

PROPAGATION OF MAGMA-FILLED CRACKS

Allan M. Rubin

Department of Geological and Geophysical Sciences, Princeton University,
Princeton, New Jersey 08544

KEY WORDS: dike propagation, magma transport, rock fracture, igneous intrusions

INTRODUCTION

The mechanism of magma transport at depth influences direction magma moves, the distance it travels before freezing, the degree to which it communicates chemically with the host rock, the form of surficial volcanism, and ultimately the growth of oceanic and continental crust. Commonly envisioned transport processes include porous flow in partially molten and deformable source rock, flow through fractures in elastic/brittle rock, and diapiric ascent (typically of granites) through viscous rock. Of these, transport in fractures, or dikes, is the most efficient means of moving magma through cold lithosphere. Porous flow is an option only if there has been sufficient advection of heat to raise the rock temperature above the solidus (a possibility beneath Hawaii, for example). Although the sheet-like form of dikes is less advantageous thermally than the equidimensional form of diapirs, except for possibly the most viscous rhyolites, this is more than offset by the fact that transport rates depend upon the magma viscosity, rather than the host rock viscosity.

Dike intrusion is also the transport mechanism that permits the most direct comparison between theory and observation. In part, this is because many dikes carry so little heat that magmatic and host rock structures produced during intrusion are preserved until exposed by erosion. The short timescale of intrusion also allows near-real-time seismic and geodetic monitoring of dikes in active volcanoes. In addition, the close kinship between dikes and artificial hydrofractures, used in the oil and geothermal industries, provides an economic incentive for monitoring field-scale experiments that are relevant to dike propagation. Finally, the nominally planar geometry of dikes permits idealizations that make a detailed theoretical treatment practical. While analogous equations describing porous flow in the source region have also been proposed, observations available

to constrain these theories are fewer and less direct. The situation is worse for granitic diapirism; because of the complexity of the processes envisioned, and the tendency for later deformations to obliterate earlier ones, even its existence is debated.

I begin this review by highlighting the observations of ancient and modern dikes that lay the groundwork for current modeling. The emphasis then shifts to the dominant physical processes involved—host rock fracture and deformation (elastic and inelastic), magma flow, and heat transfer. Previous review articles on some of these topics include those by Pollard (1987) for the solid mechanics, Lister & Kerr (1991) for coupling the fluid flow and elastic deformation, and Delaney (1987) for the heat flow. Where possible this review attempts to provide a framework for thinking about important but poorly understood processes such as dike initiation, the role of dike propagation in the ascent of granitic magmas, and earthquakes accompanying magma transport.

OBSERVATIONS OF FROZEN AND FLUID DIKES

Dikes are more or less planar sheets, with reported thickness:length aspect ratios generally in the range 10^{-2} to 10^{-4} (Figure 1). They are found in a wide variety of tectonic settings and encompass a wide variety of magma compositions, but on a regional scale mafic dikes are far more common than their silicic counterparts [but see Hutton (1992) for evidence of granitic sheeted dike complexes]. Most dikes in the upper crust are near vertical; horizontal sheets intruded parallel to bedding (sills) are also very numerous. At the outcrop or even map scale, dikes exhibit many deviations from planarity. These include en-echelon segmentation and offset margins, some of which can be used to infer propagation direction (Pollard 1987). Matching irregularities indicate that, in most cases, the opposing walls of the dike have undergone predominantly opening displacement with little relative shear.

Many dikes are found in linear or radial swarms that may extend for tens or even hundreds of kilometers. Reported average thicknesses of mafic dikes range from about 10 cm in mantle peridotites (Nicolas 1986), to 1 m in eroded Hawaiian volcanoes and the sheeted dike complexes of several ophiolites (Walker 1987, Kidd 1977), to 4 m in the Tertiary swarms of Scotland and Iceland (Speight et al 1982, Gudmundsson 1990), to about 30 m in numerous Proterozoic basaltic dike swarms in continental shields (Halls & Fahrig 1987). The Proterozoic swarms are truly remarkable in that dike thicknesses of over 100 m are common and several dikes are more than 500 km long and nearly 1 km thick. Even if the vertical extent were only a few kilometers, such dikes would contain more basalt than the largest documented lava flows. Silicic dikes are generally thicker and less extensive than their mafic counterparts.

Much has been learned about dike propagation from monitoring the intrusion of (relatively thin) basalt dikes in active volcanoes. At Krafla in Iceland, and

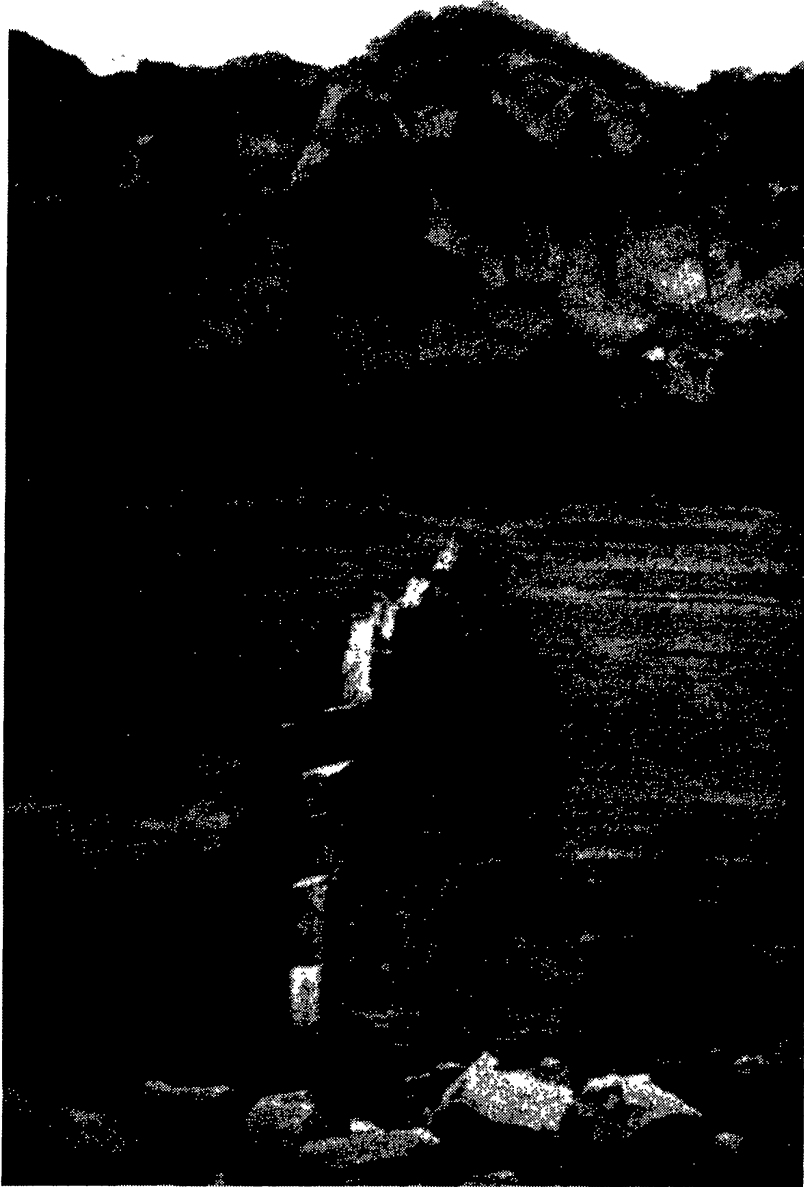


Figure 1 Photograph of a basaltic dike at tip of Reykjanes Peninsula, southwest Iceland. Dike thickness is 40 cm. Discernible radius of curvature at top indicates small-scale inelastic deformation of host (basaltic tephra). En-echelon segment is just visible at bottom. This dike did not make it to surface.

often at Kilauea in Hawaii, intrusion begins with subsidence of the caldera region at volumetric rates of 10^2 to $10^3 \text{ m}^3 \text{ s}^{-1}$, sustained harmonic tremor, and downrift migration of a swarm of earthquakes at some tens of centimeters per second (Decker 1987, Einarsson & Brandsdottir 1980, Klein et al 1987). Although the precise cause of the earthquakes is unknown, migration of the swarm is presumed to reflect the propagation of the dike at depth. Limited focal mechanism determinations for these events (Karpin & Thurber 1987),

and more numerous determinations for earthquakes associated with artificial hydrofractures (Bame & Fehler 1986, Talebi & Cornet 1987) indicate that the vast majority are typical double-couple earthquakes indicative of shear failure. The observed propagation velocities are within the range 0.01 to 10 m s⁻¹ inferred for mantle-derived dikes, based on (a) the computed settling rates of entrained xenoliths, and (b) laboratory-derived rates of depressurization reactions, such as the reversion of diamond to graphite (Spera 1984).

Shallow intrusions produce surface displacements that may be monitored geodetically and inverted for the geometry of the dike at depth (e.g. Pollard et al 1983). Displacements of tens of centimeters (somewhat less than the dike thickness) extend up to several kilometers (somewhat greater than the dike depth) from the dike plane. Recently, continuously recording tilt and strain meters have been able to provide time histories of dike growth in some cases (Okamura et al 1988, Okada & Yamamoto 1991, Linde et al 1993). Together, the seismic and geodetic data paint a picture of blade-like dikes of order 1 m thick and a few kilometers high propagating laterally for up to several tens of kilometers, while erupting only locally or not at all (see also Sigurdsson & Sparks 1978).

Evidence for subhorizontal propagation at shallow depths also comes from studies of magmatic flow fabrics in several eroded dike swarms (Knight & Walker 1988, Smith 1987). Potential flow indicators include elongate vesicles, aligned phenocrysts, and anisotropy of magnetic susceptibility that results from late crystallization of magnetic minerals in a template of aligned phenocrysts (Hargraves et al 1991). In a recent spectacular example, Ernst & Baragar (1992) used magnetic fabric to determine magma flow directions in the 1.2 Ga Mackenzie radial dike swarm, which extends over 2500 km across the Canadian shield. They found an abrupt change from predominantly vertical flow within 500 km of the source, to essentially lateral flow beyond 600 km. While flow indicators record only the flow direction at the time they are "frozen in," in the few cases that observations are available, this flow direction seems to be consistent with the fracture propagation direction inferred from striations on the fracture walls and the geometry of dike segmentation (Gartner 1986, Smith 1987).

Inelastic deformation is sometimes observed to extend from tens of centimeters to tens of meters from eroded dikes. These structures, inferred to reflect passage of the dike tip, include closely spaced dike-parallel joints (Delaney et al 1986), breccias (Johnson & Pollard 1973), faulted and folded strata (Pollard et al 1975), and fluidized clastic rocks (Baer 1991). This deformation is analogous to, but more extensive than, "process zone" deformation produced by the stress concentration at crack tips in laboratory rock fracture experiments. Still more extensive zones of inelastic deformation, consisting predominantly of vertical tension cracks and slip along existing normal faults, are produced above shallow subsurface dikes (Pollard et al 1983, Rubin 1992, and references therein). Such deformation is attributable to a combination of the stress concentration due to

the free surface and the low confining pressure. An obvious and unresolved question is the relation between the observed inelastic deformation and the earthquakes accompanying intrusion.

When a basaltic dike intersects the surface and produces a “curtain-of-fire” fissure eruption, it is generally only a matter of hours or days until the eruption has ceased or localized to one or a few point sources (Delaney & Pollard 1982). These plugs may erupt for days or (intermittently) for years. In Hawaii, effusion rates during the curtain-of-fire stage are about $100 \text{ m}^3 \text{ s}^{-1}$ per kilometer of fissure.

SOLID MECHANICS

Dike Thickness

The low in situ compressive strength of rock—about an order of magnitude less than laboratory values (Bieniawski 1984)—precludes the existence of sizable voids at depths greater than about 1 km. Therefore, dikes must widen their channels even if they don’t fracture the rock. Away from the dike tip, strains are on the order of the dike thickness:length ratio and are therefore small. This suggests that, excepting the tip region, deformation of the host rock may be largely elastic, and motivates treating dikes as cracks in a linear elastic body (see e.g. Delaney & Pollard 1981).

