Subduction of continental lithosphere: Some constraints and uncertainties

Peter Molnar, Dale Gray
Massachusetts Institute of Technology, Department of Earth and Planetary Sciences
Cambridge, Massachusetts 02139

ABSTRACT

Two effects that contribute to the subduction of continental lithosphere are the negative buoyancy of the relatively cold mantle part of continental lithosphere and the pull of a downgoing slab of oceanic lithosphere on continental lithosphere trailing behind it. The mantle part of the lithosphere could continuously subduct about 10 km of continental crust, if the upper and lower crust could be detached from one another. This estimate, however, could be in error by as much as a factor of three, given the range of plausible values for the requisite parameters. The second effect is more difficult to estimate because of our ignorance of what proportion of the gravitational force acting on the downgoing slab is transmitted to the surface lithosphere. For various assumptions, subducted oceanic lithosphere could be expected to pull a short (tens of kilometres) or long (hundreds of kilometres) length of intact continental crust into the asthenosphere. Although we favor subduction of only a short length of intact continental crust, we cannot prove that a large amount of crust is not subducted. These calculations are for continental margins that meet subduction zones flush. If continental crust can remain intact, peninsulas and microcontinents might be subducted completely.

INTRODUCTION

Because of its low density, continental crust has traditionally been assumed to be too light to be subducted. Even with the recognition of the important role of the lithosphere in subduction, earth scientists have assumed that the large thickness of continental crust provides too much buoyancy to continental lithosphere for it to be subducted (for example, Isaacs and others, 1968; McKenzie, 1969). Yet geologic evidence indicates large-scale crustal shortening by thrust faulting in most mountain belts, particularly where continents collide. Consequently, some subduction of continental crust must take place, and in general, geologic evidence places only lower bounds on the amount of thrust displacement and therefore of subduction of lower crust. For instance, Gansser (1966) estimated that the total displacement on the major thrust faults in the Himalayas is at least 300 km, excluding displacement at the Indus suture zone, where thousands of kilometres of oceanic lithosphere probably were subducted. Powell and Conaghan (1973), however, proposed as much as 1,000 km of slip on one of the Himalayan faults, the Main Central thrust. In the Norwegian Caledonides, Gee (1975) inferred at least 500 km of shortening distributed on several faults. These estimates do not necessarily require a large amount of subduction of the entire intact continental crust to depths of hundreds of kilometres in the asthenosphere. In all cases the thrust faults are exposed only within the crust and could be interpreted as detachment surfaces between the upper and lower crust. Nevertheless, a substantial amount of crustal material must be subducted if only to depths of 70 km, to account for the observed shortening.

A basic question is, If the buoyancy of continental crust does inhibit subduction of continental lithosphere, then how much crust can be subducted before convergence is halted? Resolution of this question will probably require further detailed geologic mapping of thrust belts and may require new techniques. At present, however, we can examine limits on how much crust could be subducted given basic assumptions about the physical processes and requisite parameters involved.

Two phenomena contribute to the subduction of continental lithosphere. First, the mantle part of the lithosphere is cold and therefore dense enough that its negative buoyancy would pull the bottom part of the crust down if the top of the crust could be detached from it. Second, the gravitational body force acting on the subducted oceanic lithosphere may also exert force on the leading edge of a continent following the oceanic lithosphere into the subduction zone.
GRAVITATIONAL STABILITY
OF CONTINENTAL LITHOSPHERE

Oceanic lithosphere plunges into the asthenosphere because it is
gravitationally unstable. It is colder and therefore more dense
than the asthenosphere would be if brought adiabatically to the
same depth (pressure). Although the crust is lighter than the mantle
and buoyant the lithosphere up, evidently oceanic crust is too thin
to prevent subduction. The effects of these two buoyancy forces
can be expressed simply as the difference between two terms, one
for the buoyancy of the crust and the other for the negative
buoyancy of the mantle part of the lithosphere (Fig. 1; McKenzie,
1969). For the crust, the buoyancy force per unit area inhibiting
subduction is \( (\rho_m - \rho_c) gh \), where \( \rho_m \) and \( \rho_c \) are the densities of
the mantle and crust, \( g \) is gravitational acceleration, and \( h \) is
the thickness of the crust. For the mantle part of the lithosphere,
the difference in the density between material with temperature
\( T(z) \), and that with temperature \( T_0(z) \), following the adiabatic gradient,
is simply \( \rho_m \alpha [T(z) - T_0(z)] \), where \( \alpha \) is the coefficient of
thermal expansion and \( z \) is the depth.

