[AMERICAN JOURNAL OF SCIENCE, VOL. 293, MARCH, 1993, P. 307–336]

CONVECTIVE REMOVAL OF LITHOSPHERE BENEATH MOUNTAIN BELTS: THERMAL AND MECHANICAL CONSEQUENCES

J. P. PLATT and P. C. ENGLAND
Department of Earth Sciences, Parks Road, Oxford OX1 3PR, United Kingdom

ABSTRACT. Shortening and thickening of continental lithosphere cause an increase in surface elevation, a decrease in the thermal gradient and, unless the crust is initially quite thick, a decrease in potential energy with respect to a mid-ocean-ridge lithospheric column. Continental convergence could therefore be self-sustaining. The lower part of the lithosphere (whether continental or oceanic) is probably removed intermittently by convection, however, and this process may be triggered by thickening, though delayed a few million to a few tens of million years after the start of thickening. Convective removal of the lithospheric root below regions of continental convergence will cause a rapid increase in surface elevation and potential energy and create a step in the geotherm. The large excess in potential energy with respect to its surroundings may result in indefinite extension of the continental column. In addition to extension, convective removal of the lowermost lithosphere may lead to characteristic igneous and metamorphic events. The step in the geotherm at the base of the convectively thinned lithosphere will cause partial melting if the lithospheric mantle has been previously enriched by metasomatic processes. This melting could occur before any appreciable extension of the crust took place. Extension of lithosphere immediately following convective removal of the root will decrease the thermal time constant for decay of the step in the geotherm, allowing a transient heating event to affect the crust. The increase in temperature will be synchronous with extension and decompression in upper mantle and lower crustal rocks, but if extension is rapid, it may reach the upper crust after extension has ceased and be detectable there as a post-tectonic low P/T ratio metamorphic event. Extension may be accompanied by decompression-related magmatism.

INTRODUCTION

Plate tectonics has been overwhelmingly successful in explaining the large-scale motions of the lithosphere, and the geology of the ocean basins can be interpreted in terms of plate processes down to the 10 km scale in many areas. Within the continents, however, there are areas comparable in size to some of the smaller plates that show patterns of deformation inconsistent with the premises of plate tectonics. The mechanics of the continents differ from those of the oceans in two important ways. First, the continental lithosphere is in many places weaker than the oceanic lithosphere, and so it deforms under the action of stresses that have little or no effect on oceanic lithosphere. Second, deformation of the continental lithosphere itself generates internal stresses that can drive further deformation (England and Jackson, 1989; Molnar, 1988; Molnar and Lyon-Caen, 1988; Molnar and Tapponnier, 1975). These internally generated stresses, associated in part with crustal thickness contrasts, are responsible for the horizontal extension that seems to characterize the latest phases of many mountain-building episodes.

Several lines of evidence, however, suggest that late-stage extension in mountain belts must be attributed to some process in addition to crustal thickening and the slowing down or cessation of plate convergence (England and Houseman, 1988, 1989). First, the transition from compression to extension in some Cenozoic orogenic belts took place rapidly and nearly simultaneously over large regions. Second, as in the case of Tibet, extension is taking place within a setting of overall plate convergence (Armijo and others, 1986). Third, as we show in this paper, extension is accompanied by distinctive magmatic and metamorphic effects that require abrupt changes in the thermal structure of the lithosphere.

One purpose of this paper is to unify the differing lines of argument that point to the conclusion that rapid sinking of the lowermost continental lithosphere in regions of continental convergence, and its replacement by hot asthenosphere, is the process that causes these changes in tectonic style. The second purpose is to investigate the geologically observable consequences of this process. An inevitable consequence is an attendant increase in surface height, but this can be hard to demonstrate for the past (England and Molnar, 1990, 1991a, b). We show, however, that it is also likely to lead to large amounts of extension, and that thinning of the lithosphere by a combination of convective removal and extension creates a distinctive thermal structure that can lead to characteristic relations between igneous activity and tectonic style, and to characteristic pressure-temperature-time paths for regional metamorphism.

MECHANICAL CONSIDERATIONS

The region that most clearly exemplifies both the distributed nature of continental deformation and the abrupt transition (both spatial and temporal) between compressional and extensional deformation is the Tibetan plateau and its surrounding mountain belts, where the convergence of India with Asia has been accommodated over the last 50 Ma. The distribution of seismicity demonstrates that deformation is currently distributed over an area of several million square kilometers. Furthermore the distribution of topography indicates that most, and perhaps all, of the convergence between India and Asia can be accounted for by horizontal shortening within the Asian lithosphere (England and Houseman, 1986).

Such deformation cannot be described using the methodology of plate tectonics. It is, of course, ultimately a consequence of the motions of the bounding plates, and it might reasonably be assumed that it would reflect those motions. If this assumption were correct, the region would exhibit horizontal shortening and some combination of thickening and strike-slip deformation. Much of the high Tibetan plateau, however, shows seismological and morphological evidence for normal faulting.

(Armijo and others, 1986; Molnar and Chen, 1983; Molnar and Tapponnier, 1978; Tapponnier and others, 1981). This means that the lithosphere there is currently getting thinner. The present pattern of deformation thus reflects neither the overall convergent pattern of current plate motion, nor the Cenozoic history of north-south shortening which caused the present thickened crust and elevated topography of the region (Dewey and others, 1988). Furthermore, the transition from north-south shortening to east-west extension appears to have taken place within a few million years, beginning after middle Miocene time (Armijo and others, 1986). A large-scale process that is to some extent independent of plate tectonics is at work.

Our discussion of the extension of mountain belts follows physical analyses of continental deformation carried out by a number of authors (Houseman and England, 1986; England and McKenzie, 1982, 1983; Fleitout and Froidevaux, 1982; Vilotte and others, 1986) who use continuum mechanics to analyze the deformation of idealized continental lithosphere. The use of continuum mechanics to describe the continental lithosphere often seems oversimplified, or merely wrong-headed, to the geologist whose life is spent studying the manifestly discontinuous upper crust. It is important to emphasize, however, that the continuum approach does not consist of assuming that discontinuities are absent from the continental lithosphere. Quite the reverse is the case: the application of continuum mechanics to the continental lithosphere carries with it the assumption that there are so many discontinuities that the bulk mechanical properties of the system can only be described in terms of the average mechanical properties of the discontinuous medium. In thinking of the creep of a rock sample at the hand-specimen scale, one does not take account of individual grain boundaries, subgrain boundaries, or dislocations. Rather one describes the deformation by some kind of flow law, which represents the spatially averaged properties of the discontinuous body. In the same way, investigators of continental lithosphere on a scale comparable to its thickness treat the deformation as being macroscopically ductile (Tapponnier and Molnar, 1976; England and Jackson, 1989) and take average values for its rheological properties.

Potential Energy Changes and Deformation of Continental Lithosphere

The argument that follows is based on the concept that different columns of rock, measured down to some level at which we may assume complete isostatic compensation, can have different gravitational potential energies (Frank, 1972; Molnar and Lyon-Caen, 1988). If two such columns are juxtaposed, the column with the greater gravitational potential energy is, by definition, capable of doing work upon the column of lower potential energy. Such work would, again by definition, be carried out by a force acting in the direction from the column with higher potential energy to the column with lower potential energy. Thus the column of higher potential energy exerts a horizontal compressive deviatoric stress on the column of lower potential energy. If the columns deform under these deviatoric stresses they will change shape and surface height but will not, in the absence of non-isotatic forces on their bases, depart from isostatic equilibrium.

Tectonic conclusions drawn from arguments based upon potential energy have a somewhat abstract flavor, so we summarize here an alternative way of viewing this concept. Dalmayrac and Molnar (1981) point out that because the assumption of isostatic balance is justified over horizontal length scales exceeding a few tens of kilometers, the vertical traction, σ_v, acting on a horizontal plane of depth, d, is equal to the weight per unit area of overlying rocks:

$$\sigma_v = g \int_0^d \rho(z)\, dz \tag{1}$$

where z is depth, g is the acceleration due to gravity, and we are using the geological convention for the signs of tractions.

The state of deviatoric stress within the lithosphere, when considered on length scales of a hundred kilometers or more, may be expressed in terms of the *difference* $(\sigma_v - \sigma_h)$ between σ_v and the horizontal traction σ_h on vertical planes. If σ_h varies slowly in the horizontal plane then the difference in $(\sigma_v - \sigma_h)$ between any two columns of rock may be assessed from the difference in σ_v between the two columns. The horizontal normal deviatoric stress will have a larger tensional value in the column in which σ_v is the higher. It is easy to show (Molnar and Lyon-Caen, 1988) that the difference in potential energy per unit area between two isostatically-balanced columns is physically identical to the difference in the integral of σ_v from top to bottom of the columns. For brevity we shall refer, in this paper, primarily to differences in potential energy, but it must always be borne in mind that these differences involve stress differences that lead to deformation.

One of the striking features of the Tibetan plateau is that at elevations above 4000 m the dominant modes of deformation in the upper crust are normal faulting and strike-slip on conjugate sets of faults, producing east-west extension and north-south shortening (Molnar and Deng, 1984), whereas in the Himalaya and on the eastern and northern margins of the plateau, at elevations less than 4000 m, there is a more or less radial pattern of thrust faulting normal to topographic contours (Molnar and Lyon-Caen, 1989). Much the same is true of the Andes, where the dominant mode of deformation is strike-slip and normal faulting at elevations of 3000 to 4000 m or more, and active thrusting is confined to lower elevations of the eastern flank of the range (Dalmayrac and Molnar, 1981; Sébrier and others, 1986; Suarez, Molnar, and Burchfiel, 1983). The close association of extensional faulting with high elevation and the observation that transition in tectonic style is governed by elevation strongly support the contention that potential energy differences, associated with surface height differences, play a major role in continental tectonics (Tapponnier and Molnar, 1976; England and McKenzie, 1983; Houseman and England, 1986).

Most of the studies of the mechanics of orogenic belts cited above concentrated on the changes in the potential energy (and surface height) of continental lithosphere arising from crustal thickness variations and assumed a constant density for the mantle. Under this assumption, the difference in potential energy between two columns in isostatic balance is in proportion to the difference between the squares of their respective crustal thicknesses (Frank, 1972). By these calculations, potential energy contrasts due to crustal thickness contrasts of 30 km or more, such as characterize the Tibetan and Andean plateaux, are in the range of 5 to 10×10^{12} Nm^{-1} (Newton meters per square meter of surface area) (Frank, 1972; Molnar and Lyon-Caen, 1988). This range is equivalent to a stress difference of 50 to 100 MPa averaged over a 100 km thick load-bearing layer and is comparable to the stresses associated with plate motions. This led to the idea that surface height in deforming regions might be a direct measure of the compressional stresses (Molnar and Tapponier, 1978). According to this idea, thickening of the crust will slow as the load caused by the increase in surface height counteracts the compressional stress caused by convergence. Further crustal shortening and thickening would then take place on the margins of the deformed region, or elsewhere.

