheep on a hillside traveling left to right seen at two instants (dark = now; gray = earlier). Middle (Lagrangian point of view): the pattern of sheep concentration plotted as a function of distance along the hillside, with the sharp leading edge and diffuse tail. Bottom (Eulerian point of view): pattern of sheep concentration as a function of time as seen by an observer fixed in space. At an early time, t_1, the concentration of sheep steeply rises ($dC/dt>0$) as the tightly bunched leading edge of the band comes by the view. At a later time, t_2, the concentration of sheep slowly falls ($dC/dt<0$) as the diffuse tail ambles by.

In order to go back and forth between these points of view, which in the end must describe the same system, we need advective terms in the equation describing the system. But let me do this graphically first.

We would like to be able to get back and forth between these two points of view. Clearly we need some mathematical mechanism that involves derivatives in time and derivatives in space. Let us get precise. To describe the Lagrangian point of view, we must define a derivative following along with the medium. This is called the substantial derivative, or "the derivative following the fluid" or medium. It is denoted *D/Dt*, with a big *D* instead of a partial. In the case I have illustrated, the rate of change of sheep concentration following along with the herd, as seen by the herder embedded in the herd, is zero. He sees no change in the pattern through time; it would always look like that shown in the top graph. In other words $DC/Dt = 0$, C being the concentration of sheep. The herder across the creek, however, is fixed in space. In his fixed field of view, he sees a temporal pattern of concentration, shown in the bottom diagram. He sees temporal changes while fixed in space. We use the partial derivative to denote this, or $\partial C/\partial t$. This means formally that, all other variables fixed (here we only had one spatial variable, x), this is the pattern seen. It is only changing in time. So how do we get between these points of view? It comes through the full definition of the substantial derivative. Let me define it first, and then turn back to the sheep.

The substantial derivative (also called the material or total derivative) of a hypothetical quantity, which I have called *A,* is:

$$\frac{DA}{Dt} = \frac{\partial A}{\partial t} + u\frac{\partial A}{\partial x} + v\frac{\partial A}{\partial y} + w\frac{\partial A}{\partial z} \tag{18.1}$$

Translating into words, the rate of change following the medium, the substantial derivative, depends upon the rate of change at a fixed point plus the rate of change caused by motion in any direction in which there is a gradient of the concentration. These last 3 terms are the "advective terms". Here u is the speed in the direction x, v that in the direction y and w that in the direction z. In our case, the variable of concern is the concentration of sheep, C, we have only one spatial dimension, x, and the speed of the herd moving in that direction is u. The substantial derivative then becomes:

$$\frac{DC}{Dt} = \frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} \tag{18.2}$$

In the scenario I have described, the substantial derivative is zero: there is no change following along with the herd, from Elmer's point of view. We can now translate between what Elmer sees and what Clovis sees:

$$\frac{\partial C}{\partial t} = -u\frac{\partial C}{\partial x} \tag{18.3}$$

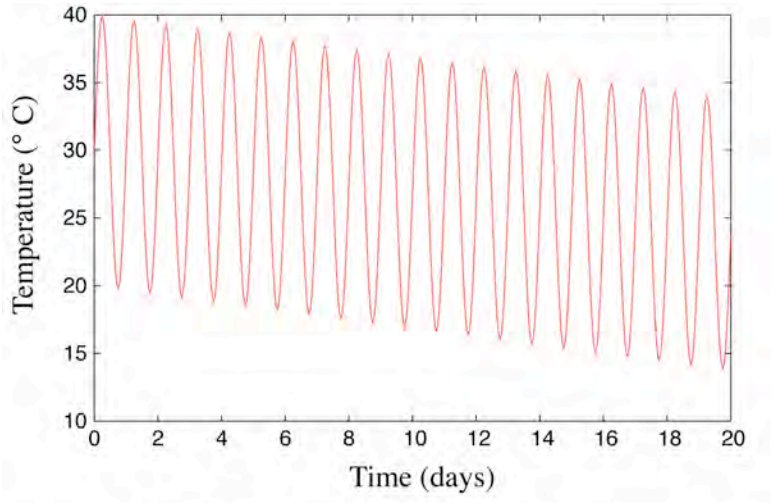
Does this work? Turning back to the figure, focus on the two times, t_1 and t_2. In the early time, the rate of change of sheep concentration with time is positive and high. It corresponds to the passage of the sharply defined front end of the band, with a strong negative concentration gradient (C declining as x increases), passing the site at a rate u. C changes because there is motion, u, in the direction in which there is a gradient, x. The minus sign assures us that concentration increases when negative concentration gradients advect by. Consider now the later time. Motion is in the same direction at the same speed, but this time the concentration gradient is positive. The negative sign assures that the concentration seen at the fixed site now declines as the pattern with a positive concentration gradient passes by. To summarize, advection of some quantity requires motion in a direction in which there is a spatial gradient in that quantity. Mathematically,

