Pergamon

PII: S0191-8141(98)00034-0

Journal of Structural Geology, Vol. 20, No. 8, pp. 1075 to 1087, 1998

0191-8141/98 $19.00 + 0.00

A healing–reloading feedback control on the growth rate of seismogenic faults

P. A. COWIE

Department of Geology & Geophysics, Edinburgh University, West Mains Road, Edinburgh, EH9 3JW, U.K.
E-mail: cowie@glg.ed.ac.uk

(*Received* 4 *September* 1997; *accepted in revised form* 25 *March* 1998)

Abstract—Spatial and temporal variations in the growth rates of faults are explained in terms of a stress feedback mechanism operating in the seismogenic upper crust. It is based on the idea that seismic rupture of a fault perturbs the surrounding stress field, advancing the occurrence of future earthquakes on some faults that are optimally oriented while relaxing stress levels on others. If post-slip healing is geologically rapid, then the earthquakes that are thus induced will contribute to reloading along the earlier rupture zone because of the symmetry of the optimal geometry. A positive feedback is set up so that, even in areas that are undergoing uniform tectonic straining, some faults develop higher displacement rates and grow more rapidly while others experience reduced rates or become inactive. Using a thin plate elastic model for lithospheric-scale faulting, it is shown that this healing–reloading feedback mechanism drives rapid localisation and the formation of major through-going faults moving at plate boundary velocities. Enhanced displacement rates (compared to an isolated fault) develop shortly after the onset of deformation along those faults which are optimally positioned in the overall fault population. Thus the formation of a new plate boundary fault zone is predetermined and is a consequence of, rather than the precursor of, preferentially high displacement rates. Also, fault segments located at points of rupture symmetry, e.g. the central portion of a fault zone, are reloaded more frequently and develop higher displacement rates and consequently have longer segment lengths and/or larger displacement to length ratios. Episodic fault movement through time is a general feature of the model. These predictions are consistent with available field observations over a wide range of scales. Thus, elastic–brittle failure and healing appear to be important rheological components of the lithosphere on long time scales (10^4–10^6 y), as well as on the time scale of earthquake recurrence.

INTRODUCTION

The concept of stress triggering of earthquakes is well known in the seismological community. Several workers have shown how an earthquake rupture induces stress changes on neighbouring faults that may advance or retard the occurrence of future earthquakes (Jaume and Sykes, 1992, 1996; Stein *et al*., 1992). An increase in stress of a few bars on optimally oriented faults may reduce the recurrence time of an earthquake by at least 10% (King *et al*. 1994; Caskey and Wesnousky, 1997). Conversely, faults that lie in the stress shadow of such an event may have an extended period of quiescence. While this phenomenon has received a lot of attention from the point of view of earthquake prediction, the long term effect on fault growth has not been considered. The aim of this paper is to show that, if a ruptured fault heals on a time scale that is short compared to interseismic periods, a feedback situation develops spontaneously in a population of actively growing faults. Healing is the process by which a fault recovers some or all of its shear strength subsequent to rupture. The combination of a positive feedback between optimally oriented faults and negative feedback in stress shadow zones is shown here to strongly influence the pattern and rate of fault development in the elastic part of the lithosphere.

Elastic interactions between simultaneously growing faults are thought to be a key ingredient in the 'self-organisation' of fault patterns in the lithosphere (Sornette, A. *et al*., 1990; Sornette, D. *et al*., 1990; Cowie *et al*., 1995). Although the spatial evolution and the size–frequency distribution of fault populations have been described by these workers, the accumulation of fault displacement in the self-organisation process has not been previously discussed. Here, the thin-plate elastic–brittle model of Cowie *et al*. (1993) is used to demonstrate specifically how stress feedback between adjacent faults affects the accumulation of displacement through time. The results are compared with field evidence for displacement rate variations along fault zones, contrasting rates of growth on adjacent structures, and alternating periods of fault activity and quiescence. The implications for interpreting the rheological behaviour of the lithosphere are also discussed.

STRESS TRIGGERING WITHIN FAULT ARRAYS

The Coulomb failure criterion has been used to define the conditions for earthquake triggering due to static stress changes (King *et al*., 1994). For any given slip event on a fault, there is a change in the shear and normal stresses on adjacent faults in the surrounding volume. The sign and relative amplitude of the shear stress and normal stress perturbations determine whether a neighbouring fault is brought closer to fail-

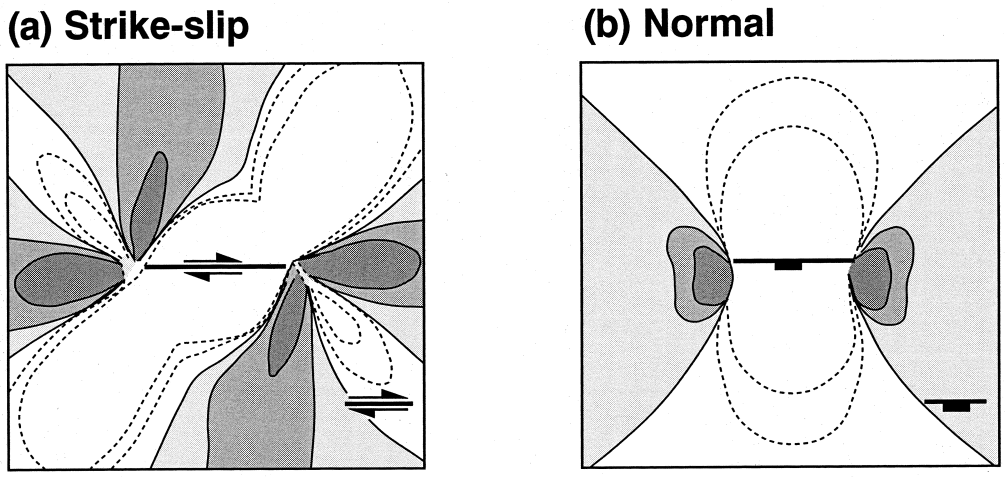

Fig. 1. Sketch showing the Coulomb stress change on optimally orientated faults due to slip on (a) a right-lateral strike slip fault, and (b) a dipping normal fault. Regions of positive stress change, i.e. stress increase, are indicated by grey shading, regions of negative stress change are indicated by white areas with dashed lines. Optimally oriented structures in each case are parallel to the fault that has ruptured, and located along strike, in either a co-planar or en échelon geometry. The stress shadow zones in both cases extend transverse to the ruptured fault, although in the strike-slip case the pattern is asymmetric. (Modified from King *et al.*, 1994; Hodgkinson *et al.*, 1996.)

ure or is relaxed. The so-called 'Coulomb Failure Function', CFF, (Caskey and Wesnousky, 1997) is used to quantify the likelihood for triggered failure on a particular fault plane. It is a function of the coefficient of friction, μ, the pore pressure, Δp, and the orientation of the plane, i.e.

$$\mathrm{CFF} = \Delta\tau_s + \mu(\Delta\sigma_n - \Delta p)$$

where $\Delta\tau_s$ is the static shear stress change (positive in the direction of failure) and $\Delta\sigma_n$ is the normal stress change (negative for compression) (see King *et al.*, 1994).

Figure 1 shows an example of the CFF calculated for planes that are strike-parallel and co-planar to a right lateral strike-slip fault rupture (Fig. 1a) and a dipping normal fault rupture (Fig. 1b). The contours of positive and negative changes in the CFF are calculated using an elastic dislocation model (King *et al.*, 1994; Hodgkinson *et al.*, 1996). In both normal and strike-slip examples there are regions transverse to fault strike where the stress level is reduced. In the strike-slip case the regions of increased failure stress are asymmetric with respect to the fault, whereas in the normal fault (or reverse fault) example the pattern is approximately symmetric for a steeply dipping fault ($\geq 60^\circ$) However, in both cases, regions of increased likelihood for failure lie along strike from the rupture zone. Thus, for an en échelon array of sub-parallel faults, rupturing individually, along-strike neighbours will feel mutual stress enhancement whereas faults lying to either side of the array will generally lie in regions of stress shadow. This is the underlying mechanism behind the enhanced localisation and self-organisation of brittle deformation described in this paper. In other words, during progressive deformation some fault array geometries will be favoured by a positive feedback during rupture and reloading and thus grow preferentially, while others experience a negative feedback.

