The brittle-ductile transition in tectonic boundary zones

Michele Dragoni
Dipartimento di Fisica, Università di Bologna, Italia

Abstract
Observations at the Earth’s surface show that tectonic boundary zones are intensively deformed belts with horizontal widths of tens or hundreds of kilometres. Most seismic activity is concentrated in a shallow «seismogenic» layer, suggesting that at larger depths the relative plate displacements are accommodated through ductile, aseismic deformation. This is confirmed by laboratory experiments indicating that, at sufficiently high temperatures and pressures, brittle behaviour in crystalline materials gives way to ductile behaviour. At the very low strain rates typical of the Earth’s tectonics, ductility is a long-term behaviour and lithospheric rocks can be modelled as viscoelastic materials, with effective viscosity dependent on depth through temperature and pressure. Theoretical models of tectonic boundary zones are able to reproduce the brittle-ductile transition and give an insight in the relationships between state of stress, rheological properties and seismic activity.

1. Introduction
The adjectives brittle and ductile denote two distinct mechanical behaviours of a solid body when it is subject to stress (Jaeger and Cook, 1976). A body behaves as brittle if, after an initial elastic deformation, it reaches soon the critical point for rupture, without any appreciable permanent deformation. A body behaves as ductile if instead it can undergo a sensible permanent deformation. The term ductility has a broad meaning and is generally referred to a rheology that is characterized by a macroscopically homogeneous deformation, regardless of the micro-mechanisms involved. Plastic materials are instead those which behave as elastic solids if the applied stress does not exceed a critical value and flow once this value has been overcome.

Laboratory experiments show that, at sufficiently high temperatures and pressures, brittle behaviour in crystalline materials gives way to ductile behaviour. There is a broad transition between these distinct regimes, where the deformation is semi-brittle or semi-ductile, involving on the microscale a mixture of brittle and ductile processes (fig. 1). As a consequence we may infer that in the Earth brittle faulting at shallow depth will give way to ductile deformation at greater depth. The transition will occur over a depth range depending largely on the lithology and the thermal conditions and must limit the depth to which earthquakes normally can occur.

Observations and theoretical models of tectonic boundary zones show in fact the existence of a preferential layer for earthquake nucleation (seismogenic layer), below which aseismic slip prevails. A stress concentration is predicted in the upper part of boundary zones (Savage and Burford, 1973; Turcotte and Spence, 1974; Savage and Prescott, 1978; Bonafede and Dragoni, 1982; Dragoni et al., 1986). The existence of a seismogenic layer has been explained by several authors in terms of the brittle-ductile transition (Meissner and Strehlau, 1982; Sibson, 1982; Doser and Kanamori, 1986; Strehlau, 1986; Dragoni, 1988).
2. Experimental observations

We have a direct knowledge of the mechanical behaviour of lithospheric material from laboratory experiments on rock samples. Experiments show that increase in temperature has a crucial role in producing ductility. However, increase in temperature alone at atmospheric pressure is normally ineffective in producing ductility in rocks. It is generally necessary to apply some confining pressure as well as raise the temperature. If a sufficiently high temperature is achieved, ductility intervenes in compression tests at confining pressures much less than 1000 MPa, which corresponds to a depth of about 30 km in the Earth (fig. 2).

Water content is an important factor in determining the ductile behaviour of silicates. Tullis and Yund (1980) found that the addition of 0.2 wt % water to granite produced a sensible weakening. The brittle-ductile transition can be induced at high temperatures by the addition of water to quartzites (Mainprice and Paterson, 1984) and clinopyroxenites (Boland and Tullis, 1986).

There are various deformation processes that may produce ductility. Which combination of processes is involved in a particular case depends on the conditions of pressure, temperature and strain rate. The following processes can be distinguished (Paterson, 1978).

*Cataclastic flow* is the deformation achieved by distributed fracturing and the relative movement of the fragments. As the confining pressure contributes to the normal forces acting on the sliding interfaces, the resistance to cataclastic flow is strongly pressure dependent.