one can always recognize advection terms as $u\frac{\partial A}{\partial x}$ or $v\frac{\partial A}{\partial y}$ or $w\frac{\partial A}{\partial z}$, where *u, v*, and *w* are the components of velocity in the *x, y* and *z* directions, respectively.

Consider a different problem, adapted from the one I originally heard to illustrate the concepts of advection and the substantial derivative. Imagine that you are a bug on a leaf, floating steadily northward in the ocean, having been delivered to the ocean from say a river draining the cloud forest of Costa Rica. What temperature history might you record in your bug-sized thermometer? You are traveling with the medium, here seawater, so any changes you experience will involve the substantial derivative of temperature. This will include two kinds of change, one associated with temporal changes at any fixed point (say associated with daily oscillations in temperature), and the other with the fact that you are in motion in the direction in which there are gradients in temperature (associated with latitudes). The equation for the rate of change of temperature will be

$$\frac{DT}{Dt} = \frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} \tag{18.4}$$

where the 1st term represents the rate of change at a point (daily cycle) and the 2nd represents the change associated with northward motion (defining *x* northward, and again taking $\underline{u}$ to be the component of velocity in the *x* direction). The first term will therefore rapidly vary with the daily temperature cycle. This oscillation will be embedded on a slow decline represented by the second term, as you slowly advect into colder climates.



Temperature history recorded by a bug on a leaf floating northward from Costa Rica. The oscillations reflect the daily cycle of temperatures. The net cooling reflects advection into colder climates.

Let us examine a couple more relevant thermal examples. In the first imagine a site on a polar glacier in which there is little melting in the summer, and only slow net accumulation, at a rate *b*. If the ice were not in motion, the temperature profile would simply be a geotherm, reflecting steady operation of the conduction process, with a uniform gradient set by the heat flux and the thermal conductivity of the ice, *Q/k*. The

net accumulation slowly buries a parcel of ice, giving it a velocity that transports it in the direction of a temperature gradient, from the relatively cold surface downward. The full equation for the evolution of the temperature is then

$$\frac{DT}{Dt} = \frac{\partial T}{\partial t} + b\frac{\partial T}{\partial z} = \kappa \frac{\partial^2 T}{\partial z^2} \tag{18.5}$$

This differs form that presented in the discussion of permafrost in the presence of the second term on the left, representing downward advection of ice. There is advection because there is motion in the direction of a temperature gradient. See discussion in both Paterson (2000) and in Hooke (2005) of the expected temperature profiles in both accumulation and ablation areas of glaciers.

A perfectly analogous situation exists when we consider the temperature profile we might expect in a landscape that is steadily eroding. Here the (cold) surface is being brought closer to the rock underground, and will lead to curvature in the expected temperature profile. This corresponds to the ablation area of a glacier, in which the surface is being brought closer to a parcel of ice at depth; the ice is being exhumed.