Figure 2 illustrates a hypothetical sequence of static stress changes due to ruptures on nearby fault segments. Between each rupture the faults are assumed to heal to their pre-rupture shear strength. Constant far-field tectonic loading is applied to all the faults in the array. As a simplified model, Fig. 2 applies to either strike-slip or dip-slip fault arrays, given the general pattern of along-strike stress enhancement and transverse stress reduction that is seen in Fig. 1. In both

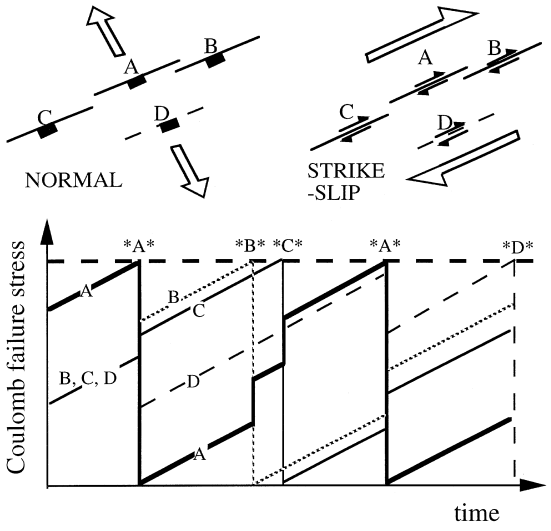

Fig. 2. Cartoon illustration of the healing–reloading feedback mechanism. Variation of Coulomb failure stress on an array of four faults, A, B, and C, aligned along strike, and D, located to one side. Both a normal fault array and strike-slip fault array are illustrated. Hollow arrows indicate the regional tectonic loading. The variation in amplitude of the static stress change due to rupture is not considered, only the sign, i.e. increase or decrease. A is the first fault segment to rupture, B and C are optimally oriented with respect to A (positive feedback) and D lies in the stress shadow of A (negative feedback) (Fig. 1). Earthquakes on each fault are indicated by *A*, *B*, etc. See text for discussion.

cases shown in Fig. 2, A, B and C are along-strike coplanar segments while D is parallel, but lies to one side of fault A. Fault A ruptures first and increases the stress on B and C, but decreases the stress on D because it is in the shadow of A. Ruptures on B and C are thus advanced in time whereas on D they are delayed. Assuming that strength recovery of ruptured faults is relatively rapid, such that they can support stress perturbations from subsequent earthquakes, then ruptures on B and C will result in the reloading of A. The stress level on D is assumed to be largely unaffected by B and C and thus D only experiences the regional tectonic loading at a constant rate during this time period. In contrast, A has experienced both the regional loading and the increase in stress due to the ruptures on B and C, and thus reaches failure once more. If A fails before D reaches the threshold for rupture, as shown in Fig. 2, then D is once more relaxed and rupture is further delayed. Thus the cycle may be repeated, with a tendency for A to reach failure more frequently because it is a point of *rupture symmetry*, and D to rupture infrequently, also because of its location.

The aim of the modelling described below is to show how the mechanism illustrated in Fig. 2 may operate during the gradual development of a fault population. This is done using a numerical model for spontaneous fault growth, i.e. fault planes are not specified *a priori*. Elastic interactions between simultaneously growing faults may thus be examined as they evolve through time. This approach is fundamentally different from static elastic models used, for example, by Aydin and Schultz (1990) and Willemse (1997) to explain varying rates of propagation. These authors have shown how local elastic interactions between two en échelon faults can enhance or impede tip propagation. While the underlying physics is the same, many of the observations discussion below cannot be explained by static models.

NUMERICAL MODELLING RESULTS

The model presented here is described in more detail by Cowie *et al.* (1993) and Sornette *et al.* (1994). It simulates antiplane shear deformation of a thin elastic–brittle plate using a two-dimensional square lattice. The lattice is made up of 180×180 elements which are oriented at 45° to the plate edges. Cyclic boundary conditions are applied in the x-direction to minimize lateral edge effects. A constant velocity is applied along one edge of the lattice ($y = 180$), while the other edge ($y = 0$) is kept fixed. This imposes a uniform antiplane shear strain across the lattice, parallel to the x-axis. Each element is assigned a stress threshold σ_c which is drawn randomly from a probability distribution chosen here to be uniform in the interval $(1 - \Delta/2,\ 1 + \Delta/2)$. In the examples shown here $\Delta = 1.9$. These threshold values remain fixed throughout a particular simulation. The elastic shear modulus for all the elements is a constant value. When an element ruptures, it undergoes an instantaneous stress drop by an amount given by $\sigma_c\beta/2$ where σ_c is the strength threshold for rupture. The parameter β equals 2.0 in the results presented here; this is equivalent to 100% stress drop. The behaviour exhibited by this model is not dependent on the choice of values used for Δ and β, although the pattern of faulting changes qualitatively when these parameters are varied.

Details of the model scaling are derived in Cowie *et al.* (1993). For a lithospheric scale model, the size of one lattice element is on the order of 10 km (the thickness of the seismogenic crust) and thus the dimensions of the lattice are of the order of 1000 km. For the simulations shown here, the applied plate boundary velocity is 10 cm y^{-1} which imposes a constant strain rate across the lattice of 10^{-7} y^{-1}. One time-step in the model equals about 1 y.

Increments of stress are applied to the whole lattice at each time step to satisfy the constant velocity boundary condition. When the strength threshold of an element is exceeded, then a rupture occurs and the rupture of one element can trigger further breaks. An earthquake is defined here as a sequence of ruptures that occur between increments of the plate boundary displacement. Some earthquakes in this model may consist of a single rupture, while others involve a cascade of triggered ruptures of many elements. The stress field throughout the lattice is recalculated after each rupture occurs such that the equation of static equilibrium is satisfied at each lattice node. This is done by solving the discretised version of the scalar elasticity equation using the conjugate gradient method (e.g. Beckman, 1960). Thus ruptured elements, once healed, support stress perturbations due to subsequent ruptures.

The stress perturbation associated with antiplane shear rupture in this model is comparable to that for a Mode III crack (e.g. Pollard and Segall, 1987). The stress pattern is qualitatively the same as that shown for a normal fault (Fig. 1b), i.e. regions of stress increase along strike from the rupture zone and regions of stress shadow in the transverse direction. The model is strictly applicable for a small amount of total strain (a few percent) because volume change and finite rotations are not taken into consideration.

Our numerical approach does not model the dynamics of earthquake rupture explicitly. In a fully dynamic model there is a net stress drop over the entire area of the earthquake rupture plane (e.g. Madariaga, 1976). The stress perturbation in the surrounding area is thus comparable to that of a crack of equivalent dimensions with a uniform stress drop along its length, e.g. Fig. 1. In the present model, the stress drop only affects one element of the model lattice at a time and other elements are healed. In detail