Concentrating on crustal thickness contrasts has the drawback, however, that it neglects the higher density of the lithospheric mantle with respect to the warmer underlying asthenosphere. Because mantle rocks of the continental lithosphere are denser than the asthenosphere, thickening of the lithosphere as a whole will produce a lithospheric root whose higher density may partly or completely offset the buoyancy of the thickened crust. Fleitout and Froidevaux (1982) have suggested that, as a result, thickening of continental lithosphere may in fact cause a decrease of potential energy, rather than an increase as suggested above. England and Houseman (1989) showed that this is true for a wide range of plausible combinations of crustal and mantle density. Furthermore, the change in potential energy of a column of lithosphere during homogeneous thickening is monotonic (England and Houseman, 1989, app. A); so if a piece of lithosphere is initially in potential energy deficit with respect to another column, it will remain so, whatever degree of thickening it undergoes, provided that the lithospheric mantle thickens in proportion to the crust.

The natural reference column to consider, in this context, is that beneath the oceanic ridges. One may imagine rigid oceanic lithosphere joining a ridge to a column of continental lithosphere. If the continental column has higher potential energy than a column measured down to the same depth beneath the ridge, then additional work is required if the oceanic lithosphere is to move toward and thicken the continental lithosphere. If, in contrast, the continental lithosphere has the lower potential energy, then no additional work is required; in fact thickening of the continental lithosphere will be a self-sustaining process.

There are few direct lines of evidence to suggest whether homogeneous thickening of continental lithosphere should be self-sustaining. England and Houseman's (1989, app. A) calculations suggest that the potential energy will decrease during thickening if the initial lithosphere thickness is greater than 120 km and the crust is less than 45 km in thickness (see also fig. 2). These are acceptable lower limits for lithospheric thickness and upper limits for crustal thickness for many regions of stable continental lithosphere (Braile, 1989; Meissner, 1986). The prevalence of thrust faulting earthquakes in tectonically quiescent continental interiors (Chen and Molnar, 1983) also suggests that these regions are in a state of lower potential energy with respect to their surroundings.

In contrast, however, the present day deformational styles of areas of distributed continental deformation, such as Tibet and the Andes, suggest that there has, at some stage, been a net increase in their potential energy. The question naturally arises, therefore, as to whether the present tectonics of those regions is the result of a set of circumstances that caused their lithospheres to be positively buoyant before deformation began, or whether some process occurred during deformation that led to an increase in their potential energy.

One such process was suggested by Houseman, McKenzie, and Molnar (1981) and Fleitout and Froidevaux (1982). The lithospheric mantle is colder than the underlying asthenosphere and, therefore, (for any given pressure) more dense. Thus, given enough time, it will sink. In its uppermost portions the viscosity of the lithospheric mantle is so great that the time required for such spontaneous sinking probably far exceeds the age of the Earth. In its lowermost portions, however, the viscosity of the lithospheric mantle approaches the viscosity of the underlying asthenosphere. Parsons and McKenzie (1978) suggested that the episodic cooling and dropping off of the lowermost lithosphere and its replacement by hot asthenosphere provide the mechanism whereby a constant heat flux is maintained through old oceanic lithosphere, and it seems entirely plausible that this mechanism also operates at the base of the continental lithosphere (Sclater and others, 1981).

Parsons and McKenzie (1978) used the term "thermal boundary layer" for the portion of the lithosphere that, though fluid at geological timescales, transfers heat primarily by conduction and is thus, thermally, part of the lithosphere. They argued that in a lithosphere with a total thickness of 130 km, the thermal boundary layer may be as much as 50 km thick. We employ that number for the purposes of illustration, but we should like to emphasize that there has been no measurement of this thickness yet. A thickness of 50 km seems likely to be close to an upper limit, because even at a thermal gradient as low as 5°C/km (equivalent to a heat flux of 15 mW m^{-2} through mantle rocks) such a thickness would encompass a temperature decrease of 250°C and, in consequence, an increase in viscosity by a couple of orders of magnitude for most plausible upper mantle rheologies.

The experimental results of Houseman, McKenzie, and Molnar (1981) show that the consequence of homogenously shortening continental lithosphere containing such a thermal boundary layer is the complete removal of that layer and its replacement by asthenosphere. Figure 1, redrawn from Houseman, McKenzie, and Molnar (1981, fig. 2) illustrates the development of this process in a numerical experiment. Several features of this process bear on the tectonics discussed in this paper. (1) There is a time lag between the horizontal shortening and the attainment of the large downward velocities that leads to the loss of the thermal boundary layer. (2) The duration of this time interval decreases with increasing Rayleigh number of the convecting fluid. At the lowest Rayleigh numbers investigated by Houseman and others, this interval is nearly 100 my. At the highest Rayleigh numbers they investigated (Ra $\sim 10^6$, which may still be lower than the Rayleigh number of the upper mantle), this interval is about 10 my. (3) Once the convective removal of the lower lithosphere reaches peak velocities, it proceeds to completion in a short time compared with the duration of the time lag in (1). At the end of this process, the thickness of the thermal boundary layer is close to, or less than, its thickness before the horizontal shortening occurred.

Replacement of the lowermost lithosphere by warmer and less dense upper mantle would cause a rapid increase in the surface height and potential energy of the continental column, in a time interval short compared with the delay between thickening and convective removal of the lowermost lithosphere. Since the latter interval may be less than 10 my, the change in surface height could take a few million years or less. This is a much shorter time than the duration of crustal thickening in most major orogenic belts. The shortening in Asia, for example, took most of the past 50 my, whereas convective removal of the lower lithosphere could take 10 my or less. This process is also fast compared with the other likely cause of regional density changes, the heating of the continental lithosphere by thermal diffusion. In the absence of convective removal, the lithosphere in regions of convergence may well be 200 km thick or more, with a thermal time constant of several hundred million years. A testable prediction of the hypothesis of convective removal of lithosphere is therefore that it will produce a rapid increase in surface elevation a few tens of million years after the onset of convergence.

In the remainder of this paper, we consider the mechanical and thermal consequences of this process and discuss how geological observations may be used to investigate it.

MECHANICAL CONSEQUENCES OF CONVECTIVE REMOVAL OF LITHOSPHERE

The amount by which the surface height changes as a result of convective removal of lower lithosphere depends upon the (unknown) fraction of the lithosphere that is involved in this process. For the purposes of this paper we make use of the estimates of England and Houseman (1989) who indicate that surface height increases in the range

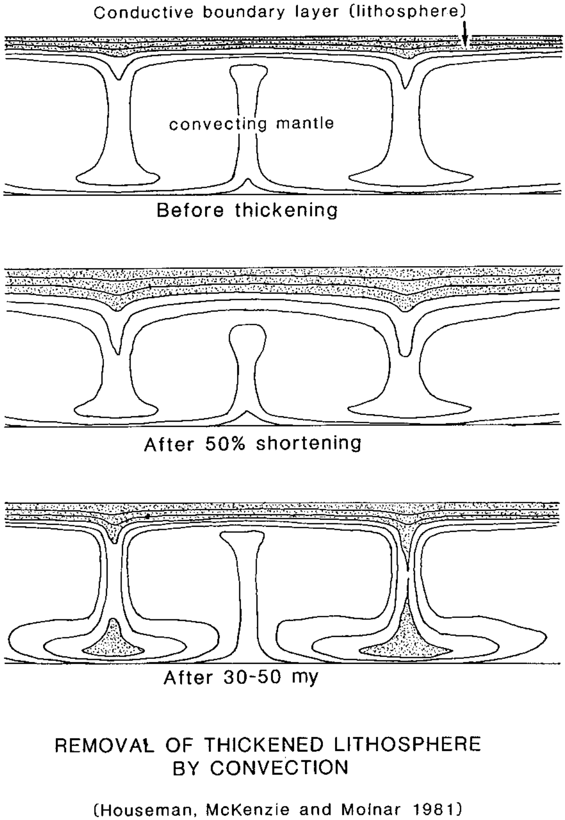

Fig. 1. Isotherms from Houseman and others (1981, fig. 2), from a two dimensional calculation of convection in a Newtonian viscous fluid whose upper part has been rapidly thickened by shortening perpendicular to the page, approximating the vertical thickening of uppermost mantle during mountain building. Contour intervals are 250°C. Thickening of the upper thermal boundary layer to the convecting mantle increases the negative buoyancy of the downwelling parts of the flow. After a delay, whose duration depends on the Rayleigh Number but which in this illustration is about 18 my, most of the remainder of the thermal boundary layer has been swept into the downwellings, to be replaced by hotter asthenosphere.

1 to 3 km are likely, if convective removal of the lower lithosphere reduces lithospheric thickness to its pre-shortening value.

Potential energy increases are difficult to assess for the same reasons, as height changes are unknown. Again, we use England and Houseman's (1989) calculations, which show that these potential energy increases (per unit surface area) amount to between 2 and 10×10^{12} Nm^{-1} for a plateau with crustal thickness of 60 to 70 km. This range is equivalent to an increase in tensional (or decrease in compressional) deviatoric stress of 10 to 50 MPa, if averaged through a load-bearing layer 100 km thick.

The degree of extension that occurs in response to such a change in the state of deviatoric stress will depend principally on the evolution of the boundary conditions, being the stresses applied by the surroundings (be they plates or other deforming regions) to the thickened lithosphere. If the continental lithosphere lies close to a steady thermal state before extension begins, then it will cool and strengthen during extension (England, 1983; Sonder and England, 1989). In all the cases we consider here, however, the geothermal gradient is depressed by thickening prior to extension, and hence temperatures rise during extension. Under these circumstances, the average viscosity of the lithosphere probably remains roughly constant or decreases during extension (England, 1987, fig. 7). It is likely, therefore, that externally applied boundary conditions, rather than internal increases in strength, provide the limits to extension of thickened continental lithosphere.

On a globe of constant radius, extension in the collisional region must be accommodated by shortening elsewhere, and unless the pattern of plate motions allows that motion to be taken up at a plate boundary, it will have to be absorbed by shortening on the boundaries of the extending region. The simplest case we can consider is that of an equant region of lithospheric thickening, beneath which the lithospheric root has been removed by convection and around which all plate motion has ceased. If this region subsequently extends, the resulting displacement will be taken up by an annular zone of shortening of the surrounding crust. Extension can, in principle, continue until the potential energy of the region equals that of the surrounding columns, although in practice this may be limited by the finite yield strength of the lithosphere.

The differences in potential energy, and therefore between the states of stress, of such an extending region and its surroundings are in some respects surprising and deserve closer investigation. A simple example suffices to show the principal features of this behavior. Let us consider a portion of continental lithosphere 125 km thick, containing crust with an average density of 2825 kgm^{-3}. Figure 2 shows the evolution of the potential energy and the surface height of this column of lithosphere as it undergoes the following sequence of strains:

A. horizontal shortening by a factor of two;
B. convective removal of the lower half of the mantle portion of the thickened lithosphere;
C. a large amount of horizontal extension.

