# An overview of gravity, seismic refraction, and heat flow in tectonics

GEOL5690

Our discussion of the Sierra Nevada highlights three geophysical techniques widely used in interpreting the tectonics of recent orogens. Some general notes on these follow.

## *Gravity*

The Earth's gravitational field is mostly a simple function of distance from the center of the Earth. It is the variations from this main field that are of interest to tectonics. We may rather brusquely summarize successive refinements of gravity anomalies as follows:

| | |
|---|---|
| Raw gravity | (about 9.81 $m/s^2$ = 981 $cm/s^2$ = 981 Gal = 981000 mGal) |
| Free air anomaly | remove both effect of latitude (reference geoid) (raw - $978031.85[1+0.005278895\sin^2\phi - 0.000023462\sin^4\phi]$ mGal, where $\phi$ is latitude) and distance from geoid (elevation above sea level) (add $0.3086z$, where $z$ is elevation in meters) |
| Bouguer anomaly | remove effect of rock under station at elevation z by subtracting $2\pi G\rho z$ (about $0.04193\ \rho\ z$ when the correction density $\rho$ is in g/cc (usually 2.65 or 2.67) and $z$ in meters) |
| Terrain-corrected Bouguer anomaly | Adjust for mountains and valleys (always a positive value added to raw Bouguer anomaly) |
| Isostatic anomaly | Assuming topography is isostatically compensated, calculate and remove the effect of the subsurface mass (usually assumed to be an Airy root) |

(you might note that we have shamelessly mixed cgs units (mGal, g/cc) with mks (meters). So be careful if you find yourself using these equations!)

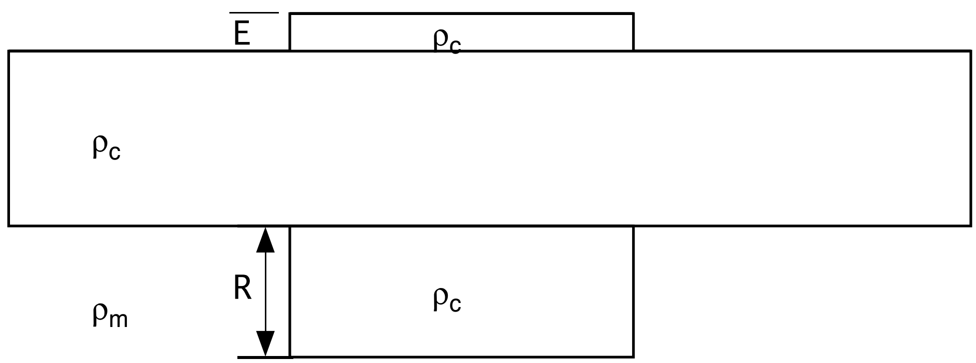
What are they good for?

**Free air anomaly**: Usually mapped at sea (where it is the measured gravity minus the latitude effects) and at long wavelengths on land. The Free Air Anomaly shows deviations from isostatic equilibrium at appropriate wavelengths. The largest free-air anomalies are associated with subduction zones; generally, the free-air anomaly at long wavelengths is small over continents, suggesting that the continents are indeed in isostatic equilibrium. Because there is no correction for topography (except in the very rarely shown terrain-corrected free-air anomaly), this is not very useful in areas with relief, such as most orogenic areas.

**Bouguer anomaly**: This removes the rock we know about above sea level; thus it would represent the gravity field of the Earth were we to bulldoze away all the mountains (assuming no elastic or isostatic rebound). It is the anomaly caused by subsurface density contrasts, such as the roots of mountains or valley fill, etc. This is probably the most frequently shown gravity anomaly. When studying near-surface features (e.g., valley fill), you will sometimes see authors subtract a "regional field" (which is usually those mountain roots) which is sometimes a polynomial fit to the Bouguer anomaly over a broad region.

**Isostatic anomaly**: This removes the surface rocks and the *presumed* support for topography. Ideally this leaves the effects of shallow density anomalies, but at times you will see interpretations of longer wavelength isostatic anomalies as "overcompensated" and "undercompensated" mountain ranges (meaning that the mountains are out of isostatic equilibrium, with an excess buoyancy at depth--mountains want to be higher--or an absence of buoyancy--mountains should sink). Although these interpretations are occasionally correct and indicate that the elastic strength of the lithosphere is important, it can also reflect an error in the isostatic model or a "dipole" effect where a shallow density anomaly is compensated by a deeper anomaly.

Let us use the gravitational acceleration of a slab as a quantitative guide to how this all works. Consider a simple mountain range in isostatic equilibrium:



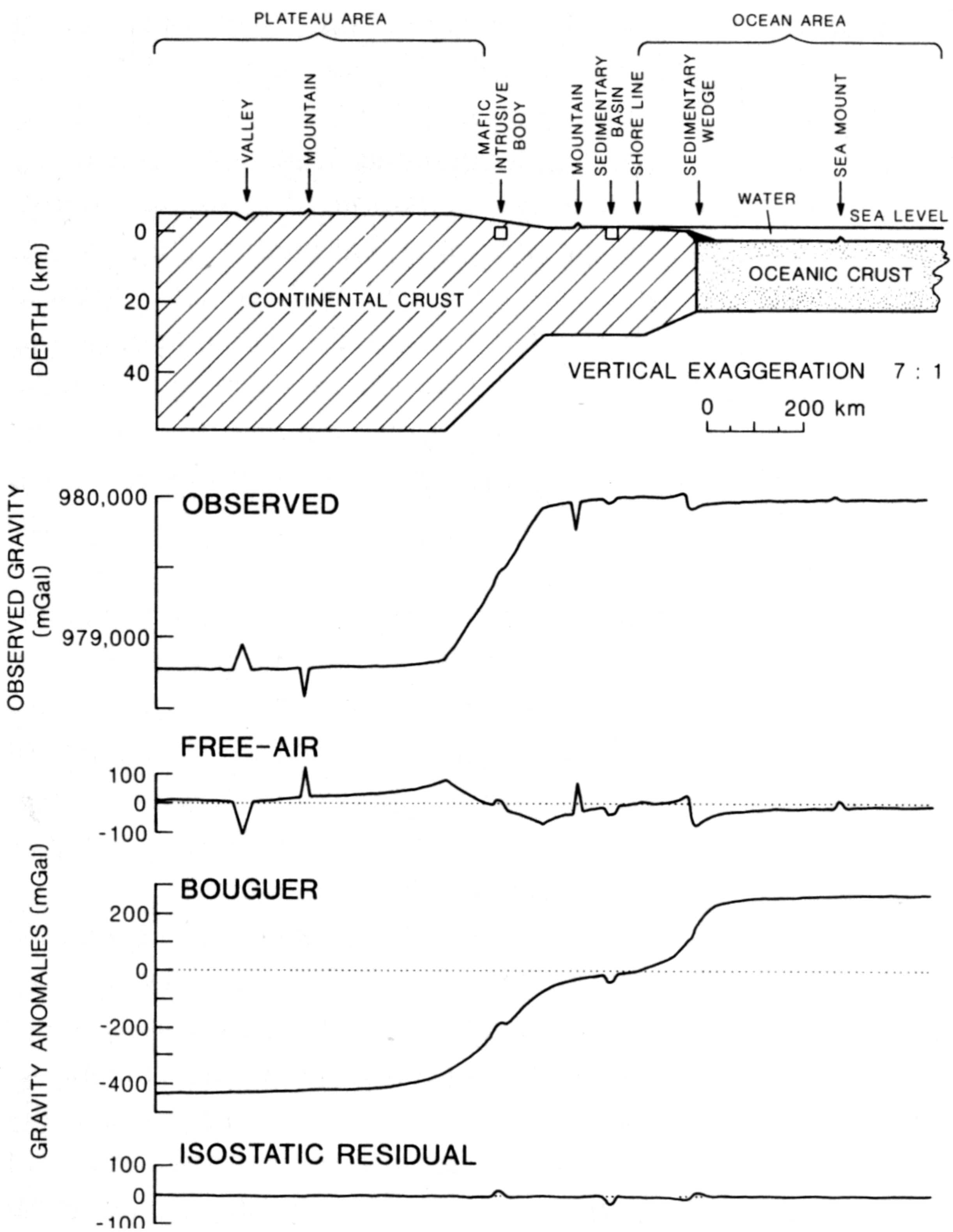

To be in equilibrium, $E\rho_c + R\rho_c = R\rho_m$, which makes $E = R(\rho_m - \rho_c)/\rho_c$. Let us presume that all our gravity anomalies are zero at the sides of the cartoon. Let's also assume our mountains are wide enough that in the center, the gravitational attraction of the topography and the root can be approximated by that of an infinite slab. The attraction of a slab of density $\rho$ and thickness t is $2\pi G\rho t$, where G is the universal gravitational constant ($6.6732\ 10^{-11}$ N $m^2$ $kg^{-2}$). What are the anomalies in the mountain?