19. Summary and discussion

We have seen throughout the book that the machinery involved in setting up and in working all of these problems is quite similar. The words may change, but the mathematical statements that grow out of the words are indeed nearly equivalent (in some cases they are identical if we only change the symbols). It is therefore the mathematics that is the bridge between the problems, and indeed between the disciplines. This too is a point made eloquently by Feynman in one of his lectures collected in *The Character of Physical Law*. We consistently arrive at a partial differential equation with a time derivative on the left hand side, and spatial derivatives of flux or discharge on the right, along with terms for sources and sinks:

$$\frac{d\psi}{dt} = -\nabla Q + B - D \tag{19.1}$$

where ψ is the concentration, and Q the flux of some quantity, and B and D are source and sink for it. I have used the ∇ (del) operator to denote the sum of spatial gradients in each dimension: $\nabla = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}$. We anticipated at the outset that the order allowed by employing these conservation equations allows us now to focus on the physics, biology and chemistry involved in the transport of the quantities of concern, the Q's in this equation. To be specific, in order to "close" our conservation equations, we need formulas, or rules, or constitutive relationships for:

- Transport of heat in rock and soil by conduction, or by turbulent fluids in the atmosphere and in the ocean (These are effectively in place in the form of Fourier's Law for heat conduction, and Reynold's extension of Newtonian viscous rheology.)
- Transport of ice
- Transport of sediment by wind or river or waves
- Transport of regolith
- Transport of water down a slope

For the sources and sinks, we need rules for

- Production of cosmogenic radionuclides
- Precipitation and evaporation of water (lakes) or of ice (glaciers)
- Production of regolith by weathering of rock
- Release of littoral sediment by backwearing of seacliffs

And so on. Several of these constitute major quests in modern geomorphology, as without them we cannot solve the larger scale problems of landscape evolution.

While this book has been about the derivation of a few representative conservation equations, and less about the solutions to the equations, I can nonetheless recommend a strategy for solution of them. As I have done in several instances, we can seek first the steady solutions. If such solutions exist, they will come from setting the left hand side of the conservation equation to zero; if the time derivative is zero, nothing can change in time, and the system is steady. This transforms a partial differential equation (a PDE) into an ordinary differential equation (an ODE), which is generally very much easier to

solve. This is why courses in ODEs are taken before courses in PDEs. I also try to work the problem in 1D before going to 2D. With these assumptions, and taking the dimension of interest to be in the direction x, the equation above transforms to

$$\frac{dQ}{dx} = B - D \tag{19.2}$$

We saw this in several of the cases described: lakes, glaciers, runoff, littoral drift, sediment transport, hillslopes. The solution, which yields the expected steady pattern of discharge, then requires one integration:

$$Q(x) = \int_0^x (B - D)\, dx \tag{19.3}$$

As long as we know the spatial pattern of sources and sinks, *B(x)* and *D(x),* we can perform this integral. While the pattern of discharge is not necessarily our final goal, it is in some cases greatly useful. For example, on hillslopes the discharge of water or of regolith at the base of the slope is that delivered to the stream. In the case of glaciers, the site at which the discharge of ice returns to zero is the terminus, meaning we can predict the terminus position of the steady glacier as long as we are given (from the meteorologists) the spatial pattern of snowfall and of melt. We did no ice physics at all. Nor, in the case of water and regolith on the hillslope, did we need to know the physics of water and regolith discharge to calculate the expected steady pattern of discharge. It almost seems like cheating!

Let us turn this on its side and look from another angle. Let us say we have a box, and we can only measure the stuff that comes out of one side of the box. The box is black, and we cannot see inside, nor can we see any of the other sides. The equation we have just described tells us that if we can measure very well what comes out this one side, and if the processes occurring inside and across the other sides of the box are steady, then we can infer from the measured discharge what the net rates of motion across all of the other invisible sides must be. This is not as silly as it sounds. Consider for example that the black box is a hillslope, and that we measure carefully the discharge of regolith at its base. This in turn provides us a first assessment of the rate at which material is being transformed from rock into regolith – and this is otherwise a very difficult thing to measure! Once we know this rate, and I can tell you that this rate is usually so slow that one would have to string many PhD's together to be able to measure it, we can begin to ask what sets the rate. Or here is another example. Let us say that we not only can measure the output from the base of the slope, and using a different tool we can measure the inputs from the transformation of bedrock, say using cosmogenic radionuclides. We find that the two measures do not match: the rate of loss from the hillslope is less than the rate of gain from weathering. We can still lean on the conservation principle. If the rate of output is less than the rate of input, there must be another sink. Something must be leaking from the system that we cannot or have not measured. Mass must be disappearing from the system in solution. Note that the chain of reasoning I have just been using involves two independent activities: generation of a theoretical framework, in this case based upon the principle of conservation, and generation of data to constrain one or another or several terms in the equation. Once the theoretical framework is in place, and has no holes in it, our ability to make statements like "mass is disappearing from the system at such and such a rate by dissolution" depends entirely upon the quality of the