Consider a two-dimensional dike, very long in one in-plane dimension relative to the other, subjected to an ambient compressive stress $\sigma_{yy}^0(x)$ perpendicular to the dike plane and an internal magma pressure $P(x)$ (Figure 2). The perturbation to the host rock stresses and displacements due to dike opening depend only upon the difference $P(x) - \sigma_{yy}^0(x)$. For uniform loading stresses, the dike thickness profile is elliptical and the half-thickness w at the center is

$$w(0) = \frac{(P - \sigma_{yy}^0)}{\mu/(1 - \nu)} l \equiv \frac{\Delta P}{M} l, \quad (1)$$

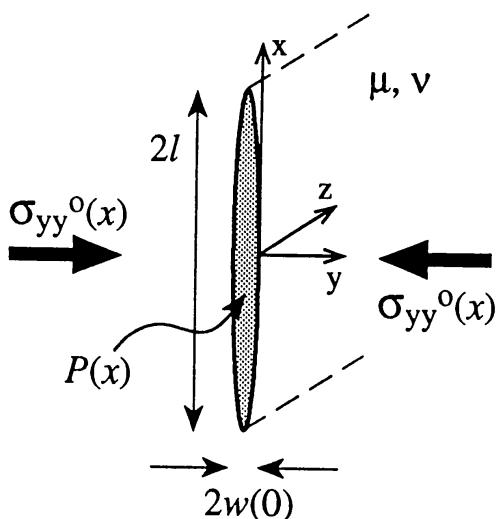


Figure 2 Thickness $2w$ of a two-dimensional dike of length $2l$. Ambient stress $\sigma_{yy}^0(x)$ acts perpendicular to dike plane; magma pressure $P(x)$ acts within the dike. Elastic displacements depend upon $\Delta P(x) \equiv P(x) - \sigma_{yy}^0(x)$.

where the pressure ($P - \sigma_{yy}^0$) available to deform the dike walls is defined as the excess or elastic pressure ΔP , and the resistance of the host rock to deformation $\mu/(1 - \nu)$ is defined as the elastic stiffness M , where μ is the elastic shear modulus and ν is Poisson's ratio (Pollard 1987). Equation (1) is approximately correct even for more complex geometries provided l is interpreted as the short in-plane dimension; for example, the height of a shallow blade-like dike. For a propagating dike the internal pressure is not uniform, and a more complete version of (1) that relates the spatially varying thickness to a spatially varying excess pressure is required (Spence & Turcotte 1985, Lister & Kerr 1991, Rubin 1993a).

Given observations of the dike thickness:length ratio, Equation (1) can be used to estimate the ratio of some average excess pressure to the elastic stiffness. From geodetic and field observations this typically gives rise to values of order 10^{-3} , assuming the measured outcrop length to be close to the relevant (minimum) in-plane dimension (Pollard 1987). For stiffnesses typical of laboratory samples of crystalline rock (~ 40 to 50 GPa), this ratio implies an excess pressure of about 40 – 50 Mpa. However, because large rock masses contain pervasive fractures not sampled in typical laboratory specimens, the in situ elastic stiffness is often significantly less, perhaps by an order of magnitude (Bieniawski 1984, Rubin 1990). This implies typical excess pressures of a few to 10 Mpa, although considerable variation is to be expected. Laboratory values of M might be approached at great depth, provided the pore pressure is not close to lithostatic values, but estimates from seismic velocities are likely to be too high.

Nonelastic processes may also contribute to the dike thickness. These include mechanical erosion of the wall rock (Delaney & Pollard 1981), thermal erosion (Bruce & Huppert 1990), distributed faulting (Rubin 1992), and viscous deformation of the host rock on the timescale of intrusion or solidification (Sleep 1988, Rubin 1993a). In a limited sense these processes may be thought of as reducing the "effective" stiffness of the host rock. Depending upon one's interest, however, it may be important to distinguish between a lower, truly elastic stiffness resulting from pervasive fractures, and a lower effective stiffness resulting from fault slip or viscous deformation on the timescale of intrusion. For example, elastic models in which a finite batch of magma rises buoyantly rely on elastic closure of the dike walls at depth to force the magma upward. Because fault slip is largely irreversible, such closure may be impossible if inelastic deformation has enlarged the dike cavity. In addition, any time-dependent deformation of the host rock provides a potential mechanism for distinguishing dikes of different viscosity (e.g. basalts vs rhyolites), because of the different timescales of intrusion.

For future reference it should also be noted that as long as it is appropriate to neglect gravitational forces on the solid in Equation (1), magma within a dike does not sense the density of the host rock, but only the ambient stresses acting on the dike plane. This is important when considering the influence of magma

buoyancy on dike orientation or propagation direction (up, down, or sideways). From (1), a measure of the elastic stress resulting from dike opening is $M(w/l)$. An estimate of the stress change due to gravity, on the length scale of the dike, is $\rho_m g w$, where ρ_m is the magma density. Equating the two stress magnitudes indicates that body forces are important when $\rho_m g l > M$, or when $l > 1000$ km for $M = 30$ GPa. This is sufficiently large for body forces (on the solid) to be negligible for dikes on the terrestrial planets, and distinguishes dike propagation from transport mechanisms involving large-scale viscous deformation of the host rock (e.g. diapirism and porous flow). Gravity is important in determining the magma pressure P that acts on the solid.

Rock Fracture

In discussions of tensile crack propagation it is useful to distinguish three regions: the crack, where the material faces are completely separated; the intact host material, where deformation is essentially elastic (or is adequately described by some other continuum model); and an intermediate region surrounding the crack tip, the “process zone,” where the strength of the material has been exceeded but where inelastic deformation has not proceeded to the point that the region has become part of the crack.

If a macroscopic crack does not exist and the material is subjected to uniaxial tension, the appropriate fracture criterion is that the tension exceed the *tensile strength* of the material. This strength is sensitive to the characteristic size of inherent (i.e. grain-scale) flaws. If the solid is under compression but contains a pore fluid, then a generally accepted tensile failure criterion is that the pore fluid pressure exceed the least compressive stress by some tensile strength (this strength could be less than in the dry case and time dependent, if the fluid and solid are chemically reactive). If a macroscopic crack exists, then the stress concentration at the tip allows propagation to occur at a lower applied tension or excess pore pressure. However, in general, it is not sufficient to specify a critical stress that must be reached at the tip, because this (concentrated) stress must act over a sufficiently large region (on the order of centimeters for rock) that the crack surfaces separate completely (Ingraffea 1987). Such details of the fracture process may be exceedingly complex. However, if the process zone is small compared to crack length, then a standard approach in fracture mechanics is to neglect this complexity, treat the host medium as linear elastic, and embody the material's resistance to fracture in a single parameter to be measured in the laboratory.

Griffith (1920) postulated that a crack would propagate if the accompanying release of potential energy was sufficient to provide the energy necessary for fracture. For dikes, the potential energy consists of elastic strain energy plus any work done on the host rock by the magma or gravity. In equation form, propagation occurs when $G = G_c$, where G is the potential energy release per

unit increase in crack length, and G_c is the critical value of G required for propagation. Griffith envisioned G_c to be the surface energy of the material, or equivalently, the work required to separate two planes of atoms (about 1 J m^{-2} for silicates). Subsequent measurements have shown that G_c for rock exceeds this value by about 2 orders of magnitude (Atkinson & Meredith 1987b). The difference is apparently due to a grain-scale zone of microcracking surrounding the crack tip (Swanson 1984). An important result of modern fracture mechanics is that G is uniquely determined by the instantaneous elastic stress field surrounding the crack tip, despite its being defined on the basis of a global energy change (Rice 1968).

A fracture criterion that is more commonly encountered in the geological literature, but that is less general than $G = G_c$, is that propagation occurs when $K = K_c$, where K is the crack-tip stress intensity factor, and K_c is the rock fracture toughness (Lawn & Wilshaw 1975). This is the realm of Linear Elastic Fracture Mechanics (LEFM). Like G , K depends upon the crack size and loading configuration; like G_c , K_c is generally considered to be a material property to be determined in the laboratory. K is a measure of the (hypothetical) stress singularity at the crack tip—a result of the mathematical idealization of cracks as sharp slits. At a distance r from the tip, the elastic stresses are proportional to $K/r^{1/2}$ plus other finite terms that depend upon the ambient stress, crack size, and $P(x)$. Because of the $1/r^{1/2}$ singularity, sufficiently near the tip the stresses are closely approximated by the $K/r^{1/2}$ term; this region is known as the *K-dominant region*. The rationale for using $K = K_c$ as a propagation criterion is that whether a crack extends is determined by the local stress field at the crack tip, and that sufficiently close to the tip this stress field is uniquely determined by K .

Real materials cannot withstand infinite stresses and instead deform inelastically near the crack tip. Nonetheless, $K = K_c$ is a valid propagation criterion, equivalent to $G = G_c$, provided that the region of inelastic deformation is small compared to the *K-dominant region* (Rice 1968, Lawn & Wilshaw 1975). A corollary is that if this restriction is met, materials have a unique fracture energy because the process zone sees only the material-dependent K_c -dominant stresses and is unaware of the stresses beyond this region that may vary from crack to crack. Because the *K-dominant region* is small compared to crack size, a necessary (but not sufficient) criterion for this restriction to be met is that the process zone be small compared to crack size.

When LEFM is applicable, K and G are related through $G = \pi K^2/2M$. If the crack in Figure 2 is subjected to a uniform excess pressure ΔP , then

$$K = \Delta P \sqrt{l}. \quad (2)$$

Again, for a propagating dike a more complete version of (2) that allows for

a spatially varying excess pressure is required (Lister & Kerr 1991, Rubin 1993a). Laboratory measurements at atmospheric pressure and temperature indicate that K_c is of order 1 MPa m^{1/2}. High-pressure experiments seem to indicate an increase in the apparent K_c with pressure (Atkinson & Meredith 1987b). However, note that the confining pressure would be irrelevant if the process zone were truly confined to the K -dominant region, because the $K_c/r^{1/2}$ stresses would dominate the ambient stress in the region surrounding the process zone.

If laboratory values of the fracture energy were applicable to dikes, then the fracture process would not significantly affect the propagation velocity of dikes of reasonable size. This is because much more energy would be consumed by magma flow than by rock fracture (e.g. Stevenson 1982, Spence & Turcotte 1985). Imagine that the dike length l in Figure 2 is such that K exceeds K_c slightly. Then the crack will propagate and magma will flow toward the tip. It cannot flow too fast, however, or the associated drop in pressure toward the tip would drop K below K_c . In this sense the propagation speed is limited by the fracture toughness. If l is much larger so that $\Delta Pl^{1/2} \gg K_c$, however, then small pressure perturbations (relative to ΔP) can produce large changes in K (relative to K_c). This in turn implies that the fracture criterion can be met with negligible changes in dike thickness and propagation velocity. For $\Delta P = 1$ MPa, $\Delta Pl^{1/2} > K_c$ if l exceeds a few meters.

These conclusions are predicted on the assumption that G_c is independent of crack size, as predicted by LEFM. Based on the observations of larger-than-lab process zones associated with some dikes, this assumption is questionable. A resolution to this apparent paradox can be found by considering a fracture model introduced by Barenblatt (1962). He idealized the fracture resistance of a material as an internal cohesive stress σ_c that resists separation of the two crack faces very near the tip. In rock, the cohesive stress seems to arise from grains that bridge the opposing surfaces (Swanson 1984); laboratory experiments suggest that it is on the order of the tensile strength—roughly 10 Mpa (Ingraffea 1987). Rubin (1993c) showed that the process zone of a propagating crack is embedded within a K_c -dominant region only if σ_c is the largest stress scale in the problem. If the confining pressure exceeds σ_c , as it does at depths of less than 1 km in the Earth, then both the ambient stress and the pressure distribution within the crack contribute significantly to the near-tip stress field. Therefore, LEFM is not applicable and one should speak of a fracture energy G_c for dikes, rather than a fracture toughness K_c . Furthermore, because inelastic deformation depends upon the near-tip stress field, it follows that G_c may depend upon the dike, as well as the rock and the ambient stress.¹

¹For historical reasons it may be difficult to abandon K_c in favor of G_c , and given the correspondence between the values of G and K computed assuming linear elasticity, it may be unnecessary as well. In this paper I use K where convenient. In the future, however, one might consider

A form of fracture that is likely to be very important in dike initiation occurs when larger-than-grain-scale cracks exist, but G is less than the nominal fracture energy G_c . In this case, propagation can proceed at a reduced rate if atomic bonds at the crack tip are weakened by chemical reaction; in a sense this amounts to a reduction in surface energy. This is the realm of subcritical crack growth (see Atkinson & Meredith 1987a for a review). Subcritical crack growth rates are very sensitive to temperature, fluid and rock chemistry, and the proximity of G to the nominal G_c . Over part of this regime the growth rate is limited by reactions at the tip; over part it is limited by transport of reactive species to and from the tip. For this reason growth rates may also depend upon fluid viscosity, diffusion rates, and confining pressure. To conduct the experiments needed to fully characterize relevant parameter space would be a daunting task. If one may judge from cartoons in the literature, subcritical propagation might be important for dikes of modest size that get stuck within the lithosphere but do not freeze completely.

Dike Propagation Direction

Understanding what controls dike propagation direction is important both for inferring ancient or modern stress fields from dike trends (e.g. Nakamura et al 1977) and for predicting dike orientation in forward models. The most basic (and classical) question is whether dikes invade existing fractures—in which case dikes of any orientation might result, or produce their own—in which case dikes are expected to propagate more or less perpendicular to the least compressive stress in the host rock (Anderson 1936). The answer appears to be that they do both, but that determining which occurred in any particular case is not always straightforward (e.g. Delaney et al 1986).