The method of calculation follows England and Houseman (1989, app. A). Figure 2 illustrates the results for crustal thicknesses between 15 and 40 km.

Let us concentrate on the curve for an original crustal thickness of 35 km. The column starts with a surface height of 700 m and with its potential energy slightly lower than that of a column measured down to the same depth at a mid-ocean ridge (point A in fig. 2). Horizontal shortening by a factor of two increases the surface height to 3700 m but reduces the potential energy of the column by 10^{12} Nm^{-1}, equivalent to a

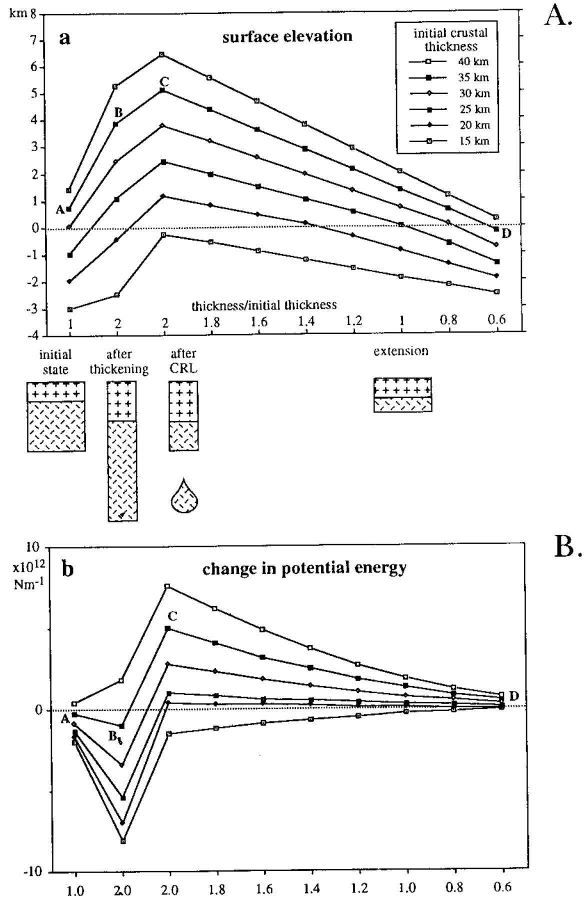

Fig. 2. Changes in surface elevation (A) and potential energy (B) as a result of thickening of lithosphere, convective removal of the lithospheric root (CRL), and subsequent extension. Units of potential energy are Newton meters per square meter of surface area. Points A, B, C, and D for a lithospheric column with an initial crustal thickness of 35 km are discussed in the text.

compressional deviatoric stress of 5 MPa averaged through a 100 km thick load-bearing layer (point B). Such a lithosphere is, therefore, prone to self-sustaining horizontal shortening of indefinite duration, as envisaged by Fleitout and Froidevaux (1982). (In fig. 2, the only lithospheric column that increases in potential energy as a result of thickening is the one with an initial crustal thickness of 40 km). Convective removal of the lower lithosphere, however, produces an increase in surface height of 1200 m and an increase in potential energy of 6×10^{12} Nm^{-1} (point C), corresponding to an extensional deviatoric stress of 30 MPa. Thereafter we imagine the lithosphere to extend horizontally until its crustal thickness is 20 km and its surface lies 200 m below sealevel. Note that although the surface height decreases during the extensional history, at no time does the potential energy of the extending column drop below the potential energy of the original, undeformed lithosphere. Thus, even when its surface lies below sealevel (point D) this extended region is capable of exerting compressional deviatoric stress on surrounding lithosphere whose surface height is greater. Indeed, as figure 2 shows, it is entirely possible that extended lithosphere, whose lower portion has been convectively removed and whose surface lies below sealevel, may exert a compressional deviatoric stress on surrounding lithosphere that retains its negatively buoyant root yet whose surface lies several kilometers above sealevel. This simple result has been emphasized by other authors (Le Pichon, 1983) but is not as widely appreciated as it deserves to be. In the last section of this paper we discuss the Alboran Sea, where extension of continental lithosphere while it lay below sealevel may be explained by this process.

The situation becomes more complicated if the relative motion of the plates bounding the thickened region adds a compressional stress during and after convective removal of the lower lithosphere, as appears to be the case in Tibet. The rate of extension in the direction of plate convergence will be governed by the excess in potential energy over that which can be supported by the compressional stress difference due to the convergence, whereas the rate of extension normal to the direction of plate motion will be governed by the excess over that of the surrounding crust. A higher rate of extension normal to the direction of plate convergence is therefore likely. This effect is seen quite clearly in the active extension of Tibet, where the dominant direction of extension has been normal to the India/Asia convergence direction since the early Pliocene (Armijo and others, 1986). The shape of the extending region may also influence the rate and direction of extension, as gradients of deviatoric stress parallel to the boundaries of elongate regions are generally smaller than gradients perpendicular to the boundaries, and strain rates parallel to the boundaries are correspondingly smaller.

IGNEOUS CONSEQUENCES OF CONVECTIVE REMOVAL OF LITHOSPHERE

The removal of the lower part of the continental lithosphere and its replacement by asthenospheric mantle has inevitable thermal consequences for the overlying crust and mantle. We distinguish here between the magmatic consequences, which may be immediate, and the metamorphic implications for the crust, which are likely to be observable only if convective removal of lithosphere leads to significant extension.

The igneous consequences of continental extension have been discussed by McKenzie and Bickle (1988) and by White and McKenzie (1989). These workers mainly discuss the melting that would occur in dry upper mantle as it is adiabatically decompressed during extension of the overlying lithosphere. McKenzie and Bickle show that large volumes of melt may be produced by extension, but that melting occurs only after the continental lithosphere has been reduced in thickness to 50 km or less.

It is somewhat surprising, therefore, to find that large outpourings of lava occurred before, or at the very onset of, extension in the eastern Basin and Range Province (Axen, Taylor, and Bartley, 1992; Gans, Mahood, and Schermer, 1989), and that there is considerable igneous activity pre-dating and contemporaneous with the beginning of active extension in Tibet and Anatolia (Arnaud and others, 1992; Pearce and Mei, 1988; Pearce and others, 1990).

A recent proposal by McKenzie (1989) suggests a mechanism for this igneous activity. He argues that small degrees of partial melt are constantly present beneath the lithosphere and that, even if they are present in fractions as low as 10^{-5}, these will separate from their matrix and percolate upward until they cool below their solidus to form a metasomatic layer within the lithosphere. Since these melts are likely to be rich in carbonates and water, their solidus lies at a considerably lower temperature than the dry solidus for peridotite. If, as McKenzie argues, this metasomatic layer lies above the thermal boundary layer that forms the lowermost part of the lithosphere, it does not participate in convection and remains isotopically distinct until some tectonic activity causes it to melt. McKenzie shows that such a metasomatic layer could be partially remelted to produce large volumes of high-potassium basic melt either by adiabatic decompression during lithospheric extension or by conductive heating of the lower lithosphere by a plume whose temperature lay a couple of hundred degrees above the average for the asthenosphere. We suggest that it would also melt if convective removal of the thermal boundary layer took place, as envisaged by Houseman, McKenzie, and Molnar (1981) and discussed in this paper. Removal of the thermal boundary layer would juxtapose asthenosphere material against much cooler material of the mechanically strong portion of the lithosphere. As McKenzie (1989) points out, the metasomatic layer within the mechanical boundary layer would lie close to its solidus and would readily melt if heated slightly. The temperature of the asthenosphere lies close to 1300°C, whereas the solidus of the metasomatic layer may lie at about 900°C. The convective removal of much, if not most, of the material separating these two isotherms in the continental lithosphere (fig 1) would surely lead to rapid partial melting of the metasomatic layer.

We should therefore expect that, in addition to a significant increase in surface height, convective removal of the lower lithosphere would cause the generation of appreciable quantities of melt before significant extension occurs. These melts, formed within the metasomatic layer inside the lithosphere, would be chemically and isotopically distinct from melts formed by the adiabatic decompression of asthenospheric upper mantle: in particular, they would be highly potassic (McKenzie, 1989). In addition, the composition of such melts could well be modified by assimilation of crustal material as they make their way to the surface through thick hot crust (Gans, Mahood, and Schermer, 1989). Subsequent extension of the lithosphere would cause adiabatic decompression of the asthenosphere and, hence, melting in the fashion investigated by McKenzie and Bickle (1988). Substantial extensional strain would be required, however, for this asthenospheric melting to take place (see McKenzie and Bickle, 1988).

METAMORPHIC CONSEQUENCES OF CONVECTIVE REMOVAL OF LITHOSPHERE

The upwelling of hot asthenosphere to replace the sinking lowermost lithosphere would sharply raise the temperature at the base of the remaining lithosphere. This change in boundary condition would send a wave of temperature increase through the overlying rocks. The first result of this temperature increase is likely to be the partial melting discussed in the previous section, but if the temperature rises at the base of the mechanical boundary layer by several hundred degrees as suggested above, then the thermal pulse will also contribute to metamorphism in the thickened crust.

Several studies of extension in the Basin and Range province have concluded that extension by a factor of two or more occurred, at particular locations, within a few million years. In contrast, if the present regional rate of extension in Tibet ($\sim 5 \times 10^{-9}$/yr; Molnar and Deng, 1984) were to continue, it would require over one hundred million years to thin the crust there to half its present thickness. The latter time interval is comparable with the time for thermal relaxation of lithosphere thickened during continental convergence (England and Thompson, 1984; Oxburgh and Turcotte, 1974). We should expect, therefore, that if extension were to take place over such a time interval, the heating due to the change in condition at the base of the lithosphere would be indistinguishable from prograde metamorphism occurring in the absence of this change in basal condition (see also, England and Thompson, 1984, app. B; Thompson and England, 1984, sec. 3.1). If, however, the duration of extension is much less than the time for thermal relaxation of the lithosphere, the influence of the change in lower boundary condition may not be felt until after the major tectonic activity has ceased. The result of this thermal history would be that the latest phase of metamorphism would be one of isobaric heating, followed by isobaric cooling.

To illustrate this argument, we present a series of one-dimensional calculations of the temperature distribution in a simple model lithosphere subjected to thickening, followed by convective removal of the thermal boundary layer, and then immediately by thinning.