The temperature in the asthenosphere presumably closely
follows the adiabatic gradient (\( \approx 0.5^\circ/\text{km} \)). Because of the small
amount of radioactivity in the mantle, the geotherm in the lithosphere
is essentially linear (Fig. 1). If the base of the lithosphere
is defined by a particular isotherm, \( T_\alpha \), then in oceanic lithosphere,

\[
T(z) \approx \frac{T_\alpha}{a} z,
\]

where \( a \) is the thickness of the lithosphere. As the temperature in
only the mantle part of the lithosphere is important, a more precise
description of the temperature profile across it is

\[
T(z) = (T_\alpha - T_C) \left( \frac{z}{a} \right) + T_C,
\]

where \( T_C \) is the temperature at the base of the crust and \( h \) is the
depth of the Moho (Fig. 1). For continental lithosphere, equation
2 is slightly better than equation 1, because radioactivity in the
continental crust makes an important contribution to the geotherm.
The approximation of the geotherm in the mantle as two
essentially straight lines—one due to conduction through the
lithosphere and the other as an adiabat maintained by convection—is, of course, a simplification. More realistic geotherms,
discussed by Parsons and McKenzie (1978), however, are not
warranted here.

The negative buoyancy force per unit area due to the low
temperature of the lithosphere is \( \rho_m \alpha g \int_0^h [T_\alpha(z) - T(z)]dz \).
This can be written directly by recalling the formula for a triangle
(see Fig. 1): \( \frac{1}{2} \rho_m \alpha g [T_\alpha(h) - T(h)](a - h) \). Because the adiabatic
gradient is so small, \( T_\alpha(z) \) is nearly constant, so that \( T_\alpha(h) \)
\( = T_\alpha(a) = \) temperature of the asthenosphere.

For an instability to exist, the mantle part of the lithosphere
must be heavier than the crust, or

\[
(\rho_m - \rho_c) gh - \frac{1}{2} \rho_m \alpha [T_\alpha(a) - T(h)](a - h) < 0.
\]

Reasonable values for the various parameters are

- \( \rho_m - \rho_c = 3.35 - 2.8 = 0.55 \pm 0.1 \text{ g/cm}^3 \)
- \( \rho_m = 3.35 \pm 0.05 \text{ g/cm}^3 \)
- \( \alpha = 3 \pm 1 \times 10^{-5} \text{ °C}^{-1} \)
- \( T_\alpha = 1200 \pm 150 \text{ °C} \)
- \( T(h_C) = 500 \pm 100 \text{ °C} \) (for continents)
- \( a - h_C = 150 \pm 50 \text{ km} \) (for continents)
- \( a - h_0 = 100 \text{ km} \) (for old ocean floor)

For oceanic lithosphere, where \( h_o \ll a \approx 100 \text{ km}, T(h_o) \equiv 0 \)
° C, and \( \rho_c = 2.9 \text{ g/cm}^3 \), and instability exists when

\[
h_o < \frac{\rho_m \alpha [T_\alpha(a) - T(h_o)](a - h_o)}{2(\rho_m - \rho_o)} < 13.4 \text{ km}.
\]

This condition is easily fulfilled in oceanic regions.

Expression 3 allows an estimate of what fraction of the lower
continental crust, \( h_C \), would be unstable if the upper part could be
detached from it. Using the values given above, \( h_C < 10 \text{ km} \), approxi-
ately 30% of the crust. This estimate is very uncertain,
however. Note that if we choose the extreme values given above,
\( h_C < 2 \text{ km} \) or \( h_C < 29 \text{ km} \). In the former case, a small fraction
of the crust would be carried down, whereas in the latter nearly
the entire continental lithosphere would be unstable! These ex-
trême estimates arise from a conspiracy of extreme values for
each of the parameters given above, and therefore both estimates
are unlikely to apply in general. The calculations do show, how-
ever, that we cannot eliminate the possibility of subducting a large
fraction of continental crust, if it could be detached from the
upper part.