*Crystal plasticity* is the deformation of crystalline material by slip and twinning of the grains. These are basically shearing processes without volume change and are very slightly dependent on the confining pressure.

*Diffusional flow* is a deformation due to diffusion of material from one part of the body to another. Nabarro-Herring and Coble creeps are simple models considering paths through the crystals or along grain boundaries, respectively (Nabarro, 1948; Herring, 1950; Coble, 1963). Other possibilities exist, involving other diffusion mechanisms. In porous rocks, diffusion may occur through a fluid phase in intergranular spaces. A mechanism of pressure solution, based on the change in solubility of the solid at the sites of high normal stress, has been proposed (Rutter, 1976).

The change in deformation mode is gradual and occurs through an intermediate regime (Carter and Kirby, 1978; Kirby, 1983). The transition occurs under conditions where mechanisms producing ductility (e.g., crystal dislocation motion) come in competition with fracturing. Since fracturing involves both a volume increase and a
frictional work, it is inhibited by increasing pressure. Dislocation glide instead involves no volume change or friction and is insensitive to pressure. On the other hand, the mobility of dislocations is enhanced by thermal activation, so dislocation motion is favoured by increasing temperature, while stress-induced cracking is fairly insensitive to temperature. Hence the ductile behaviour is favoured by both high temperature and pressure (Scholz, 1990).

Diffusional and grain-sliding flows are both expected to be strongly temperature dependent, because of the underlying thermally activated processes, and to be relatively more important at low stresses and low strain rates, conditions which are both present in the Earth’s lithosphere. Dependence on the confining pressure will probably be relatively small (Paterson, 1978).

The average value of the geothermal gradient in the crust is 30 °C/km, but it may vary from 10 to 50 °C/km according to the zone. Assuming a geothermal profile and an average composition for the crust, the depth at which ductile deformation prevails can be easily estimated. The necessary temperature in order that the rocks of the upper crust become ductile is between 300 and 450 °C; this is the temperature interval in which quartz and feldspars, the main mineralogical components, start to deform by grain boundary sliding or dislocation glide or show marginal recrystallization (Ashby and Verrall, 1977; Tullis et al., 1982). These temperatures correspond to depths approximately ranging from 10 to 30 km, according to the geothermal gradient.

3. Constitutive equation

The rheological behaviour of any material is described by a constitutive equation, relating strain, stress and their time derivatives. It is a matter of fact that crustal and mantle rocks behave elastically in the short term, since they propagate seismic shear waves. Their response is ductile only in the long term: in this case, rocks may undergo large, permanent shear strain, while normal strain is still found to be elastic. Different constitutive equations have been proposed to describe such a ductile behaviour. In the classical Nabarro-Herring or diffusion creep, strain rate is proportional to stress. Laboratory experiments done on crustal materials at high temperatures and stresses in the kbar range indicate a power-law creep, where the strain rate is proportional to a power $n$ of stress, with $n$ ranging from 1 to about 3 (Kirby, 1977; Kock et al., 1980). Power-law creep with higher values of $n$ has been observed only in experiments at larger stresses (Carter and Ave Lallemand, 1970; Kirby and Raleigh, 1973).

The greater relative importance of diffusional and grain-sliding flow at the low stress and low strain rate conditions which are present in the Earth’s crust suggests that the Newtonian creep law ($n = 1$) is a reasonable approximation of the long-term response of rocks (Carter, 1976; Christie et al., 1979; Brace and Kohlstedt, 1980). The hydration degree of rocks is another factor favouring Newtonian deformation, since the exponent $n$ decreases passing from anhydrous to hydrated rocks (Jaoul et al., 1984). Also theoretical models (Dragoni and Pondrelli, 1991) show that $n = 1$ yields results consistent with observations for the depth of brittle-ductile transition, while higher values of $n$ give in most cases larger depths (several tens of kilometres), in disagreement with observations.