The essential physics of the problem can be described as follows. We assume an equilibrium thermal gradient in lithosphere of initial thickness L_0 (fig. 3A). Thickening the lithosphere rapidly by a factor S_1 reduces the gradient by the same factor (fig. 3B). Reduced gradients such as this are associated with high-P/T ratio metamorphism in areas of continental convergence (Oxburgh and Turcotte, 1974). Note that the thermal time constant for re-establishment of a normal thermal gradient in this thick-

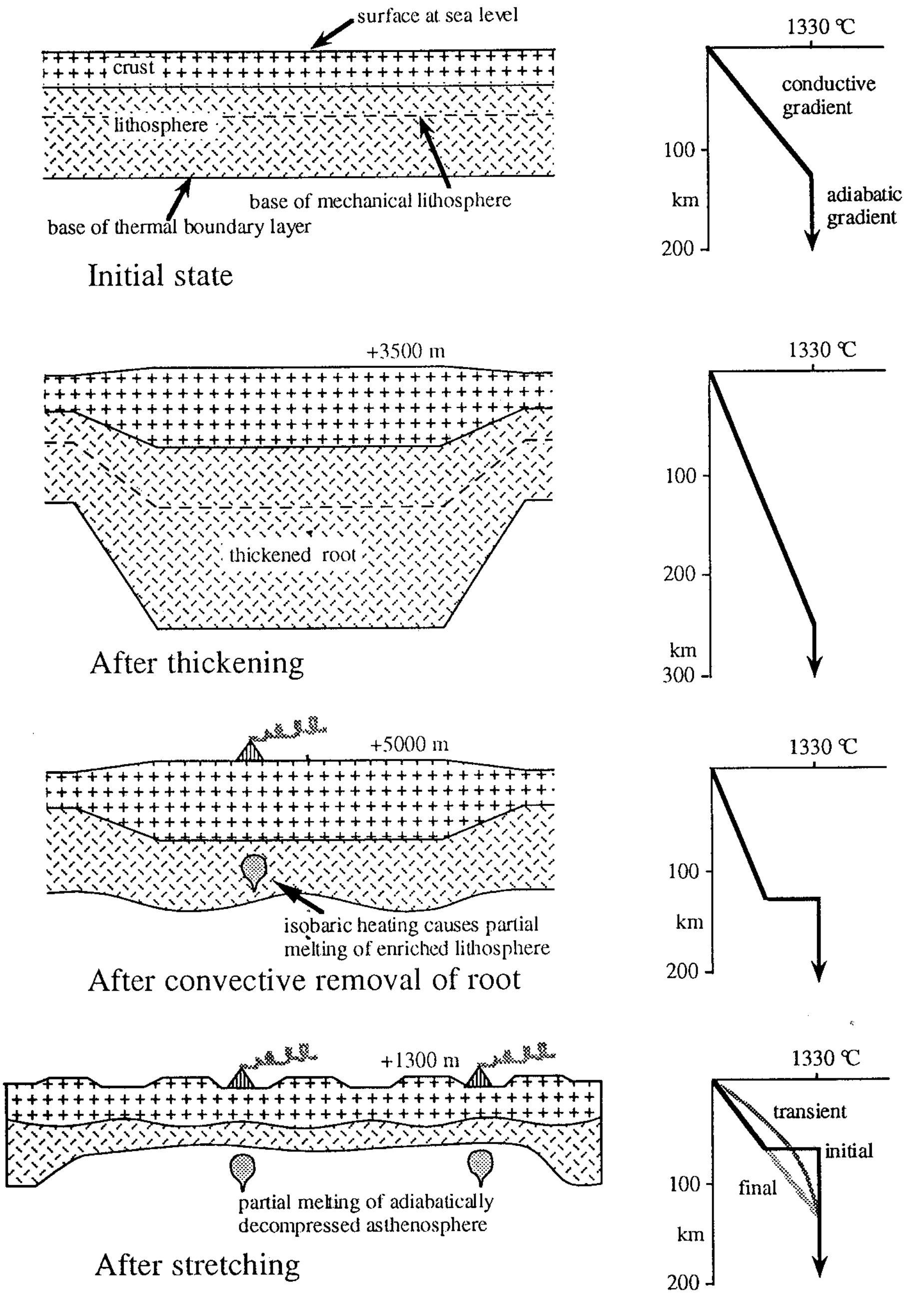

Fig. 3. Thermal effects of thickening lithosphere, convective removal of the lithospheric root, and subsequent extension. Diagrams on the left show the geometrical configuration of the lithosphere at each stage and the possible sources of magma; those on the right show the thermal gradients in simplified form, with the same vertical scale.

ened lithosphere is very large: about 240 my for a vertical stretch of 2 (England and Thompson, 1984). Before appreciable thermal relaxation can occur, the lower part of the thickened lithosphere is removed by convection. A step in the thermal gradient is introduced, as asthenospheric mantle rises to some level, L_2, above the base of the previously thickened lithosphere (fig. 3C). Diffusive decay of this step in the geotherm would eventually restore the thermal gradient in the lithosphere to its equilibrium value, but the thermal time constant for this is still fairly large (about 60 my if $L_2 = L_0 = 120$ km). We have argued that such thinning of the lithosphere is accompanied by a large increase in its potential energy, and that extension of the lithosphere would follow. For our calculations, we suppose that thinning begins immediately.

Thinning by a factor S_3 increases the thermal gradient by the same factor (fig. 3D) and brings the step in the geotherm closer to the surface. The result is a transient increase in temperature of the overlying lithosphere, as the thermal gradient decays back to its initial (equilibrium) value, with a time constant of about 15 my if $L_3 = 60$ km. It is this transient increase in the model that produces distinctive thermal histories that can be tested against observation.

With the aim of clarifying the underlying process, we have deliberately kept the model simple and have resisted the temptation to build in a large number of potentially interesting variables. We illustrate the results of a single set of calculations, but the simplicity of the model permits us to draw fairly general conclusions.

In the calculations illustrated in figures 4 and 5, the starting value for the lithosphere thickness (L_0) is 125 km. After thickening the lithosphere has thickness $L_1 = 250$ km. This thickness is reduced to $L_2 = 125$ km by convective removal of the lower lithosphere, then to $L_3 = 62.5$ km by extension. We take no account of erosion or sedimentation. Thus at the end of the extension, the crust is at its initial thickness, while the lithosphere is at half its original thickness. Four calculations are illustrated, which differ only in the duration of the extensional phase, which lasts either 3, 10, 30, or 100 my.

Figure 4 displays pressure-temperature (PT) paths for rocks initially at depths of 1/4, 1/8, and 1/16 L_0, which are buried to twice their original depths, and then returned by extension to their original depths. Figure 5 shows the temperature time paths for these same rocks. In each case the thickening phase is treated as instantaneous and is not illustrated.

We assume an initially linear geotherm and make no attempt to generate "realistic" temperatures by the appropriate choice of thermal conductivity or radioactive heating. Accordingly, the temperatures and pressures in figure 4 are scaled to their values at the base of the lithosphere. Thus the shapes of the PTt paths are the issue here, rather than any individual value of (dimensional) pressure and temperature along them.

All PTt paths, except those for extension of duration 3 my, show heating during extension (decompression). The amount of this heating is

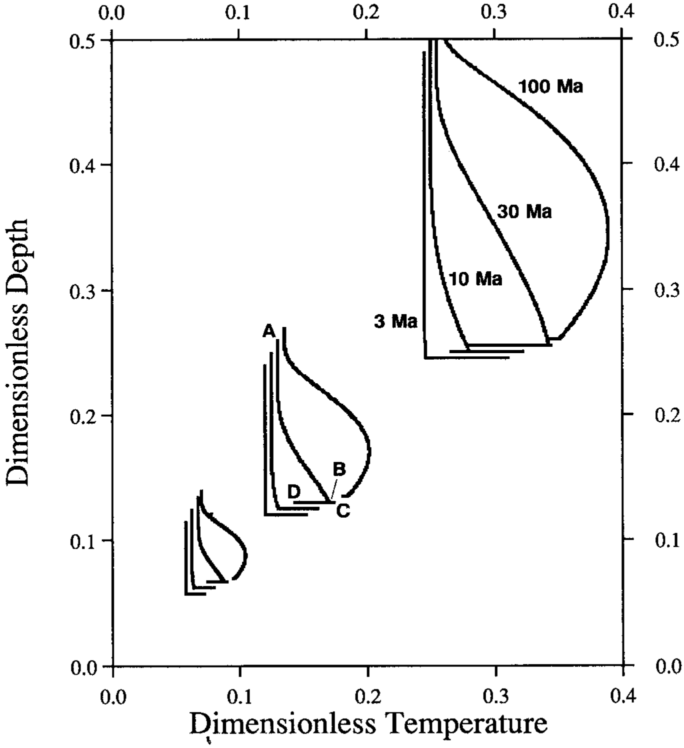

Fig. 4. Pressure-temperature (PT) paths for calculations described in the text. A model lithosphere is thickened homogenously and instantaneously by a factor of two from 125 km to 250 km in thickness, and its lower half is then replaced by asthenosphere. The lithosphere is then thinned by a factor of two so that its crust returns to its original thickness (fig. 3). Paths are shown for three sets of rocks, originally buried at depths of 1/4, 1/8, and 1/16 of the thickness of the lithosphere, and for four durations of extension, 3, 10, 30, and 100 my, as indicated for the deepest group of PT paths. Each PT path begins at its greatest depth and at the beginning of the phase of extension (point A indicated on one of the paths). Point B marks the end of extension, point C the end of heating, and point D is the end of the calculation. The PT paths for the different durations of extension are slightly offset vertically from one another in order to show the differences between their post-extensional forms (see also fig. 5).

greater for more deeply buried rocks than for shallower rocks and is greater for longer durations of extension. All PTt paths, except for those for extension lasting 100 my, show a phase of isobaric heating after the end of extension. When extension lasts 30 my, the heating during this phase is small in comparison to that experienced during decompression, but in the two faster cases, heating during this phase greatly exceeds heating during decompression.