GRAVITATIONAL FORCE ACTING ON DOWNGOING
SLAB OF OCEANIC LITHOSPHERE

If a continental mass is part of an oceanic plate being
consumed at a subduction zone, gravitational body forces acting
on the excess mass in the slab might be able to pull continental
crust into the asthenosphere (see Fig. 2). Given plausible param-
eters that govern the various forces, we can place an upper bound
on what length (\( d \)) of continental crust could be subducted. There
are basically two questions here: What is the gravitational force
acting on the slab? How much of it actually pulls on the part
Figure 2. Gravitational force acting on oceanic lithosphere and pulling continental lithosphere into asthenosphere. Force is opposed by buoyancy of continental crust of length $d$.

of the lithosphere at the surface? Given our inability to answer the latter question, the uncertainty in the estimate of the amount of continental crust that can be subducted is very large. Nevertheless, we present the results for an assortment of assumptions and then state our prejudices about them.

The gravitational body force (per unit length of subduction zone) acting on the slab is due to the temperature of the slab being lower than that of the surrounding asthenosphere, and it can be estimated from calculations of the temperature in the slab, which depends primarily upon the subduction rate and the plate thickness (for example, McKenzie, 1969). Numerical calculations show that radioactivity in the oceanic crust and mantle does not contribute much to the temperature in the slab (Minear and Toksöz, 1970). The role of frictional heating remains controversial, but since we seek an upper bound on the amount of crust subducted, here we ignore its contributions both as a heat source and as a retarding force. Once a convergence stops, friction ceases to be either a heat source or a retarding force. We discuss phase changes below. The principal source that heats the slab is conduction from the surrounding asthenosphere (McKenzie, 1969; Minear and Toksöz, 1970).

McKenzie (1969) derived an analytic solution for the temperature in the slab and calculated the gravitational force acting on the slab due to its excess mass. This calculation is given in the Appendix 1, but it is extended to include different finite depths of penetration of the slab, and it is evaluated using more reasonable parameters for the plate thickness and surrounding temperatures than McKenzie used. From equation A5 and the parameters in the appendix, we obtain forces (per unit length along the arc) of $1.8 \times 10^{16}$ dyne/cm for a slab extending to a depth of 350 km (that is, 500 km long and dipping at 45°) and of $3.2 \times 10^{16}$ dyne/cm for a slab extending to a depth of 700 km (that is, 1,000 km long and dipping at 45°). For comparison, we note that Fitch (1977) integrated the force per unit length calculated by Schubert and others (1975) and obtained a maximum stress of 3.2 kb acting on a plate 100 km thick extending to 700 km depth, corresponding to a force per unit length of $3.2 \times 10^{16}$ dyne/cm. A crude estimate of the excess mass calculated by Toksöz and others (1973) yields a similar value.

The principal sources of error in all of these calculations are the values of the thermal expansion, $\alpha$, and the thicknesses of the plate, $a_0$ (because the force per unit length is proportional to $a_0^2$). Because of the uncertainties in the parameters inserted into these formulas, the resulting calculations could be in error by as much as a factor of two.

Another possible force is the excess mass in the slab due to the possible occurrence of the olivine-spinel phase change at a shallower depth in the slab than in the surrounding asthenosphere (for example, Richter, 1973; Schubert and others, 1975). Fitch (1977) integrated Schubert and others’ calculation to obtain an additional force of $1.9 \times 10^{16}$ dyne/km. There is a contrary suggestion, however, that the reaction is inhibited by its kinetics. In the cold central part of the slab the reaction would be so slow that this part of the slab would penetrate well below the equilibrium depth before the reaction would occur (Sung and Burns, 1976a, 1976b). In this case, it would not add a force pulling the slab down; instead, it would oppose subduction (Sung and Burns, 1976a, 1976b).