data. This is what links observational and theoretical aspects of geomorphology. Kepler was only able to discard a hypothesis that the planets orbited in circular orbits by knowing that the orbital observations of Tycho Brahe would not permit an error of 8 arc seconds (see Feynman 1965) in the position of Mars – sending him on his hunt to arrive at a better shape of orbit, the ellipse. We too must link high quality observations with a sound theoretical backdrop in order to continue to advance our science.

In his book about the evolution of everyday things, like pins and tape and forks and zippers, Henry Petroski (1992) strongly advocates the hypothesis that it is the failure of these items to live up to our expectations that drives innovation. The good engineer, he says, is a good critic of the things around him, always asking how they came about, and how they fail to work as well as we might wish. The scientist's currency is not things but ideas or, more grandly, theories. We must be the critics of those ideas, asking how they fail. But we cannot assess success or failure unless we have both a theory and data against which to test the theory. This Little Book is designed to provide the geomorphologist a firm foundation from which to launch theories of how our world works, and how particular features of the surface of the world, geomorphic features that is, have come to be.

References

Ahnert, F., 1970, An approach towards a descriptive classification of slopes. *Zeitschrift fur Geomorphologie,* N. F., Suppl., 9: 71-84.

Ahnert F., 1976, Brief description of a comprehensive three-dimensional model of landform development. *Zeitschrift fur Geomorphologie*, NF Supplement Band 25, 29-49.

Anderson, M., 2007. Introducing groundwater physics. *Physics Today* May 2007.

Anderson, R. S., 1987, A theoretical model for aeolian impact ripples. *Sedimentology* 34: 943-956.

Anderson, R. S., 1987, Eolian sediment transport as a stochastic process: The effects of a fluctuating wind on particle trajectories. *Journal of Geology* 95: 497-512.

Anderson, R. S., Repka, J. L. and Dick, G. S., 1996, Dating depositional surfaces using in situ produced cosmogenic radionuclides. *Geology* 24: 47-51.

Anderson, R. S. and Bunas, K. L., 1993, The mechanics of aeolian ripple sorting and stratigraphy as visualized through a cellular automaton model. *Nature* 365: 740-743.

Anderson, R. S. and Haff, P. K., 1988. Simulation of eolian saltation. *Science* 241: 820-823.

Anderson, R. S., 2002, Modeling of tor-dotted crests, bedrock edges and parabolic profiles of the high alpine surfaces of the Wind River Range, Wyoming. *Geomorphology* 46: 35-58.

Anderson, R. S. and Hallet, B., 1996, Simulating magnetic susceptibility profiles in loess as an aid in quantifying rates of dust deposition and pedogenic development. *Quaternary Research* 45: 1-16.

Anderson, R. S., P. Molnar, and M. A. Kessler, 2006, Features of glacial valley profiles simply explained, *J. Geophys. Res.*, 111, F01004, doi:10.1029/2005JF000344.

Andrews, J. T., 1968. Postglacial rebound in Arctic Canada: similarity and prediction of uplift curves. *Can. J. of Sci.* 5: 39

Attal, M., and Lavé, J., 2006, Changes of bedload characteristics along the Marsyandi River (central Nepal): Implications for understanding hillslope sediment supply, sediment load evolution along fluvial networks, and denudation in active orogenic belts, *in* Willett, S.D., Hovius, N., Brandon, M.T., and Fisher, D., eds., Tectonics, Climate, and Landscape Evolution: Geological Society of America Special Paper 398, p. 143–171, doi: 10.1130/2006.2398(09).