We may interpret these results in a simple fashion by recognizing that the heating of the lithosphere during or after extension results from the diffusion of heat from the "knee" in the temperature profile shown in figure 3D. (Recall that, in a medium not generating heat internally,

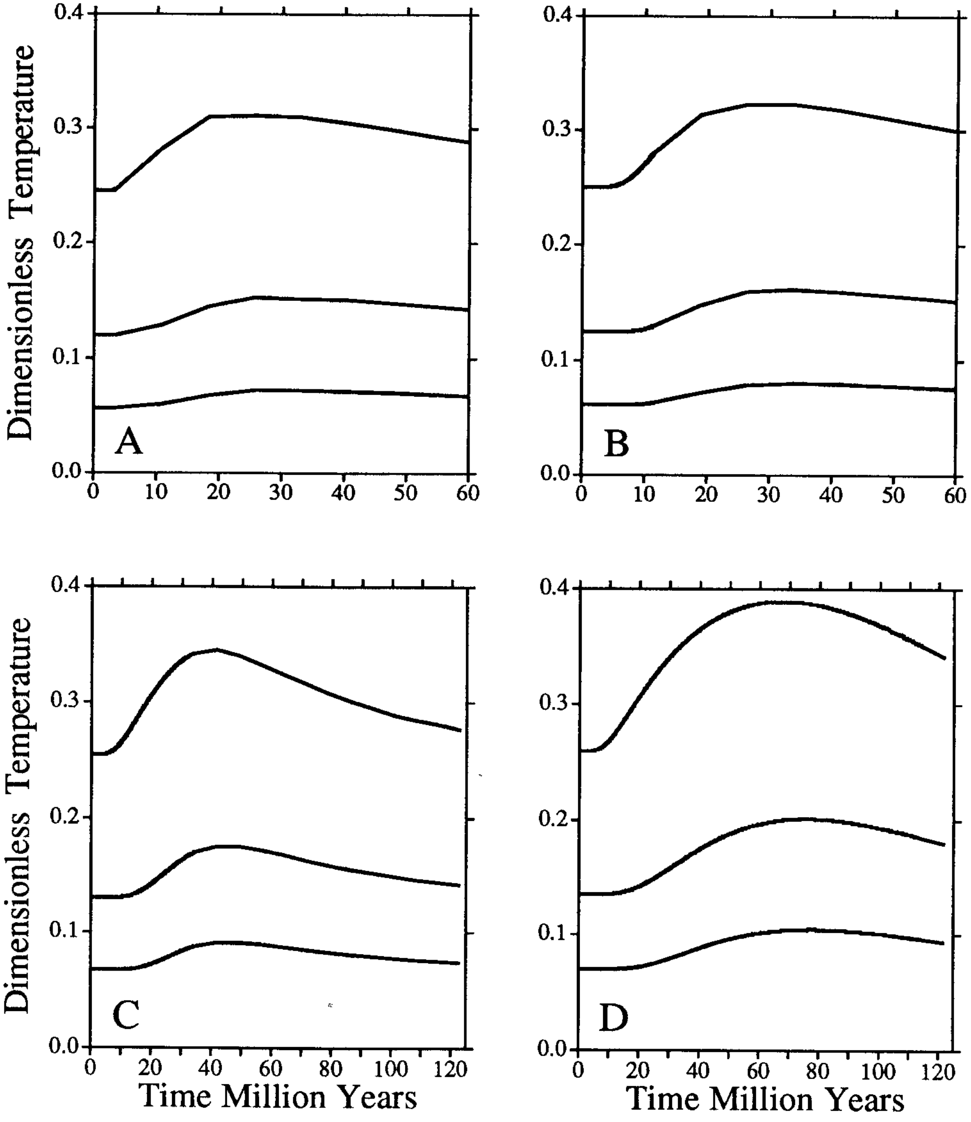

Fig. 5. Temperature-time paths for the calculations illustrated in figure 4. Duration of extension is 3 my (A), 10 my (B), 30 my (C), and 100 my (D). In each figure the upper (middle, lower) curve is for a rock initially buried at a depth of 1/4 (1/8, 1/16) the lithospheric thickness.

temperature changes due to diffusion occur only where there is curvature in the geotherm). There are three obvious timescales associated with the development of metamorphism during the extension we discuss here. The first is the duration of the extension itself, t_e. The other two timescales are those associated with thermal diffusion through the lithosphere at its pre-extensional thickness:

$$t_1 = L_2^2/\pi^2\kappa \tag{2}$$

and with thermal diffusion through lithosphere at its final thickness:

$$t_2 = L_3^2/\pi^2\kappa \tag{3}$$

If the duration of extension, t_e, is less than the smaller of these diffusional timescales, t_2, the extension must be essentially isothermal, and the heating pulse will pass through the crust after the end of extension. If, however, t_e exceeds t_1, most of the heating will occur during extension, and there will be little temperature change in the crust after the end of tectonic activity. This behavior is seen in the temperature-time paths of figure 5.

For the values we have chosen ($L_1 = 125$ km, $L_3 = 62.5$ km, $\kappa = 8 \times 10^{-7}$ m^{2}s^{-1}), $t_1 \sim 60$ my and $t_2 \sim 15$ my. For the shortest duration of extension we illustrate (3 my, fig. 5A) there is, indeed, no appreciable change in temperature during extension, and peak temperatures are reached 15 to 20 my after the end of extension; that is, after a time interval approximately equal to t_2. Similarly, when the duration of extension is 10 my (fig. 5B), peak temperatures are attained at about 25 my after the beginning of extension, which is again roughly comparable to t_2 after its end. In figure 5C the duration of extension, 30 my, is intermediate between t_1 and t_2, thus diffusion and extension operate over roughly the same timescales, and the peak is reached just before the end of extension. Finally, in figure 5D, we illustrate extension taking 100 my, where $t_e > t_1$. Here peak temperatures are attained at roughly 60 my ($\sim t_1$), and temperatures are falling at the end of extension.

Closer inspection of figure 5 shows that, in each case, the more deeply buried rocks experience their peak temperatures before the more shallow ones. This effect is to be expected because the temperature wave propagates upward from the base of the lithosphere. It would, in principle, be possible to define a time scale for the heating of each level in the crust, just like those in eqs (2) and (3) above, but with the appropriate length scale being the distance from that level to the base of the lithosphere. Indeed for the simple experiments we present here, such an analysis yields a very satisfactory approximation to the calculated results. We do not present this analysis, however, because we regard it as an over-elaboration in view of the likely contributions to the thermal development of real orogenic belts from radioactive heating. Radioactive heating, being generally concentrated toward the top of continental crust, heats the upper layers more rapidly than the deeper lithosphere, to some extent compensating for the more rapid heating of the lower layers by diffusion from the base of the lithosphere. Additionally, the influence of any erosion during metamorphism would be to cool shallower rocks before deeper ones (England and Richardson, 1977). Thus it, would be difficult, or impossible, to use variations with depth of the timing of the metamorphic peak to say anything definite about the tectonic processes that caused it.

Variations with depth of the timing of metamorphism are, however, second order features of the process we investigate here. The first order effect is relative timing of metamorphism and the extension. If extension of the lithosphere happens rapidly without convective removal of the lower lithosphere, PTt paths may be recorded that are indicative of large amounts of isothermal decompression (England and Thompson 1984; Thompson and England, 1984), but at the end of extension, the rocks should cool. In contrast, if convective removal of lower lithosphere precedes rapid extension, the rocks will heat up after the end of extension. That fact that rapid extension following convective removal of the lower lithosphere can cause a phase of post-tectonic isobaric heating is important to the tectonic interpretation of metamorphic PTt paths. Isobaric heating has previously been interpreted as implying important magmatic contribution to the heat budget (England and Richardson, 1977; Thompson and England, 1984).

The pressure-temperature curves in figure 4 should only be compared with observational data with considerable circumspection. First, we have ignored radiogenic heating, which will increase overall temperatures somewhat, particularly for lower extension rates. Second, we have assumed that convective removal of the lower lithosphere does not occur until thickening has ceased, and that thinning then starts immediately. This process may, in fact, start during thickening, and this will cause some heating at depth before extension starts. (Note that Looseveld, 1989 proposed that convective removal of the lower lithosphere during lithospheric thickening, but without subsequent extension, could be responsible for anti-clockwise P-T-t paths). These two effects would tend to diminish the initial increment of isobaric decompression shown by many of the calculated curves. Thirdly, we have ignored the effects of erosion or sedimentation, which could cause changes in pressure independently of the effects of thickening and thinning. Finally, we have ignored the effects of faulting in the upper crust, which by juxtaposing rocks of differing temperature, will cause local and very short-lived thermal perturbations.

GEOLOGICAL EXAMPLES

In the preceding sections we marshalled arguments that support the contention of Houseman, McKenzie, and Molnar (1981) that convective removal of lower lithosphere may play an important role in the tectonics of mountain belts. We then investigated the likely geologically observable consequences of such a process. In this section we review observations from several different settings consistent with convective removal of lower lithosphere. We wish to make it clear that we are not trying to prove the hypothesis advanced in this paper. We simply want to show that there is enough supporting evidence to make it worthy of further investigation.

The onset of extension in the Tibetan Plateau.—The present thinning of the crust in Tibet has long been explained by the loads produced by the high surface elevation of the region (Molnar and Tapponier, 1978). This explanation is not, however, sufficient on its own. The extension is occurring because the potential energy of the plateau is greater than can be supported by the compressional stresses applied to it by its surroundings. Clearly, compressional stresses cannot build up a plateau of higher potential energy than they can support, so some change in the boundary conditions must have caused a relative increase in potential energy above the level that can be supported by the boundary conditions. Such a relative increase could have occurred either by a reduction in the compressional stresses or by some additional factor causing an increase in the potential energy of the plateau (England and Houseman 1988, 1989).

England and Houseman (1989) show that the necessary reduction in compressional stress could come from one, or a combination of two, processes. The north-south compressional stress acting on the region would have been reduced if India's northward motion had slowed down, but net extension would follow only if that motion had virtually ceased. (This is because the bulk rheology of the lithosphere is dominated by materials with power-law, plastic, or frictional behavior, for which the stress required to cause deformation is relatively insensitive to the rate of strain). Alternatively, an east-west extensional deviatoric stress applied at a distant boundary (the eastern margin of Asia) could raise extensional strain rates within the plateau.

It seems likely that the onset of widespread extension in the plateau and, by inference, the relative increase in potential energy took place in late Miocene or Pliocene time (Armijo and others, 1986). There was no detectable reduction in the rate of India's convergence with Asia around this time, so the removal of a north-south directed compressional stress seems implausible as an explanation for the extension—the more so since the principal extension direction in the plateau lies roughly east-west. The strain-rate fields to the west and east of the plateau appear, from the active faulting, to consist of a combination of strike-slip with shortening perpendicular to the margins of the plateau (Molnar and Lyon-Caen, 1989; Molnar and Tapponier, 1975; Molnar and Deng, 1984), so an externally applied east-west extensional deviatoric stress is also unlikely to be the cause of the extension.

Another mechanism that seems at first sight to be a plausible cause of extension is weakening of the plateau by heating during thermal relaxation. England and Houseman (1988) show that such weakening would have caused increased compression, and not extension, in the plateau. It thus seems extremely probable that the mechanical basis for the present extension of the plateau is a rapid increase in its potential energy, such as would result from convective removal of the lower lithosphere.

The most direct evidence of an increase in the plateau's potential energy would come from measurements of its surface height over the interval spanning the onset of extension. Such measurements have not been made. Several claims of rapid recent increase in the height of the plateau have been made, but all are based on incorrect inference (see England and Molnar, 1990; Molnar and England, 1990 for review, and England and Molnar 1991a, b; Hatfield, 1991; and Pinter and Keller, 1991 for further discussion of this and other claims of rapid recent uplift).

Relations among extension and volcanism in the Basin and Range.—The Basin and Range province of western North America is a region of broadly distributed continental extension imposed on the locus of substantial intracontinental shortening in Mesozoic to early Tertiary time. Extension probably began in Eocene to Oligocene time and has continued, episodically or continuously, to the present day. The region has also been the locus of substantial volcanic activity during the same period, although there has been considerable debate about the degree of correlation, in space and time, between extension and magmatism, and whether one can be regarded as the cause of the other (Christiansen and Lipman, 1972; Snyder, Dickinson, and Silberman, 1976; Gans, Mahood, and Schermer, 1989; Bartley and others, 1988; Best and Christiansen, 1991).