A necessary condition for a dike to invade an existing crack (on the length scale of the dike) is that the magma pressure exceed the ambient compression perpendicular to the crack. If the magma pressure is only slightly greater than the least compressive stress, then only cracks nearly perpendicular to the least compressive stress can be dilated; if the magma pressure exceeds the greatest compressive stress, then cracks of any orientation can be dilated (for details see Delaney et al 1986). However, a dike may not be able to follow such a fracture for long. If the fracture is not aligned with the principal stresses, then there is an ambient resolved shear stress on the crack that is reduced to essentially zero

emphasizing the parameter with physical significance (G_c) rather than the one whose units include the square-root of meters (K_c). This still leaves open the question of what value to adopt at high confining pressure. High-pressure experiments on samples saturated with fluid at the pressure of the fluid within the crack seem to show no dependence of K_c on confining pressure (Atkinson & Meredith 1987b). This is what one might expect if the pore pressure acts within the process zone, and this is probably the most pertinent loading configuration for dikes that are on the verge of propagation; that is, dikes in a partial melt for which $\Delta Pl^{1/2} \approx K_c$. The issue of the appropriate fracture energy of dikes for which $\Delta Pl^{1/2} \gg K_c$ can then be addressed separately.

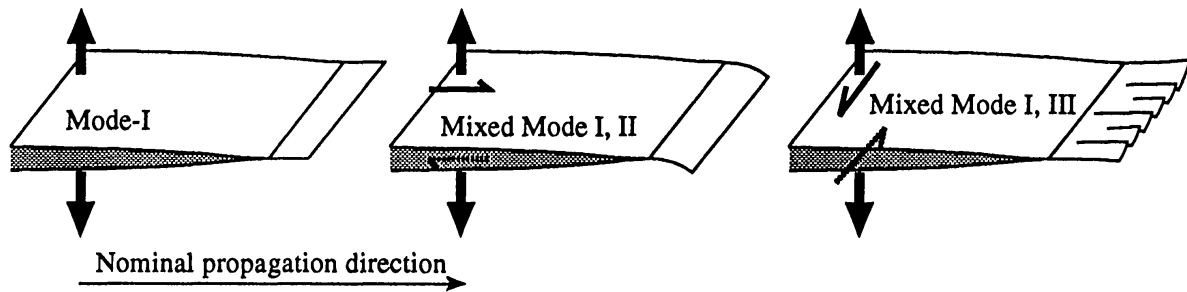


Figure 3 Schematic illustration of crack propagation paths under mixed mode loading. Mode I (tensile) loading: Crack propagates in plane. Mixed Mode I, II loading (shear perpendicular to local crack front): Crack front propagates perpendicular to local tension direction. Mixed Mode I, III loading (shear parallel to local crack front): Crack front breaks down into en-echelon segments.

by intrusion. This produces a shear stress concentration at the dike tip, and slip for some distance along the closed crack ahead of the dike. If the frictional resistance to this slip is sufficient, the resulting stress concentration at the dike tip could cause the rock to break in tension along oblique fractures, in much the same way that tensile “wing cracks” frequently form at fault tips. The dike could then hop out of the existing fracture and into a direction controlled by the ambient stresses.

If a dike fractures intact rock, then the near-tip stress field determines the direction of propagation, as well as whether or not propagation occurs. Abundant experimental evidence demonstrates that a straight crack perpendicular to an applied uniaxial tension propagates in its own plane. Similarly, hydrofractures within pressurized laboratory samples propagate as planar features perpendicular to the least compressive stress (Haimson & Fairhurst 1969). If there is some resolved shear stress on the crack plane, it is convenient to distinguish between *in-plane* (Mode-II) shear stresses perpendicular to the crack front, and *out-of-plane* (Mode-III) stresses parallel to the crack front (Figure 3).

Various theories have been proposed for the propagation of cracks subjected to a mixture of in-plane normal and shear loads. Each predicts that a crack propagating in a smoothly varying stress field follows a smooth trajectory along which the Mode-II stress intensity factor, K_{II} , is zero (Cotterell & Rice 1980). That is, with the exception of abrupt kinks that result from abrupt temporal or spatial changes in load, the chosen propagation path is that which leads to zero shear displacement of the crack walls at the crack tip. This is also this path that maximizes the potential energy release rate G , and is the direction of maximum circumferential tension at the crack tip (conceivably G_c could be minimized along this path, but this proposition is difficult to test). This leads to crack propagation paths that mimic trajectories of the ambient greatest compressive stress in the host rock, although if these trajectories are curved the correspondence is not exact because growth of a curved crack alters the principal stress

directions at the tip. The expected qualitative agreement between dike trend and ambient stress was exploited by Muller & Pollard (1977), who interpreted the curved dikes in the Spanish Peaks swarm, Colorado, as resulting from the superposition of a radial stress field due to inflation of the magma source, and an anisotropic remote stress. A few calculations of crack path have accounted for how crack growth modifies the stress state [e.g. the development of the typical "hooked" pattern as two en-echelon cracks grow and overlap (Olson & Pollard 1989)]. Because these calculations assume a uniform internal pressure, rather than a constant K_I , they are most appropriate for subcritical crack growth.

Because the propagation direction that maximizes G is that for which $K_{II} = 0$, the direction of gravity influences dike orientation only insofar as it contributes to the shear displacement of the dike walls. It was noted previously that gravitational body forces on the solid are generally negligible during dike propagation. However, sufficiently small differential stresses can also be negligible. Given a differential tension $\Delta\sigma$, misalignment of a dike with respect to the principal stresses gives rise to a resolved shear stress along the dike of order $\Delta\sigma$. Misalignment with respect to gravity (an inclined dike) gives rise to an unbalanced normal stress on the opposite dike walls of order $\rho_m g w$. Comparing the two stresses and writing $(\Delta P/M)l$ for w indicates that the contribution to K_{II} from gravity could be important if $l > (M/\rho_m g)(\Delta\sigma/\Delta P)$. This differs from the earlier estimate of 1000 km by the factor $(\Delta\sigma/\Delta P)$, indicating that gravity could influence dike orientation if the differential stress were many orders of magnitude less than the excess pressure. Differential stresses so low do not seem worth considering.

The substantial independence of dike propagation direction from the direction of gravity distinguishes dike ascent from porous flow in the source region and diapirism; in both these cases the energy for magma ascent is provided by the larger decrease in gravitational potential as the host flows around the magma.² Sleep (1984) and Davies & Stevenson (1992) suggest that the stresses associated

²The mechanical energy release rate G for dikes depends upon the instantaneous stress field at the tip, and not upon how the elastic or gravitational energies individually evolve during growth. However, it is noteworthy that the energy for dike ascent is provided by a decrease in gravitational potential of the magma-rock system only if Poisson's ratio ν is 0.5, that is, if the rock is incompressible. For each increment of dike rise in this case, the mass of rock displaced downward exceeds the mass of magma displaced upward by the ratio of the host rock to magma densities. For a Poisson's ratio typical of rock (0.25), on the other hand, a high density halo of compressed rock rises with the dike. In this case, the individual contributions to the energy balance depend upon the nature of the far-field boundary conditions; however, in all cases I have examined, the mass of rock displaced downward is less than the mass of magma displaced upward. Thus, the gravitational potential of the magma-rock system actually increases during dike ascent; ascent is driven by the larger reduction in stored elastic strain energy in the rock as the dike moves from a region of higher to lower confining pressure. Ultimately, of course, the increased confining pressure at depth is due to the weight of the rock; in this indirect fashion dike ascent may be thought of as resulting from magma buoyancy.

with mantle flow beneath mid-ocean ridges and above subducting slabs lead to nonvertical ascent of dikes, such that off-axis seamounts and subduction zone volcanoes may be far removed laterally from their sources. In principal, nonvertical ascent of melt by porous flow can be also driven by tectonically induced nonhydrostatic pressure gradients in the host rock; however, beneath mid-ocean ridges these pressure gradients are believed to be too small by 2 orders of magnitude to be significant (Phipps Morgan 1987).

Crack propagation under mixed Mode-I/Mode-III loading is more complex than under mixed Mode-I/Mode-II loading, because the resulting fracture geometry is inherently three dimensional. In this case the fracture front breaks down into a series of shorter segments, each of which twists into a direction more nearly orthogonal to the local least compressive stress. This process is responsible for the common en-echelon segmentation of dikes and fissure eruptions (Pollard et al 1982). It is believed to result when a Mode-I dike propagates into a region where the least compressive stress is rotated slightly about an axis parallel to the propagation direction. Nicholson & Pollard (1985) showed that if the bridges of intact rock between segments are long and narrow, the average dike thickness is not much reduced by the segmentation, and with continued dilation these bridges often break, rendering them less influential still. Slickensided surfaces have been observed within failing bridges (P Delaney, personal communication); thus the segmentation may provide a source of seismicity and xenoliths during dike propagation.

FLUID MECHANICS

Flow Through a Slot

Computation of dike propagation rates and the ultimate in-plane size and shape of dikes requires consideration of magma flow. Laboratory experiments indicate that crystal- and bubble-free magmas may be treated as Newtonian viscous fluids. Because dikes are much longer than they are thick, the mean flow velocity perpendicular to the dike walls is small, pressure differences across the dike thickness may be neglected, and for purposes of coupling the magma flow to the host rock deformation, it is sufficient to consider the magma flux averaged over the channel thickness. Different relations exist for laminar and turbulent flow. Laminar flow may be assumed if the Reynold's number $Re \equiv \rho_m \bar{u} w / \eta$ is less than roughly 10^3 , where \bar{u} is the average flow velocity and η is magma viscosity. Often-quoted magma viscosities range from about 10^1 – 10^2 Pa s for basalts (Murase & McBirney 1973) to 10^4 – 10^8 Pa s for rhyolites (Petford et al 1993). For a 1-m-thick basalt dike with $\eta = 50$ Pa s and $\bar{u} = 0.5$ m s⁻¹, Re is approximately 10. Thus, flow in thin dikes is likely to be laminar, although the dependence of \bar{u} on w^2 (and thus the dependence of Re on w^3) means that moderate increases in thickness might induce turbulence.

For laminar flow, the velocity is zero at the dike walls and varies parabolically across the channel thickness. If x is the strike direction and z the dip direction of the dike, then the local mean velocity $\bar{\mathbf{u}}(x, z)$ averaged across the channel thickness has components

$$\bar{u}_x = -\frac{1}{3\eta} w^2 \frac{dP}{dx} \quad (3)$$

and

$$\bar{u}_z = -\frac{1}{3\eta} w^2 \left(\frac{dP}{dz} - \rho_m g_z \right) \quad (4)$$

(Batchelor 1967), where $w(x, z)$ is the dike half-thickness, $P(x, z)$ is the magma pressure, and g_z is the component of gravity acting in the positive z direction. The leading negative signs arise because flow is in the direction of decreasing magma pressure. Lister & Kerr (1991, equation 8c) give a comparable relation for turbulent flow; here I consider laminar flow only. Most semianalytic treatments of dike propagation consider one-dimensional flow of magma, \bar{u}_x or \bar{u}_z , in a two-dimensional dike; in some cases two such solutions are married through mass balance to approximate a three-dimensional dike. In the following I write g for g_z with the understanding that the latter is appropriate for nonvertical dikes.

For an incompressible magma, mass balance requires that the net flux into or out of a section of the dike be accommodated by a local change in dike thickness. This leads to the statement that the time derivative of the thickness equals (minus) the spatial derivative of the flux:

$$\frac{dw}{dt} = -\nabla \cdot (w\bar{\mathbf{u}}) = -\frac{d}{dx}(w\bar{u}_x) - \frac{d}{dz}(w\bar{u}_z). \quad (5)$$

Magma flow and host rock deformation are coupled because dw/dt must be consistent with both mass balance within the fluid and changes in the elastic thickness due to the changing dike length and excess pressure distribution.

The Tip Cavity

Near the dike tip (and, in the two-dimensional case, along the entire dike), the average flow velocity in the propagation direction is close to the tip velocity. From (3), this implies that the pressure gradient dP/dx varies approximately as one over the local thickness squared. For this reason the magma runs out of pressure before it makes it to the very narrow crack tip, and there is a gap between the tip and the magma front (Barenblatt 1962). Most likely this tip cavity is filled at some low pressure P_t by volatiles exsolving from the magma (Lister 1990a) or by pore fluids infiltrating from the host rock (Rubin 1993c). Under some conditions volatile exsolution or absorption might maintain the tip

pressure at the equilibrium vapor pressure of the magma, but kinetic factors or magma cooling may interfere with this simple scenario. Possible examples of relict tip cavities include magma-less cracks extending beyond the magma front (e.g. Figure 13), and pegmatite zones at the tips of some granite dikes (Rubin 1993a, figure 8a).

For dikes long enough that $\Delta Pl^{1/2} \gg K_c$, P_t must be less than the ambient compression σ_{yy}^0 if the crack tip is to propagate at less than elastic wave speeds. I define the tip suction or underpressure as $p_t = P_t - \sigma_{yy}^0$. The greater the magnitude of tip suction, the shorter the tip cavity. Both the tip suction and the intrinsic rock fracture resistance inhibit propagation of the dike tip, but if $\Delta Pl^{1/2} \gg K_c$, the intrinsic fracture resistance is secondary.

If the tip cavity occupies only a small fraction of the dike, then in most calculations to date it has little influence on global properties such as dike thickness and propagation velocity, even when magma freezing is included (Lister 1990a, Rubin 1993b, c). In fact, Spence & Turcotte (1985) ignored the existence of the cavity. Such models, which imply infinite magma suctions at the tip, are similar in spirit to linear elastic models with infinite stresses at the tip, in that they are unphysical in a sufficiently small region to remain useful for many purposes. They also share the property that they are inadequate for addressing near-tip processes. The tip cavity may play an essential role in generating inelastic deformation and seismic radiation during dike intrusion.