The *free air anomaly* only corrects for elevation, so it actually will include the attraction from the mountains ($2\pi G\rho_c E$ in the center of the mountains) and the root (which has a density anomaly of $\rho_c - \rho_m$, and thus an attraction of $2\pi G(\rho_c - \rho_m)R$). Since $E = R(\rho_m - \rho_c)/\rho_c$, the total attraction is $2\pi G\rho_c R(\rho_m - \rho_c)/\rho_c + 2\pi G(\rho_c - \rho_m)R = 0$. Thus our statement above that at long wavelengths, the FAA is 0 when isostasy is operating.

The *Bouguer anomaly* subtracts out the mass of the mountains; thus the Bouguer anomaly should be just the root term, $2\pi G(\rho_c - \rho_m)R$. Note that for isostatically compensated mountains, this is equal to $-2\pi G\rho_c E$, or about -112 mGal for each kilometer of elevation. This tends to be much larger than effects of small, near-surface variations in density.

The *isostatic anomaly* should correct for both the mountains and the root, and thus there should be no isostatic anomaly. Note that a key difference from the free-air anomaly is that to calculate the isostatic anomaly we had to assume both a density for the mountains and a compensation model. An error in either will produce a non-zero isostatic anomaly despite a range being in isostatic equilibrium. Such an error is not possible with the free air anomaly. On the other hand, edges of compensating roots will produce free air anomalies that a correct isostatic anomaly will not have (compare the anomalies under the edge of the Plateau Area in the figure on the next page).

A greater combination of influences was illustrated by Simpson & Jachens (Simpson, R. W., and R. C. Jachens, Gravity methods in regional studies, *Mem. Geol. Soc. Am.*, *172*, Geol. Soc. Am., 35-44, 1989.):



The problems with the isostatic anomaly were discussed by Jachens et al. (Jachens, R. C., R. W. Simpson, R. J. Blakely, and R. W. Saltus, Isostatic residual gravity and crustal geology of the United States, *Mem. Geol. Soc. Am.*, *172*, Geol. Soc. Am., 405-424, 1989):

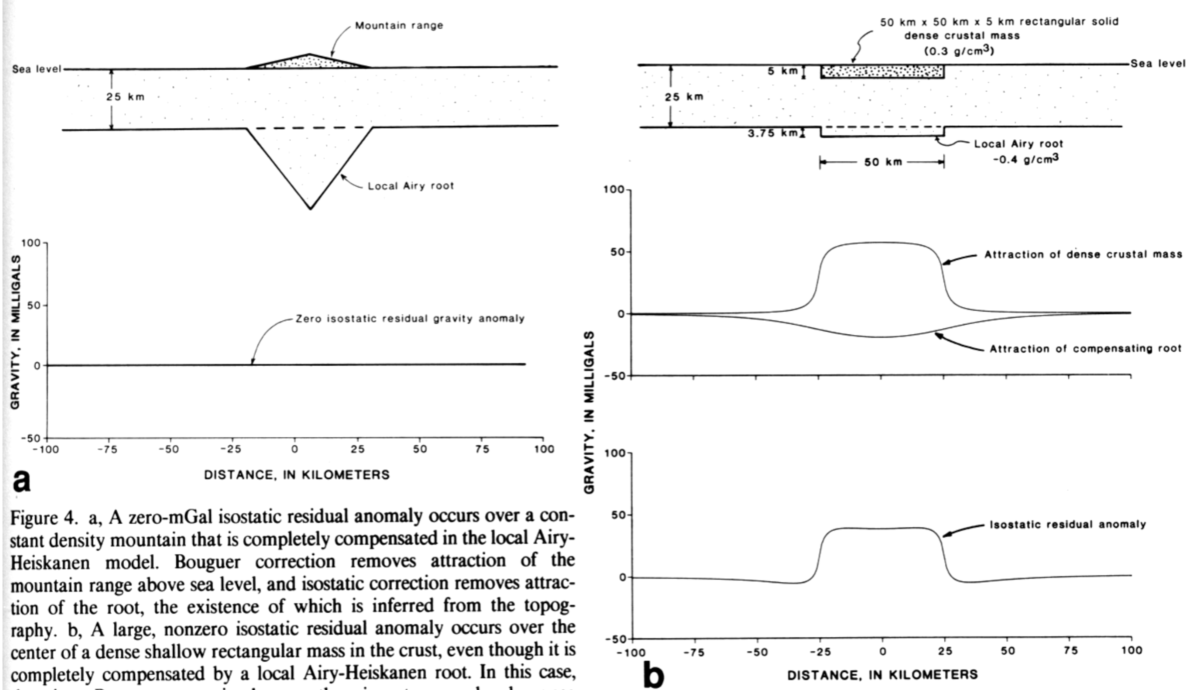



Figure 4. a, A zero-mGal isostatic residual anomaly occurs over a constant density mountain that is completely compensated in the local Airy-Heiskanen model. Bouguer correction removes attraction of the mountain range above sea level, and isostatic correction removes attraction of the root, the existence of which is inferred from the topography. b, A large, nonzero isostatic residual anomaly occurs over the center of a dense shallow rectangular mass in the crust, even though it is completely compensated by a local Airy-Heiskanen root. In this case, there is no Bouguer correction because there is no topography above sea level. Also, there is no isostatic correction for a root, even though one may exist, because the usual isostatic correction only considers the com-

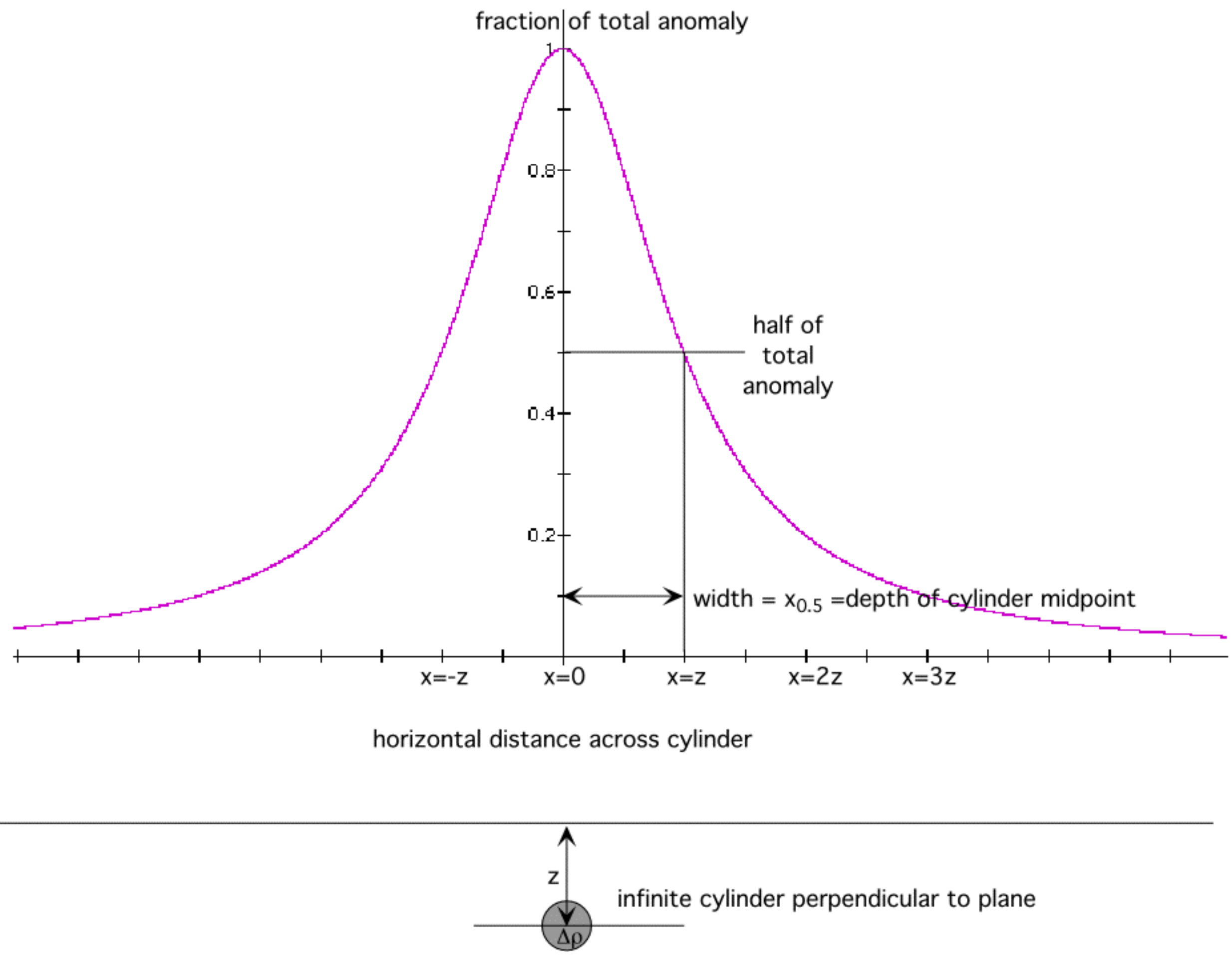
The cartoon on the right illustrates how a disc-shaped subsurface density anomaly could produce a large isostatic anomaly without any area being out of isostatic equilibrium. What is not shown is how an error in the presumed isostatic model can also produce an isostatic anomaly.