The complete constitutive equation must include both the elastic and the ductile behaviour. Such an equation for a medium which is elastic with respect to normal stress, but is able to relax deviatoric stress, can be written as

$$\dot{\sigma}_{ij} + \mu \frac{1}{\eta} \left( \sigma_{ij} - \frac{1}{3} \sigma_{kk} \delta_{ij} \right) = \lambda \dot{e}_{kk} \delta_{ij} + 2 \mu \dot{e}_{ij}$$ (1)

where $\sigma_{ij}$ is the stress tensor, $e_{ij}$ is the strain tensor, $\eta$ is viscosity, $\lambda$ and $\mu$ are the Lamé parameters. Dots indicate differentiation with respect to time.

This rheological model was introduced by Maxwell to describe the behaviour of substances such as pitch, which show instantaneous elasticity, but flow under small stresses, if they are applied for a long time. The same model can be applied to the Earth’s crust and mantle, which behave elastically for short times, but are able to flow under prolonged stresses. The upper cold crust differs in that its viscosity is many orders of magnitude larger than the viscosity of the mantle.

39
4. Boundary zone model

Tectonic boundary zones are belts of intensively deformed material (shear zones), included between two tectonic plates: their width can reach tens of kilometers or more (Yuen et al., 1978; Prescott and Nur, 1981; Tullis et al., 1982). The rheological properties of boundary zones are of primary importance in determining the conditions which produce earthquakes (Meissner and Strehlau, 1982; Sibson, 1982; Dragoni et al., 1986).

According to the commonly accepted view, the upper part of a boundary zone (the seismogenic layer) deforms in an elastic-brittle regime and is able to accumulate relatively large amounts of stress: it is characterized by friction-controlled slip on fault planes. The lower part, where temperature and pressure are higher, deforms instead in an elastic-ductile regime. These vertical variations, in particular the existence of a seismogenic layer, can be reproduced by an analytical model in which the time evolution of shear stress in a transcurrent boundary zone is calculated (Dragoni, 1988, 1989, 1990).

An earthquake source can be described as a dislocation imposing a sudden strain on the system. As a consequence, a sudden elastic change in the stress field takes place. Then stress relaxation starts, since the medium is viscoelastic, governed by one or more characteristic times. The case of a transcurrent plate boundary is particularly simple, since only one stress component and one relaxation time are relevant. Similar results as stated below can be obtained for compressional and extensional boundaries.

Consider a transcurrent boundary zone of width $d$ included between two tectonic plates moving with relative velocity $v$ (fig. 3). The shear-strain rate applied to the boundary zone is therefore

$$\dot{\varepsilon} = \frac{v}{2d}$$  \hspace{1cm} (2)

If $\sigma$ is the corresponding shear-stress component, the constitutive equation (1) reduces to

$$\frac{\dot{\sigma}}{\mu} + \frac{\sigma}{\eta} = 2\dot{\varepsilon}$$  \hspace{1cm} (3)

where both rigidity $\mu$ and viscosity $\eta$ may depend on depth $z$. In particular, viscosity can be written as

$$\eta = \frac{1}{2A} \exp\left[\frac{Q}{RT}\right]$$  \hspace{1cm} (4)

where $Q$ is the activation energy, $R$ is the gas constant, $T$ is the absolute temperature and $A$ is an empirical parameter. Temperature $T$ changes with depth, while $A$, $Q$ and $\mu$ may change with petrological composition, so that equation (3) is able to describe any local change in the rheological behaviour. The rate of stress relaxation, at a given depth, depends on the value of the Maxwell characteristic time

$$\tau(z) = \eta / \mu$$  \hspace{1cm} (5)

and, following (4), on temperature $T$. We assume that the strain rate is uniform and constant:

$$\dot{\varepsilon} = \dot{\varepsilon} \hspace{1cm} (6)$$

Then equation (3) can be solved for $\sigma$ and provides the stress evolution as a function of time and depth (Dragoni, 1988):