Bagnold, R. A., 1941, *The Physics of Blown Sand and Desert Dunes*, Chapman and Hall, 265 pp.

Bear, J. 1988, *Dynamics of Fluids in Porous Media*. New York: Dover Press, 784 pp.

Bear, J. 1979, *Hydraulics of Groundwater*. Dover Press. 569 pp.

Bird, R. B., Stewart, W. E., and Lightfoot, E. N., 2002, *Transport Phenomena, 2nd edition,* Wiley, 920 pp.

Carslaw, H. S. and J. C. Jaeger, 1967. *Conduction of Heat in Solids*, Oxford University Press, New York.

Carson, M. A. and M. J. Kirkby, *Hillslope Form and Process*, 475 p., Cambridge University Press, Cambridge, 1972

Cerling, T. and H. Craig, 1994, Geomorphology and in-situ cosmogenic isotopes, *Annu. Rev. Earth Planet. Sci.*, 22, 273-317.

Clark, M. K., and Royden, L. H., 2000. Topographic ooze: Building the eastern margin of Tibet by lower crustal flow, *Geology* 28: 703–706.

Cornish, V., 1914, *Waves of Sand and Snow and the Eddies Which Make Them*, Open Court publishing company, 383 pp.

Culling, W. E. H., 1960, Analytical theory of erosion, *Journal of Geology* 68(3): 336–344.

Culling, W. E. H., 1965, Theory of erosion on soil-covered slopes, *Journal of Geology* 73: 230–245.

Davies, G. 2000, *Dynamic Earth: Plates, Plumes and Mantle Convection*, Cambridge University Press, 458 pp.

de Mestre, Neville, 1990. *The Mathematics of Projectiles in Sport* (Australian Mathematical Society Lecture Series), Cambridge University Press, 187 pp.

Dietrich, W. E., 1982, Settling velocity of natural particles, *Water Resources Research*, 18(6): 1615-1626.

Dietrich, W. E., Bellugi, D., Heimsath, A. M., Roering, J. J., Sklar, L., and Stock, J. D., 2003, Geomorphic transport laws for predicting the form and evolution of landscapes, in P. Wilcock and R. Iverson, editors, *Prediction in Geomorphology*, AGU Geophysical Monograph Series, V. 135, p. 103-132.

Edwards, B.F. *et al* 2001 Dynamics of falling raindrops, *Eur. J. Phys.* 22 113-118 doi:10.1088/0143-0807/22/2/302

Emanuel, K., 2005. *Divine Wind, the History and Science of Hurricanes*, Oxford Press, 295 p.

Exner, 1920, Zur physik der dunen, Sitz.-Ber. Akad. Wiss. Wien, Math-naturw. Klasse, Abt. IIa, Band 129.

Fernandes, N. F., and Dietrich, W. E., 1997, Hillslope evolution by diffusive processes: the timescale for equilibrium adjustments. *Water Resources Research* 33(6): 1307–1318.

Feynman, R., 1965, *The Character of Physical Law*, MIT press.

Freeze, R. A. and Cherry, J. A., 1979, *Groundwater*, Prentice Hall, 604 pp.

Furbish, D. J., 1997, *Fluid Physics in Geology: An Introduction to Fluid Motions on Earth's Surface and Within its Crust*. Oxford University Press, 496 pp.

Furbish, D. J., K. K. Hamner, M. Schmeeckle, M. N. Borosund, and S. M. Mudd, 2007, Rain splash of dry sand revealed by high-speed imaging and sticky paper splash targets, *J. Geophys. Res.*, 112, F01001, doi:10.1029/2006JF000498.

Gabet, E. J., O. J. Reichman, and E. Seabloom, 2003, The effects of bioturbation on soil processes and hillslope evolution. *Annual Review of Earth and Planetary Sciences* 31: 249-273.

Gilbert, G. K., 1877, *Report on the geology of the Henry Mountains*, U.S. Geographical and Geological Survey of the Rocky Mountain Region, Washington, D.C.