**Constraints from gravity**

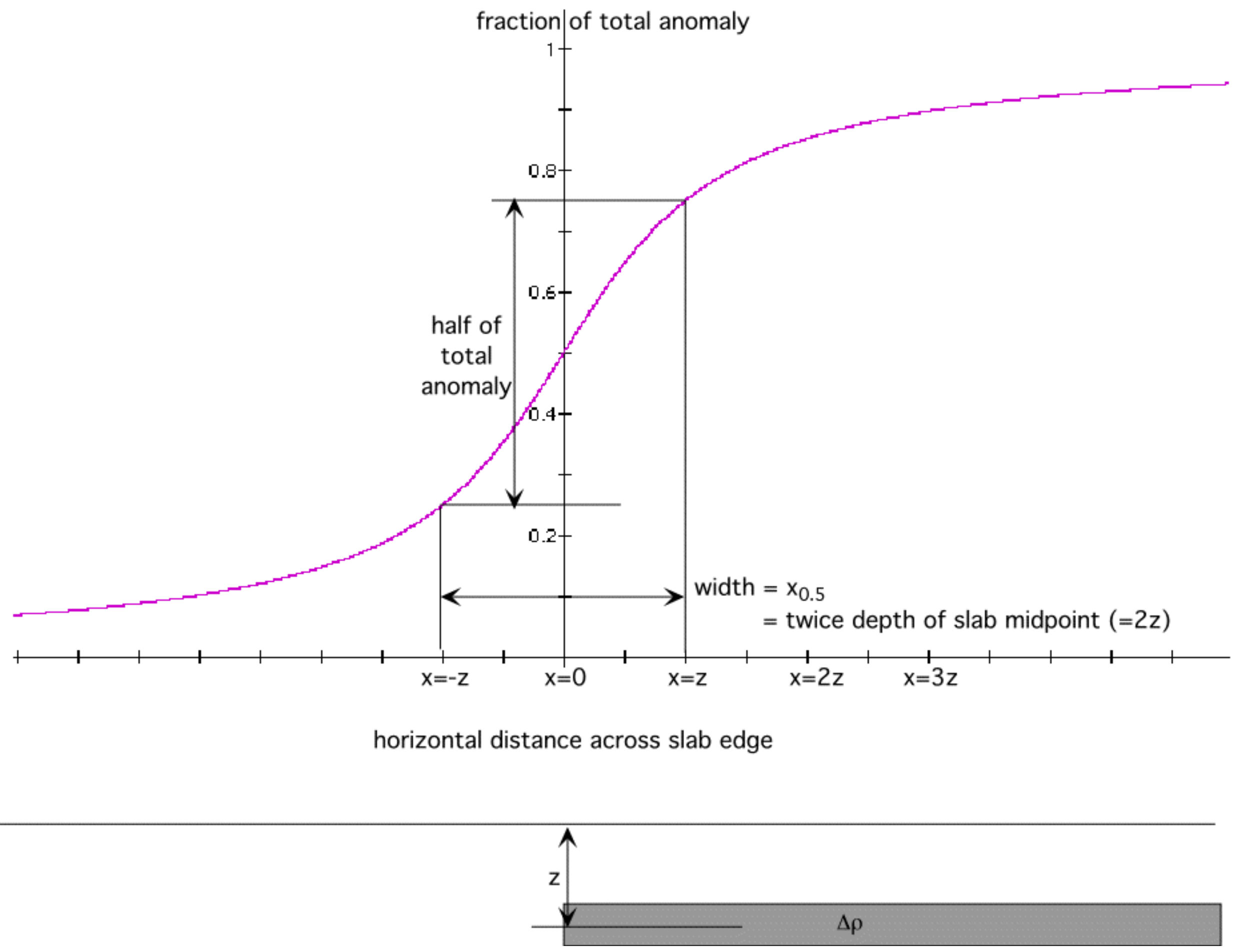
Frequently one will hear scientists say that gravity is non-unique, meaning that many possible density structures can produce a given set of gravity measurements. There are some important constraints that frequently get ignored in the rush to declare gravity non-unique. The first is that the integral of the gravity anomaly will yield the total mass anomaly within the region. Thus in the case of the isostatic anomaly pictured above (at right), if you integrated the isostatic anomaly over the Earth' s surface in two dimensions, it would come to zero. This is because the anomaly from the deeper "root" will cancel that of the shallower load. Were there no deeper "root", the integral would be positive. In many cases such an integral is impossible in practice because of interference from other features unrelated to the feature to be studied. Even so, it is often possible to discern something of the full magnitude of the mass anomaly through an approximate integration.

The more common constraint is that of depth. The *amplitude* of a gravity anomaly usually depends upon the product of the density contrast and the thickness of the body (see the equation for the infinite slab, above). Thus there is frequently ambiguity in the density contrast and the thickness (though the product might be well-constrained); this is a source of the "non-uniqueness" of gravity. However, the *slope* of a gravity anomaly depends on the depth of the midpoint of the density anomaly. A very steep gradient in gravity *requires* a shallow mean depth for the density anomalies causing the gradient. The reverse is not true; a gentle gradient merely permits the density anomaly to be deep, but it can also be shallow. Thus the gradient places a maximum depth upon a density anomaly. Occasionally bad gravity interpretations can be identified by their systematic failure to fit a gradient.

The very simplest rule is that the smallest width that covers half the gravity anomaly is about equal to twice the mean depth of the anomalous density. For instance, for a horizontal cylinder of infinite length, the horizontal distance from the anomaly's peak to a point where the magnitude of the anomaly is half its peak ($x_{0.5}$)is equal to the mean depth of the density anomaly (z). (The width covering half the anomaly might then be from $-x_{0.5}$ to $+x_{0.5}$). A shallower density anomaly could mimic this pattern, but a deeper one cannot.




For the edge of a sheet or slab, the depth of the midpoint of the slab (z) is about half of the width of the steepest part of the anomaly covering half of the total magnitude of the anomaly ($x_{0.5}$) (see figure below). As before, this is the *deepest* the anomaly could be; a shallower dip on the edge of the slab would smear out the anomaly and suggest the possibility of a deeper, vertically dipping slab edge.