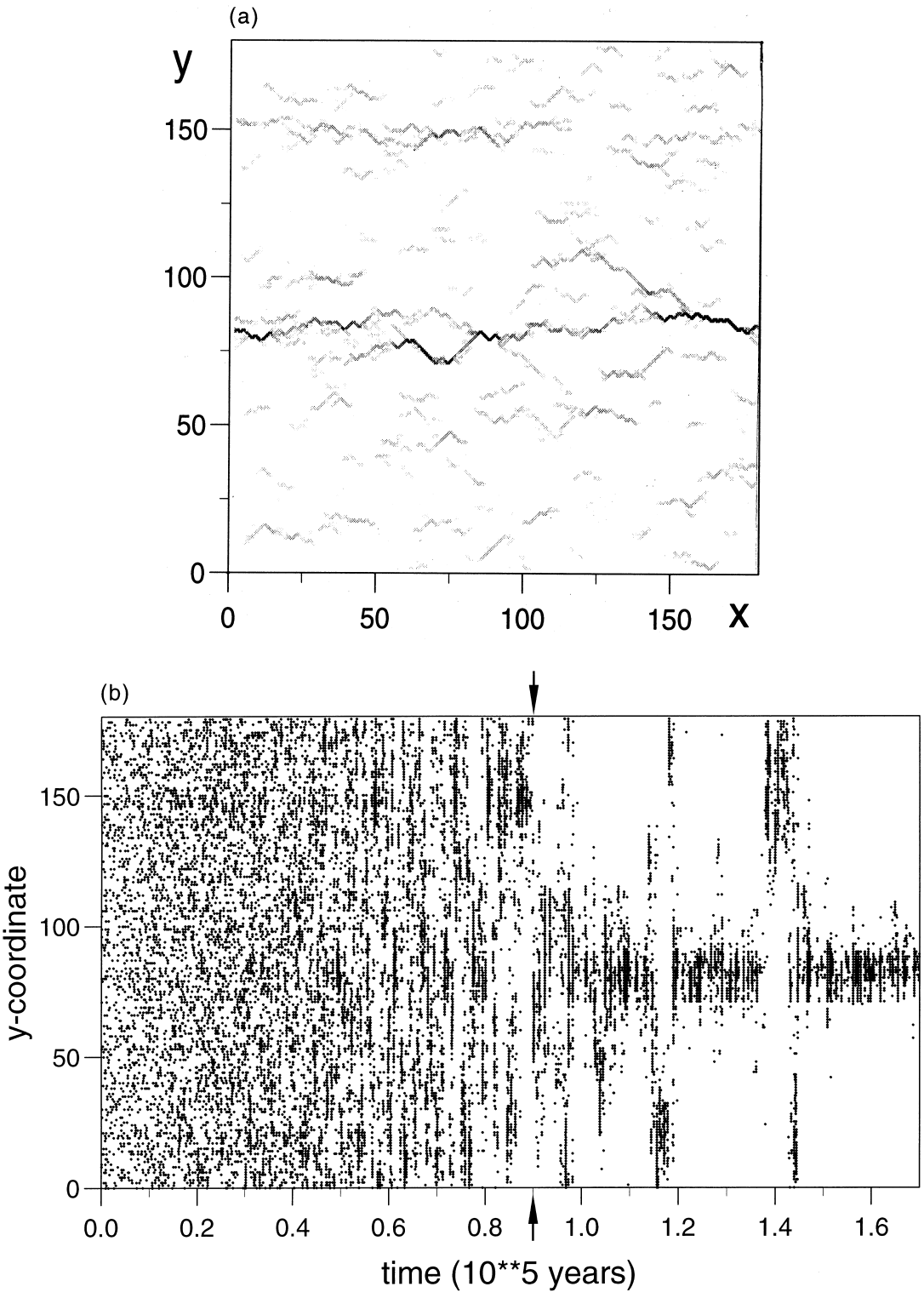

Fig. 3. Example of rupture development exhibited by the thin-plate elastic–brittle rupture model of Cowie *et al.* (1993) and Sornette *et al.* (1994). A square lattice is used of 180×180 elements with uniformly random material properties and no weakening. (a) Normalised grey-scale fault displacement map at time $t = 115{,}000$ y. (b) y-coordinate of ruptured elements as a function of time. The vertical arrows mark the transition from distributed to localised deformation at time $t = 90{,}000$ y.

there will be differences in the scale and amplitude of the stress perturbations between the two types of earthquake representation. These differences are not discussed because the reloading phenomenon illustrated here only requires the rate of healing to be short compared to the time interval between consecutive ruptures it does not depend on the coseismic rupture dimensions.

Figure 3(a) shows a grey scale map of a rupture pattern generated using this model. The level of grey indi-

cates the accumulated displacements on the ruptured elements from white (unruptured areas) to black (maximum displacement). This snap shot is taken at time t = 115,000 y. Figure 3(b) shows the y-coordinate of all the breaking events in the lattice for $0 \leq t \leq 170{,}000$ y, including re-rupturing of existing broken points. Note that the single prominent fault zone that extends across the centre of the lattice is the locus of most of the rupture events that occur after t = 90,000 y; prior to this time the pattern of rupture is distributed throughout the lattice. The approximate point of transition between the distributed and localised regimes is indicated by arrows in Fig. 3(b), and these two regimes are analysed separately in Fig. 4.

In this particular simulation only one major fault forms. In other simulations, using identical parameters (except for a different random number seed for the strength heterogeneity) up to 2 or 3 major faults may form in a lattice of this size. The present realisation was chosen for the sake of simplicity.

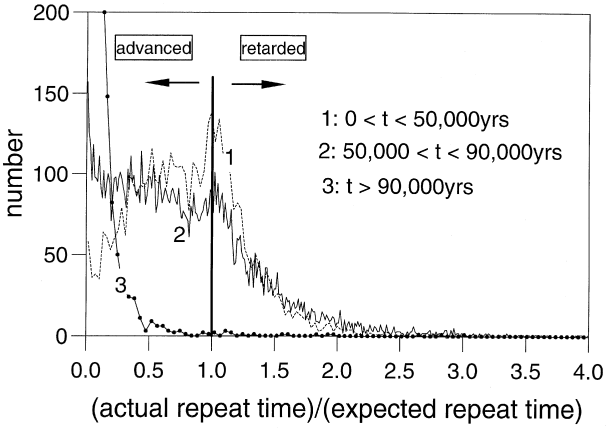

Fig. 4. Histogram of normalised rupture repeat times for all elements for the periods of activity $0 \leq t \leq 50{,}000$ (dashed line) $50{,}000 \leq t \leq 90{,}000$ (solid line) and $90{,}000 < t = 170{,}000$ (line with dots) (see Fig. 3b). The vertical axis is the number of elements. On the horizontal axis, the value 1 indicates the rupture repeat time for an element that is rupturing at the rate predicted by a uniform strain rate across the whole lattice, i.e. $\Delta t_a = \Delta t_e$, defined in Fig. 5.

Enhanced vs inhibited growth

In order to investigate the idea outlined in Fig. 2 in more detail, the time intervals between consecutive ruptures on individual lattice elements have been analysed for the simulation shown in Fig. 3. Histograms of the actual time interval (Δt_a) between ruptures, normalised by the expected repeat time (Δt_e) have been calculated for each ruptured element (Fig. 4). Three time periods have been analysed: $0 \leq t \leq 50{,}000$ y, $50{,}000 \leq t \leq 90{,}000$ y and $t > 90{,}000$ y. Δt_e is calculated by dividing the stress threshold for each element by the increment of stress per time step imposed by the boundary condition. These data include all values of Δt_a for all ruptured elements (as opposed to the average value of Δt_a which is a lot less informative). For the three time periods there is a large spread along the horizontal axis with many elements rupturing more frequently than predicted ($\Delta t_a/\Delta t_e < 1.0$) and many rupturing less frequently ($\Delta t_a/\Delta t_e > 1.0$) (Fig. 4). There is a peak at $\Delta t_a/\Delta t_e = 1.0$ for the period $0 \leq t \leq 50{,}000$ y, indicating that some elements are rupturing with the predicted time interval based on the constant strain rate imposed. However, even during this period of distributed deformation, some repeat times are more than two times longer than expected, whereas some are more than ten times shorter. For the time period $50{,}000 \leq t \leq 90{,}000$ y the peak at $\Delta t_a/\Delta t_e = 1.0$ is much less pronounced and the spread of values about the mean value is greater. For $t > 90{,}000$ there is no peak at $\Delta t_a/\Delta t_e = 1.0$ and the histogram shows an approximately power-law shape. In this period of time ($t > 90{,}000$) some parts of the lattice have become completely inactive (see Fig. 3b). For the elements that have remained active, the repeat times are generally either strongly retarded ($\Delta t_a/\Delta t_e \gg 1.0$) or strongly advanced ($\Delta t_a/\Delta t_e \ll 1.0$). This is consistent with localisation of all the deformation onto a single major fault zone even though the imposed loading of the plate is uniform.

The explanation for the advanced and retarded repeat times is found in Fig. 5 where the stress levels, as a function of time for elements P and Q, are shown. The locations of P and Q relative to the main fault traces in Fig. 3(a) are shown in the inset. Figure 5(a) refers to point P which is one of the elements that make up the main through-going fault. Figure 5(b) is for point Q which is located to one side of this fault in an area that has undergone relatively little deformation. Both plots are normalised by the strength of the respective elements. Also shown by a dotted line is the expected rate of stress increase, imposed by the constant velocity boundary condition, which defines Δt_e. Element P experiences significant and rapid fluctuations in stress and $\Delta t_a/\Delta t_e \ll 1.0$ with ruptures occurring frequently, whereas Q experiences low amplitude fluctuations and $\Delta t_a/\Delta t_e > 1.0$. In both Fig. 5(a) and (b) the gradual increase due to far field loading of both elements is visible. In addition, both P and Q experience stress perturbations due to ruptures occurring elsewhere in the lattice, indicated by abrupt step changes. In the case of P these stress perturbations generally raise the stress significantly, while for Q the change is generally to reduce the stress. This demonstrates the positive stress feedback, for P, and negative stress feedback, for Q. The fluctuations in stress at P are most noticeable just prior to or at the time of rupture, i.e. P is loaded to failure, undergoes a 100% stress drop and is then partially reloaded all within the same earthquake (see definition above). This indicates that more than one element adjacent to P ruptured during these earthquakes. In contrast, Q rarely reaches failure because it is being continually unloaded counteracting the externally applied loading of the