It is clear the the uncertainty in the gravitational body force acting on a downgoing slab is large. In addition, there are few data constraining how much of this force is transmitted to the surficial part of the lithosphere. Some excess mass in the slab will be supported by stress in the surrounding asthenosphere. For instance, the fault-plane solutions of deep earthquakes suggest that the excess mass of downgoing slabs below about 350 km does not pull the shallow part down (Isacks and Molnar, 1969, 1971). Perhaps the force acting only on the part shallower than 350 km pulls the surficial lithosphere. In fact, the possibility exists that the downgoing slabs do not exert a substantial force on the surface lithosphere. Kanamori (1971) suggested that at least some slabs break off at the trench and sink freely into the asthenosphere. As the rate that the lithosphere sinks into the asthenosphere is comparable to the rate at which it would sink into a fluid with the viscosity of the asthenosphere, Richter (1977) inferred a small force exerted by the slab on the rest of the plate. Richter and McKenzie (1978) suggested that the downgoing slab exerts a maximum stress of only 100 bar on the remaining surficial lithosphere. For a plate 100 km thick, this leads to a force per unit length of $10^{15}$ dyne/cm. Like the others, this estimate is surely uncertain by a factor of two or more.

Thus, the force per unit length can be very different depending upon the assumptions made and can be summarized as follows:

\[ F = 1.0 \times 10^{15} \text{ dyne/cm}: \text{little direct force applied by the sinking slab to the surface plate} \]
\[ F = 1.8 \times 10^{16} \text{ dyne/cm}: \text{pull from a slab to a depth of 350 km} \]
\[ F = 3.2 \times 10^{16} \text{ dyne/cm}: \text{pull from a slab to a depth of 700 km} \]
\[ F = 5.1 \times 10^{16} \text{ dyne/cm}: \text{pull from a slab to a depth of 700 km with excess mass in an elevated olivine-spinel phase change} \]

The variation in the values underscores our uncertainty in the processes occurring.

LENGTH OF CONTINENTAL CRUST SUBDUCTED

The pull of the slab is counteracted by the buoyancy of the light continental crust. This force per unit length is ($\rho_m$
- ρ_0 gh_c d, where d is the length of the continental crust that is subducted (Fig. 2). The mantle part of the continental lithosphere will contribute to the subduction by pulling down a fraction of the crust, h_3. The heavy oceanic slab is therefore resisted by (ρ_m - ρ_o)g(h_c - h_3)d. Subduction of continental lithosphere should cease when the force pulling it down is balanced by buoyancy of the continental crust. We seek a value of

$$d = \frac{F}{(\rho_m - \rho_o)g(h_c - h_3)},$$

the length of the continent subducted. Using (ρ_m - ρ_o) = 0.55 g/cm³, g = 10^3 cm/s², and (h_c - h_3) = 35 - 10 = 25 km, we obtain for F = 1.0 \times 10^{15} \text{ dyne/cm}: d = 7 \text{ km}; F = 1.8 \times 10^{14} \text{ dyne/cm}: d = 131 \text{ km}; F = 3.2 \times 10^{14} \text{ dyne/cm}: d = 237 \text{ km}; and F = 5.1 \times 10^{14} \text{ dyne/cm}: d = 298 \text{ km}. Note that an error of 0.1 g/cm³ in (ρ_m - ρ_o) contributes a 19% error in the length of the crust subducted. Recall also that the uncertainty in h_3 calculated in the previous section would nearly allow steady-state subduction of continental lithosphere. If h_3 = 20 km instead of 10 km, the amount of crust to be subducted would be about 40% greater than that calculated above. As in the previous calculation, the 33% uncertainty of α in the parameters listed above leads to another 33% uncertainty in d. Each of these estimates for the amount of continental crust subducted is therefore uncertain by at least a factor of two, but notice that the uncertainty due to our ignorance of the processes involved leads to much greater uncertainty in d.

These calculations are all based on the implicit assumption that the continent meets the subduction zone flush and that throughout the length of the subduction zone there is continental material. Such a situation is highly unlikely, for surely peninsulas will meet the subduction zone first. Moreover, continental fragments need not have the same width as the length of the subduction zone. The buoyancy of a long peninsula or of a small continental fragment, such as the Seychelles Islands, is surely negligible compared with the negative buoyancy of the oceanic lithosphere; subduction of them ought to be complete. In the case of a continent half as long as the subduction zone, the amount of continent that could be subducted should be twice the values given above.