$$\sigma(z,t) = \bar{\sigma}(1 - \exp[-t/\tau(z)]), \hspace{1cm} t > 0$$  \hspace{1cm} (7)

where $\bar{\sigma}$ is the maximum shear stress associated with each depth $z$:

$$\bar{\sigma}(z) = 2 \dot{\varepsilon} \eta(z)$$  \hspace{1cm} (8)
Assume that the boundary zone is crossed by vertical fault planes and that the frictional resistance of a fault can be expressed as

\[
\sigma_f(z) = \kappa(\sigma_n - p)
\]  
(9)

where \(\kappa\) is the coefficient of friction, \(\sigma_n\) is the normal stress and \(p\) is the pore pressure. Normal stress can be written as

\[
\sigma_n(z) = \rho g z + T_n(z)
\]  
(10)

where \(\rho\) is the density of rocks, \(g\) is the acceleration of gravity and \(T_n\) is the normal stress exerted by plate motion: \(T_n\) is positive for a compressional boundary and negative for an extensional boundary, while it vanishes for a transcurrent boundary. Hydrostatic fluid pressure is assumed

\[
p(z) = \rho_{w}gZ
\]  
(11)

where \(\rho_{w}\) is the density of water. At any depth, if \(\bar{\sigma}\) is less than \(\sigma_f\), earthquake nucleation cannot take place. Since \(\bar{\sigma}\) is a decreasing function of depth, while \(\sigma_f\) is usually an increasing function of depth, a maximum depth \(z_0\) for earthquake nucleation on the fault will exist, such that

\[
\bar{\sigma}(z_0) = \sigma_f(z_0)
\]  
(12)

Hence the base of the seismogenic layer can be defined as the horizontal plane crossing the intersection of the \(\bar{\sigma}(z)\) and \(\sigma_f(z)\) curves (fig. 4).

Assume that a dislocation takes place at \(t = t_0\) and adds a shear stress \(\sigma_d\). Then the total shear stress on the fault becomes

\[
\sigma(z,t) = \bar{\sigma}(1 - \exp[-(t - t_0)/\tau]) + \\
+ \sigma_d \exp[-(t - t_0)/\tau], \ t > t_0
\]  
(13)

If we define \(\sigma_0\) as the total shear stress at \(t = t_0\):

\[
\sigma_0(z) = \sigma(z, t_0+)
\]  
(14)

we can also write

\[
\sigma(z,t) = \bar{\sigma}[1 - \exp[-(t - t_0)/\tau]] + \\
+ \sigma_0 \exp[-(t - t_0)/\tau], \ t > t_0
\]  
(15)

Equation (15) describes the post-seismic stress evolution in the boundary zone. If we calculate the time derivative of \(\sigma(z,t)\),

\[
\frac{\partial \sigma}{\partial t} = \frac{\bar{\sigma} - \sigma_0}{\tau} \exp[-(t - t_0)/\tau], \ t > t_0
\]  
(16)

we see that it is positive if \(\sigma_0 < \bar{\sigma}\) and negative if \(\sigma_0 > \bar{\sigma}\). This means that shear stress will accumulate at points where \(\sigma_0 < \bar{\sigma}\), while it will relax where \(\sigma_0 > \bar{\sigma}\). From the mathematical point of view, the curve \(\sigma_d(z) = \bar{\sigma}(z)\) is the set of points for which the temporal derivative of stress is equal to zero: we choose this curve as indicating the brittle-ductile transition. In fig. 5, the region where \(\sigma_0 < \bar{\sigma}\) is «brittle», while the region where \(\sigma_0 > \bar{\sigma}\) is «ductile» (Dragoni, 1990).