Dikes Driven by an Excess Source Pressure

Three sources of pressure are available to drive magma flow in dikes: the excess magma pressure at the source, magma buoyancy, and gradients of the “tectonic” stress normal to the dike plane. Defining the tectonic stress as $\Delta\sigma_y = \sigma_{yy}^0 - S_v$, where S_v is the vertical stress, and writing the magma pressure as $P = \sigma_{yy}^0 + \Delta P$, the pressure gradients driving viscous flow in (3) and (4) may be written (taking z to be positive downward)

$$\frac{dP}{dx} = \left(\frac{dS_v}{dx} + \frac{d\Delta\sigma_y}{dx} \right) + \frac{d\Delta P}{dx} \quad (6)$$

$$\left(\frac{dP}{dz} - \rho_m g \right) = \left(\Delta\rho g + \frac{d\Delta\sigma_y}{dz} \right) + \frac{d\Delta P}{dz}, \quad (7)$$

where $\Delta\rho = (\rho_t - \rho_m)$ and it is assumed that $dS_v/dz = \rho_r g$. Because $\Delta\rho g$ and the ambient host rock stress gradients are fixed, the terms in parentheses may be treated as an “effective” magma buoyancy. For convenience I refer only to magma buoyancy here and consider tectonic stresses later. Bear in mind, however, that within the brittle crust $d(\Delta\sigma_y)/dz$ can be several times $\Delta\rho g$. In addition, the significance of the $d(\Delta\sigma_y)/dx$ term is suggested by the widening of rift-zone dikes with distance from the magma source (Speight et al

1982, Walker 1987, Gudmundsson 1990); presumably, less frequent intrusion farther from the source allows more time for regional extension to reduce σ_{yy}^0 (Rubin 1992). Even the dS_v/dx term can be significant in regions of moderate topographic relief (e.g. volcanoes).

If one defines the excess pressure at the dike entrance $\Delta P(0)$ to be p_0 , and the dike length to be l , the excess source pressure gives rise to a global excess pressure gradient for flow of order p_0/l . A uniform excess pressure along the dike, on the other hand, implies a pressure gradient for vertical flow of $\Delta\rho g$. Because p_0/l exceeds $\Delta\rho g$ for sufficiently small l , this leads to a progression from excess pressure-dominated flows for near-source dikes to buoyancy-dominated flows for large dikes. For $p_0 = 3$ Mpa and $\Delta\rho = 300$ kg m⁻³, the excess source pressure dominates for $l < \sim 1$ km.

A schematic diagram of a dike driven by an excess source pressure is shown in Figure 4. The dike is assumed to be two dimensional, which is probably adequate provided l is small compared to the reservoir dimensions. When the excess source pressure and tip suction remain constant, and the fracture energy is insignificant, dike growth is self-similar (Rubin 1993a)—i.e. new variables

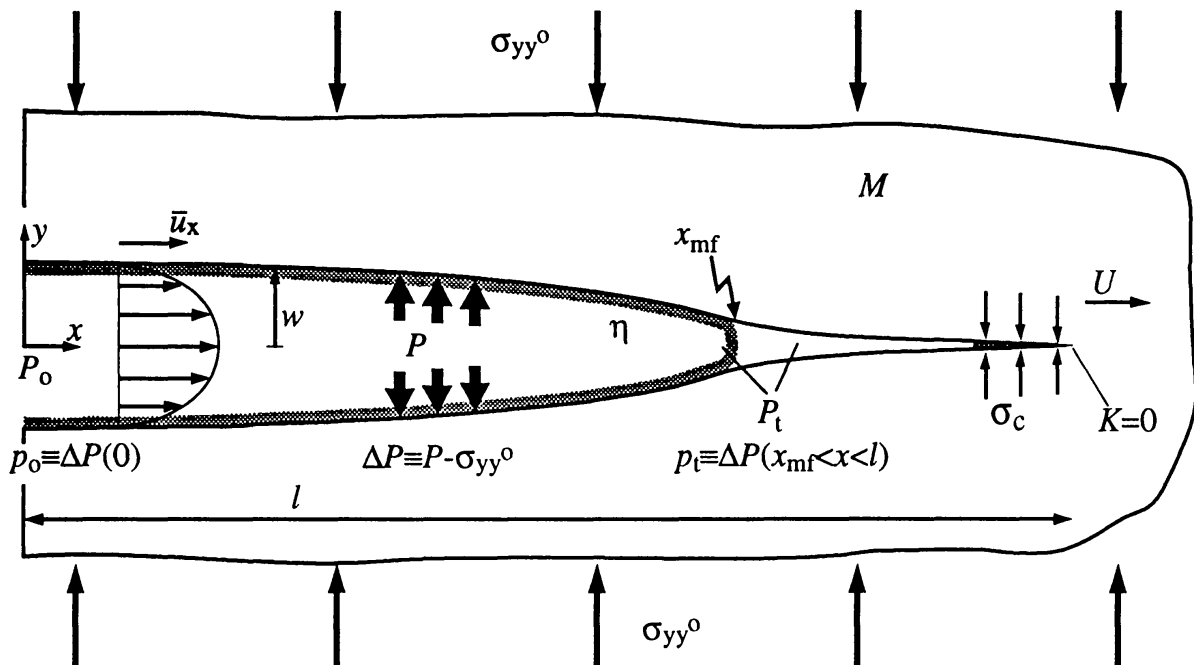


Figure 4 Schematic diagram of a two-dimensional, pressure-driven, propagating dike. Material properties: magma viscosity η , host rock elastic stiffness M , cohesive stress σ_c acting over critical crack-wall separation δ_c (negligible for large dikes). Boundary conditions: ambient dike-perpendicular compression σ_{yy}^0 , magma pressure at dike entrance P_0 , pressure in tip cavity P_t . Variables to be computed: magma pressure P , dike thickness w , position of magma front x_{mf} , average flow velocity \bar{u}_x , propagation velocity U . Elastic thickness depends upon excess pressure ΔP . Propagation criterion is $K = 0$ at tip of Barenblatt cohesive zone, where dike walls close as cusp.

can be defined with a particular dependence upon time such that time no longer appears explicitly in the governing equations. This turns the initial value problem into a boundary value problem whose solution can be scaled for all time. For the case at hand the dike thickness and propagation velocity both increase linearly with dike length.

A dimensional estimate of the propagation velocity U may be obtained by substituting p_0/l for dP/dx and $w(0)$ (Equation 1) for w in Equation (3). This yields

$$U \approx \frac{1}{3\eta} \left(\frac{p_0}{M} l \right)^2 \left(\frac{p_0}{l} \right) = \frac{1}{3\eta} \frac{p_0^3}{M^2} l. \quad (8)$$

The velocity approaches zero with decreasing l because the thickness squared decreases faster than the pressure gradient increases. The full solution requires combining Equations (3) and (5) with the more complete versions of (1) and (2); it shows that (8) overestimates the velocity by a factor of about 4. The dike propagation velocity and thickness, and tip cavity length and thickness, are shown in Figure 5 as a function of p_t/p_0 . The dike shape and pressure

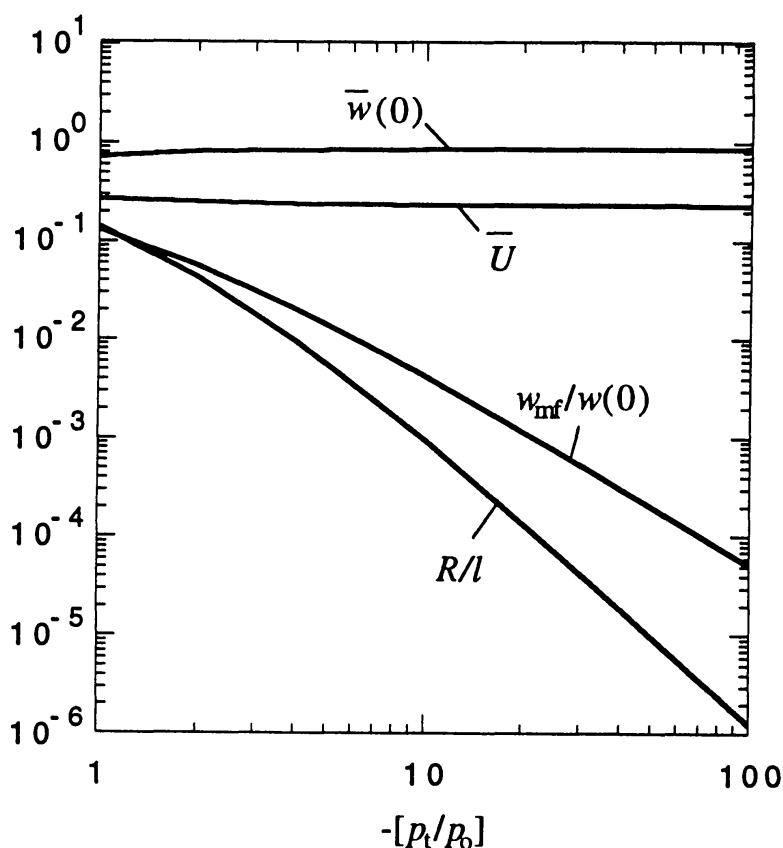


Figure 5 Computed properties of pressure-driven dikes as a function of ratio of tip suction p_t to excess source pressure p_0 . R/l , (tip cavity length)/(dike length); $w_{mf}/w(0)$, (thickness at magma front)/(thickness at entrance); $\bar{w}(0)$, thickness at entrance normalized by Equation (1); \bar{U} , propagation velocity normalized by Equation (8).

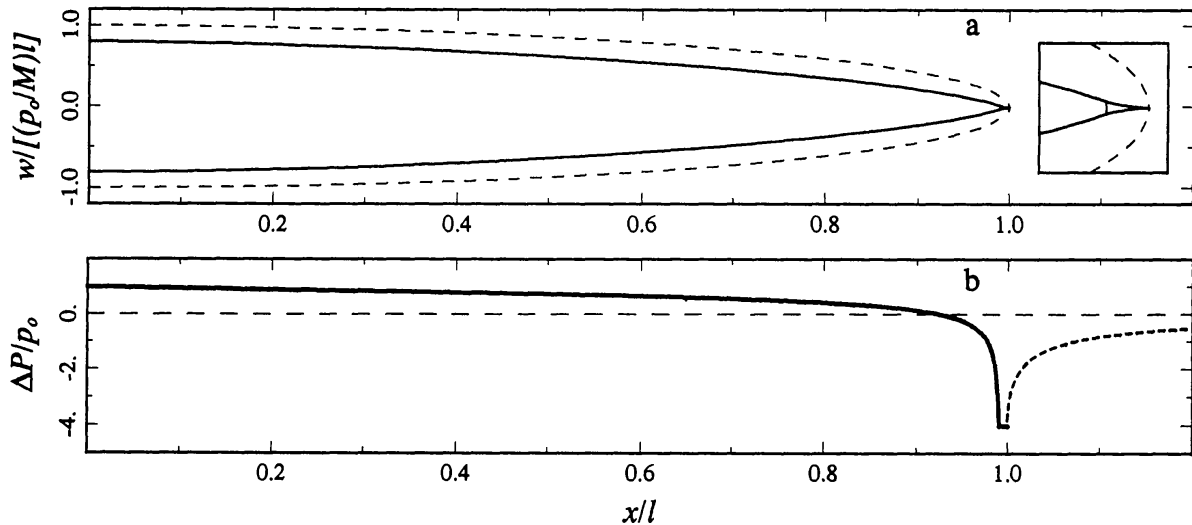


Figure 6 (a) Dimensionless thickness profile of a pressure-driven dike for $p_t/p_0 = -4$ (solid curve). Dashed curve indicates elliptical shape for uniform excess pressure p_0 . Inset (right) shows tip region magnified by $5\times$; vertical line indicates magma front. Length of tip cavity $\sim 0.01l$. (b) Computed excess magma pressure ΔP within dike, normalized by p_0 , for $p_t/p_0 = -4$ (bold solid curve). Bold dashed curve indicates change in elastic stress σ_{yy} normal to dike plane.

distribution are shown in Figure 6 for $p_t/p_0 = -4$. The shear stresses along most of the dike wall are of order $w/l \times \Delta P$, and therefore are negligible.

Spence & Turcotte (1985) and Emerman et al (1986) derived analogous results for dikes driven by a specified source flux, rather than a specified source pressure. The two solutions are comparable in the sense that any particular flux history defines a unique source pressure history, and vice versa. However, the specified source pressure seems more realistic geologically. For plausible boundary conditions, it is reasonable to assume that a dike does not deplete the source pressure by a significant fraction of p_0 until the dike length is a significant fraction of the reservoir dimensions. The paradox that the velocity is zero when the dike length is zero is resolved if the dike starts with some finite nonelastic thickness (in a partial melt, for example). A constant source flux, on the other hand, requires an infinite source pressure at zero dike length. In addition, comparison at the same source flux can disguise the influence of magma viscosity on propagation rate. For a specified source flux the velocity varies as $\eta^{-1/6}$; for a specified source pressure it varies as η^{-1} . The difference is due to the fact that a specified source flux requires a higher source pressure for a higher viscosity magma.