Gilbert, G. K., 1909, The convexity of hilltops. *Journal of Geology* 17: 344–350.

Gold L. W. and A. H. Lachenbruch, 1973, Thermal conditions in permafrost—a review of North American literature. In: *Proceedings Second International Conference on Permafrost, Yakutsk, U.S.S.R., North American Contribution*, National Academy of Sciences, Washington, US, pp. 3–25.

Gosse, J. C., and Phillips, F. M., 2001, Terrestrial in situ cosmogenic nuclides: theory and application: Quaternary Science Reviews, v. 20, p. 1475-1560.

Granger, D. E., Kirchner, J. W., and Finkel, R. C., 1997, Quaternary downcutting rate of the New River, Virginia, from differential decay of 26Al and 10Be in cave-deposited sediment, Geology, v. 25, p. 107-110.

Granger, D. E., and Muzikar, P., 2001, Dating sediment burial with cosmogenic nuclides: Theory, techniques, and limitations, Earth and Planetary Science Letters, v. 188, no. 1-2, p. 269-281.

Haff, P.K. and Anderson, R.S., 1993. Grain-scale simulations of loose sedimentary beds: The example of grain-bed impacts in aeolian saltation. *Sedimentology* 40: 175-189.

Hancock, G.S., Anderson, R.S., Chadwick, O. A., and Finkel, R. C., 1999, Dating fluvial terraces with ^{10}Be and ^{26}Al profiles, Wind River, Wyoming. *Geomorphology* 27: 41-60.

Harte, J., 1988, *Consider a Spherical Cow: a Course in Environmental Problem Solving*, University Science Books, Sausilito, 283 pp.

Hooke, R. leB., 2005, *Principles of Glacier Mechanics*, Cambridge University Press, 448 pp.

Hornberger, G. M., Raffensperger, J. P., Wiberg, P. L., and. Eshleman, K. N., 1998, *Elements of Physical Hydrology*, The Johns Hopkins University Press, 312 pp.

Horton, R.E., 1945, The erosional development of streams and their drainage basins: hydrophysical approach to quantitative morphology, *GSA Bulletin* 56: 275-370.

Hostetler, S. and Bartlein, P., 1990, Simulation of lake evaporation with application to modeling lake level variations of Harney-Malheur Lake, Oregon, *Water Resources Research*, 26(10): 2603-2612.

Kamb, B., C. F. Raymond, W. D. Harrison, H. Englehardt, K. A. Echelmeyer, N. F. Humphrey, M. M. Brugman and W. T. Pfeffer, 1985, Glacier surge mechanism: 1982-83 surge of Variegated Glacier, Alaska. *Science* 227: 469-479.

Kirkby, M. J., 1971, Hillslope process-response models based on the continuity equation. *Institute of British Geographers, Special Publication*, 3: 15–30.

Kirkby, M. J., 1985, A model for the evolution of regolith-mantled slopes. In Models in Geomorphology, Woldenberg M. J. (ed.). George Allen & Unwin: London, 213–238.

Lachenbruch, A. H., and Marshall, B. V., 1986, Changing Climate: Geothermal Evidence

from Permafrost in the Alaskan Arctic, *Science* 234 (4777): 689-696.

Lachenbruch, A. H., Sass, J. H., and Marshall, B. V., 1982, Permafrost, heat-flow, and the geothermal regime at Purdue Bay, Alaska, *J. Geophys. Res.*, 87 (NB11): 9301-9316.

Lal, D., 1988, In-situ produced cosmogenic isotopes in terrestrial rocks. *Ann. Rev. Earth Planet. Sci.* 16:355-388.

Logan, G. A., Hayes, J. M., Heishma, G. B., and Summons, R. E., 1995, Terminal Proterozoic reorganization of biogeochemical cycles, *Nature* 376: 53-56.

Loso, M. G., Doak, D. F., and Anderson, R. S., Application of a new, process-based lichenometric technique to the dating of Little Ice Age glacier moraines, in review, *Arctic, Antarctic and Alpine Research.*

Manga, M. and H.A. Stone, 1994, Interactions between bubbles in magmas and lavas: Effects of deformation, *Journal of Volcanology and Geothermal Research*, vol. 63, 269-281.