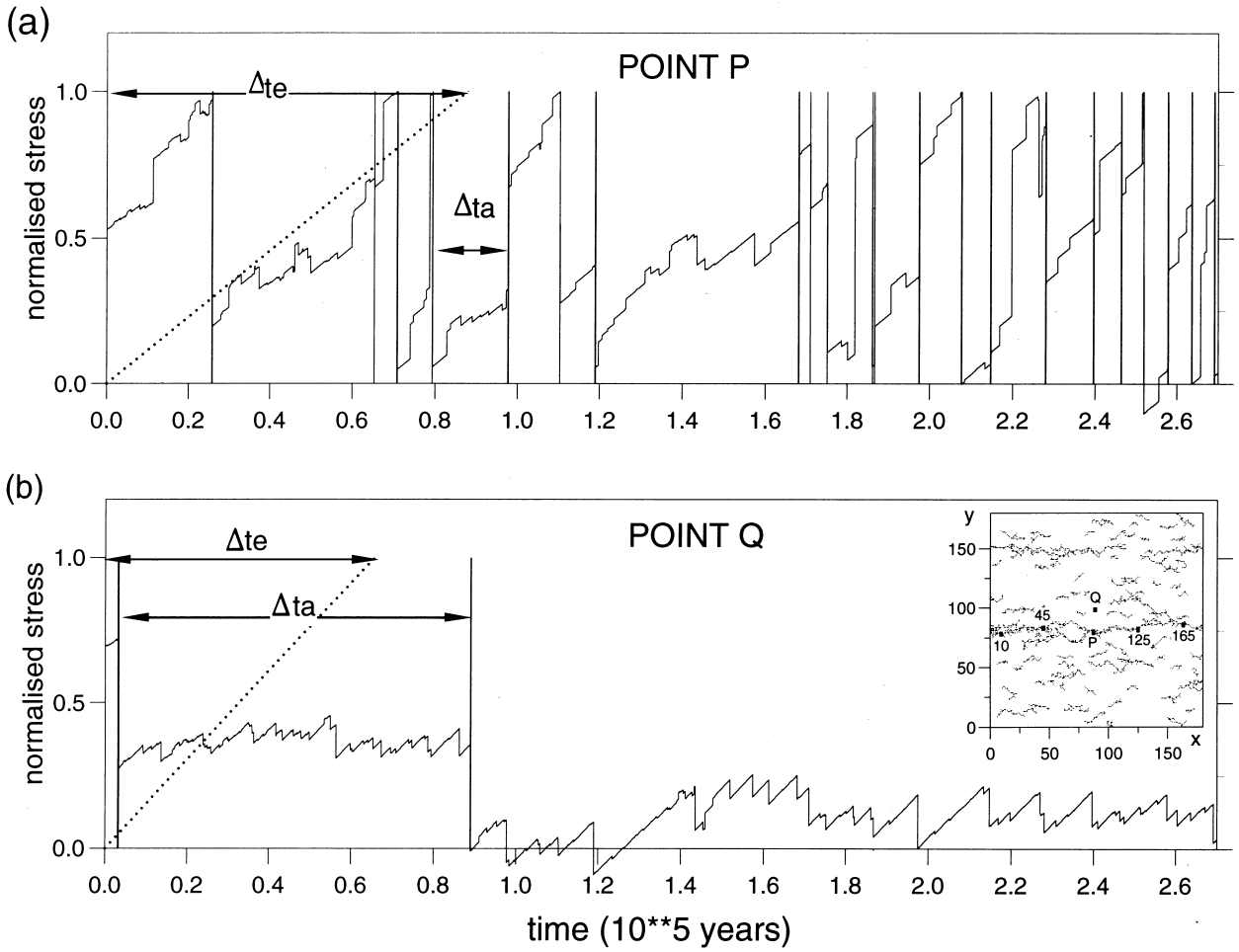

Fig. 5. Stress level as a function of time for two lattice elements, (a) P and (b) Q, location given in inset. Stress level is normalised by the strength of the element in each case. Dotted line indicates uniform loading rate imposed by the boundary condition. Δt_e is the expected rupture repeat time, and Δt_a is the actual repeat time (see Fig. 4). Vertical lines indicate ruptures.

plate. For Q in particular, the effect of negative feedback can cause the stress level to become negative, i.e. the sense of shear stress felt by the element is opposite to the imposed boundary condition.

Figure 6 shows accumulated displacement as a function of time at five different points along the main fault shown in Fig. 3(a) (for location see inset to Fig. 5). Three reference lines are also shown. The line with the lowest slope corresponds to the displacement rate imposed on each lattice element by the boundary condition. The line with the greatest slope shows the total displacement rate of one edge of the plate relative to the other. For a given displacement rate, the difference between the slope of these two reference lines depends directly on the dimensions of the lattice, in this case 180 [i.e. (10 cm y^{-1})/180 $\sim$ 0.6 mm y^{-1}). The cumulative displacement curves for each of the five points along the fault show a distinct transition from relatively low displacement rates for t < 90,000 y, to much higher rates for t > 90,000 y. However, even prior to the onset of localisation (t < 90,000 y; Fig. 3b) the displacement rates are substantially higher than expected for uniform loading (i.e. 0.6 mm y^{-1}). For example, at some points along the fault the displacement rate is almost two orders of magnitude greater than expected. Moreover, some points develop high rates right from the moment when a fault initiates, e.g. at point 165. In contrast, some earlier formed sections of the fault show a more gradual increase in displacement rate, e.g. at point 125. The displacement rate after localisation (t > 90,000 y) is equal, on average, to the total plate deformation rate, i.e. the fault is acting like a new plate boundary accommodating virtually all of the subsequent displacement, while the adjacent areas of the plate behave almost like rigid blocks.

The enhanced displacement rates shown in Fig. 6, for the period t < 90,000 y, are due to the significant advance of rupture repeat times (i.e. $\Delta t_a/\Delta t_e < 1.0$) even though the rupture activity is distributed throughout the lattice. Note that new faults are continuing to nucleate during the period of distributed deformation so that the total number of faults is actually increasing, not decreasing (Cowie *et al.*, 1995). The positive stress feedback between elements in the lattice is

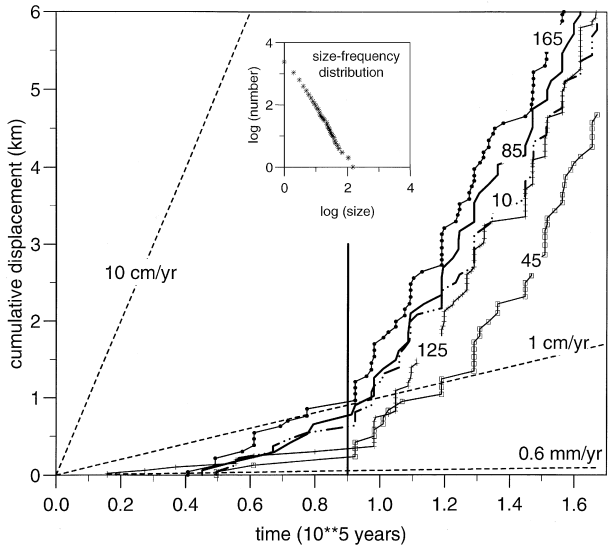

Fig. 6. Accumulated displacement as a function of time at five points (x = 10, 45, 85, 125, 165) along the prominent fault shown in Fig. 3(a) (locations shown in inset to Fig. 5). Point P in Fig. 5 corresponds to x = 85. The vertical line is at t = 90,000 y. Three dashed lines indicate: displacement rate applied to individual elements (0.6 mm y^{-1}), total displacement rate across the lattice (10 cm y^{-1}), and an intermediate rate for reference (1 cm y^{-1}). Inset shows the cumulative size-frequency distribution of fault sizes at t = 90,000 y; a line-fit to the distribution has a slope of approximately −1.6.

leading to more rapid reloading and thus higher displacement rates on some faults in the developing population. These faults ultimately link to form the new 'plate boundary' that divides the plate into two after t = 90,000 y. The initial strength heterogeneity plays a role in this process because it influences where faults nucleate and therefore which ones are optimally positioned within the earliest stages of deformation. However, the strength of the elements remains fixed through time in these simulations. Thus, even though it might be inferred that the transition from distributed deformation to localisation is due to mechanical weakening, it is in reality a consequence of post-rupture healing and positive stress feedback.

Finally, the inset in Fig. 6 shows the size frequency distribution of the fault population at t = 90,000 y. The distribution is well described by a power law with a slope of −1.6. This plots shows that a wide range of fault sizes (more than two orders of magnitude) has developed prior to full localisation of the deformation. Again this demonstrates the large differential rates of fault growth even when the deformation appears to be relatively homogeneously distributed.

Variable displacement rates

Figure 7 shows the temporal evolution of an en échelon fault array produced during a different simulation. The only modification compared to the previous example is that a different random number seed was used to generate the strength variation across the lattice. The trace of the fault at three different stages in its development is shown, together with profiles of the total displacement along the zone (Fig. 7b). A clustering algorithm is used to define groups of connected ruptured elements at each time (Cowie *et al.*, 1993). The only clusters shown here are those which ultimately link together; adjacent faults that remain isolated are not shown for reasons of clarity. Bold lines are used to define the largest clusters, or longest segments, at each stage of development. Figure 8 shows cumulative displacement through time at six points along the fault zone (A–F). Where the fault trace has more than one strand or segment at the same along-strike position, the total displacement is the sum of the values on all the strands.