Dikes Driven by Magma Buoyancy (Two Dimensions)

Once the dike height h grows sufficiently that $\Delta\rho gh > p_0$, magma buoyancy begins to provide the dominant pressure gradient for flow, and the dominant excess pressure for dike widening. The general case of a buoyant dike with a

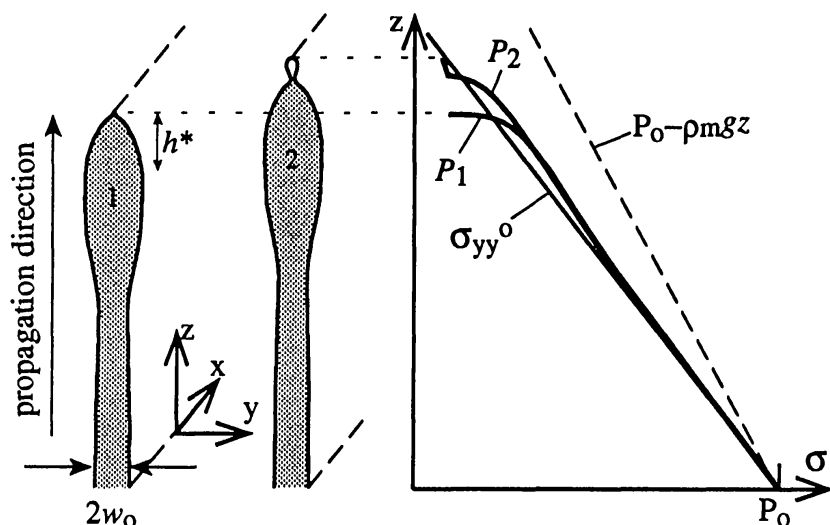


Figure 7 Schematic diagram showing two stages in the growth of a steady, two-dimensional, buoyant dike driven by a constant-flux line source. A thickness $2w_0$ is asymptotically approached with increasing depth. Magma pressure P deviates significantly from ambient dike-perpendicular stress σ_{yy}^0 only over a length scale h^* below tip (Equation 9). Dashed line indicates static magma pressure curve for reference. At stage 1, tip suction is large and tip cavity (not visible) is small; at stage 2, magma front is closer to volatile saturation depth and tip cavity may become unstable.

constant p_0 has not been addressed. In two dimensions this would lead to a dike thickness and velocity that increased with time, so eventually the source would be unable to supply the increasing flux. Lister (1990a) considered two-dimensional dikes driven by a constant flux q from a line source, in the limit that the dike tip is “far” from the source. A steady state was sought in which the dike propagates with constant velocity, and in which the dike shape as seen in the reference frame of the tip is unchanging (Figure 7).

From Equation (1), the excess pressure within a dike of thickness w is of order $(w/h)M$. If $\Delta P \sim (w/h)M \ll \Delta\rho gh$, the elastic resistance of the dike walls to deformation is negligible in the sense that large variations in dike thickness (relative to w) can be produced by small variations in excess pressure (relative to $\Delta\rho gh$). From the above inequality this condition is met when

$$h \gg \sqrt{\frac{wM}{\Delta\rho g}} \equiv h^*. \quad (9)$$

If $w = 1$ m, h^* is roughly 10 km. The height h^* is also a dimensional estimate of the distance from the dike tip over which the elastic resistance to dike opening is significant. Note that on physical grounds w should be determined as part of the solution and not stipulated a priori.

In the constant flux formulation, far from the tip compared to h^* , the elastic pressure approaches zero, so the pressure gradient for flow is $\Delta\rho g$ and the thickness approaches a constant value of $2w_0$. One can then determine \bar{u}_z and

w_0 from $q = 2w_0\bar{u}_z$ and Equation (4). Reasonable results are obtained in that $\bar{u}_z = 1 \text{ m s}^{-1}$ for $2w_0 = 1 \text{ m}$, $\Delta\rho = 300 \text{ kg m}^{-3}$, and $\eta = 25 \text{ Pa s}$. Within (3–4) h^* from the dike tip [defined by replacing w in (9) by w_0], ΔP is no longer negligible and the dike develops a bulbous nose with a maximum thickness some 25% greater than $2w_0$ (Lister & Kerr 1991). The nose arises because if the dike narrowed monotonically towards the tip, the pressure gradient for flow would increase monotonically towards the tip, generating a magma underpressure everywhere along the dike and (mathematically) interpenetration of the dike walls. An increased thickness implies a reduced pressure drop, consistent with a local excess pressure that generates the widening.

For this solution to be applicable the dike height should be several times h^* ; for mantle-derived basalts this requirement can be met if the dike thickness is no more than a few meters. A more limiting requirement is that the lateral extent of the source exceed the growing dike height. One very interesting result that should be applicable even if these conditions are not met concerns the possible fate of the tip cavity. Because the tip phase is much less viscous than the magma that follows, if it travels at the same speed then the viscous pressure drop in the cavity is small. Together with the low density of the tip phase, at least at crustal depths, this gives rise to an underpressure at the crack tip that is less than that at the magma front (Figure 7, curve P_2). For small tip cavities this pressure variation is irrelevant, but Lister & Kerr (1991) found that for sufficiently large cavities (plausibly $\sim 100 \text{ m}$), the steady solution implies interpenetration of the dike walls slightly above the magma front. One interpretation is that at this point steady solutions are not feasible and the tip cavity propagates upward at a rate controlled by either the tip fluid viscosity or elastic wave speeds, leaving the magma behind (see also Crawford & Stevenson 1988). Because large tip cavities result from small tip suction, such behavior becomes more likely as the dike approaches the depth at which the magma nominally becomes saturated with volatiles and exsolution occurs at a small underpressure. This raises the possibility of the repeated rise of CO_2 pulses from dikes in the mantle.

Dikes Driven by Magma Buoyancy (Three Dimensions)

Lister (1990b) derived approximate solutions for the growth of buoyant dikes fed by a point source. An analogy is drawn with a film of water on a slope: The water runs primarily downhill under the influence of gravity, but also spreads in the cross-stream direction because the greater thickness at the center produces pressure gradients that drive lateral flow. For a constant source flux Q , the lateral extent of the flow approaches a steady value after the leading edge has passed (Figure 8). While on a slope the lateral flow is driven by gravity, in a dike it is driven by the elastic pressure required to prop the dike walls open.

The dike shape well behind the leading edge can be determined from mass balance and the assumption that the flow is steady. It is assumed that $\Delta P \ll$

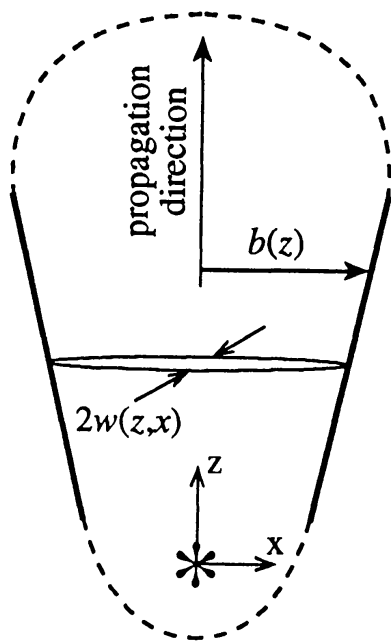


Figure 8 In-plane shape of a buoyant dike driven by a constant-flux point source (*star*). Above magma source and below dike top, lateral extent of dike $b(z)$ approaches a steady value.

$\Delta\rho gh$, so that the pressure gradient for vertical flow is $\Delta\rho g$, and that the lateral extent (width) b is the short in-plane dimension of the dike, so that the pressure gradient driving lateral flow $\sim \Delta P/l$ is $[M(w/b)]/b$. Both assumptions are satisfied sufficiently far from the source. Assuming that db/dz is given by the ratio of the lateral to vertical flow velocities, the lateral extent of the dike is

$$b(z) \sim \left(\frac{Q\eta M^3}{(\Delta\rho g)^4} \right)^{1/10} z^{3/10}. \quad (10)$$

The lateral extent decreases with increasing $\Delta\rho g$, just as it would for a film on a slope. It increases with source flux, magma viscosity, and elastic stiffness, because each increases the pressure gradient for lateral flow, while the pressure gradient for vertical flow remains fixed. For $\Delta\rho = 300 \text{ kg m}^{-3}$, $M = 20 \text{ GPa}$, $\eta = 100 \text{ Pa s}$, and $Q = 10^4 \text{ m}^3 \text{ s}^{-1}$, b reaches 8 km at 10 km above the source and 16 km at 100 km above the source (Lister & Kerr 1991, figure 12).

Finite Magma Batches

The rise of isolated volumes of magma in dikes may be the most common ascent mechanism portrayed in cartoons of “magma plumbing systems” beneath volcanoes. Weertman (1971) and Secor & Pollard (1975) outlined a scenario for how such behavior might occur. The stress intensity factors at the dike top and bottom of a stationary, two-dimensional crack of height $2h$ are

given by

$$K^{\pm} = \left(\Delta P(h) \pm \frac{1}{2} \Delta \rho g h \right) h^{1/2}, \quad (11)$$

where $\Delta P(h)$ is the excess pressure at the dike center and the plus and minus refer to the upper and lower dike tips, respectively (Figure 9). If magma is pumped slowly into a short dike, $\Delta P(h)$ decreases while maintaining $K^+ = K_c$ as the dike top rises. Eventually h increases to the point that K^- drops to zero. If any more magma is added, extension of the dike top drops the magma pressure allowing the walls near the bottom to squeeze inward, forcing the magma upward. A very slow, nearly shape-preserving ascent results. Given a stagnant column of magma (no viscous pressure drop), the half-height and (center) half-thickness of the critical dike that simultaneously satisfies $K^- = 0$ and $K^+ = K_c$ are

$$h_{\text{crit}} = \left(\frac{K_c}{\Delta \rho g} \right)^{2/3}, \quad w_{\text{crit}} = \frac{1}{2} \left(\frac{K_c^4}{\Delta \rho g M^3} \right)^{1/3}. \quad (12)$$

For $K_c = 1 \text{ Mpa m}^{1/2}$, $\Delta \rho g = 300 \text{ kg/m}^3$, and $M = 20 \text{ GPa}$, the critical dike height and thickness are about 100 m and 0.7 mm, respectively.

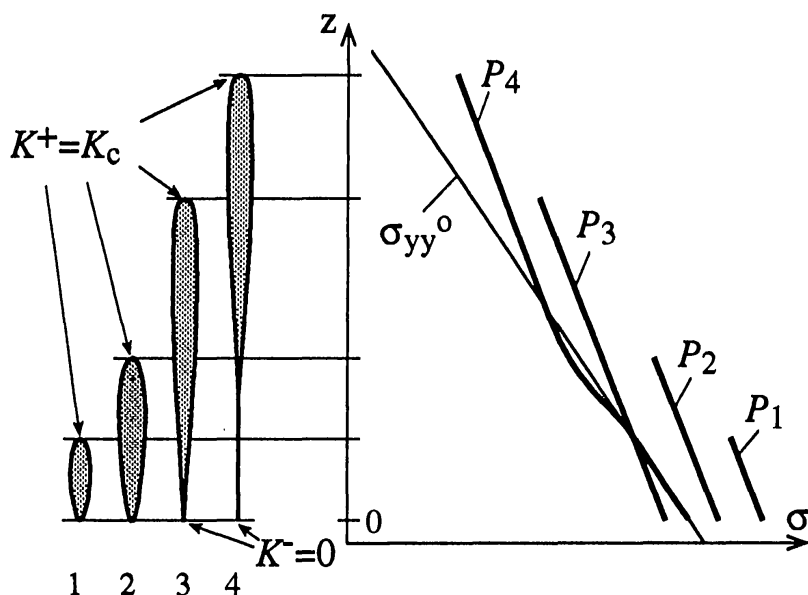


Figure 9 Growth of a buoyant two-dimensional “tadpole” dike. As magma is added from stages 1 to 3, magma pressure is reduced while maintaining $K^+ = K_c$. At stage 3, K^- has reached zero. In stage 4, an infinitesimal addition of magma allows the dike to propagate slowly upward, maintaining a nearly constant shape as it leaves a tail behind. The progression shown is governed by Equation (11), assuming K_c constant. In rock, such growth is more likely to occur in partial melt at a constant magma pressure, by accelerating subcritical crack growth as K^+ approaches K_c . Note that $\Delta \rho$ implied by these curves ($\sim 50\%$ of the rock density) is greatly exaggerated.

This need not be what dikes look like. Dikes of the critical height carry little magma and would freeze before propagating far into subsolidus rock. In addition, the ascent velocity of dikes of the critical height is infinitesimal (or K^+ would drop below K_c); thus there is no reason to expect that the magma supply cuts off at this point. Because the ascent velocity increases as more magma is pumped into the dike, eventually the dike may propagate rapidly upward with insignificant additional input from the source. As pointed out by Stevenson (1982), however, the magma in the dike “tail” cannot be entirely squeezed out because of the very large viscous resistance as the crack thins.