Manga, M., 1997, Interactions between mantle diapirs, *Geophysical Research Letters*, 24: 1871-1974.

Marshall, J., D. Ferreira, J-M Campin and D. Enderton, 2007, Mean climate and variability of the atmosphere and ocean on an aqua-planet: in press, *J. Atmos. Sci.*

Menking, K. M., Anderson, R. Y., Shafike, N. G., Syed, K. H., and Allen, B., 2004, Wetter of colder during the Last Glacial Maximum? Revisiting the pluvial lake question in southwestern North America. *Quaternary Research*, v. 62, p. 280-288.

Middleton, G.V. and P.R. Wilcock, 1994. *Mechanics in the Earth and Environmental Sciences*, Cambridge University Press.

Mono Basin Ecosystem Study Committee, 1987, *The Mono Basin Ecosystem: Effects of Changing Lake Level*, National Academy Press, Washington D.C.

Morsi, S.A. and A. J. Alexander, 1972, An investigation of particle trajectories in two-phase flow systems, *J. Fluid Mech.*, 55, 193-208.

Mudd, S. M. and Furbish, D. J. 2006. Using chemical tracers in hillslope soils to estimate the importance of chemical denudation under conditions of downslope sediment transport. *Journal of Geophysical Research – Earth Surface*, 111, F02021, doi: 10.1029/2005JF000343.

Oerlemans J., 2001, *Glaciers and Climate Change*. A.A. Balkema Publishers, 148 pp. ISBN 9026518137

Oerlemans J., 2005, Extracting a climate signal from 169 glacier records. *Science* 308, 675-677; 10.1126/science.1107046

Paola, C. and V. R. Voller, 2005, A generalized Exner equation for sediment mass balance, *J. Geophys. Res.*, *110*, F04014, doi:10.1029/2004JF000274.

Parsons, B. and J. G. Sclater, 1977, An analysis of the variation of ocean floor bathymetry and heat flow with age. *J. Geophys. Res.* **82**, 803-827.

Paterson, W. S. B., 2000, *The Physics of Glaciers*, 3rd edition. Elsevier, 496 pp.

Peltier, W. R. and Andrews, J. T., 1976. Glacial-isostatic adjustment I. The forward problem. *Geophys. J. R. Astr. Soc.* 46: 605-646.

Perg, L. A., Anderson, R. S., and Finkel, R. C., 2001, Young ages of the Santa Cruz marine terraces determined using 10Be and 26Al. *Geology* 29 (10): 879 – 882.

Perg, L. A., Anderson, R. S., and Finkel, R. C., 2003, Use of cosmogenic radionuclides as a sediment input tracer: Constraints on the long-term sediment budget of the Santa Cruz littoral cell, California. *Geology* 31(4): 299-302.

Petroski, H., 1992, *The Evolution of Useful Things*, Alfred A. Knopf, New York, 288 pp.

Repka, J. L., Anderson, R. S., and Finkel, R. C., 1997, Cosmogenic dating of fluvial terraces, Fremont River, Utah. *Earth and Planetary Science Letters* 152: 59-73.

Roering, J. J., J. W. Kirchner, L. S. Sklar, and W. E. Dietrich, 2001, Experimental hillslope evolution by nonlinear creep and landsliding, *Geology* 29: 143-146.

Rosenbloom, N. A. and Anderson, R. S., 1994, Evolution of the marine terraced landscape, Santa Cruz, California. *J. Geophys. Res*. 99: 14,013-14,030.

Saffer, D. M. and Bekins, B. A., 1998, Episodic fluid flow in the Nankai Acceretionary Complex: Timescale, geochemistry, flow rates, and fluid budget: *J. Geophys. Res.*, 103 (B12): 30,351-30,370

Saffer, D. M., and Bekins, B. A., 1999, Fluid budgets at convergent plate margins: Implications for the extent and duration of fault zone dilation, *Geology* 27 (12): 1095-1098.