During the early stages of fault development, numerous short, separate fault segments have nucleated (t = 70,000 y). Even at this stage, there is a variation in segment length along strike. The displacement profile (dashed line), however, shows little variation in amplitude along strike. At time t = 90,000 y, several fault segments have grown and linked together. The displacement profile (solid line) now shows a systematic variation with the maximum displacement towards the centre of the array. During the time period 70,000 < t < 90,000 y the main change in the displacement profile has been along the central portion of the zone, whereas the tip regions change very little. At t = 150,000 y, a fully linked fault has developed. This has a simple displacement profile comparable to a single fault with a maximum displacement in the centre diminishing to zero at the tips. This overall profile shape started to develop prior to complete linkage along the zone. This indicates that positive feedback between the individual segments causes gradual re-adjustment towards the displacement distribution of a single larger fault, even before a through-going structure has fully formed.

This re-adjustment process is evident in Fig. 8. Displacement rates at different points along the fault zone vary markedly. The highest average rates are found towards the centre of the zone, points B, C, D and E. In contrast, the average rates near the fault tips are significantly lower (A and F). Even though some of the distal segments formed early, they accumulate relatively small amounts of displacement whereas central sections of the fault that formed later, for example D and E, have much high displacement rates because of their optimal position. Throughout the evolution of the fault array, the main process has been linkage between existing segments and there has been little change in the total length of this array. Subsequent to the stages of development discussed here, linkage with another fault in the population occurs and the process continues at a larger scale.

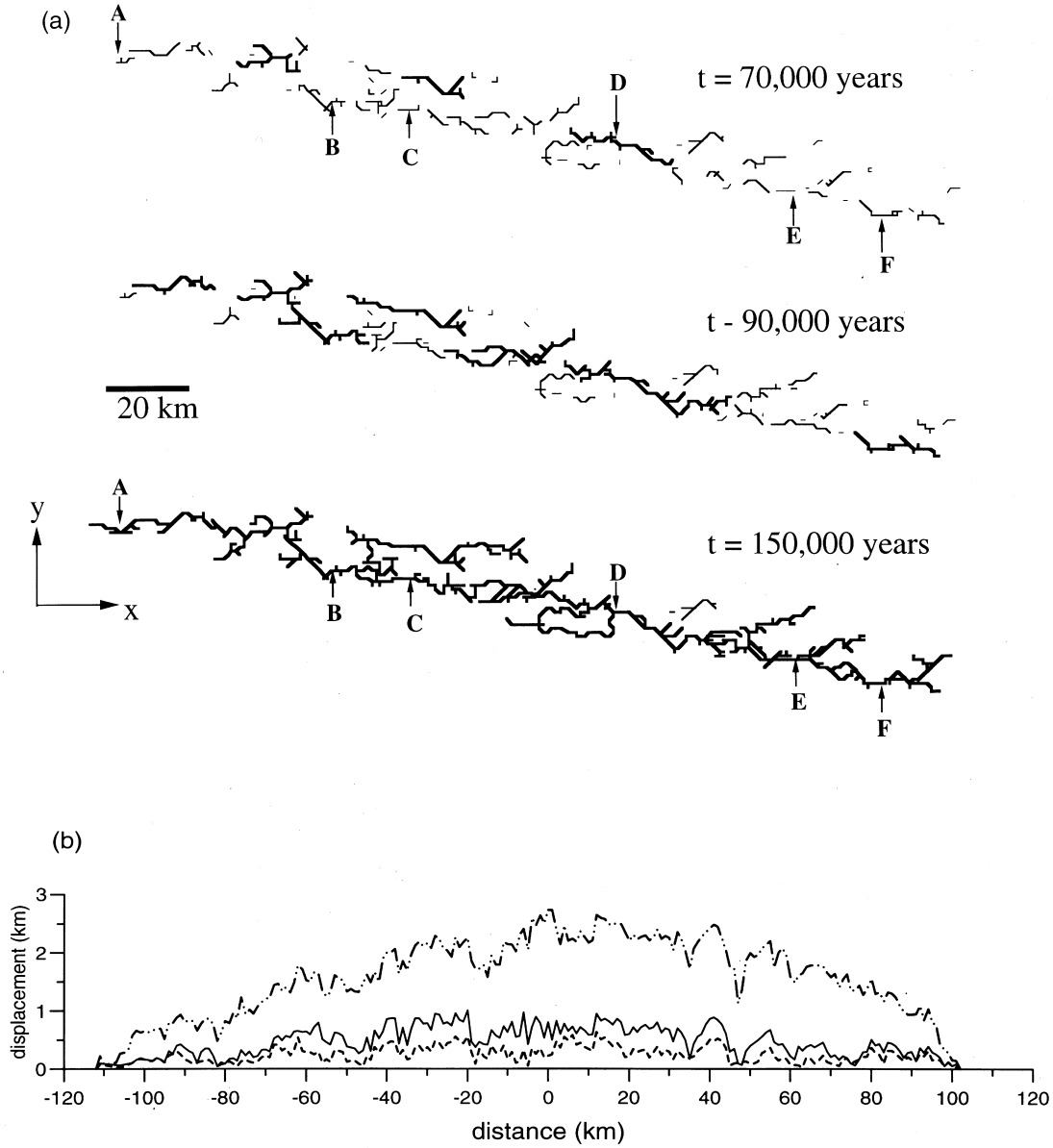

Fig. 7. Evolution through time of a fault zone produced using the numerical rupture model. (a) Fault trace map at three different stages of the evolution with bold lines indicating the largest fault segments at each stage. (b) Displacement variation as a function of distance along the fault zone for the same three times shown in (a). Where there is more than one fault strand at any position the displacement is the sum of the values on each strand.

DISCUSSION

Fault healing

A key assumption in the feedback mechanism described here is that ruptured faults recover some or all of their strength and are thus able to support stress perturbations resulting from subsequent ruptures. Heimpel (1997) shows, using a numerical model, that frictional shear strength of a rough fault gradually increases after earthquake rupture due to interlocking of asperities. Cataclastic deformation inherently leads to an evolving population of asperities as the fault moves. Although complete strength recovery is not found to be instantaneous in Heimpel's (1997) model, the most significant increase occurs shortly after rupture has stopped. Recent experimental results (Marone, 1998) show that immediate post-seismic healing may be retarded for a period of several hundred days due to compaction effects, but on longer time scales the rate of strength increase is of the order of 3–6 MPa per decade. Chemically assisted healing mechanisms may also play a role in natural faults, such as the precipitation of quartz and calcite and annealing of microcracks (Bruhn *et al.*, 1994). Annealing under geological conditions is thought to occur relatively rapidly, i.e. within a few months (Smith and Evans, 1984; Brantley *et al.*, 1990).

These strength recovery times are long compared to the interval between most triggered seismic events, for example, compound earthquakes which may be separated by only a few minutes or a few hours (e.g.