Spence & Turcotte (1990) extended these results to include viscous flow within the dike. They obtained solutions for dikes containing fixed volumes of magma, in the limit that elasticity can be neglected. That is, ΔP is assumed to be zero and the pressure gradient for vertical flow is therefore $\Delta\rho g$, and the dike thickness responds passively to spatial gradients in mass flux [w is governed by Equations (4) and (5) but not (1)]. Because the flow rate varies as the thickness squared, the thicker parts of the dike catch up with the thinner parts, resulting in a “shock wave” (jump in thickness) at the dike top. However, neglect of elasticity means that the solution is not accurate within a distance of h^* (Equation 9) from the dike top, so this region should be ignored. The solution should be applicable to two-dimensional dikes once the dike height is a few times that at which the magma supply shuts off. The dike height h and thickness w (well below the top) are given by

$$h(t) \sim \left(A^2 \frac{\Delta\rho g}{\eta} t \right)^{1/3}, \quad w(z, t) \sim \left(\frac{\eta}{\Delta\rho g} \frac{z}{t} \right)^{1/2} \sim \left(\frac{A^2 z}{h^3} \right)^{1/2}, \quad (13)$$

where A is the cross-sectional area of the dike and z is height above the source. The dike tail thins for all time. Similar (but sketchier) results were obtained by Stevenson (1982).

The three-dimensional case has not been addressed theoretically. In gelatin experiments, constant-volume batches have been observed to ascend as approximately circular dikes maintaining a nearly constant diameter (Takada 1990, Heimpel & Olson 1994). The low ascent velocities of these “dikes” indicates that they were rate limited by the gelatin fracture toughness as well as the fluid viscosity, so these experiments may be relevant for dike heights slightly greater than (the three-dimensional) h_{crit} .

Lateral Propagation at the “Level of Neutral Buoyancy”

It is reasonable to expect that most dikes propagate upward through most of the lithosphere. This is because (a) rock is usually more dense than magma and (b) temperatures are most often high enough, and hence rock weak enough, that the horizontal stress is close to the vertical stress. However, sediments

within the upper crust, and fractured and vesicular lava flows in volcanoes, are often less dense than magma. This has led to the intuitive notion that the observed lateral propagation at shallow depths occurs at the level of “neutral buoyancy.” Ryan (1987, 1993) gives examples of magmatic systems that seem to be located near this level, and discusses how the depth of the neutrally buoyant zone may be maintained during the growth of volcanoes by elastic closure and mineralization of cracks at depth. However, I reiterate that magma within a dike senses the density of the host rock only insofar as that density contributes to the ambient stress normal to the dike plane. For simplicity I continue to use the term buoyancy here.

Rubin & Pollard (1987) considered the lateral growth of dikes from a shallow source of constant pressure. Once the lateral extent exceeds the dike height, magma flow is predominantly horizontal, the vertical magma pressure gradient is approximately $\rho_m g$, and the dike thickness in any vertical cross section (away from the leading edge) is determined largely by the excess pressure within that cross section. The requirement for lateral propagation is that the excess pressure near the depth of the dike center exceed that at the dike top and bottom (Figure 10). For laboratory (negligible) fracture toughness, the dike height stabilizes when the underpressure at the dike top and bottom reduces K to near zero. Rubin & Pollard used geodetic estimates of dike height and thickness from Kilauea to place limits on the values of the magma pressure and horizontal stress within the rift. While they did not consider lateral propagation rates explicitly, they noted that the along-strike slope of the East Rift Zone of Kilauea (0.025) would provide

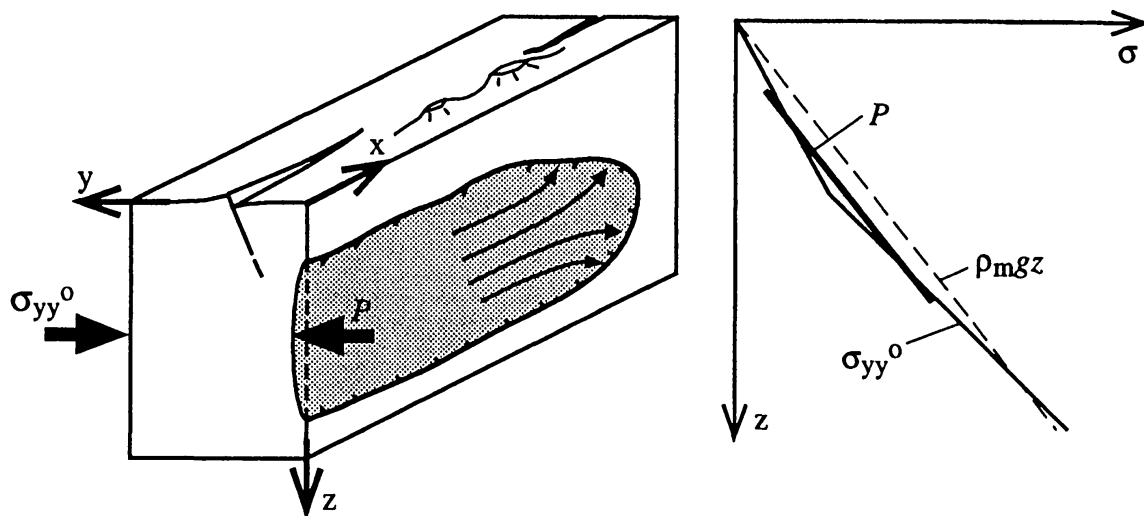


Figure 10 (Left) Growth of a blade-like dike from a shallow source. (Right) Stress state required for lateral propagation. Kink in σ_{yy}^0 may be due to a density step or to variations in tectonic stress.

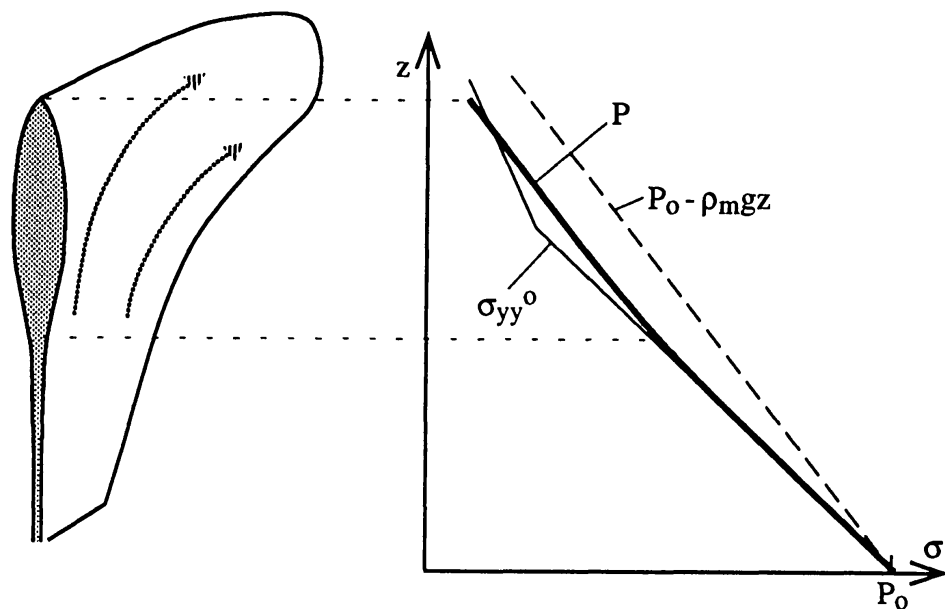


Figure 11 Lateral propagation of a buoyant three-dimensional dike encountering level of “neutral buoyancy.” Amount of overshoot increases with excess pressure at this level, which for given material properties increases with source flux.

a pressure gradient for downrift flow of ~ 0.6 Mpa/km [dS_v/dx term in Equation (6)], sufficient to generate propagation rates close to those observed (0.3 m s^{-1} for $w = 1 \text{ m}$ and $\eta = 50 \text{ Pa s}$). This suggests that viscous pressure drops associated with downrift flow could influence the long-term slope of the rift zone.

When the level of neutral buoyancy is encountered by a vertically propagating dike, the situation is more complex. The ascent velocity decreases, which decreases the viscous pressure drop and increases the excess pressure within the dike. This widens the dike and enables the magma to ascend further. For a two-dimensional dike fed by a line source (a case that is more instructive than relevant), the dike will ultimately erupt if the source is deep enough and is able to supply magma at a pressure close to the overburden. Alternatively, it may come to rest at depth if the source is too shallow or becomes depleted. In either case, the dike fate may be ascertained by recourse to static elastic solutions; the local density contrast with the host rock is irrelevant. For a source of finite lateral extent, however, this equilibrium height might never be attained because magma will bleed off laterally due to the greater excess pressure at the level of neutral buoyancy (Figure 11). The source of this excess pressure is the flux from below, and it seems reasonable to approximate the lateral growth as being driven by a point source of flux at this level (Lister 1990b, Lister & Kerr 1991). Note that even though the excess pressure is greatest at this depth, there is no incentive (at the tip, which is where it matters) for the

dike to turn into a sill, provided the ambient least compressive stress remains horizontal.

In the specified flux model of Lister & Kerr (1991), lateral dike growth is driven by the lateral variation in excess pressure [$d\Delta P/dx$ term in (6)]. Together with mass balance and the above approximations of no vertical viscous pressure drop and two-dimensional elasticity, this assumption enabled them to compute the dike height, length, and thickness as functions of time. For a constant flux Q , the dike length increases with time as $t^{8/11}$ and the height as $t^{1/11}$; for given material properties, the height increases with flux as $Q^{2/11}$. The vertical growth rate is of interest because it determines the extent to which a dike may overshoot the level of neutral buoyancy, which in turn may determine whether the dike erupts. A spectacular example of overshoot might be represented by the dense lunar mare basalts, which erupted through perhaps 60 km of low-density anorthositic crust.

Viscous Deformation of the Host Rock

The propagation velocity of dikes driven by an excess source pressure is inversely proportional to the magma viscosity. This raises the possibility that dikes of sufficiently viscous magma could produce significant viscous deformation of the host rock on the timescale of intrusion. To assess the influence of the magma–host rock viscosity contrast on the transport of granites, Rubin (1993a) examined the growth of pressure-driven dikes in linear viscoelastic rock. For ratios of p_0/M of 10^{-3} to 10^{-4} , the viscosity contrast must exceed 11 to 14 orders of magnitude for the viscous contribution to dike thickness to be negligible. While these contrasts seem very large, they are met by all basalts and most granites. However, very viscous granites ($\eta = 10^8$ Pa s) might have viscosities only 9 orders of magnitude less than hot thermal aureoles ($\eta = 10^{17}$ Pa s, say). In this case, the viscous dike wall displacements away from the tip can exceed the elastic displacements by 1 to 2 orders of magnitude. Nonetheless, such dikes are still narrow, with thickness:length ratios of no more than a few times 10^{-2} . Viscous deformation is important in these cases not because the viscous strains are large, but because the elastic strains are so small. Elastic strains continue to dominate near the tip.

Larger thickness:length ratios could result for propagating granite dikes fed by porous flow in the source region, since propagation rates are probably less under these conditions than for a constant pressure source. Viscous deformation could dominate in the source region even for basalts if the crack tip did not propagate (Sleep 1988). Exploring this issue is difficult in the absence of information on crack growth in partial melts; observations of the aspect ratios of dikes and veins within presumed source regions could shed some light on this matter.

HEAT TRANSFER

Because dikes carry hot magma into cold rock, heat flow is an important aspect of intrusion. Freezing of magma might prevent dikes from depleting the source region or erupting, and melting of the host rock might alter the magma composition. Energy sources available to retard freezing include an excess magma temperature above some (depth-dependent) effective solidus, release of latent heat during crystallization, and viscous dissipation. Heat-transfer mechanisms include conduction out the dike walls, downstream and cross-stream advection within the magma (the latter due to turbulence or changes in dike thickness), and perhaps hydrothermal effects in the host rock or even radiation in iron-poor, crystal-free magmas. Formulations coupling all of these processes to the fluid and solid mechanical parts of the problem have not been attempted.

Conduction

Early studies of freezing in dikes considered the instantaneous emplacement of magma into a slot of specified thickness. For similar magma and host-rock thermal properties and no latent heat release, the contact initially adopts the average of the magma and host rock temperatures. Thermal boundary layers develop in both the magma and host rock, with thicknesses that increase with time as $(\kappa t)^{1/2}$, where κ is thermal diffusivity (Turcotte & Schubert 1982). This relation is exact for two juxtaposed half-spaces; however, once the thermal boundary layer in the dike approaches the dike center, the magma temperature falls below that of the half-space solution because from symmetry no heat is available to flow across the dike centerline.