Schoenbohm, L. M., Burchfiel, B. C., Liangzhong, C., 2006, Propagation of surface uplift, lower crustal flow, and Cenozoic tectonics of the southeast margin of the Tibetan Plateau, *Geology* 34 (10): 813–816; doi: 10.1130/G22679.1

Sharp, M., W. Lawson, and R. S. Anderson, 1988, Tectonic processes in a surge-type glacier -- an analogue for the emplacement of thrust sheets by gravity tectonics. *Journal of Structural Geology* 10: 499-515.

Shelton, J. and Shapiro, A. B., 1976, The Wood Burner's Encyclopedia, Vermont Crossroads Press, 155 pp.

Shen, F., Royden, L. H., and Burchfiel, B. C., 2001, Large-scale crustal deformation of the Tibetan Plateau, *J. Geophys. Res.*, 106 (B4): 6793-6816.

Shreve, R. L., 1979, Models for prediction in fluvial geomorphology, *Journal of the International Association for Mathematical Geology*, 11(2): 165-174.

Small, E. E., L. C. Sloan, S. Hostetler, and F. Giorgi, 1999, Simulating the water balance of the Aral Sea with a coupled regional climate-lake model, *J. Geophys. Res.*, *104* (D6), 6583–6602.

Sparks, R.S.J., Bursik, M.I., Carey, S.N., Gilbert, J.S., Glaze, L.S., Sigurdsson, H., Woods, A.W. (1997) *Volcanic Plumes*, John Wiley & Sons Ltd, Chichester, UK, 574pp.

Stein, C. A. and S. Stein, 1992, A model for the global variation in oceanic depth and heat flow with lithospheric age. *Nature*, 359, 123-129.

Stock, G. M., Anderson, R. S., and Finkel, R. C., 2004, Cave Sediments Reveal Pace of Landscape Evolution in the Sierra Nevada, California, *Geology*, March 2004, v. 32, no. 3, p. 193-196; doi: 10.1130/G20197.1

Stock, G.M., Granger, D.E., Anderson, R.S., Sasowsky, I.D., and Finkel, R.C., 2005, Dating cave deposits for use in landscape evolution studies: Insights from caves in

the Sierra Nevada, California. *Earth and Planetary Science Letters* 236: 388-403

Syvitski, J.P.M., Murray, J.W., 1981. Particle interaction in fjord suspended sediment. Marine Geology 39 (3–4), 215–242.

Syvitsky, J.P.M. (ed.), 1991. Principles, Methods, and Application of Particle Size Analysis. University of Cambridge Press

Syvitski, J. P. M. and Saito, Y., 2007. Morphodynamics of deltas under the influence of humans. *Global and Planetary Change* 57: 261-282.

Trenberth, K. E., and J. M. Caron, 2001, Estimates of meridional atmosphere and ocean heat transports. *J. Climate,* 14, 3433–3443.

Turcotte, D. L. and Schubert, G., 2002, *Geodynamics*, 2nd edition, Cambridge University Press.

Van de Wal, R. S. W. and J. Oerlemans, 1995, Response of valley glaciers to climate change and kinematic waves: a study with a numerical ice-flow model. *Journal of Glaciology* 45 (137), 142-152.

Van Rijn, L. C., 1993, Principles of sediment transport in rivers, estuaries, and coastal seas, Section 3.2.5: Particle fall velocities. Aqua Publications.

Vogel, S., 1996. *Life in Moving Fluids*: *The Physical Biology of Flow*, 2nd edition. Princeton University Press, 484 pp.

Walcott, R. I., 1972, Late Quaternary vertical movements in eastern North America: Quantitative evidence of glacio-isostatic rebound. *Rev. Geophys. Space Phys.* 10: 849-884.

Walker, G. P. L., 1983, Ignimbrite types and ignimbrite problems. *Journal of Volcanology and Geothermal Research 17*: 65–88.

Wilson, C. J. N.; Walker, G. P. L. 1985: The Taupo eruption, New Zealand. I. General aspects. *Philosophical Transactions of the Royal Society of London A314*: 199–228.