This is a «mechanical» definition of the brittle-ductile transition. According to this definition, the depth of the transition is not unique, because it depends on stress. Since the shear stress on and around a fault is highly inhomogeneous after a dislocation, there is an interval of depths within which the transition may occur. For the higher values of \(\sigma_0\), the transition will be
shallower than for the lower values. The base of the seismogenic layer, as defined above, is included in the interval of the brittle-ductile transition: this accounts for the fact that the base of the seismogenic layer is often identified with the rheological transition.

Of course, the depth of the transition can change considerably depending on the values of model parameters, such as geothermal profile, strain rate, activation energy and elastic parameters (Dragoni and Pondrelli, 1991). Moreover, it may be different, in the same seismogenic region, for earthquakes of different size, producing different distributions of σ. These variations may explain the different depths of brittle-ductile transition inferred from seismic activity in different areas (Chen and Molnar, 1983).

In particular, strain rate R can vary over a few orders of magnitude as a function of the tectonic activity and the width of the shear zone, according to (2). A typical value of $10^{-14} \text{ s}^{-1}$ has been estimated for the San Andreas Fault (Prescott and Nur, 1981; Meissner and Strehlau, 1982; Sibson, 1982). The activation energy Q depends on petrological composition. It decreases passing from basic to silicic rocks and also passing from anhydrous to hydrated rocks. For granites and quartzites, average values of Q have been found varying from about 100 to 200 kJ mol$^{-1}$ (Kirby and Kronenberg, 1987; Jaoul et al., 1984). Moreover, Q is strongly influenced by temperature (Jaoul et al., 1984; Kronenberg and Tullis, 1984). In general, Q can be assumed as increasing with depth. As to the elastic parameters, their greatest variations take place generally in the uppermost layer of the crust and may not influence the depth of the brittle-ductile transition.

5. Conclusions

An objective of theoretical models, such as the one described above, is to show which is the relative influence of the model parameters on the depth of the brittle-ductile transition. Assuming that a boundary zone deforms as a linear viscoelastic material, with temperature-dependent viscosity, a mechanical definition of the brittle-ductile transition can be given. Such a transition takes place in an interval of depths depending on physical parameters, as geothermal gradient, strain rate and activation energy. The base of the seismogenic layer is placed within the transition interval.

The geothermal gradient is the parameter that affects the brittle-ductile transition depth in the most sensible way: as may be expected, the transition deepens as the gradient decreases. The relaxation time τ is also strongly dependent on temperature. A remarkable dependence of the depth of the brittle-ductile transition on strain rate is also found. The transition depth increases when R increases. The dependence on the activation energy Q is relatively less pronounced. The brittle-ductile transition becomes deeper as Q increases. Through Q, the composition and the hydration degree of rocks influence their rheology and may explain variations of the transition depth as a consequence of heterogeneities in crustal rocks.

The viscoelastic behaviour of rocks must be underlined: their response depends crucially on the timescale of the process considered. Ductile behaviour is a long-term property of lithospheric rocks, that is it becomes effective only for $t \gg \tau$. For $t < \tau$, rocks behave elastically. This means that processes typical of the brittle layer may take place in the ductile layer, if the involved time-
scale is sufficiently short. Typical cases are the downward propagation of seismic dislocations nucleated in the seismogenic zone and the occurrence of aftershocks.

A dependence of the depth of brittle-ductile transition on stress is predicted, the transition being deeper for smaller stresses. This is important in particular for the distributions of aftershock foci. Aftershocks are a consequence of the stress redistribution connected with the main shock and are due to the failure of smaller fault asperities. They involve lower stress values than the main shock and thus may occur at sensibly larger depths than greater earthquakes, as is sometimes observed (Strehlau, 1986). The short timescale of aftershock processes also concurs to this effect, as explained above.

Since earthquakes are the consequence of a frictional instability, one may argue that their depth limitation must be controlled by a stability transition, rather than by a transition from brittle to ductile behaviour. In fact it can be shown (e.g., Scholz, 1990) that ductility as a great influence on stability conditions, determining a lower stability transition in the Earth’s crust at similar depths as the brittle ductile transition.

REFERENCES