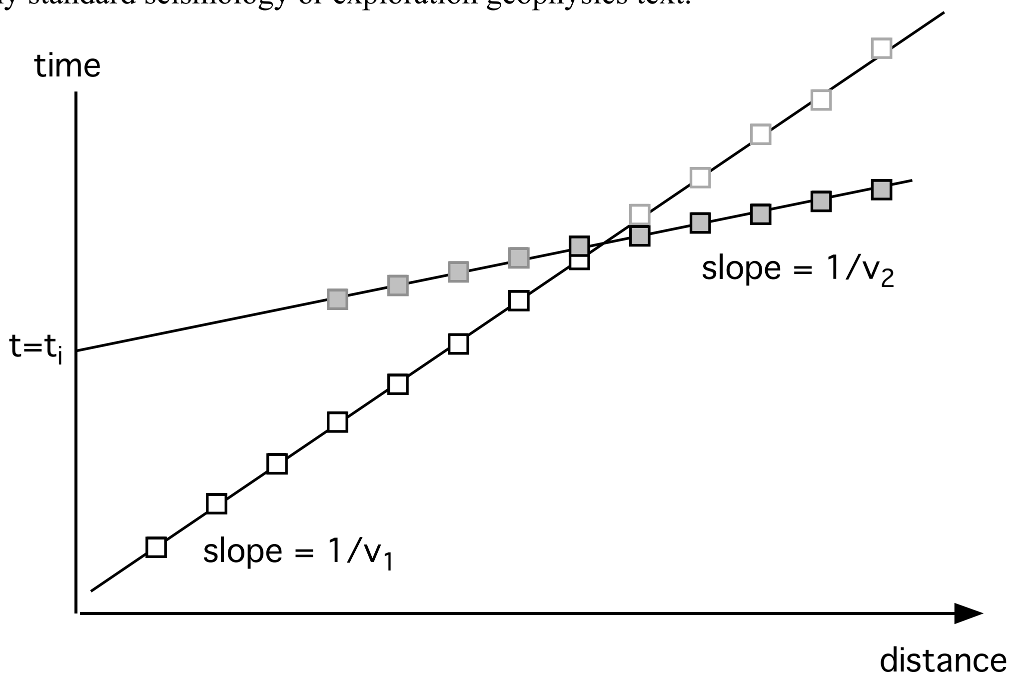


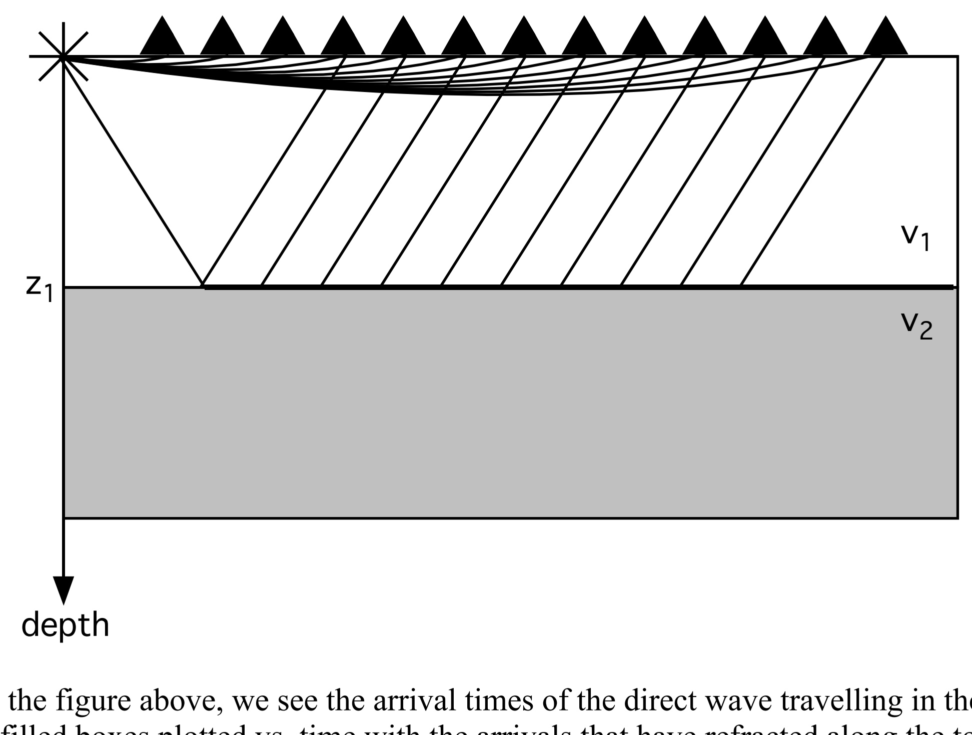

If you applied this lesson to the first diagram of the different kinds of gravity anomalies, you would suspect from the Bouguer anomaly that the depth of the density anomaly at the continent-ocean transition should be shallower than that at the edge of the plateau area. (The latter depth you would derive—about 75 km or so—would be too great because the edge of the plateau's root is not vertical).

## ***Seismic Refraction***

The most popular active source seismological tools for tectonic analysis are seismic refraction and seismic reflection. Reflection yields an image of the crust, usually showing gently dipping sediments and the faults that cut them, sills, and fabrics caused by shearing of deep faults. Reflection is sensitive to rapid changes in the seismological properties of rocks; it is relatively insensitive to overall bulk characteristics. When trying to determine the bulk character of the crust (usually the seismic P velocity), seismologists usually turn to refraction.

As usually practiced, refraction consist of an explosion recorded by a line of seismometers (or geophones). Usually practitioners will try and reverse the line by shooting a charge at each end and work in the last 30-40 years will generally include a large number of intermediate shots for longer profiles (such as those imaging the Moho). For a flat structure, the results are the same from each end; for a dipping structure, the results differ; the average of the two should give a reliable estimate of the average depth and average velocity. Detailed equations can be found in nearly any standard seismology or exploration geophysics text.



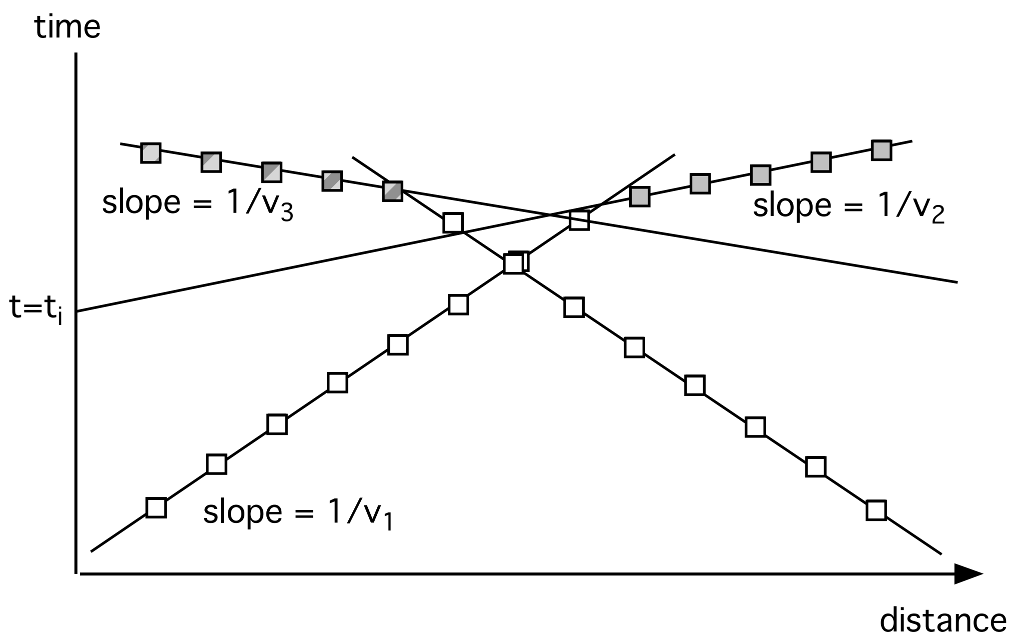



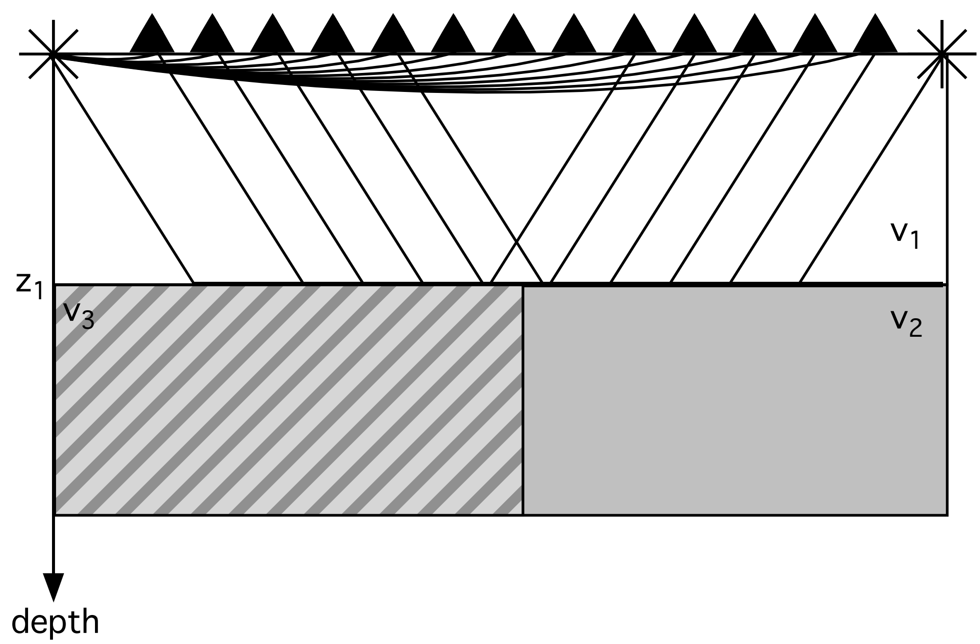

At top in the figure above, we see the arrival times of the direct wave travelling in the top layer as white filled boxes plotted vs. time with the arrivals that have refracted along the top of the bottom layer as gray filled boxes. At any given station, the first arrival (which is usually the most clearly identified in the field) is outlined with black and later (secondary) arrivals are outlined in gray. The bottom part of the figure is a cross section showing the paths the seismic energy follows from the explosion on the left to the stations (solid triangles). Each arrival (the "direct wave" that travels in the top layer and the "refracted wave" that has travelled along the top of the bottom layer) produces a linear array of points, the slope of which is inverse to the velocity of the layer. The intercept time found by extrapolating the second layer's arrivals back to a distance = 0 tells us of the thickness of the top layer through the relation $t_i = 2 z_1(v_1^2 + v_2^2)^{0.5}/(v_2v_1) = 2 z_1 \cos(\theta_c)/v_1$, where $\theta_c$ is the angle of the downgoing ray measured from the vertical for the refracted wave (this is determined by Snell's law).