We think that except for such cases, however, only very short lengths (d = 50 km) of intact continental crust can be subducted. The gravity anomalies across island arcs (Kogan, 1975; Watts and Taiwani, 1975) do not show evidence for the excess mass of the elevated olivine-spinel phase change. Because the fault-plane solutions of earthquakes below 350 km imply a resistance to the sinking of the slab, this lower part probably does not exert a pull on upper parts. Therefore, the shortest length (131 km) seems the most plausible of the three largest values given above.

Beyond this, the evidence that large events beneath the trenches seem to rupture through the entire lithosphere (Abe, 1972; Kanamori, 1971; Stewart, 1975) and the theoretical arguments of Richter (1977) and Richter and McKenzie (1978) suggest that only a small fraction of the negative buoyancy of the sinking slab is transferred directly to the surface. Moreover, the outer topographic rises seaward of deep-sea trenches do not seem to require large forces or moments that would be applied by the weight of the downgoing slab. Using laboratory data to constrain the strength of the lithosphere, C. Geotze and B. Evans (in prep.) have concluded that a bending moment per unit length of about 10^{22} \text{ dyne} could cause a flexure of the lithosphere and thereby account for the observed topography. If this moment were caused by the negative buoyancy force per unit length of the downwelling slab, applied 100 km landward of the trench axis, a force per unit length of 10^{15} \text{ dyne/cm} would be adequate. A much larger force could deflect the oceanic lithosphere much more than is observed. If this logic is correct, then only a short length of intact continental crust would be subducted.

Subduction of a short length of intact crust is not necessarily incompatible with the large crustal shortening observed in orogenic belts. The geologic evidence implies detachment of the upper and lower crust along shallow-dipping thrust faults. If the upper crust can be detached, the negative buoyancy of the mantle part of continental lithosphere is likely to be large enough to sustain subduction of it and the lower crust. Moreover, if the crust has a low melting temperature, it might melt before penetrating deep into the asthenosphere and thereby further reduce the resistance to subduction.

A DIGRESSION ON PRECAMBRIAN BLUESCHISTS

Blueschists are most common in Mesozoic and Cenozoic terranes; there seem to be no Precambrian examples (de Roever, 1956; Ernst, 1972). Greater heat generation and higher geothermal gradients in Precambrian time are thought to have prevented the high-pressure-low-temperature stability field for blueschists at that time (de Roever, 1956; Ernst, 1972). Another possibility is that Precambrian lithosphere was too thin and too buoyant to be able to pull cold crustal material down to the depths required for formation of blueschists.

SUMMARY AND CONCLUSIONS

Two factors contribute to subduction of continental lithosphere: the negative buoyancy of the mantle part of the lithosphere and the force exerted on the surface lithosphere by a downgoing slab of oceanic lithosphere. If there were a mechanism for detaching the upper crust from the lithosphere, the negative buoyancy of the mantle part would make it and the lower crust gravitationally unstable. The mantle lithosphere could subduct about 10 km of crust with it, but this estimate of 10 km is very uncertain and could be nearly three times larger if the values of the requisite parameters conspired. Gravity acting on the excess mass of a downgoing slab of lithosphere exerts a downward force on the surficial lithosphere. This force should be able to pull some continental lithosphere into the asthenosphere. If the crust remains intact, peninsulas and microcontinents might be subducted completely. The greatest uncertainty in the calculation of this force arises not in estimating the excess mass but in deciding how much of its gravitational body force is transmitted to the surface. If all of the force were transmitted, then hundreds of kilometres of continent could be subducted, as noted also by Bird and others (1975). If only a small fraction were transmitted, then only a negligible length of intact continental crust could be subducted. Although we think that the latter is the case, at present we cannot demonstrate it conclusively. If a substantial amount of continental crust can be subducted, this clearly will cause pronounced chemical heterogeneity of the mantle and will profoundly affect the crustal evolution of subduction zones such as in Tibet.
If the oceanic lithosphere has an approximately linear temperature gradient \( \frac{dT}{dz} = -\frac{T_0}{a} \), where \( T_0 \) is temperature in the asthenosphere and \( a \) is thickness of the lithosphere, then the temperature difference between the downgoing slab and the surrounding asthenosphere is given by (McKeeen, 1969)