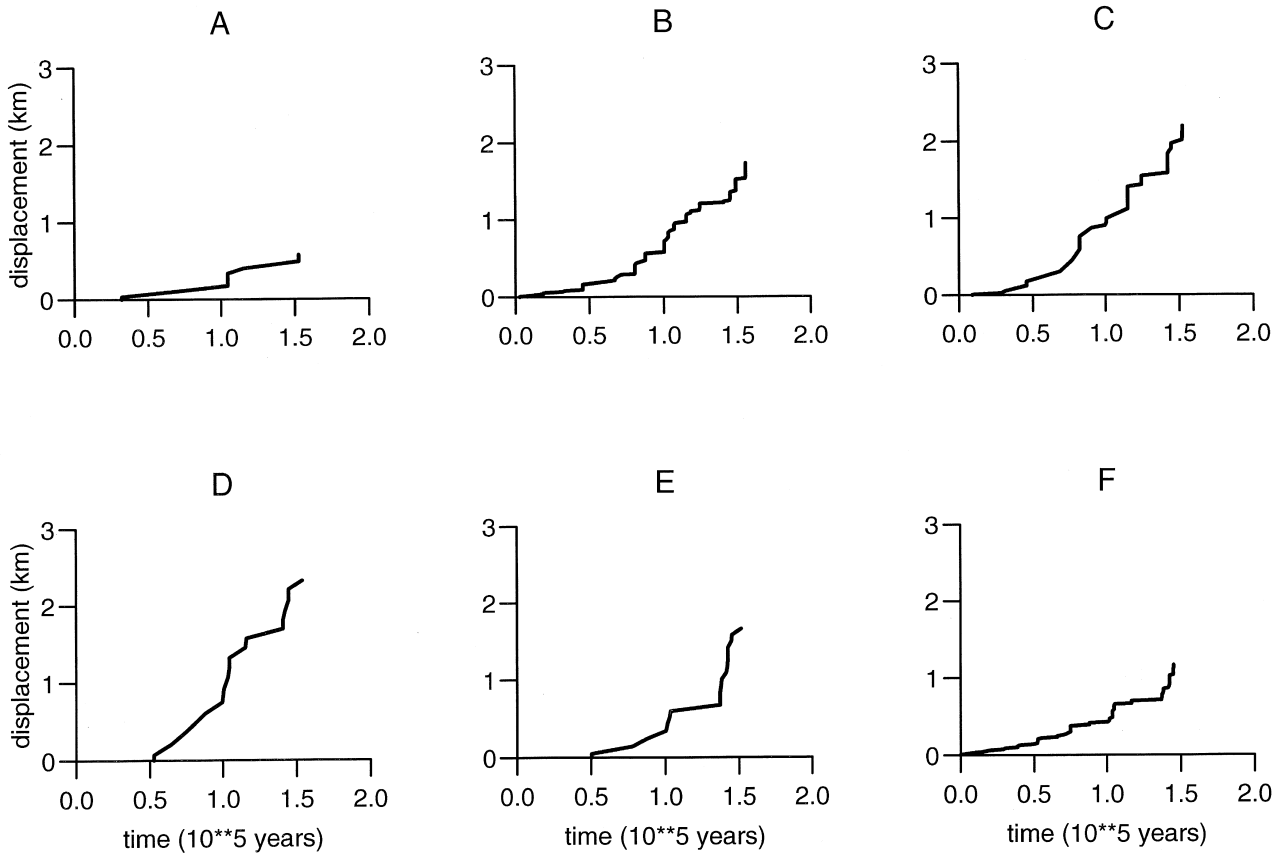

Fig. 8. Cumulative displacement as a function of time at six different points along the fault zone shown in Fig. 7. See text for discussion.

Caskey and Wesnousky, 1997). Other sequences of events, thought to be triggered by neighbouring ruptures, are separated by time intervals on the order of years (King *et al.*, 1994). Although some fault healing may have occurred over this time interval, the effects of reloading within a triggered sequence of earthquakes are not considered. This is because the amplitude of any stress recovery due to stress feedback is very small compared to the drop in stress which occurs during the earthquake. The impact of the healing–reloading feedback mechanism requires a longer time period to manifest itself, and thus data on the intermediate- and long-term history of fault displacement accumulation are required. Several lines of field evidence for this mechanism are presented and discussed below.

Enhanced vs inhibited growth

The positive correlation between displacement rate and fault size, documented by Nicol *et al.* (1997) for extensional faults, appears to provide support for the stress feedback mechanism. These authors analysed displacement rates on normal faults using stratigraphic data from six rift basins around the world. They found that large faults exhibit high displacement rates throughout their development, whereas smaller faults, which initiated at the same time, never achieved high rates. The explanation proposed here, and illustrated in Fig. 9, is that the high displacement rate faults are those that initiated in an optimal location and remained optimally positioned with respect to other active faults as the deformation progressed.

In Stage 1 of Fig. 9, several isolated faults begin to nucleate at approximately the same time at points where the crust is locally weak. The growth rates of the individual faults are similar in this first stage (Fig. 9a). During Stage 2, enhanced displacement rates develop on those faults which are optimally located with respect to neighbouring faults and thus preferentially reloaded (W, X, Y and Z, Fig. 9b). These faults grow more rapidly, even while remaining relatively isolated. Ultimately linkage leads to larger through-going faults. As deformation progresses, this pattern of enhanced growth rates and linkage may be repeated at larger scales as larger structures interact. During Stages 1 and 2 new faults may continue to nucleate but, depending on their location relative to other faults, their growth may be strongly inhibited and/or they soon become inactive. Some faults in the population that initially experience enhanced growth may subsequently become inactive when they cease to be optimally located within the overall population (X and W, Fig. 9c). They become inhibited by stress shadows

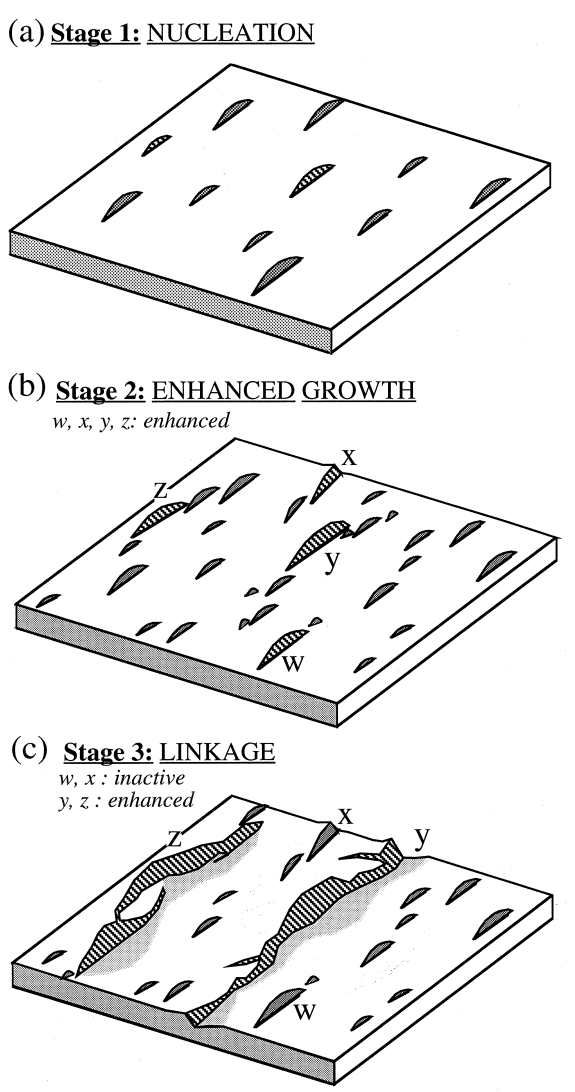

Fig. 9. Block diagram illustrating the evolution of a fault pattern for which the healing–reloading stress feedback mechanism plays a dominant role, based on results of the numerical modelling. (a) Stage 1: initial nucleation of many isolated faults. (b) Stage 2: enhanced growth of those faults which are optimally located with respect to neighbouring structures (W, X, Y, Z). Suppressed growth of faults which are not optimally located. (c) Stage 3: localisation of deformation onto a few through-going faults (Z, Y) and activity ceasing on faults in intervening regions (X, W).

generated by more optimally located fault arrays. This interpretation modifies the traditional view that it is the orientation of a fault with respect to the regional principal stress directions that determines whether it continues to move. The location of a fault with respect to its neighbours also strongly influences whether a fault is active and the rate at which it grows. It also explains why faults of all sizes develop in regions of active deformation, even in simple tectonic settings.

Displacement rate variations along linked fault arrays

Several studies of segmented normal faults have shown that the maximum displacement of the entire zone often occurs along the central segment. In some cases the segments have similar displacement–length ratios, but the central segment may be the longest. In other cases, perhaps more commonly, the central segments have displacement–length ratios that can be as much as a factor of four larger than the distal segments (e.g. Peacock and Sanderson, 1991; Dawers and Anders, 1995; Willemse *et al.*, 1996). Willemse *et al.* (1996) explained the higher displacement–length ratios in terms of static elastic interaction. Using the present model, it is possible to show how, and at what stage of the evolution, these patterns emerge as a fault system grows.

Fault growth by linkage, as described by Cartwright *et al.* (1995), assumes that most of the re-adjustment of the displacement profile occurs subsequent to complete ('hard') linkage. According to their model, recently linked faults are 'under-displaced' relative to isolated faults or those that linked at an earlier point in time. The present work suggests that significant adjustment of displacement profiles takes place while faults still remain relatively isolated. This is achieved via preferential growth of those faults located at points of *rupture symmetry* in the evolving population. The central segment of an en échelon fault array is by definition a point of rupture symmetry (Fig. 9). This segment may achieve a greater length and thus have a larger displacement before any structural linkage starts to take place (Fig. 9). Further modification will occur if tip interaction inhibits further propagation, resulting in higher displacement–length ratios for central segments and asymmetric profile shapes (Peacock and Sanderson, 1991; Willemse *et al.*, 1996).

The fault array mapped by Dawers and Anders (1995) supports the idea that displacement profiles begin to re-adjust prior to linkage. They found that the overall displacement–length ratio for the entire array, which one could argue is in the early stages of linkage, is similar to that expected for self-similar growth of an isolated fault in the same rock type, i.e. it is not noticeably 'under-displaced'. Even if, subsequent to hard linkage, a fault is relatively 'under-displaced', the displacement profile will gradually readjust by variable rates of movement along strike. Clearly variable growth rates are likely to be an integral part of growth by linkage and thus a widespread phenomenon.