On the surface, this is straightforward: we determine $v_1$ from the first part of the arrivals that are observed, $v_2$ from the second part, and the thickness from $t_i$ and the velocities we've already found.

**Real life wrinkles**. We'll consider the more common kinds of problems that might affect a tectonic interpretation. First, consider a slow increase in seismic velocity with depth in the top layer. It can happen that no first arrival will appear from this layer; when using our equation for getting this thickness of the layer $z_1$, we will need a *mean* velocity of the top layer but in fact will have only the velocity of the upper part of the layer, which in this case will be too low. We will tend to simultaneously underestimate the velocity of the top layer and overestimate its thickness. The reverse is also possible: if velocities decrease with depth (as can happen with overthrusts or in the presence of magmatic systems), we will overestimate the mean velocity and underestimate the thickness. Thus refraction usually provides a constraint on (roughly speaking) the ratio of the velocity to the thickness of the top layer(s).

*Lateral heterogeneity*. The other main wrinkle is that things change along the profile. The seismic velocity of the lower layer is determined from the right 5 stations in the cartoon above; this corresponds to the velocity in the second layer only *between* where the ray comes up from the second layer to the fifth station from the right and where the ray comes up from the second layer to the last station on the right. Imagine that the velocity of the second layer changes abruptly, and that the profile has been reversed with a shot at the right side.



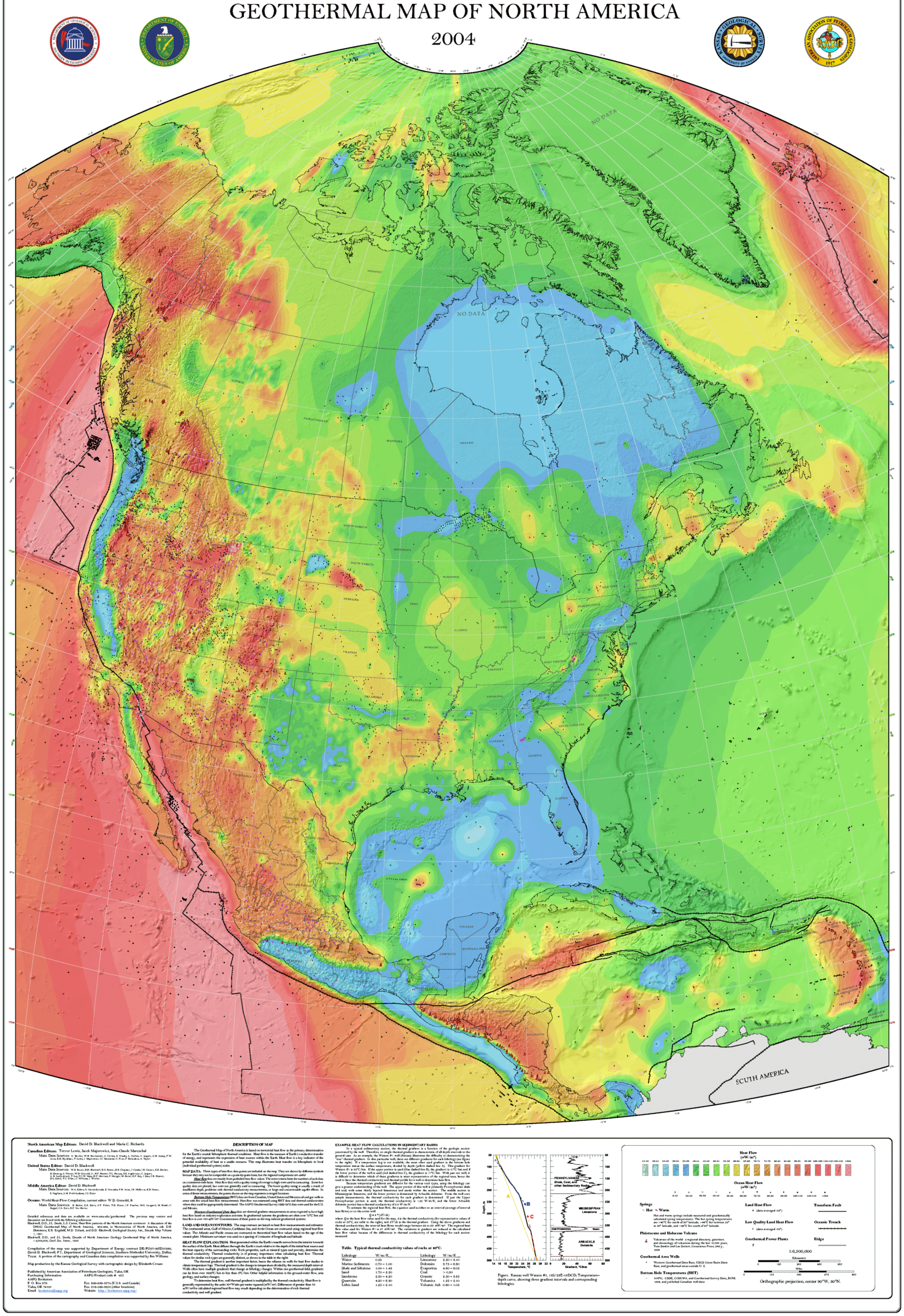

We have dropped the refracted rays that are not first arrivals; because we measure velocities from the slope, we would find from the left shot a velocity of $v_2$ and from the right shot a velocity of $v_3$. This would normally be interpreted as the result of a dip on the bottom of layer 1 and that the velocity of the bottom layer was uniform and about the geometric mean of $v_2$ and $v_3$. In some cases, the error could be profound. Similar sorts of problems can arise from a non-planar interface between the two layers and from lateral variations within the upper layer. In this case, the mean velocity and mean thickness of the layer would be correct, but the inference of a dip incorrect. (A charming example of a non-planar interface came from Mesquite Valley in southern Nevada; a refraction experiment there yielded apparent velocities of about 7 km/s from a refractor at the base of a ~2 km thick sedimentary basin from shots at both ends of the line. A naïve interpretation would have been that there was indeed material with a velocity of 7 km/s under the basin. In fact the basin was bowl shaped and the refracted arrivals were always observed coming updip, which produces higher than true velocities).

A less commonly seen problem is when the seismic energy travels out of the plane of the section. This might be most easily envisioned for shooting down a narrow valley: if the depth of the sediment is greater than the distance to the walls of the valley, the observed refracted arrival will be that from the valley wall and not the basement beneath the valley. Many of the older seismic profiles sought to parallel structural trends so that they could consider the structure underneath to be one-dimensional or a simple 2-D structure (constant velocity layers with dipping interfaces); these profiles are potentially combining arrivals from under their profile with arrivals from the sides of their profiles.

Because in tectonics we frequently are interested in areas with potentially profound lateral differences in physical properties, these difficulties need to be kept in mind. More recent refraction profiles usually have a large number of shots to try and constrain velocities better; in addition, they frequently try to identify secondary arrivals to reduce ambiguity in their interpretation. Older profiles are frequently still of interest both because of the absence of more expensive "modern" profiles and the longer lengths and larger shots used in the past; interpretations of these profiles are potentially loaded with these problems.

### *Heat Flow*

A technique rarely used in commercial geophysics is heat flow. As we have seen, the changes in the thermal structure of the lithosphere are responsible for the topography of most of the ocean basins and dictate the sedimentary history of passive margins. The input of heat into the lithosphere is thus of great interest. It is measured geophysically as the conductive heat flow $q$, where $q = k\frac{dT}{dz}$, where $T$ is the temperature with depth and $k$ is the coefficient of thermal conductivity. For typical rocks, $k$ is 2-3 W $m^{-1}$ $°C^{-1}$. In most areas $dT/dz$ = 20 – 30 °C/km, yielding a heat flow of 40-90 mW $m^{-2}$. Another unit currently falling out of favor is the heat flow unit (HFU): 1 HFU = 41.84 mW $m^{-2}$. The continental average heat flow is about 65 mW $m^{-2}$.


GEOTHERMAL MAP OF NORTH AMERICA
2004
NO DATA
NO DATA
SOUTH AMERICA

Actual measurements of the heat flow are usually made in boreholes. Difficulties that have to be overcome include thermal transients associated with drilling the hole, hydrological effects (hot water bringing heat in or cold water taking heat out), and topographic effects. Most of the practitioners of heat flow are well aware of these problems and usually will have avoided holes that will not yield reasonably representative values.