\[
\Delta T(x, z) = 2T_0 \sum \frac{(-1)^n}{n!} \exp \left( \frac{(R - R^2 + n^2 \pi^2)}{a} \right) \sin n\pi \frac{z}{a} \tag{A1}
\]

where \( x \) and \( z \) denote an orthogonal coordinate system with \( x \) the distance downwarp along the slab, \( z \) the distance into the slab measured from its base, and \( R = \frac{\rho C_p \mu a}{2k} \), where \( \rho \) is density, \( C_p \) is heat capacity, \( \nu \) is rate of subduction, and \( k \) is thermal conductivity. \( T(x, z) \) does not include adiabatic heating, but we use only the difference between the temperatures in the slab and in the asthenosphere, for which the adiabatic effect can be neglected (McKeeen, 1970). Thus, material in the slab is heavier than that in the asthenosphere by

\[
\Delta \rho = \rho_0 \Delta T = 2\rho_0 T_0 \sum \frac{(-1)^n}{n!} \exp \left( \frac{(R - R^2 + n^2 \pi^2)}{a} \right) \sin n\pi \frac{z}{a} \tag{A2}
\]

The force per unit length along the arc due to the excess mass of the slab is

\[
F = \sin \phi \int_0^{\theta} d\theta \int_0^{\alpha} \Delta \rho \rho \ d\gamma \ dx.
\]

where \( \phi \) is the dip of the slab and \( x_0 \) is the length along the slab of interest. Integrating equation A3 with A2, we get

\[
F = 4\pi g \alpha T_0 \rho^2 \sin \phi \sum \frac{i}{i} \exp \left( \frac{(R - R^2 + (2l + 1)^2 \pi^2)}{2R} \right) \frac{x_n}{a} \tag{A4}
\]

Using the values \( \rho = 3.35 \text{ g cm}^{-3}, C_p = 0.28 \text{ cal g}^{-1} \text{ C}^{-1}, \nu = 10 \text{ cm yr}^{-1}, k = 7.5 \times 10^5 \text{ cal cm}^{-1} \text{ s}^{-1} \text{ C}^{-1}, \) and \( a = 100 \text{ km} \), we obtain \( R \approx 200 \). Because \( R >> \pi \) and only the terms with \( l < \approx 5 \) turn out to be important, we may use

\[
\sqrt{R^2 + (2l + 1)^2 \pi^2} - R \approx (2l + 1)^2 \pi^2 / 2R
\]

Then equation A4 becomes

\[
F = 8\pi g \alpha T_0 \rho^2 R \sin \phi \sum \frac{i}{i} \left( \frac{2l + 1}{2\pi} \right) \frac{x_n}{a} \tag{A5}
\]

As

\[
\sum \frac{1}{(2l + 1)^4} = \frac{\pi^4}{64\pi^4},
\]

we have

\[
F = \frac{g \alpha T_0 \rho^2 R \sin \phi}{12} \left( \frac{1}{1} - 96 \sum \frac{1}{(2l + 1)^4} \right) \tag{A5}
\]

This derivation was given by McKeeen (1969), but he let \( x_0 = \infty \), so that the sum in equation A5 vanishes. The number, \( m \), of terms used in sums depends upon the depth of penetration of the slab, but for \( x_\phi >> a \), fewer than five terms are adequate, and for \( x_\phi \approx 5a \), only the first term is important.

For \( x_\phi = 5a \) or \( 10a \), the factor in parentheses is 0.13 or 0.23, respectively. With \( g = 10^5 \text{ cm}^2 \text{ s}^{-2}, T_0 = 100^\circ \text{C}, a = 100 \text{ km}, \) and \( \phi = 45^\circ, F = 1.8 \times 10^1 \text{ dyne/cm and } 3.2 \times 10^4 \text{ dyne/cm for these cases.}

## REFERENCES CITED


## ACKNOWLEDGMENTS

Reviewed by B. C. Burchfiel and F. M. Richter. Supported by National Science Foundation Grant 77-23017EAR and by National Aeronautics and Space Administration Grant NGS5260.

MANUSCRIPT RECEIVED AUGUST 14, 1978
MANUSCRIPT ACCEPTED OCTOBER 31, 1978