In the Basin and Range province of the USA the Wasatch fault zone shows a gradual increase in Holocene displacement rates, by a factor of at least five, from the distal segments towards the centre (Machette *et al.*, 1991). The Teton fault shows a similar pattern with a rate, averaged over the Quaternary, that increases by a factor of 2.5 along the central portion of the fault (Smith *et al.*, 1990; Byrd *et al.*, 1994). These data come from fault scarps formed by multiple earthquakes and thus do not simply reflect the slip distribution of an individual event. In both cases, the overall topography of the range front, which is related

to the cumulative displacement on the fault, mimics the variation in Holocene or Quaternary displacement rates. There is also a systematic pattern to the segment lengths, well documented for the Wasatch fault zone, with the longest segments being located along the central portion (Machette *et al.*, 1991). The comparison with model results (Figs 7 & 8) is striking, in particular the displacement profiles which are analogous to the topography of the range fronts (Fig. 7b). These patterns have not previously been explained in terms of a physical mechanism.

A rough estimate of the relative enhancement of the displacement rate at the mid-point of a fault zone may be made by dividing the total length of the zone by the length of the central segment. The enhancement is defined relative to an isolated fault with a length equal to the central segment. This calculation is based on the work of Dawers and Anders (1995) who showed that the displacement profile of a linked fault array conforms to that expected for a single fault of equivalent length. Thus for the Teton fault (total length ~60 km), which consists of three main segments, the expected enhancement factor is about ×2.4 (60 km/25 km) (based on data from Smith *et al.*, 1990). For the Wasatch fault (total length ~360 km), Parry and Bruhn (1987) found that the maximum displacement on the Wasatch fault (>11 km) probably occurs near the southern end of the Salt Lake City segment which is estimated to be 46 km long (Machette *et al.*, 1991). Thus the enhancement factor for this fault is estimated to be about ×8 (360 km/46 km). These estimates are approximately equal to the magnitude of the variation that is observed.

The geologic history of the Wasatch and Teton areas are clearly much more complex than I have implied here. The general pattern of the displacement rate increase along the central segments is, however, captured by using this simple approach. The estimates of the enhancement factor require independent criteria for defining fault zone length; for normal faults this is not too difficult because of the dramatic topography they produce. The Teton and Wasatch faults zones are some of the few for which displacement rate data along strike are available, but such patterns are likely to be widespread.

Episodic fault growth and earthquake clustering

Episodic displacement accumulation through time is a general feature of this model (Figs 6 & 8). It is a temporal manifestation of the healing–reloading feedback mechanism, producing periods of earthquake clustering and rapid growth, followed by more quiescent intervals. Such episodic behaviour is consistent with stratigraphic observations of pulsed fault-controlled subsidence (inferred on time scales of several millions of years) in extensional basins (Underhill, 1991; Ravnas and Bondevik, 1997). Using radiometric age constraints, Dorsey *et al.* (1997) interpreted a sequence of delta progradation cycles in terms of earthquake clustering with a frequency of several thousand years over a 100,000 y time period. In contrast, paleoseismological studies of active or recently active faults provide a much shorter time scale picture of individual earthquakes over periods of, at most, a few thousand years (e.g. Machette *et al.*, 1991). It is very difficult to resolve the underlying patterns of earthquake clustering using paleoseismology alone (Weldon, 1997). If clustering is significant, as suggested here, then fault displacement rates based on averaging over a few earthquake cycles may not be representative of the longer term rate of movement.

Interpreting the rheological properties of the lithosphere

The discussion above refers exclusively to elastic stress transfer horizontally within a thin plate. While this might be an adequate simplification for small upper crustal faults, large scale fault growth is influenced by the layered rheological structure of the lithosphere. Coupling between the near surface elastic–brittle layer and underlying viscous layers has been shown to play a major role in influencing the spatial pattern of faulting in continental regions (Davy *et al.*, 1995; Heimpel and Olsen, 1996). Davy *et al.* (1995) showed that the relative 'strength' of the brittle and ductile layers is the controlling parameter. The strength of the brittle layer depends on its density and its thickness, equivalent to the seismogenic thickness. The strength of the lower layer depends on its thickness, the viscosity (assumed to be linear), and the imposed strain rate. If the ductile layer is comparatively 'strong', it acts to confine the regions of stress reduction to the vicinity of a ruptured fault because part of the stress is absorbed by the lower layer. This preserves a high stress level over adjacent areas of the lithosphere allowing additional faults to nucleate (Heimpel and Olsen, 1996). As the total amount of strain increases the number of active faults increases, linkage is a subsidiary process and large faults do not develop.

In contrast, if the lower layer is 'weak' compared to the elastic–brittle layer then the deformation localises onto a few major fault zones (Davy *et al.*, 1995). The results I present, where viscous flow is not considered, are most analogous to this case. The magnitude of the stress feedback effect, in natural fault arrays, will be a function of the thickness of the seismogenic layer and also the rate of fault healing. If stress feedback is an important controlling mechanism, as it is here, then as strain increases linkage of existing faults predominates over nucleation. Linkage events are associated with a marked increase in displacement rates on active faults (Fig. 6). At the same time many other faults become inactive. Eventual localisation of all the deformation onto one or a few major through-going (plate bound-

ary) fault zones occurs after only a small amount of total strain, typically a few percent. For example, when the applied plate velocity is 10 cm y^{-1}, the deformation localises after only 1% strain, or approximately 0.1 my (Figs 3 & 6).

The brittle-to-ductile strength ratio of the lithosphere varies with heat flow and strain rate, and thus different regions are expected to exhibit different styles of fault development. Davy *et al.* (1995) have already shown that the spatial pattern of faulting and size frequency distributions may be used to elucidate this ratio. Based on the present work, it is suggested that fault displacement rate variations and the relative timing of activity (or cessation of activity) on neighbouring structures may provide more critical information. Extensional basins are most likely to yield this type of information, particularly if thick sequences of sediments with biostratigraphic age constraints are available.

Finally, a continuum mechanics approach to modelling lithospheric deformation, as used by England and Molnar (1991), is unable to account for the observations summarised above. Such models are based on the argument that the elastic–brittle part of the lithosphere responds passively to viscous flow at depth. An entirely passive role for the upper layer is incompatible with stress feedback between faults. One of the key pieces of evidence given as support for the continuum representation, is the apparent agreement between geodetically determined fault displacement rates and rates determined from geological observations (Bourne *et al.*, 1998). The latter are estimated from fault offsets accrued during recent major earthquakes. If earthquake clustering and episodic fault movement are ubiquitous, as found with this model and supported by field observations, then these rates may not be representative of longer-term displacement rates. The agreement between the geologic and geodetic measurements could be fortuitous and possibly misleading (see also Haines, 1998).

CONCLUSIONS

The healing and strength recovery of ruptured faults in zones of active deformation plays an important role in driving the localisation of deformation in the lithosphere. This is demonstrated here using a simple numerical rupture model for antiplane shear of an elastic–brittle plate undergoing uniform tectonic loading. The analogy between the model and real faults is based on the idea that static stress changes due to earthquake rupture can advance, or delay, the occurrence of earthquakes on neighbouring faults. For this process to operate, ruptured faults must recover some or all of their strength on a time scale that is geologically rapid. Experimental evidence and field data suggest that healing occurs over a few years or decades.

The numerical model shows that, during progressive elastic–brittle deformation, positive and negative stress feedbacks develop spontaneously between faults. Positive feedback occurs between faults that have mutually overlapping zones where the stress is increased due to rupture. Negative feedback occurs between faults with mutually overlapping zones in which the stress is reduced, so called 'stress shadow zones'. As a consequence of both positive and negative feedback, some faults grow rapidly exhibiting high displacement rates even during the early stages of the deformation, while other faults experience low displacement rates and may eventually become inactive. This result is born out by displacement rate studies in extensional basins (Nicol *et al.*, 1997). The high displacement rate structures in the model eventually link up to form major through-going faults which accommodate the bulk of the subsequent deformation. In this sense, the formation of a new plate boundary fault zone is a consequence of, rather than the precursor of, preferentially high displacement rates.

Agreement between natural fault arrays and the results presented here suggests that the healing–reloading feedback mechanism is an important ingredient for the development of large faults. In particular, variable displacement rates through time and episodic growth are a general feature of this model and natural faults for which observations are available. Predicted lateral variations in displacement rates of from $\times 2$ to $\times 5$ between the distal and central segments of fault zones are also consistent with field measurements from Basin and Range normal faults. Furthermore, the healing–reloading feedback mechanism provides an explanation for longer segment lengths and/or higher displacement–length ratios along the central portion of fault arrays. Such lateral and temporal variations are predicted to be a widespread phenomenon, based on the results of this work.