Heat in the Earth comes both from heat being transferred out of the mantle (the end result of both gravitational energy from formation of the core and radioactive decay in the Earth's interior) and from radioactive decay in the crust. Because most radioisotopes are concentrated in the crust, this last term needs to be considered before we can infer the infusions of heat into the deeper levels of the lithosphere. It turns out, somewhat surprisingly, that in many areas the surface heat flow is linearly related to the heat production of the surface rocks (measured as a volumetric production in $\mu W\ m^{-3}$ or by mass as $\mu W\ kg^{-1}$). The Sierra Nevada turns out to be a classic example:

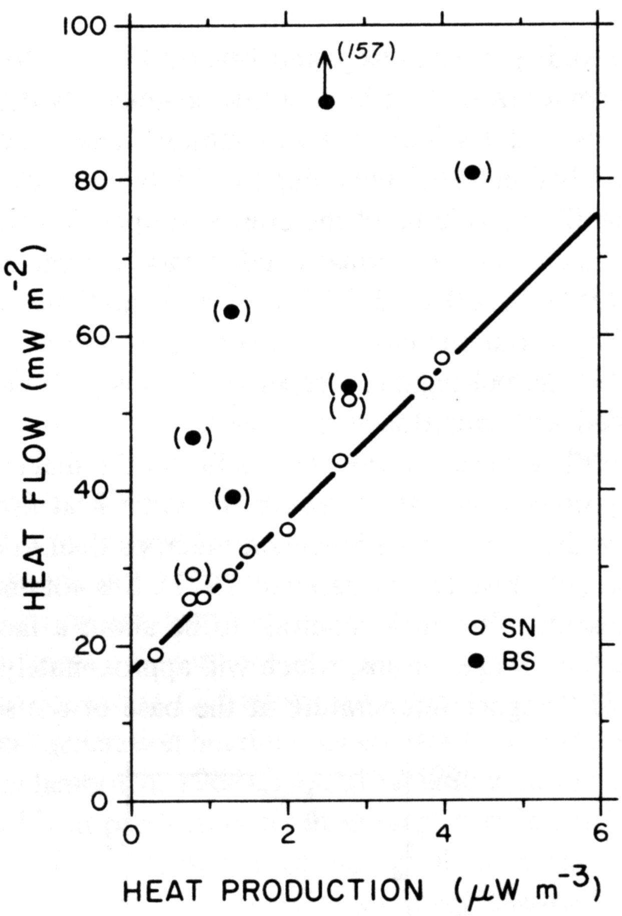



**Figure 8. Heat-flow–heat-production data pairs for Sierra Nevada (SN) and Sierra Nevada–Basin and Range transition (BS).**

The linear fit allows us to remove the effect of heat production in the crust by seeing where the line hits the axis of 0 heat production; this value is called the **reduced heat flow** $q_r$. This is frequently interpreted to be the heat coming from the mantle although it can represent other convective heat sources as well (e.g., emplacement of a sill) or the heat coming from a more uniform body below the surface layer with varying heat production.

The very simplest explanation for the linear relationship of heat flow to heat production is that the radiogenic isotopes are in a layer of some thickness $D$. In this case, $q = q_r + DA$, where

$A$ is the volumetric heat production. The slope of the line above then determines the thickness of the slab. An alternative that will produce the same result is if the heat production decreases exponentially with depth (discussed in Saltus and Lachenbruch's Sierra Nevada paper and in Turcotte and Schubert, pp. 145-147). While the Sierra and many parts of the eastern U.S. have the linear behavior with heat production, other areas do not, including the Rocky Mtns. and the Basin and Range.

Deriving the geotherm with radioactivity isn't too hard. If we are at steady state, so there is no change in temperature with time, then the heat flow out of a volume of rock has to equal that in plus the heat produced in the rock. If heat is flowing upward along the $z$ axis and our volume is a slab of thickness $dz$ with a heat production $A$ (expressed as W m$^{-3}$), then we find

$$q(z-dz) = q(z) + Adz$$
$$\frac{\mathrm{dq}}{\mathrm{dz}} = -A \tag{1}$$

We have made a mild change in definitions so that $q = k\frac{dT}{dz}$ for $z$ positive down but heat flow positive up, so

$$\frac{dq}{dz} = \frac{d}{dz}\left(k\frac{dT}{dz}\right) = k\frac{d^2T}{dz^2} = -A \tag{2}$$

Over a slab where $A$ and $k$ are constant, we can simply integrate this to get

$$\frac{dT}{dz} = -Az + c_1 \tag{3}$$

This is the heat flow, and if $z$= 0 is the surface, then $c_1 = q_o$, the observed surface heat flow. We integrate again to get (within the radioactive body)

$$T = -\frac{A}{2k}z^2 + \frac{q_0}{k}z + T_s \tag{4}$$

where the second constant of integration falls out easily as the temperature at the surface. This is a very helpful equation, though, as we shall see, it has had some misleading implications. We can rewrite this in terms of the heat flow entering from below the slab, as we noted that $q_0 = q_r + DA$:

$$\begin{aligned} T &= \frac{A}{2k}z(2D-z) + \frac{q_r}{k}z + T_s \qquad z < D \\ &= \frac{A}{2k}D^2 + \frac{q_r}{k}z + T_s \qquad z > D \end{aligned} \tag{4a}$$

Note that the two righthand terms are the linear conductive geotherm in the crust without radioactivity. Thus this would require the temperatures to be hotter beneath the radioactive material than to the sides.

Although useful, the derivation above has a problem encountered in many places, such as the Sierra, is that the depth of erosion of plutons is variable, suggesting that the thickness $D$ should vary, but the heat flows still plot on the nice regression line. Although there are a number of possible solutions, an elegant one suggested by Art Lachenbruch some years ago is an exponential decay of heat, $A = A_0 e^{-z/h_r}$, where $h_r$ is a length scale of decay of radioactive heat production with depth. We can redo the integral from (2) to (3):

$$q = k\frac{dT}{dz} = h_r A + c_1 \tag{5}$$

In this case, we might specify the heat flow at great depth be a constraint, which we will designate $q_m$, and so $c_1 = q_m$. Thus we see that $q$ will depend linearly on $A$ independent of the depth $z$, satisfying our observations in some areas. We also get (once again) that the heat flow at the surface when $A$ equals zero is the heat flow at depth. Another integration yields

$$\begin{aligned} T &= -\frac{h_r^2}{k}A_0 e^{-\frac{z}{h_r}} + \frac{q_m}{k}z + c_2 \\ &= \frac{h_r^2}{k}A_0\left(1 - e^{-\frac{z}{h_r}}\right) + \frac{q_m}{k}z + T_s \\ &= T_s + \frac{q_m}{k}z + \frac{(q_0 - q_m)h_r}{k}\left(1 - e^{-\frac{z}{h_r}}\right) \end{aligned} \tag{6}$$

This alternative geotherm isn't very different from what we would get with a slab, but the geotherms will diverge with greater depth such that an exponential decay will have higher temperatures than the slab below the slab thickness.

Neither a slab of constant heat production nor an exponential decay of heat production is consistent with what we observe in the field (usually it seems to be somewhere in between). These two models provide us some bounds for the shallow subsurface and for greater depths under certain circumstances.

In a steady-state, equilibrium system, the reduced heat flow is generally interpreted as the heat flux coming from the mantle across the Moho (and indeed the heat flux coming through the base of the conducting lithosphere). Again, in tectonics we frequently want to examine areas undergoing change, and so a valid question is, how long does it take for the heat flow to change? The plot below from Saltus and Lachenbruch (1991) helps understand this:

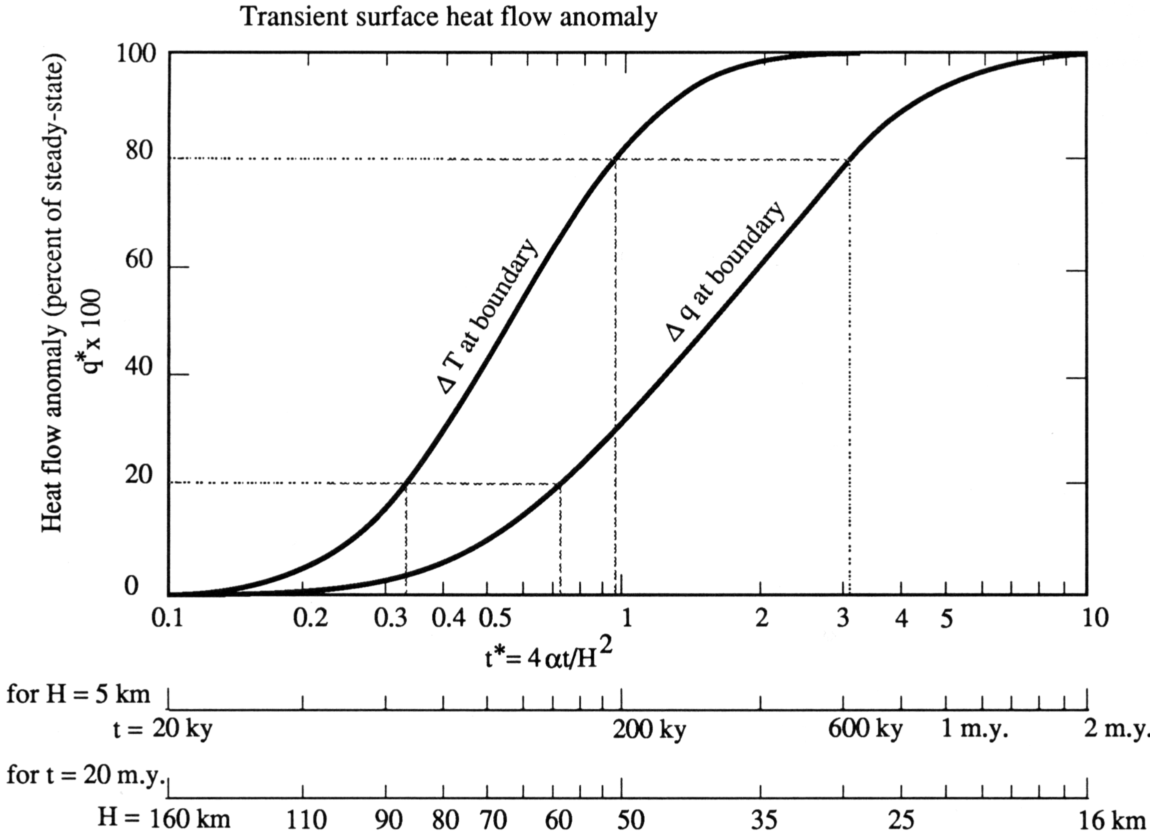



Fig. 7. Dimensionless heat flow $q^*$ versus time $t^*$ for the conduction of a thermal disturbance through a finite slab of thickness $H$. Dimensionless heat flow $q^* = q/q_\infty$ is the ratio of the transient surface heat flow $q$ to the steady state (equilibrium) heat flow $q_\infty$. The dimensionless time is $t^* = 4\alpha t/H^2$ where $\alpha$ is thermal diffusivity and $t$ is time since the initiation of the basal thermal disturbance. Two thermal boundary conditions for the disturbance are used: $\Delta T$ is a change in basal temperature, and $\Delta q$ a change in basal heat flux. The time axis has been labeled dimensionally for two different situations as discussed in the text. If $H$ is taken to be 5 km, the depth to active seismicity in the southeastern Sierra Nevada, the axis may be labeled in time required to conduct a thermal disturbance at that depth to the surface. If $t$ is taken to be 20 Ma, the time since the onset of uplift in the southern Sierra Nevada, the axis may be labeled in lithospheric thickness. The horizontal and vertical dotted lines show the intersections of the curves with the 20% and 80% surface heat flow conditions as discussed in the text.

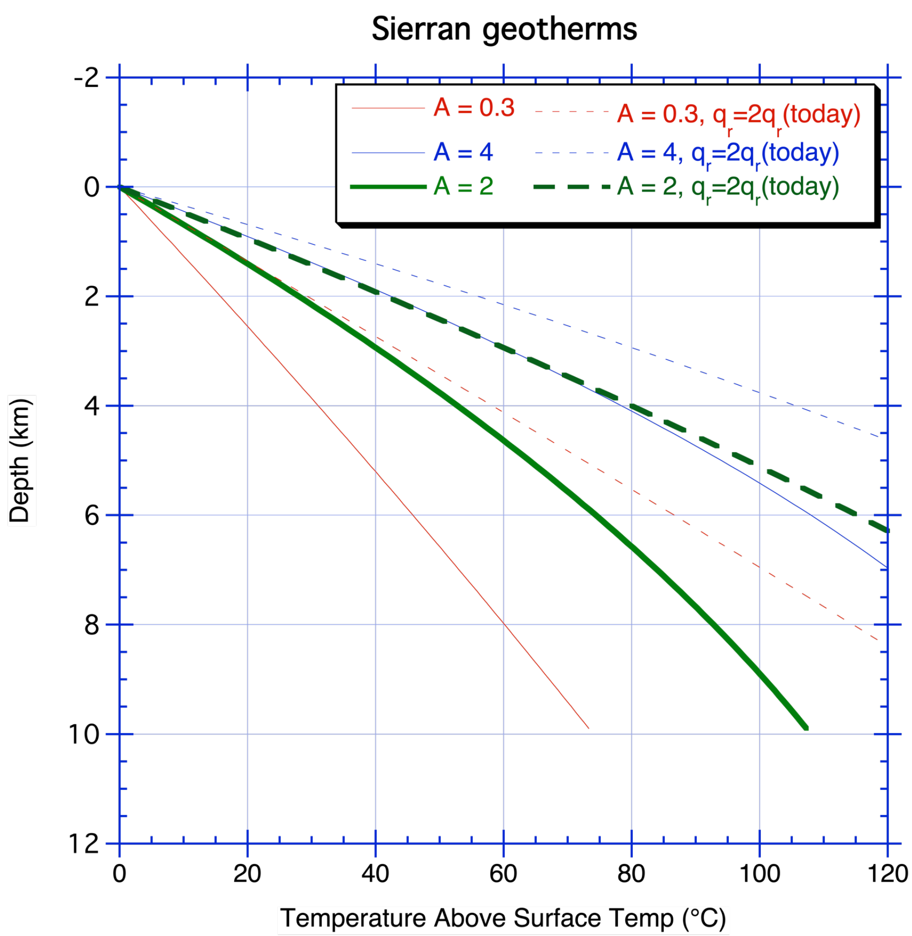
The idea is that if a new fixed temperature is emplaced at some depth $H$, then the heat flow with time will change as shown by the "$\Delta T$ at boundary" curve above. If instead of a fixed temperature we change the heat flux, then the surface heat flow will follow the "$\Delta q$ at boundary" curve, where the original heat flow is at the 0 position on the y-axis and the new steady-state is at 100. The two extra axes below are for emplacement of a body at 5 km depth and for the depth of a body emplaced 20 m.y. ago and producing the changes. Using the bottom curve, for instance, if we suspect that the surface heat flow is exactly halfway from old values to new, and we believe this was caused by a change in heat flow, the depth of that change would be about 40 km. Thus we can see that changes to the thermal structure at the base of the lithosphere take tens of millions of years to show up at the surface.

The above assumes conductive transport of heat. Unfortunately, convective transport occurs at many levels in the Earth. Of course convection in the mantle exists, but this isn't a serious problem here as we are not assuming that conduction extends into this area. Convection in the crust is most spectacular near mid-ocean ridges. For a long time it was a puzzle that heat flow observed near mid-ocean ridges was fairly low; it turns out that most of the heat comes out through hydrothermal vents and much of the sea floor is actually the recharge area for these vents and thus has abnormally low heat flow as cold seawater is moving down into the rocks.

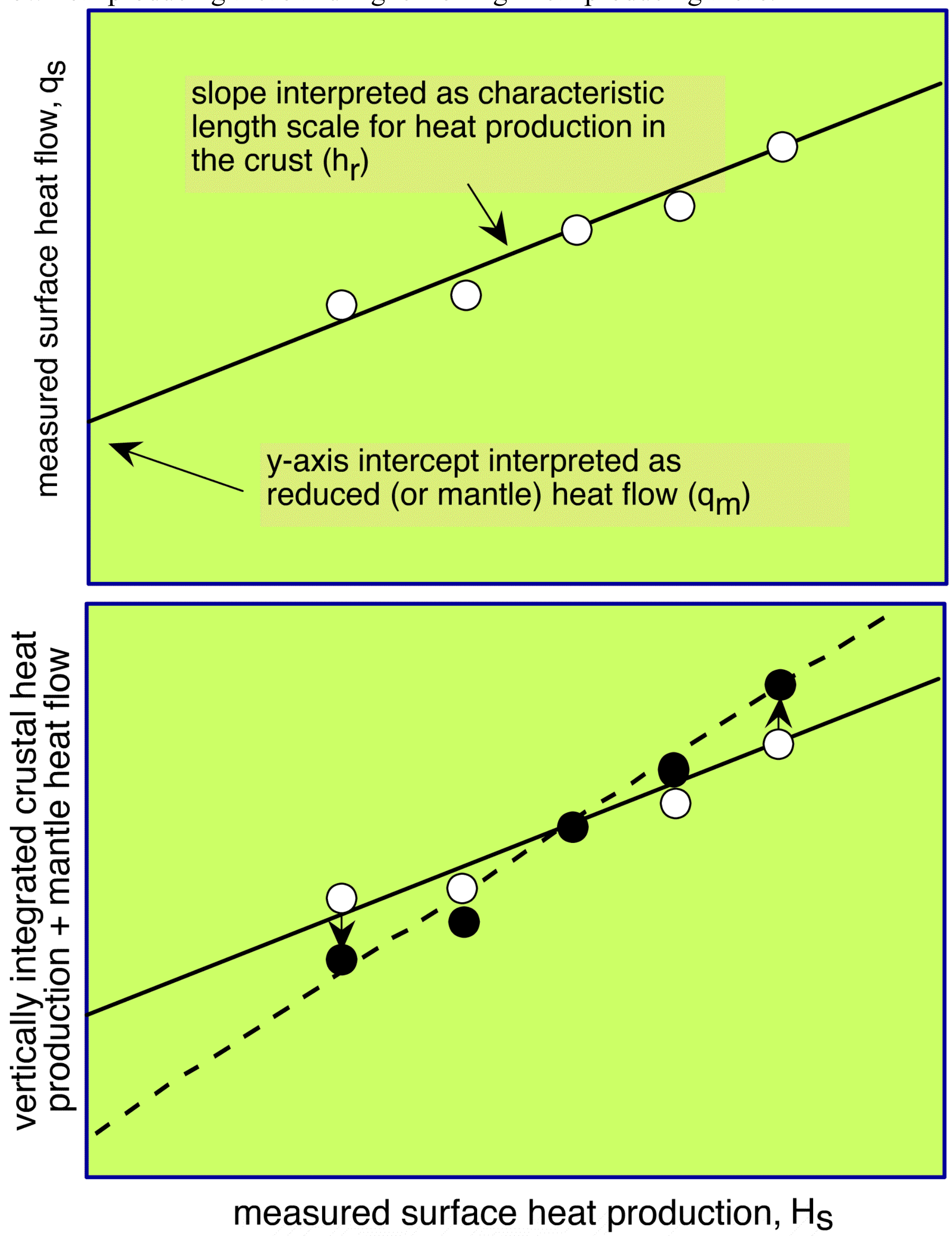
On continents the most obvious conductive transport of heat is magmatism and associated motion of hydrothermal fluids. A more subtle effect, locally quite significant, is hydrological. In many areas there are regional aquifers that have fairly rapid transport of water horizontally. These waters act to transport the heat convectively out to the discharge areas; some examples include the Snake River Plain in Idaho and parts of the Colorado Plateau. Low heat flow measured in these areas appears to be an artifact of the groundwater hydrology.