The magmatic and tectonic history of the province has been widely discussed in terms of plate boundary processes. Lipman, Christiansen, and Protska (1972), Christiansen and Lipman (1972), and Glazner and Bartley (1984) for example, suggested that the dominantly calc-alkaline mid-Tertiary volcanism was related to east-dipping subduction along the North American margin and that the change to a bimodal basalt-rhyolite suite in Neogene time, associated with regional extension, was a consequence of the establishment of a strike-slip regime. These concepts face some problems, notably (A) the extreme breadth of the region of volcanic activity and extension, and their distance from the active margin; (B) the well-documented southward sweep of volcanic activity (Stewart and Carlson, 1976), which does not obviously correlate with plate boundary processes; and (C) the fact that extension appears to have commenced in many areas before subduction ceased along the adjacent part of the plate margin (Wernicke and others, 1987; Gans, Mahood, and Schermer, 1989). As a result, there have been several recent proposals relating extension and volcanism in the Basin and Range to changes in the thermal structure and potential energy of the underlying crust and mantle. Factors that might have contributed to an increase in potential energy of the region include Laramide crustal thickening (Molnar and Chen, 1983; Coney and Harms, 1984; Sonder and others, 1987; Wernicke and others, 1987), thermal erosion by a mantle plume (Fitton, James, and Leeman, 1988), and the steepening or removal of a proposed subhorizontal slab of subducted oceanic lithosphere beneath the region (Best and Christiansen, 1991; Severinghaus and Atwater, 1990).

We suggest that convective removal of lithosphere is a viable alternative explanation for Basin and Range extension and volcanism, and that it has the advantage that it is a physically predictable process that lacks *ad hoc* assumptions specific to the region. It also leads to predictions that can be tested against observation. Laramide and earlier convergence in the Cordillera was distributed over a region 500 to more than 1000 km wide and is likely to have resulted in thickening of the entire continental lithosphere. As discussed earlier, convective removal of the thickened thermal boundary layer is likely to occur a few to a few tens of million years after thickening. It should cause: (A) partial melting of a metasomatically enriched layer within the lithosphere, if one exists; (B) a rapid increase in surface elevation; (C) an increase in potential energy that may drive extension; and (D) if extension occurs, a detectable syn- to post-extension thermal pulse in the crust. At present there appears to be insufficient published data to test the second and fourth of these predictions, but there does appear to be evidence in favor of the first.

We are not aware of a systematic study of the relative timing of the earliest volcanic activity and extension in the Basin and Range Province, but the syntheses of Gans, Mahood, and Schermer (1989, fig. 19), Best and Christiansen (1991, fig. 10), and Axen, Taylor, and Bartley (1992) show that the onset of volcanism either preceded or was synchronous with the start of extension in many parts of the Great Basin. This earliest phase of magmatism cannot therefore be attributed to decompression of asthenosphere consequent upon extension. Significant partial melting of upper mantle within or beneath lithosphere of normal thickness (greater than about 100 km) requires either that temperatures in the asthenosphere be elevated, perhaps by a plume (White and McKenzie, 1989), or that the solidus of the source region lies substantially below 1200°C (McKenzie, 1989). If the lithosphere contains a metasomatic layer in which volatile-rich partial melts from the asthenosphere have accumulated, then convective removal of the lowermost lithosphere would almost certainly raise temperatures within this layer sufficiently to remelt its volatile-rich component and produce potassium-rich melts. We note that Eocene potassic magmatism in an extensional setting is recognized in Montana (O'Brien, Irving, and McCallum, 1991) and southwestern Idaho (Norman and Meertzman, 1991), and Miller and Miller (1991) describe synchronous extension and potassic volcanism in the Old Woman mountains of California at ~23 to 19 Ma. All these authors attribute a significant role in the volcanism to the melting of a metasomatically enriched layer within the continental lithosphere; indeed Norman and Meertzman (1991) regard this layer as the sole source for the rocks they describe and suggest convective removal of the lower lithosphere as the cause of the melting. Similarly, Bradshaw, Hawkesworth, and Gallagher (1992) conclude that most of the basalts of the Colorado River Trough, including all those erupted before or at the onset of extension in early Miocene time, are derived from the melting of metasomatically enriched lithosphere.

Relations among extension and volcanism in eastern Anatolia and Tibet.—A late Cenozoic volcanic province covers a belt, about 900 km long and 350 km wide at its greatest, running from eastern Turkey to Northern Iran, These volcanic rocks were erupted between about 6 Ma and the present, following the onset of crustal thickening in the region which began no later than 12 Ma and perhaps 20 Ma or more ago (Pearce and others, 1990). Studies of strontium and neodymium isotopes and of trace element systematics indicate that the lavas come from a source within the lithosphere. Pearce and others (1990) argue that the volcanism is unconnected with subduction of oceanic lithosphere and that convective removal of the lower lithosphere is the simplest hypothesis that explains the geochemistry of the volcanic rocks, almost simultaneous initiation of volcanism over the entire region, and the relation of volcanism to extension in the region.

Arnaud and others (1992) make a similar suggestion for the cause of late Cenozoic volcanism in northern Tibet, though they also consider other mechanisms, such as the subduction of continental lithosphere beneath the northern edge of the plateau, to be plausible candidates.

Post-collisional thermal metamorphism in the Betic/Rif Arc.—The Internal Zones of the Betic Cordillera of southern Spain and the Rif mountains of Morocco show evidence for a late magmatic and thermal event roughly coeval with the onset of extension in the collisional mountain chain (Platt and Vissers, 1989). The region was the locus of convergence between Africa and Iberia during Early Tertiary time, and glaucophane schist and eclogite facies metamorphism to pressures of the order of 10 to 15 kb suggest that crustal thicknesses reached at least 50 km (Bakker and others, 1989; Tubia and Gil Ibarguchi, 1991). In early Miocene time the collisional edifice was cut by a set of major low-angle extensional faults, which rapidly exhumed the high-pressure metamorphic rocks and, locally, parts of the underlying subcontinental mantle, now exposed as the Ronda and Beni Bousera ultramafic complexes. Crustal thickness in the central part of the system has been reduced to as little as 15 km (Banda and others, 1983), and this area has subsided to form the Alboran Sea between Iberia and North Africa. Extension was accompanied by potassic, basaltic, and calc-alkaline volcanism (Torres-Roldán, Poli, and Peccerillo, 1986; Hernandez and others, 1987). The upper mantle beneath the whole region shows anomalously low P-wave velocities, characteristic of asthenospheric mantle (Banda and others, 1983). Two seismic events at 600 km depth beneath the region have been interpreted as occurring in a detached fragment of lithospheric mantle (Grimison and Chen, 1986).

Extension on the site of the Early Tertiary collisional orogen continued during Early and Middle Miocene time, while much of the region subsided below sealevel (Watts and others, in press). Africa-Iberia convergence continued during this period; this convergence, together with the extension in the Alboran Sea area, was accommodated by crustal shortening and thickening in the peripheral mountain belts (the External Zones of the Betic Cordillera and the Rif, which connect across the Gibraltar arc). The driving force for extension in the Alboran Sea must have been internally generated and was sufficient to overcome the compressional forces exerted by the bounding plates, as well as those produced by the topographic elevation of the surrounding mountain chains.

Radiometric dating provides clear evidence for a widespread thermal event in the Betic Cordillera between 25 and 16 Ma (Zeck and others, 1989, Monié and others, 1991, De Jong, 1991). High-grade gneiss and associated anatectic granitic dikes yield Rb/Sr ages clustering around 20 to 22 Ma, and palaeontological dating of the overlying Neogene sediments constrains their exhumation rate to 5 to 10 km/my (Zeck and others, 1992). Some of the highest grade rocks are spatially associated with the ultramafic complexes, which themselves show evidence for high temperature re-equililibration during exhumation from a considerable depth (Obata, 1980). Crustal rocks closest to the peridotite reached temperatures of the order of 800°C at pressures of 800 MPa or more and show assemblages characteristic of the granulite facies (Loomis, 1972; Westerhof, 1977). They subsequently experienced nearly isothermal decompression, so that early kyanite was progressively overprinted by sillimanite and then by andalusite. Metamorphism was synchronous with intense deformation and mylonitization in these rocks (Westerhof, 1977), which was presumably associated with the process of extension and exhumation (Torres-Roldán, 1981). Rocks showing somewhat similar metamorphic evolutions are distributed along the southern margin of the Betic Cordillera (Torres-Roldán, 1974; Westra, 1969). Elsewhere there are approximately coeval occurrences of rocks that were apparently never deeply buried and show a simple near isobaric heating path at low pressures, producing largely post-deformational porphyroblasts of cordierite, andalusite, and biotite (Westerhof 1977; Torres-Roldán 1981; Akkerman, Maier, and Simon, 1980; De Jong 1991).

The contrast between the history of the originally more deep seated rocks, which show evidence for a thermal event during exhumation-related deformation, and the higher level rocks where the thermal pulse is largely post-deformational, accords qualitatively with our predictions of the thermal effects of rapid thinning of the lithosphere by convective removal of the thermal boundary layer followed by extension (fig. 4). We suggest that this area could be suitable for a quantitative test of these predictions.

DISCUSSION AND CONCLUSIONS

The evidence for late orogenic extension in both ancient and currently active collisional mountain belts is strong. The most widely accepted explanation for the phenomenon is that it reflects changes in the relative magnitudes of the horizontal loads associated with plate convergence and the vertical loads produced by surface topography. Although one way to remove support from a mountain belt is to stop the convergence that caused it, there are several regions, whose tectonics are summarized above, in which convergence continued during the extensional phase or in which no simple relation is seen between changes at the plate boundary and changes in tectonic style in the interior. For these regions, it is natural to look for an increase in the potential energy of the mountain belt as a possible cause of extension. Although they cause an increase in surface elevation, shortening and thickening of continental lithosphere are not certain to increase potential energy, because of the negative buoyancy of the mantle component of the lithosphere. Unless the crust is of the order of 40 km thick or more, the potential energy will be decreased by thickening, and convergence may therefore be a self-sustaining process.

Some form of convective removal of the thickened lithospheric root below regions of continental convergence is likely, and this will lead to substantial increases in the potential energy of the lithospheric column. This is one plausible cause of late orogenic extension, particularly in situations where other possible changes in boundary conditions, such as a decrease in the rate of convergence, can be eliminated.