Temporal development of an individual fault is found to depend on whether or not it maintains an optimal geometry, i.e. positive feedback, with respect to the overall active fault population. A fault may become inactive, or be re-activated, at any stage due to its position and orientation relative to other active faults. This interpretation significantly modifies the traditional view that factors such as mechanical weakening and fault orientation relative to regional principal stresses exclusively control fault activity. Furthermore, these results explain why it is often difficult to distinguish truly 'quiescent' from 'active' faults for earthquake hazard purposes. Finally, measuring displacement rate variations along faults, on time scales of 10^4–10^6 y, provides additional constraints on the rheological properties of the lithosphere in regions of active deformation.

Acknowledgements—I thank The Royal Society of London for their continued support. Many useful discussions and criticisms of an earlier article prompted the ideas presented here, in particular comments from Bertrand Maillot, Rick Sibson, Javier Escartin and Fernando Niño. Joe Cartwright, Sanjeev Gupta, Nancye Dawers, Moritz Heimpel and Francis Nimmo made useful comments which considerably improved the discussion. The Fault Analysis Group (Liverpool) kindly provided a preprint of their Nature paper. Ron Bruhn and an anonymous reviewer are thanked for their constructive comments on the manuscript.

REFERENCES

Aydin, A. and Schultz, R. A. (1990) Effect of mechanical interaction on the development of strike-slip faults with en échelon patterns. *Journal of Structural Geology* **12,** 123–129.

Beckman, F. S. (1960) The solution of linear equations by the conjugate gradient method. In *Mathematical Methods for Digital Computers*, eds A. Ralston and H. S. Wilf, John Wiley and Sons Inc., New York, **62**.

Bourne, S. J., England, P. C. and Parsons, B. (1998) The motion of crustal blocks driven by flow of the lower lithosphere and implications for slip rates of continental strike-slip faults. *Nature* **391,** 655–659.

Brantley, S. L., Evans, B., Hickman, S. H. and Crerar, D. A. (1990) Healing of microcracks in quartz: implications for fluid flow. *Geology* **18,** 136–139.

Bruhn, R. L., Parry, W. T., Yonkee, W. A. and Thompson, T. (1994) Fracturing and hydrothermal alteration in normal fault zones. *Pure and Applied Geophysics* **142,** 609–644.

Byrd, J. O. D., Smith, R. B. and Geissman, J. W. (1994) The Teton fault, Wyoming: topographic signature, neotectonics, and mechanisms of deformation. *Journal of Geophysical Research* **99,** 20,095–20,122.

Cartwright, J., Mansfield, C. and Trudgil, B. (1995) The growth of faults by segment linkage: evidence from the Canyonlands grabens of S. E. Utah. *Journal of Structural Geology* **17,** 1319–1326.

Caskey, S. J. and Wesnousky, S. G. (1997) Static stress changes and earthquake triggering during the 1954 Fairview Peak and Dixie Valley Earthquakes, Central Nevada. *Bulletin of the Seismological Society of America* **87,** 521–527.

Cowie, P. A., Vanneste, C. and Sornette, D. (1993) Statistical physics model for the spatio-temporal evolution of faults. *Journal of Geophysical Research* **98,** 21,809–21,822.

Cowie, P. A., Sornette, D. and Vanneste, C. (1995) Multifractal scaling properties of a growing fault population. *Geophysical Journal International* **122,** 457–469.

Davy, P., Hansen, A., Bonnet, E. and Zhang, S.-Z. (1995) Localisation and fault growth in layered brittle–ductile systems: implications for deformation of the continental lithosphere. *Journal of Geophysical Research* **100,** 6281–6289.

Dawers, N. H. and Anders, M. H. (1995) Displacement–length scaling and fault linkage. *Journal of Structural Geology* **17,** 607–614.

Dorsey, R. J., Umhoefer, P. J. and Falk, P. D. (1997) Earthquake clustering inferred from Pliocene Gilbert-type fan deltas in the Loreto basin, Baja California Sur, Mexico. *Geology* **25,** 679–682.

England, P. C. and Molnar, P. (1991) Inferences of deviatoric stress in actively deforming belts from simple physical models. *Philosophical Transactions of the Royal Society of London* **A337,** 151–164.

Haines, J. (1998) Continental mechanics. *Nature* **391,** 634–635.

Heimpel, M. and Olsen, P. (1996) A seismodynamical model of lithospheric deformation: Development of continental and oceanic rift networks. *Journal of Geophysical Research* **101,** 16,155–16,176.

Heimpel, M. (1997) Critical behaviour and the evolution of fault strength during earthquake cycles. *Nature* **388,** 865–868.

Hodgkinson, K. M., Stein, R. S. and King, G. C. P. (1996) The 1954 Rainbow Mountain–Fairview Peak–Dixie Valley earthquakes: a triggered normal faulting sequence. *Journal of Geophysical Research* **101,** 25,459–25,471.

Jaume, S. C. and Sykes, L. R. (1992) Change in the state of stress on the southern San Andreas fault resulting from the California earthquake sequence of April to June 1992. *Science* **258,** 1325–1328.

Jaume, S. C. and Sykes, L. R. (1996) Evolution of moderate seismicity in the San Francisco Bay region, 1850 to 1993: seismicity changes related to the occurrence of large and great earthquakes. *Journal of Geophysical Research* **101,** 765–789.

King, G. C. P., Stein, R. S. and Lin, J. (1994) Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America* **84,** 935–953.

Machette, M. N., Personius, S. F., Nelson, A. R., Schwartz, D. P. and Lund, W. R. (1991) The Wasatch fault zone, Utah—segmentation and history of earthquakes. *Journal of Structural Geology* **13,** 137–149.

Madariaga, R. (1976) Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America* **66,** 639–666.

Marone, C. (1998) The effect of loading rate on static friction and the rate of fault healing. *Nature* **391,** 69–72.

Nicol, A., Walsh, J. J., Watterson, J. and Underhill, J. R. (1997) Displacement rates of normal faults. *Nature* **390,** 157–159.

Parry, W. T. and Bruhn, R. L. (1987) Fluid inclusion evidence for minimum 11 km vertical offset on the Wasatch fault, Utah. *Geology* **15,** 67–70.

Peacock, D. C. P. and Sanderson, D. J. (1991) Displacements, segment linkage and relay ramps in normal fault zones. *Journal of Structural Geology* **13,** 721–733.

Pollard, D. D. and Segall, P. (1987) Theoretical displacements and stresses near fractures in rock: with applications to faults, joints, veins, dikes, and solution surfaces. In *Fracture Mechanics of Rock 277–349*, ed. B. K. Atkinson. Academic Press.

Ravnas, M. and Bondevik, K. (1997) Architecture and controls on Bathonian–Kimmeridgian shallow-marine synrift wedges of the Oseberg–Brage area, northern North Sea. *Basin Research* **9,** 197–226.

Smith, D. L. and Evans, B. (1984) Diffusional crack healing in quartz. *Journal of Geophysical Research* **89,** 4125–4135.

Smith, R. B., Byrd, J. O. D. and Susong, D. D. (1990) Neotectonics and structural evolution of the Teton fault. In *Geologic Field Tours of Western Wyoming and parts of adjacent Idaho, Montana, and Utah*, ed. S. Roberts, pp. 126–138. Public Information Circular, Geological Survey of Wyoming, **29**.

Sornette, A., Davy, P. and Sornette, D. (1990) Growth of fractal fault patterns. *Physical Review Letters* **65,** 2266–2269.

Sornette, D., Davy, P. and Sornette, A. (1990) Structuration of the lithosphere as a self-organised critical phenomenon. *Journal of Geophysical Research* **95,** 17,353–17,361.

Sornette, D., Miltenberger, P. and Vanneste, C. (1994) Statistical physics of fault patterns self-organised by repeated earthquakes. *Pure and Applied Geophysics* **142,** 491–527.

Stein, R. S., King, G. C. P. and Lin, J (1992) Change in failure stress in the southern San Andreas Fault System caused by the 1992 magnitude 7.4 Landers Earthquake. *Science* **258,** 1328–1332.

Underhill, J. R. (1991) Controls on Late Jurassic seismic sequences, Inner Moray Firth, UK North Sea: a critical test of a key segment of Exxon's original global cycle chart. *Basin Research* **3,** 79–98.

Weldon, R. J. (1997) Is there clustering in the paleoseismic data for the southern San Andreas fault? *Geological Society of America Abstracts with Programs* **29,** 73.

Willemse, E. J. M., Pollard, D. D. and Aydin, A. (1996) Three-dimensional analysis of slip distributions on normal fault arrays with consequences for fault scaling. *Journal of Structural Geology* **18,** 295–310.

Willemse, E. J. M. (1997) Segmented normal faults: Correspondence between three-dimensional models and field data. *Journal of Geophysical Research* **102,** 675–692.