**Real heat flow from the mantle**

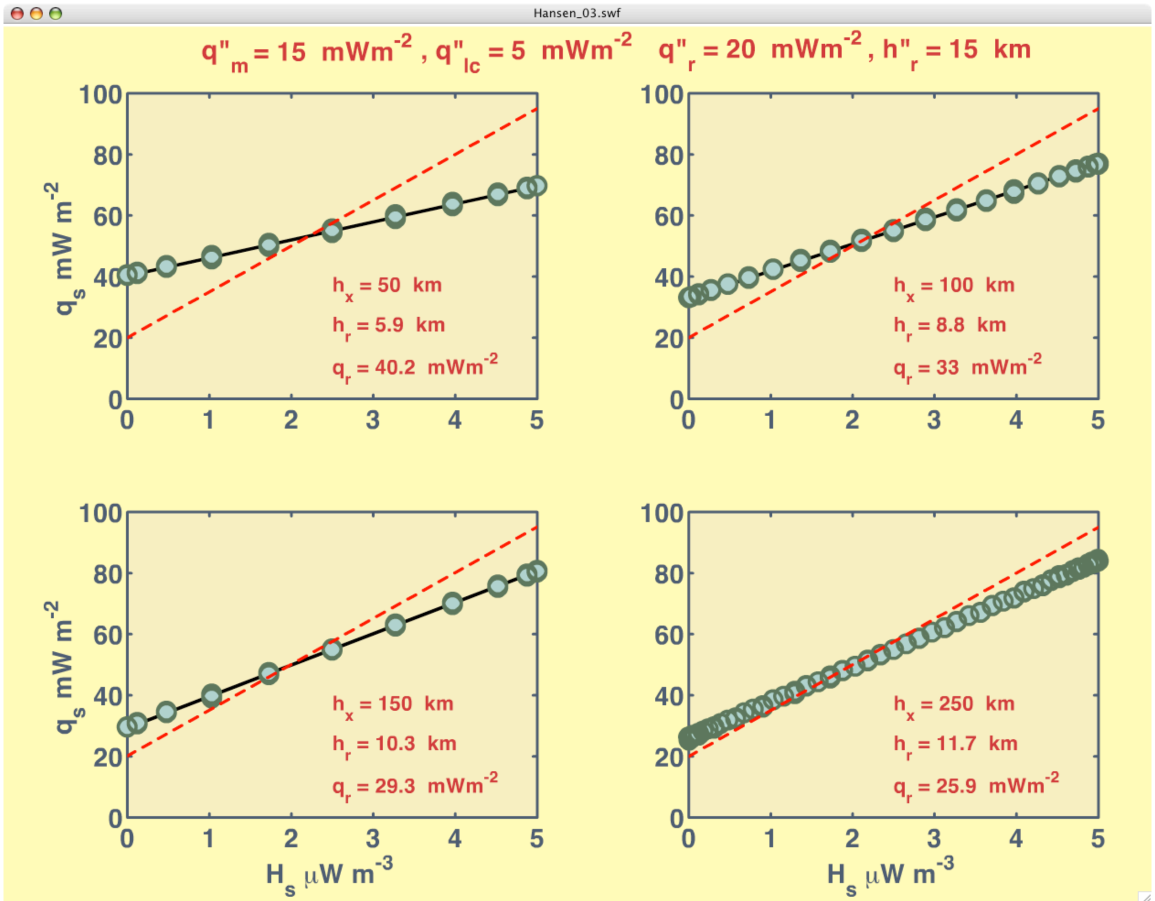
We return to what the real deep thermal structure looks like. Problems with our 1-D interpretation can most easily be seen by calculating one-dimensional geotherms from the surface down assuming a constant deep heat flow ($q_r$) but differing heat productions in a slab using equation (4):




As you go down from a fixed surface temperature, the geotherms diverge, and while they bend back to being parallel as you emerge from the bottom of the radioactive part of the crust (see eqn 4a), the temperatures would seem to be higher under the more radioactive crust. At first blush, this would seem to require the lithosphere to vary in thickness with variations in heat production, which is really rather ridiculous unless the length scales of radioactive heat production are on the order of the thickness of the lithosphere. For much shorter variations (which is true in the Sierra to a large degree) we suspect that the heat from the mantle under high heat producing regions will instead flow into the colder regions under lower heat producing regions. This means that the actual heat being produced under a given measurement is actually lower for low heat producing areas and higher for high heat producing areas:



Thus the reduced heat flow is no better than the maximum heat flow from the mantle. Mike Sandiford has made some nice plots of this effect (http://jaeger.earthsci.unimelb.edu.au/msandifo/Essays/Hansen/Hansen.html). The summary is illustrated here:




Here the lateral length scale of variations in heat production is $h_x$ and the actual heat flow from depth is 20 mW m$^{-3}$ (the dashed line is the relationship you would expect if there was no lateral conduction of heat).

From this analysis, the thickness of the lithosphere calculated from a conductive geotherm assuming a heat flow equal to the reduced heat flow will be too thin. What you really need is the properly averaged heat production over the region (notice that the dashed lines cross the “observed” curves above at some point likely to represent the average heat production) with a proper (lower) mantle heat flow.

**Assumptions, assumptions, assumptions**

The discussion above includes a number of assumptions we should be very careful about. One is that the thermal conductivity $k$ is constant with depth. This is not true in the mantle (there is a depth dependence) and it is not true for crustal rocks. Most notably, this is not true for shales, which tend to have a lower $k$ than most other crustal rocks. This means that a pile of shale will act as a blanket, and temperatures will tend to be higher under them than other rocks. It turns out this is a big issue in interpreting fission track ages in the Rocky Mountains (see Kelley and Chapin. Denudation history and internal structure of the Front Range and Wet

Mountains, Colorado, based on apatite-fission-track thermochronology. New Mexico Bureau of Geology and Mineral Resources Bulletin (2004) vol. 160 pp. 41-78).

Another huge assumption is that things are in equilibrium. You can go out of equilibrium in two easy ways: change the temperature or heat flow at the bottom of the lithosphere, or add or remove material from the top by sedimentation or erosion, from the middle through tectonism, or from the bottom through convective processes.

The final big assumption is that heat is conducted. In fact, within the crust there are processes like hydrothermal systems, regional aquifers, and magmatic systems that all are capable of moving heat through advection. We will not explicitly consider this issue here, but you should be aware that these can be large-scale effects. For instance, low heat flow over much of the Colorado Plateau has been attributed to water flow transporting heat sideways in aquifers such as the Coconino sandstone; a similar anomaly is along the Snake River Plain, where surface waters dive into porous basalts and emerge as large springs into the Snake River. Perhaps the most dramatic example is the collection of low heat flow measurements along midocean ridges: these areas are recharging the hydrothermal systems that have more localized upwelling limbs at black smokers and similar hydrothermal vents.