If the magma possesses a unique freezing temperature, all crystallization occurs at a migrating solidification front. When the thermal boundary layer thickness is much less than the dike thickness, the frozen margin thickness δ grows with time as

$$\delta = 2\lambda(\kappa t)^{1/2}, \quad (14)$$

where λ depends upon the initial magma and host rock temperatures T_m and T_0 , the freezing temperature T_s , the latent heat of crystallization L , and the heat capacity c (Carslaw & Jaeger 1959). For $(T_m - T_s) = 100^\circ\text{C}$, $(T_s - T_0) = 1000^\circ\text{C}$, $L = 500 \text{ kJ kg}^{-1}$, and $c = 1 \text{ kJ kg}^{-1} \text{ }^\circ\text{C}$, $\lambda = 0.45$; for $\kappa = 0.5 \times 10^{-6} \text{ m}^2 \text{ s}^{-1}$, the solidified margin grows to 0.5 m in about 7 days. This is several times the lifetime of the curtain-of-fire stage of observed eruptions in Hawaii and Iceland, perhaps because flow becomes choked long before all the latent heat is released. For $T_m = T_s$, the growth rate is only 20% faster because more energy is carried in as latent heat in this example (500 kJ kg^{-1}) than as excess temperature [$c(T_m - T_s) = 100 \text{ kJ kg}^{-1}$]. The opposite dike wall also enhances the freezing rate, but by less than the $T_m = T_s$ case, because

if $T_m = T_s$ there is no heat conduction in the magma and the opposite wall is not felt. Carslaw & Jaeger (1959) provide equations for dissimilar magma and host rock thermal properties and for continuous release of latent heat between T_m and T_s ; Delaney (1987) gives numerical results for temperature-dependent thermal properties.

A quick estimate of the distance that dikes may propagate before freezing can be obtained by multiplying the solidification time by the flow velocity. Because both increase as w^2 , the propagation distance l_f increases as w^4 ; combining (14) with (3) yields

$$l_f \sim \left(\frac{w^2}{4\lambda^2\kappa} \right) \left(\frac{w^2}{3\eta} \frac{dP}{dx} \right) = \frac{1}{12\lambda^2} \frac{w^4}{\kappa\eta} \frac{dP}{dx}. \quad (15)$$

This implies, for example, that if a 2 m-thick dike in Iceland can propagate 30 km, a 6 m-thick dike on the Canadian shield can propagate nearly 2500 km (neglecting any decrease in dP/dx with l). Equation (15) also implies that a rhyolite 4 orders of magnitude more viscous than a basalt would have to be 10 times thicker in order to propagate as far (neglecting any dependence of w on dP/dx through elasticity). More sophisticated treatments account for how growth of the frozen margin reduces the aperture for flow and hence the flow velocity with time [w in Equations (3) and (4) must be replaced by $w - \delta(t)$]; integrating the total flux over time and dividing by the thickness produces an estimate lower than that given in (15) by a factor of 10 (Delaney 1987). Even more sophisticated treatments could account for the variation of dP/dx in space and time; however, given that (15) neglects advection of heat, such sophistication might be unwarranted, and the conclusion that propagation distance is proportional to w^4/η would be unchanged.

Coupled Conduction and Advection

Magma flow may increase the lifetime of a dike, because magma is everywhere being displaced by hotter magma that has left the source more recently. This advection ensures that simple analytic solutions to the thermal problem do not exist. Delaney & Pollard (1982) and Bruce & Huppert (1990) consider magma flow through a slot of specified thickness, in host rock of initially uniform temperature. The magma is idealized as constant viscosity until it freezes at T_s . Immediately after magma comes into contact with rock, the large temperature gradient ensures that the conductive heat flux out the dike walls dominates the advective flux, and magma freezes along the dike wall according to Equation (14). As the thermal boundary layer grows, the temperature gradient and conductive flux decrease as $(\kappa t)^{-1/2}$, and downstream advection of heat in the thermal boundary layer adjacent to the frozen margin also becomes important. While the formulation of Delaney & Pollard is accurate only for as long as the solidification front moves into the dike, Bruce & Huppert consider subsequent

melting of the dike walls (meltback) as well. They find that for dikes less than a critical thickness the dike walls grow inward until solidification is complete, but that for greater initial thicknesses advection of heat is sufficient for the magma to thermally erode the previously solidified magma, and eventually the host rock.

Bruce & Huppert (1990) considered vertical flow of basalt from a source at depth h , driven by a constant pressure $\Delta\rho gh$ in excess of hydrostatic. The critical dike thickness for meltback was found to be about 1.3 to 1.7 m for h from 2 to 5 km; it increases with the parameter $\alpha = [\eta\kappa h/(w^4\Delta\rho g)]^{1/3}$ (essentially the ratio of a characteristic thermal boundary layer thickness to the initial dike thickness, where the time for boundary layer growth is h/\bar{u}_z). This thickness seems rather low in that, to my knowledge, many thicker basalt dikes show no signs of melting the host rock. Several factors, in addition to neglect of temperature-dependent viscosity, may contribute to this apparently low estimate. The adopted parameters imply a prefrozen flow velocity of 3 m s^{-1} for a dike 1.3 m thick, which is 3 to 10 times observed lateral propagation rates; lateral propagation implies vertical velocities that are considerably lower still. Applying the model to laterally propagating dikes is not straightforward. It may be permissible to interpret h as the local distance from the source, implying (through the dependence of α on h) that dike thicknesses large enough to permit meltback near the source are not sufficient at distances at which eroded dikes are typically observed. However, this interpretation does not allow for the possibilities that (a) solidification further downstream impedes flow locally, (b) the pressure gradient driving flow might decrease as the dike length increases, or (c) the implied magma flux prior to the onset of meltback might be large enough that the constant source pressure assumption is often violated.

Given the potential importance of meltback, a greater effort to identify it in the field is certainly warranted. Platten & Watterson (1987) describe discontinuities in internal structure within dikes that may be indicative of meltback that was arrested before it reached the dike walls. Macleod & Rothery (1992) describe meter-wide sheeted dikes in the Oman ophiolite that widen to 10 m as they grade imperceptibly into the gabbroic source rock below. This may reflect melting of the host rock; however, it occurs only within a few tens of meters of the source, where initial temperatures were presumably quite high.

Viscous Dissipation

The rate of thermal energy generation within a unit volume that spans the dike half-thickness can be computed from the rate at which work is done on that volume; from force balance it is the half-thickness times the pressure gradient times the average flow velocity. In the absence of freezing this is (Equation 3)

$$\frac{\text{Work rate}}{\text{unit breadth and length}} = w \left(\frac{dP}{dx} \right) \bar{u}_x = \frac{1}{3\eta} w^3 \left(\frac{dP}{dx} \right)^2. \quad (16)$$

An estimate of the rate of conductive heat loss as the magma solidifies is given by the sum of latent and excess heats divided by the cooling time (Equation 14):

$$\frac{\text{Rate of heat loss}}{\text{unit breadth and length}} \approx \frac{\rho w [L + c(T_m - T_s)]}{[w^2 / (4\lambda^2 \kappa)]} \quad (17)$$

The ratio of (16) to (17) is a measure of the importance of viscous heat generation to conductive heat loss; it is essentially the Brinkman number (Bird et al 1960, p. 278). For $\eta = 100 \text{ Pa s}$ and $dP/dx = 1 \text{ MPa km}^{-1}$ (implying a flow velocity of 85 cm s^{-1} in a 1-m dike), $\rho = 2.6 \times 10^3 \text{ kg m}^{-3}$, $(T_m - T_s) = 100^\circ\text{C}$, and L , c , and κ appropriate for basalts (500 kJ kg^{-1} , $1 \text{ kJ kg}^{-1} \text{ }^\circ\text{C}^{-1}$, and $0.5 \times 10^{-6} \text{ m}^2 \text{ s}^{-1}$), (16) and (17) are comparable for a dike thickness of only 1.4 m. Given the potential range of these variables, and the dependence on w^4/η , viscous dissipation may be significant for many basaltic dikes.

Considerations of viscous dissipation generally have been back-of-the-envelope style, similar in spirit to that above. McConnell (1967) also estimated the dike thickness required for a gross balance between viscous dissipation and conductive heat loss to be about 1 m for basalts. He suggested that this accounts for the observation that many basalt dikes are about 1 m thick; the implication is that narrower dikes would freeze quickly, while thicker dikes would continue to propagate (and thin, given a finite magma volume) for as long as dissipation kept the magma fluid. This is interesting in that it provides an alternative to the explanation that rhyolite dikes are thicker than basalts because otherwise they would not propagate for enough to be seen (Equation 15). However, simple steady-state calculations of this sort may be deceptive. As with Equation (15), if dP/dx is constant and the frozen margin grows as \sqrt{t} , the average work rate over the lifetime of a dike that freezes is less than that obtained from (16) by a factor of 10. Establishing when viscous dissipation is important may require time-dependent numerical experiments with realistic boundary conditions.

The same may be true for thermal runaway. Fujii & Uyeda (1974) considered the flow of magma under a fixed pressure gradient through walls at a fixed temperature T_w , and determined the conditions under which viscous heating and the decrease in viscosity with temperature would lead to a flux that increased without bound. Their critical parameter differs from the Brinkman number essentially by the ratio of the viscosity of magma at T_m to that of magma at T_w . (Clearly, there is some ambiguity in defining the viscosity at T_w when T_w is the solidus.) They concluded that thermal runaway would be common for basaltic dikes greater than about 1 m thick, and that the onset time for the instability was a few weeks. This is greater than the expected lifetime of meter-thick dikes in cool rock; however, Fujii & Uyeda's initial flow velocity was only one tenth that of meter-thick, hours-old dikes in Iceland and Hawaii, and increasing it would greatly decrease the onset time for instability (their figure 1).

Coupled Conduction, Advection, and Host Deformation

Formulations in which the dike thickness is stipulated a priori cannot address the questions of how freezing near the tip might inhibit propagation, or whether a particular dike is capable of reaching the stipulated thickness in the first place. In an effort to understand the apparent preponderance of basaltic dikes and granitic plutons, Rubin (1993b) considered the thermal fate of two-dimensional dikes growing from a magma chamber into elastic host rock. It was assumed that the initial host rock temperature equaled the magma temperature at the source and decreased with distance from the source with a uniform gradient dT_0/dx . From Equations (1) and (8), short dikes are thin and move slowly. Whether a dike can successfully transport magma away from a source region depends upon the competition between the rate at which the dike thickness increases due to propagation and the rate at which the aperture for flow decreases due to solidification.

The simplest case occurs when $T_m = T_s$, because then the magma temperature equals T_m everywhere it is flowing. For constant source and tip pressures, the growth of short dikes is (again) self-similar in that the dike thickness, frozen margin thickness, and propagation velocity are all proportional to dike length. The thermal state of the dike is determined by a dimensionless parameter β that is essentially a ratio of the frozen margin thickness to the elastic thickness; from Equations (14) and (1) this is

$$\frac{\delta}{w} \sim \frac{2\lambda\sqrt{\kappa t}}{(p_0/M)l} \sim \frac{2\sqrt{3}}{\sqrt{\pi}} \frac{c|dT_0/dx|\sqrt{\kappa\eta}}{L(p_0/M)^2\sqrt{p_0}} \equiv \beta. \quad (18)$$

To obtain the third expression the relation $\lambda = c[T_m - T_0(x)]/(\pi^{1/2}L)$, accurate for small λ , has been used for the frozen margin thickness, and the time for frozen margin growth t is taken to be the dike length divided by the velocity in (8). The full solution [Equations (1)–(3), (5), and (14)] shows that propagation is impossible for $\beta > \sim 0.15$, because the frozen margin would grow faster than the dike could widen. For $p_0 = 3$ MPa, $|dT_0/dx| = 0.1^\circ\text{C km}^{-1}$, and expected values of the other variables, typical basalts have values 1 to 2 orders of magnitude less than this critical value, while rhyolites have values 1 to 2 orders of magnitude greater (in addition to the difference in viscosity, L for rhyolites is approximately 200 kJ kg^{-3} , compared to 500 kJ kg^{-3} for basalts). However, the dependence of β on $(\eta/p_0^5)^{1/2}$ means that an order-of-magnitude increase in excess source pressure can offset a 5 order-of-magnitude increase in magma viscosity.

Taken at face value, the self-similar solution implies that dike intrusion is impossible for β greater than the critical value. However, if the magma temperature is above the solidus, then any dike can propagate at least a short distance before freezing. In this case, propagation is not self-similar because a length scale has been introduced—the distance from the source at which the host rock temperature drops below the magma solidus. Nearer the source than

(dimensionally) the product of the flow velocity and the thermal diffusion time $\bar{u}_x w^2 / \kappa$ (Equation 15), most of the heat entering the dike still resides in the magma. This region is known as the *thermal entrance length* (Delaney & Pollard 1982). For a growing dike, this length is not constant because both dike velocity and thickness are proportional to dike length (Equations 1 and 8), so the thermal entrance length increases as the dike length cubed. This means that short dikes are always longer than the thermal entrance length, and magma reaching the dike tip has cooled to the temperature of the host rock. In sufficiently large dikes, on the other hand, the magma temperature is essentially unchanged from the source. In order for a dike to escape the source region, the rapidly increasing thermal entrance length must exceed the dike length before the dike tip encounters sub-solidus rock.