Convective removal of the lower lithosphere has predictable consequences that provide possible tests for its occurrence. (1) Rapid replacement of lower lithosphere by asthenosphere would produce a rapid increase in surface height, which would coincide with the start of extension. (2) The step in the geotherm resulting from the juxtaposition of asthenospheric mantle with middle levels of the continental lithosphere is likely to cause partial melting and the production of high-K_2O magmas if the lithosphere has previously been enriched by metasomatic processes. This distinctive type of volcanism would precede extension. (3) Perceptible thermal effects within the crustal part of the lithosphere are unlikely if there is no extension, because of the long thermal time constant of the system, but the combination of convective removal of the lithospheric root and subsequent extension reduce the thermal time constant and hence may allow a transient thermal pulse to migrate upward through the crust. If extension occurs on a timescale of a few tens of millions of years or less, the thermal pulse will be synchronous with extension at deep levels in the crust but will postdate extension in the upper crust, causing a static thermal metamorphic event.

An alternative mechanism has been suggested as a trigger for extension, at least in the Basin and Range province. The rapid steepening or removal of a sub-horizontal slab, imagined to have underlain the region during Late Mesozoic and Tertiary time, might have caused the closely associated volcanism and extension of the region (Best and Christiansen, 1991; Severinghaus and Atwater, 1990). The removal of a slab from beneath the overlying mechanical boundary layer of another plate could presumably cause the rapid increase in surface height, the extension, and the pre-extensional melting in the upper plate which we have discussed in the context of convective removal of the lower portion of a single piece of thickened continental lithosphere. These two mechanisms are, therefore, qualitatively similar in their likely outcomes and qualitatively different from the 'delamination' of the entire mantle portion of the lithosphere envisaged by Bird (1979) and Bird and Baumgardner (1981).

Such delamination would juxtapose asthenospheric mantle against the base of the crust and produce volumes and compositions of melt inconsistent with the observed volcanism (Pearce and others, 1990).

Despite their qualitative similarities, convective removal of the lower lithosphere and the steepening of a sub-horizontal slab are distinct kinematically (the first involves a single, thickened, piece of lithosphere, the other involves two separate plates), and physically (the first involves distributed flow in a ductile lithosphere, and the second involves the rigid-body motion of at least one plate). We should, in consequence, expect there to be important quantitative differences between the surface height histories, and perhaps between the states of stress, of regions subjected to these two mechanisms. One aim of this paper is to stimulate further investigations that could resolve this issue.

A critical measurement to make in any extended terrain is of the surface height before, during, and after extension. If the surface lay close to sealevel before and until the initiation of extension, then crustal thickening, associated thickening of the lithosphere, and convective removal of its lower portion, can all be confidently ruled out. If, on the other hand, surface heights before extension are consistent with thickened crust, it is reasonable, though not necessary, to assume concomitant thickening of the whole lithosphere. Subsequent convective removal of the lower lithosphere would cause an abrupt increase in elevation preceding or synchronous with extension. The pattern of such uplift, if produced by the distributed convective process envisaged in this paper, would differ markedly from the migration of a roughly linear trend that would presumably be associated with the sinking of a slab.

Acknowledgments

We are grateful to John Bartley and Leigh Royden for critical and helpful reviews.

References

Akkerman, J. H., Maier, G., and Simon, O. J., 1980, On the geology of the Alpujarride Complex in the western Sierra de las Estancias (Betic Cordilleras, SE Spain): Geologie en Mijnbouw, v. 59, p. 363–374.

Armijo, R., Tapponier, T., Mercier, J. L., and Han, T-L., 1986, Quaternary extension in Southern Tibet: field observations and tectonic implications: Journal of Geophysical Research, v. 91, p. 13802–13872.

Arnaud, N. O., Vidal, Ph., Tapponier, P., Matte, Ph., and Deng, W. M., 1992, The high K_2O volcanism of northwestern Tibet: Geochemistry and tectonic implications: Earth and Planetary Science Letters, v. 111, p. 351–367.

Axen, G. J., Taylor, W. J., and Bartley, J. M., 1993, Space-time patterns and tectonic controls of Tertiary extension and magmatism in the Great Basin of the western United States: Geological Society of America Bulletin, v. 105, p. 56–76.

Bakker, H. E., De Jong, K., Helmers, H., and Biermann, C. 1989, The geodynamic evolution of the internal zone of the Betic Cordilleras (south-east Spain): A model based on structural analysis and geothermobarometry: Journal of Metamorphic Geology, v. 7, p. 359–381.

Banda, E., Udías, A., Mueller, S., Mezcua, J., Boloix, M., Qallart, J., and Aparicio, A., 1983, Crustal structure beneath Spain from deep seismic sounding experiments: Physics of the Earth and Planetary Interiors, v. 31, p. 277–280.

Bartley, J. M., G. J. Axen, W. J. Taylor, and Fryxell, J. E., 1988, Cenozoic tectonics of a transect through eastern Nevada near 38°N latitude: in Weide, D. L. and Faber, M. L., editors, Cordilleran Section Fieldtrip Guidebook, 1–20: Boulder, Colorado, Geological Society of America.

Best, M. G., and Christiansen, E. H., 1991, Limited extension during peak Tertiary volcanism, Great Basin of Nevada and Utah: Journal of Geophysical Research, v. 96, p. 13,509–13,528.

Bird, P., 1979, Continental delamination and the Colorado plateau: Journal of Geophysical Research, v. 84, p. 7561–7571.

Bird, P., and Baumgardner J., 1981, Steady propagation of delamination events: Journal of Geophysical Research, v. 86, p. 4891–4903.

Bradshaw, T. K., Hawkesworth, C. J., and Gallagher, K., 1993, Basaltic volcanism in the southern Basin and Range: No role for a mantle plume: Earth andPlanetary Science Letters: v. 116, p. 45–62.

Braile, L. W., 1989, Crustal structure of the continental interior, *in* Pakiser, L. C. and Mooney, W. D., editors, Geophysical Framework of the Continental United States: Geological Society of America Memoir 172, p. 285–316.

Chen, W.-P., and Molnar, P., 1983, Focal depths of intra-continental and intraplate earthquakes and their implications for the thermal and mechanical properties of the lithosphere: Journal of Geophysical Research, v. 88, p. 4183–4214.

Christiansen, R. L., and Lipman, P. W., 1972, Cenozoic volcanism and plate-tectonic evolution of the western United States, II Late Cenozoic, Royal Society of London Philosophical Transactions, v. A271, 249–284.

Coney, P. J., and Harms, T. A., 1984, Cordilleran metamorphic core complexes: Cenozoic extensional relics of Mesozoic compression: Geology, v. 12, p. 550–554.

Dalmayrac, B., and Molnar, P. 1981, Parallel thrust and normal faulting in Peru and constraints on the state of stress: Earth and Planetary Science Letters, v. 55, p. 473–481.

De Jong K., ms, 1991, Tectono-metamorphic studies and radiometric dating in the Betic Cordilleras (SE Spain): Thesis, University of Utrecht, 204 p.

Dewey, J. F., Shackleton, R. M., Chengfa, C., and Yiyin, S., 1988, The tectonic evolution of the Tibetan plateau: Royal Society of London Philosophical Transactions, v. A327, p. 379–413.

England, P. C., 1983, Constraints on extension of continental lithosphere: Journal of Geophysical Research, v. 88, p. 1145–1152.

——— 1987, Diffuse continental deformation: length scales, rates and metamorphic evolution: Royal Society of London Philosophical Transactions, v. 327, p. 3–22.

England, P. C., and Houseman, G. A. 1986, Finite strain calculations of continental deformation. 2. Comparison with the India-Asia collision zone: Journal of Geophysical Research, v. 91, p. 3664–3676.

——— 1988, The mechanics of the Tibetan plateau: Royal Society of London Philosophical Transactions, v. A326, p. 301–319.

——— 1989, Extension during continental convergence, with application to the Tibetan plateau: Journal of Geophysical Research, v. 94, p. 17561–17579.

England, P. C., and Jackson, J. A., 1989, Active deformation of the continents: Annual Review of Earth Science: v. 17, p. 197–226.

England, P. C., and McKenzie, D. P., 1982, A thin viscous sheet model for continental deformation: Geophysical Journal of the Royal Astronomical Society, v. 70, p. 295–32.

——— 1983, Correction to: A thin viscous sheet model for continental deformation: Geophys. Journal of the Royal Astronomical Society, v. 73, p. 523–532.

England, P., and Molnar, P., 1990, Surface uplift, uplift of rocks, and the exhumation of rocks: Geology, v. 18, p. 1173–1177.

——— 1991a, Reply to comment on "Surface uplift, uplift of rocks, and the exhumation of rocks" by C. B. Hatfield: Geology, v. 19, p. 1052.

——— 1991b, Reply to comment on "Surface uplift, uplift of rocks, and the exhumation of rocks" by N. Pinter and E. A. Keller: Geology, v. 19, p. 1053–1054.

England, P., and Richardson, S. W., 1977, The influence of erosion upon the mineral facies of rock from different metamorphic environments: Journal of the Geological Society of London, v. 134, p. 201–213.

England, P. C., and Thompson, A. B., 1984, Pressure-temperature-time paths of regional metamorphism I. Heat transfer during the evolution of regions of thickened crust: Journal of Petrology, v. 25, p. 894–92.

Fitton, J. G., James, D., and Leeman, W. P., 1991, Basic magmatism associated with late Cenozoic extension in the western United States: compositional variations in space and time: Journal of Geophysical Research, v. 96, p. 13693–13711.
Fleitout, L., and Froidevaux, C., 1982, Tectonics and topography for a lithosphere containing density heterogeneities: Tectonics, v. 1, p. 21–56.
Frank, F. C., 1972, Plate tectonics, the analogy with glacier flow and isostasy: *in* Heard, H. C., Borg, I. Y., Carter, N. I., and Raleigh C. B., editors, Flow and Fracture of Rocks: Washington, D.C., American Geophysical Union, Geophysical Monograph Series 16, p. 285–292.
Gans, P. B., Mahood, G. A., and Schermer E., 1989, Syn-extensional magmatism in the Basin and Range province; a case study from the eastern Great Basin: Geological Society of America Special Paper 233, p. 1-53.
Glazner, A. F., and Bartley, J. M., 1984, Timing and tectonic setting of Tertiary low-angle normal faulting and associated magmatism in the southwestern United States: Tectonics, v. 3, p. 385–396.
Grimison, N. L., and Chen, W. P., 1986, The Azores-Gibraltar plate boundary: focal mechanisms, depths of earthquakes, and their tectonic implications: Journal of Geophysical Research, v. 91, p. 2029–2047.
Hatfield, C. B., 1991, Comment on "Surface uplift, uplift of rocks, and exhumation of rocks": Geology, v. 19, p. 1051.
Hernandez, J., de Larouzière, F. D., Bolze, J., and Bordet, P., 1987, Le magmatisme néogène bético-rifain et le couloir de décrochement trans-Alboran: Société Géologique de France Bulletin, v. 3, p. 257–267.
Houseman, G. A., and England, P. C., 1986, Finite strain calculations of continental deformation. I. Method and general results for convergent zones: Journal of Geophysical Research, v. 91, p. 3651–3663.
Houseman, G. A., McKenzie, D. P., and Molnar, P., 1981, Convective instability of a thickened boundary layer and its relevance for the thermal evolution of continental convergence belts: Journal of Geophysical Research, v. 86, p. 6115–6132.
Le Pichon, X., 1983, Land-locked ocean basins and continental collision: the eastern Mediterranean as a case example, *in* Hsu, K. J., editor, Mountain Building Processes, London, Academic Press, p. 201–213.
Lipman, P. W., Christiansen, R. L., and Protska, H. J., 1971, Evolving subduction zones in the western United States, as interpreted from igneous rocks: Science, v. 143, p. 821–825.
Loomis, T. P. 1972, Contact metamorphism of pelitic rock by the Ronda ultramafic intrusion, southern Spain: Geological Society of America Bulletin, v. 83, p. 2449–2474.
Loosveld, R. J. H., 1989, The synchronism of crustal thickening and low-pressure facies metamorphism in the Mount Isa inlier, Australia. 2. Fast convective thinning of mantle lithosphere during crustal thickening: Tectonophysics, v. 165, p. 191–218.
McKenzie, D. P., 1989, Some remarks on the movement of small melt fractions in the mantle: Earth and Planetary Science Letter, v. 95, p. 53–72.
McKenzie, D. P., and Bickle, M. J., 1988. The volume and composition of melt generated by extension of the lithosphere: Journal of Petrology, v. 29, p. 625–667.
Meissner, R., 1986, The continental crust: A geophysical approach. Orlando, Florida, Academic Press.
Miller, J. S., and Miller, C. F., 1991, Tertiary extension-related volcanism, Old Woman Mountains area, eastern Mojave Desert, California: Journal of Geophysical Research, v. 96, p. 13,629–13,644.
Molnar, P., 1988, Continental tectonics in the aftermath of plate tectonics: Nature, v. 335, p. 131–137.
Molnar, P., and Chen, W-P., 1983, Focal depths and fault plane solutions of earthquakes under the Tibetan plateau: Journal of Geophysical Research, v. 88, p. 1180–1196.
Molnar, P., and Deng, Q.-D., 1984, Faulting associated with large earthquakes and the average rate of deformation in central and eastern Asia: Journal of Geophysical Research, v. 89, p. 6203–6228.
Molnar, P., and England, P. C., 1990, Late Cenozoic uplift of mountain ranges and global climatic change: chicken or egg?: Nature, v. 346, p. 29–34.
Molnar, P., and Lyon-Caen, H., 1988, Some simple physical aspects of the support, structure and evolution of mountain belts: in Clark, S. P., Jr, editor, Processes in continental and lithospheric deformation: Geological Society of America Special Paper 218.
——— 1989, Fault plane solutions of earthquakes and active tectonics of the Tibetan Plateau and its margins: Geophysical Journal International, v. 99, p. 123–153.

Molnar, P., and Tapponnier, P., 1975, Cenozoic tectonics of Asia: Effects of a continental collision: Science, v. 189, p. 419–426.
——— 1978, Active tectonics of Tibet: Journal of Geophysical Research, v. 83, p. 5361–5375.
Monié, P., Galindo-Zaldivar, J., Gonzalez Lodeiro, F., Goffé, B., and Jabaloy, A., 1991, 40Ar/39Ar geochronology of Alpine tectonism in the Betic Cordilleras (southern Spain): Journal of the Geological Society of London, v. 148, p. 289–297.
Norman, M. D., and Meertzman, S. A., 1991, Petrogenesis of Challis volcanics from central and southwestern Idaho: Trace element and Pb isotopic evidence: Journal of Geophysical Research, v. 96, p. 13,279–13,294.
Obata, M., 1980, The Ronda peridotite: garnet, spinel, and plagioclase lherzolite facies and the P-T trajectories of a high-temperature mantle intrusion: Journal of Petrology, v. 21, p. 533–572.
O'Brien, H. E., Irving, A. J., and McCallum, I. S., 1991, Eocene potassic magmatism in the Highwood Mountains, Montana, and tectonic models for the Montana Alkalic Province: Journal of Geophysical Research, v. 96, p. 13,261–13,278.
Oxburgh, E. R., and Turcotte, D. L., 1974, Thermal gradients and regional metamorphism in overthrust terrains with special reference to the Eastern Alps: Schweizer. Mineralogic Petrographic Mitterlunger, v. 54, p. 641–66.
Parsons, B., and Mckenzie, D., 1978, Mantle convection and the thermal structure of the plates, Journal of Geophysical Research, v. 83, p. 4485–4496.
Pearce, J. A., Bender, J. F., De Long, S. E., Kidd, W. S. F., Low, P. J., Güner, Y., Saroglu, F., Yilmaz, Y., Moorbath, S., and Mitchell, J. G., 1990, Genesis of collision volcanism in Eastern Anatolia, Turkey: Journal of Volcanology and Geothermal Research, v. 44, p. 189–229.
Pearce, J. A., and Mei, H., 1988. Volcanic rocks of the 1985 geotraverse: Lhasa to Golmud: Royal Society of London Philosophical Transactions, v. A327, p. 169–201.
Pinter, N. and Keller, E. A., 1991, Comment on "Surface uplift, uplift of rocks, and exhumation of rocks," Geology, v. 19, p. 1053.
Platt, J. P. and Vissers, R. L. M., 1989, Extensional collapse of thickened continental lithosphere: a working hypothesis for the Alboran Sea and the Gibraltar arc: Geology, v. 17, p. 540–543.
Sclater, J. G., Parson, B., and Jaupart, C., 1981, Oceans and continents: Similarities and differences in the mechanisms of heat transport: Journal of Geophysical Research, v. 86, p. 11535–11552.
Sébrier, M., Mercier, J. L., Megard, F., Laubacher, G., and Carey-Gailhardis, E., 1986, Quaternary normal and reverse faulting and the state of stress in the central Andes of Peru: Tectonics, v. 4, p. 739–780.
Severinghaus, J. and Atwater, T., 1990, Cenozoic geometry and thermal state of the subducting slabs beneath western North America, in Wernicke, B. P. (editor), Basin and Range extensional tectonics near the latitude of Las Vegas, Nevada: Geological Society of America Memoir 176, p. 1–22.
Snyder, W. S., Dickinson, W. R., and Silberman, M., 1976, Tectonic implications of space-time patterns of Cenozoic magmatism in the western United States: Earth and Planetary Science Letters, v. 32, p. 91–106.
Sonder, L. J., and England, P. C., 1989, Effects of a temperature-dependent rheology on large scale continental extension: Journal of Geophysical Research, v. 94, p. 7603–7619.
Sonder, L. J., England, P. C., Wernicke, B. P., and Christiansen, R. L., 1987, A physical model for Cenozoic extension of western North America, in Coward, M. P., Dewey, J. F., and Hancock, P. L., editors, Continental Extensional Tectonics: Geological Society of London Special Publications, v. 28, p. 187–201.
Stewart, J. H., and Carlson, J. E., 1976, Cenozoic rocks of Nevada; four maps and brief descriptions of distribution, lithology, age, and centers of volcanism: Nevada Bureau of Mines and Geology, Map 52, scale 1:1,000,000.
Suarez, G., Molnar, P., and Burchfiel, C., 1983, Seismicity, fault plane solutions, depth of faulting and active tectonics of the Andes of Peru, Ecuador and southern Columbia: Journal of Geophysical Research, v. 88, p. 10,403–10,428.
Tapponnier, P., and Molnar, P., 1976, Slip-line field theory and large-scale continental tectonics: Nature, v. 264, p. 319–324.
Tapponnier, P., Mercier, J. L., Armijo, R., Han, T., and Zhou, J., 1981, Field evidence for active normal faulting in Tibet: Nature, v. 294, p. 410–414.
Thompson, A. B., and England, P. C., 1984, Pressure-temperature-time paths of regional metamorphism II. Their inference and interpretation using mineral assemblages in metamorphic rocks: Journal of Petrology, v. 25, p. 929–955.

Torres-Roldán, R. L., 1974, El metamorfismo progresivo y la evolución de la serie de facies en las metapelitas alpujárrides al SE de la Sierra Almijara (sector central de las Cordilleras Béticas, Sur de España): Cuadernos de Geología, v. 5, p. 21–77.

——— 1981, Plurifacial metamorphic evolution of the Sierra Bermeja peridotite aureole (Southern Spain): Estudios Geológicos, v. 37, p. 115–133.

Torres-Roldán, R. L., Poli, G., and Peccerillo, A., 1986, An Early Miocene arc-tholeiitic magmatic dike event from the Alboran-Sea—Evidence for precollisional subduction and back-arc crustal extension in the westernmost Mediterranean: Geologische Rundschau, v. 75, p. 219–234.

Tubia, J. M., and Gil Ibarguchi, J. L., 1991, Eclogites of the Ojen nappe: a record of subduction in the Alpujárride complex (Betic Cordilleras, southern Spain). Journal of the Geological Society of London, v. 148, p. 801–805.

Vilotte, J. P., Madariaga, R., Dainieres, M., and Zienkiewicz, O. C., 1986, Numerical study of continental collision: Influence of buoyancy forces and a stiff inclusion: Royal Astronomical Society Geophysical Journal, v. 84, p. 279–310.

Watts, A. B., Platt, J. P., and Buhl, P., in press. Tectonic evolution of the Alboran Sea basin: Basin Research.

Wernicke, B. P., Christiansen, R. L., England, P. C., and Sonder, L. J., 1987, Tectonomagmatic evolution of Cenozoic extension in the North American Cordillera, in Coward, M. P., Dewey, J. F., and Hancock, P. L., editors, Continental Extensional Tectonics: Geological Society of London Special Publications, v. 28, p. 203–221.

Westerhof, A. B., 1977, On the contact relations of high temperature peridotites in the Serranía de Ronda, Southern Spain: Tectonophysics, v. 39, p. 579–592.

Westra, G., 1969, Petrogenesis of a composite metamorphic facies series in an intricate fault-zone in the southeastern Sierra Cabrera, S.E. Spain: GUA Series in Geology, University of Amsterdam, 166 p.

White, R., and McKenzie, D., 1989, Magmatism at rift zones: the generation of volcanic continental margins and flood basalts: Journal of Geophysical Research v. 94, p. 7685–7729.

Zeck, H. P., Albat, F., Hansen, B. T., Torres-Roldán, R. L., García-Casco, A., and Martín-Algarra, A., 1989, A 21 ± 2 Ma age for the termination of the ductile Alpine deformation in the internal zone of the Betic Cordilleras, S. Spain: Tectonophysics, v. 169, p. 215–220.

Zeck, H. P., Monié, P., Villa, I. M., and Hansen, B. T., 1992, Very high rates of cooling and uplift in the Alpine belt of the Betic Cordilleras, southern Spain: Geology, v. 20, p. 79–82.