Numerical results indicate that, for $p_0 = 3$ Mpa, a basalt dike with $\eta = 50$ Pa s needs a partially molten envelope only several meters thick to escape a magma reservoir, but a rhyolite dike with $\eta = 10^6$ Pa s might require an envelope a few kilometers thick (Rubin 1995). This can explain the observation that granite dikes are quite common near granitic plutons, but rare elsewhere. At the same time, the dependence of this critical distance on $(\eta/p_0^5)^{1/2}$ means that occasional granitic dikes with a large excess source pressure will be able to survive thermally. Recently, it has been proposed that granitic magmas are transported through the crust in dikes and then “balloon” into more equidimensional plutons at the observed level of exposure (Clemens & Mawer 1992, Petford et al 1993). If so, then the process that stops the dikes and ultimately leads to the “ballooning” (for rhyolites, but apparently not basalts) remains a major mystery.

Once a dike grows sufficiently that magma buoyancy becomes important, the two-dimensional formulation is no longer appropriate because propagation occurs primarily in one direction and does not continue to widen the dike. In addition, for sufficiently large dikes it is unrealistic to assume a constant source pressure. Lister (1994) considers the effect of magma freezing on the buoyant rise of two-dimensional dikes of finite volume. He uses the formulation of Spence & Turcotte (1990) that neglects elasticity, which is applicable farther from the dike tip than the length scale in (9). In this case, freezing impedes the flow by removing magma from circulation and by reducing the channel aperture; however, the dike walls are very flexible in that they respond passively to mass balance of the flow.

For $T_m = T_s$, the maximum height of ascent is

$$h_f \approx 0.48 \lambda^{-2/5} \left(\frac{3}{4} A \right)^{4/5} \left(\frac{\Delta \rho g}{\eta \kappa} \right)^{1/5}, \quad (19)$$

where A is the cross-sectional area and λ is calculated on the basis of a uniform T_0 . For $(T_m - T_0) = 500^\circ\text{C}$, $\Delta \rho = 300 \text{ kg m}^{-3}$, $\eta = 100 \text{ Pa s}$, and thermal properties as above, a basalt dike could grow to a height of 130 km, and be

nearly a meter wide, for $A = 10^5 \text{ m}^2$. More complex results, which allow for meltback near the source, are given for the case $T_m > T_s$. If the source depth is shallower than h_f , the analysis can be used to estimate the fraction of basalt that erupts. Because elasticity is neglected, the possible role of freezing near the narrow dike tip in limiting propagation cannot be evaluated.

These two-dimensional pressure-dominated and buoyancy-dominated solutions leave several questions concerning real dikes unanswered. Lister & Kerr (1991) and Lister et al (1991) point out that freezing in three-dimensional dikes is likely to limit flow in the secondary direction; that is, the lateral extent of buoyant dikes or the height of laterally propagating dikes. In addition, three-dimensional flows may evolve even from two-dimensional boundary conditions. Delaney & Pollard (1982) and Bruce & Huppert (1990) concluded that, due to the sensitivity of heat flux to w^3 , small variations in dike thickness could amplify with time, with narrower regions freezing quickly and thicker regions freezing slowly or melting the host rock. This is consistent with the observation that curtain-of-fire fissure eruptions typically localize to one or a few plugs within a matter of hours to days.

Help From Field Observations

It may be apparent from the above paragraphs that our understanding of the thermal aspects of dike intrusion is less complete than our understanding of the purely mechanical aspects. In part, this results from conceptual difficulties in devising simple scaling relations that include multiple interacting processes, variation both along the dike length and with time, and sufficiently general boundary conditions. In part, it results from computational difficulties encountered in carrying out full numerical solutions. Even the existing semi-analytic solutions contain enough variables raised to enough powers that small uncertainties in the appropriate values can lead to substantial differences in predicted behavior; this makes observations of active dikes only very loose constraints on the theory. Finally, we have limited understanding of several processes of importance, e.g. the interaction between solidification and flow for real magmas that crystallize both along a stagnant margin and internally.

Given this state of affairs, it would be useful for future work to have more guidance from field observations. One such observation could be the variation in crystal size distribution both across and along dikes that can be followed for much of their length. Temperature indicators such as metamorphism or remagnetization within host rock may also be used to determine if advection has raised the contact temperature to well above $(T_m + T_0)/2$ (e.g. McClelland-Brown 1981). In addition, observations that are as yet difficult to interpret can be enlightening. Platten & Watterson (1987) describe crystal segregation within phenocryst-bearing dikes that may be indicative of the time-varying relative strengths of frozen margin growth and shear flow. Knight & Walker (1988) and

Smith (1987) describe slickenside-like lineations within dike margins that are presumably indicative of shear failure during solidification.

ROCK FRACTURE REVISITED

Inelastic Deformation at Dike Tips and Induced Seismicity

As noted earlier, the large process zones adjacent to some dikes indicate that fracture energy depends upon the dike, as well as the host rock. To the extent that earthquakes associated with dike propagation are important for eruption forecasting or for providing clues to the mechanisms of magma transport at depth, it is important to understand the causes of this deformation. To do so, one must use models that avoid the crack-tip stress singularity obtained in LEFM. Rubin (1993c) adopted the Barenblatt model, which (a) ensures that the tip propagation criterion is exactly met and (b) assumes that all inelastic deformation is restricted to the dike plane. The resulting stress field can then be examined to see if reasonable failure criteria are exceeded off the plane.

For propagating dikes, the near-tip stress field is dominated by the large suction generated by viscous flow of magma within the dike. Perturbations to the ambient stress are on the order of the tip suction and act over regions on the order of the tip cavity length. Elastic dilation of the host rock produces a drop in pore pressure near the dike tip that depends upon the dike speed and host rock poroelastic properties. If this reduced pressure exceeds the vapor pressure of the magma, then pore fluids flow into the tip cavity, the pore fluid pressure off the dike plane is greater than the local least compressive stress, and inelastic deformation is enhanced (Figure 12). On the other hand, if exsolving volatiles maintain the tip cavity pressure above the adjacent pore fluid pressure, the local pore pressure is less than the least compressive stress and inelastic deformation may be suppressed. The sensitivity of the pore pressure field to dike velocity suggests that basalt and rhyolite dikes may behave differently in this regard.

Both tensile and shear failure off the dike plane are possible. If pore fluids flow toward the tip cavity, then within most of the region of excess pore pressure the greatest compression is dike-parallel, and dike-parallel joints may form even while the tip propagates at equilibrium (Rubin 1993c, figure 9). For sills, the shear strength of bedding surfaces close to the fracture tip and magma front is exceeded; the resulting slip is likely responsible for the blunt terminations of many sills (Pollard et al 1975, Rubin 1993c, figure 12). If inelastic deformation of any sort extends over the scale of the tip cavity, then because the tip cavity grows with dike size (Figure 5), it is probable that the fracture energy increases with dike size. Furthermore, the total fracture energy may be rather insensitive to whether the dike dilates an existing fracture or creates its own.

The generation of shear failure on planes oblique to the dike is more complex. Tentatively, it appears that: 1. It is difficult for dikes to produce shear failure of

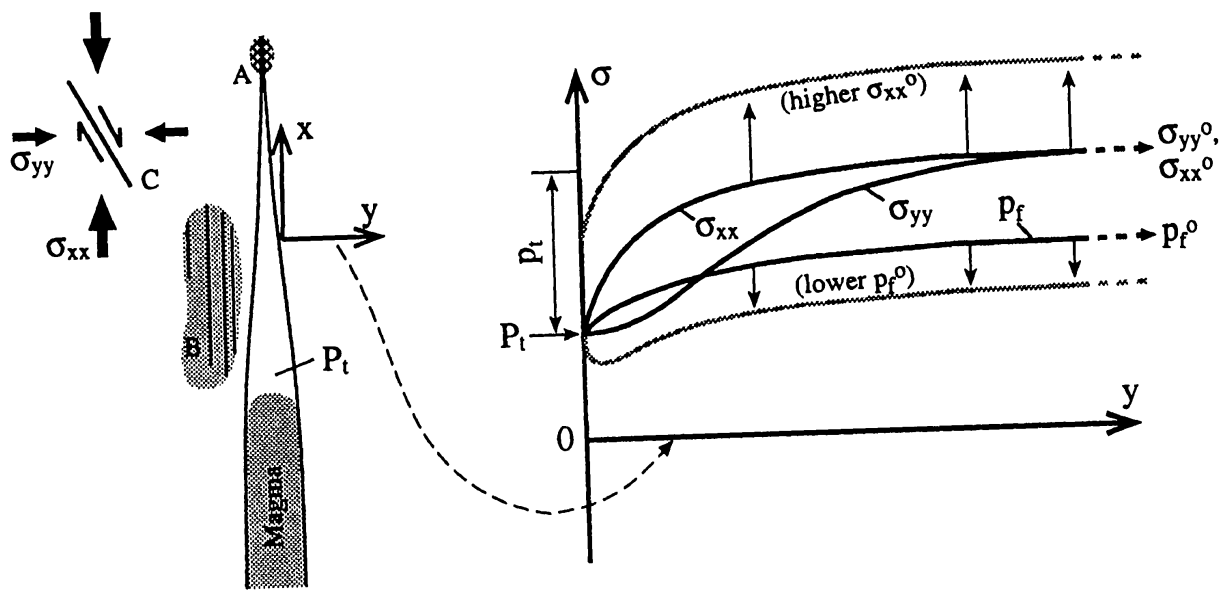


Figure 12 Illustration of some of the many possible stress states surrounding a dike tip cavity. σ_{yy}^0 and σ_{xx}^0 , ambient dike-perpendicular and dike-parallel stresses; p_f^0 , ambient pore fluid pressure; P_t , pressure in tip cavity; p_t , tip suction. Curves on right are schematic stresses along perpendicular transect near middle of tip cavity. Large stress changes occur on the length scale of the tip cavity; principal stresses are close to dike-parallel and dike-perpendicular over most of this region. If the tip cavity is fed by influx of pore fluids (black p_f curve), p_f locally exceeds σ_{yy} because $dp_f/dy > 0$ is implied, but from elasticity $d\sigma_{yy}/dy = 0$ at the wall. If p_f^0 is less but the tip pressure is maintained at P_t by magmatic volatiles, p_f may be everywhere less than the least compressive stress (and far from the free surface). Note that if the dike is perpendicular to the least compressive stress (and far from the free surface), $\Delta\sigma_{xx} = \Delta\sigma_{yy}$ along the dike plane, so $(\sigma_{xx} - \sigma_{yy}) = (\sigma_{xx}^0 - \sigma_{yy}^0)$ along the cavity wall. Diagram on left illustrates likely styles of failure off the dike plane: (A) Intergranular damage at centimeter scale or less at crack tip; (B) dike-parallel joints on the scale of the tip cavity when the cavity is fed by pore fluids; (C) slip along preexisting fractures oblique to dike on the scale of the tip cavity when $\sigma_{xx}^0 = \sigma_{yy}^0$, or on a larger scale for large $\sigma_{xx}^0 - \sigma_{yy}^0$.

intact rock, but slip along suitably oriented existing fractures is likely. 2. If the ambient stress is isotropic, then the stress perturbation due to intrusion is sufficient to drive faulting only on the scale of the tip cavity. 3. The production of earthquakes larger than magnitude 1 to 2 probably requires a large enough ambient differential stress that the earthquake focal mechanisms should be consistent with the ambient stress field. 4. If the ambient differential tension is so large that existing faults are near failure, then the stress perturbation due to intrusion can induce failure over distances that scale with the dike length, rather than the tip cavity length. Such deformation is analogous to the "large-scale yielding" conditions encountered during fracture growth in ductile metals. Rubin (1992) found that such a stress state was required to generate subsidence of a graben several kilometers across during dike intrusion north of the Krafla caldera.

Shallow intrusions at Kilauea generally produce several magnitude 2 earthquakes with the largest approaching magnitude 3; at Krafla numerous magnitude

3 earthquakes were produced with the largest approaching magnitude 4. Assuming standard moment-magnitude relations and stress drops, magnitude 2 to 3 earthquakes imply fault dimensions of 100 to several hundred meters. The earthquakes are concentrated at the leading edge of the migrating swarm (see, in particular, Einarsson & Brandsdottir 1980), but because considerable seismicity often occurs well behind the front (Klein et al 1987), the precise relation to the dike tip process zone is uncertain (thermally induced pore pressure increases behind the magma front are also a possible cause). The migrating swarms sweep out regions 1 to 2 km thick; this could reflect the true extent of the failure zone or relative location error. Karpin & Thurber (1987) found that magnitude 2 to 3 earthquakes during small intrusions at Kilauea were strike-slip events with the greatest compression aligned approximately parallel to the dike, consistent with conclusion 3. above. The model of Hill (1977) also predicts such focal mechanisms; however, that model supposes that this slip occurs on fractures linking en-echelon dike segments.

The stress calculations indicate that shear failure near dikes is very sensitive to the ambient differential stress. The July 1978 intrusion at Krafla produced nearly 100 earthquakes above magnitude 3, over a nominal dike area of 100 km² (Einarsson & Brandsdottir 1980). In contrast, the 1991 intrusion at Hekla probably covered about 16 km² (Linde et al 1993), yet produced no earthquakes above magnitude 2 and only three above magnitude 1 (S Jakobsdottir, personal communication). The possibility that dikes propagate through deep (hot) regions of low differential stress without producing any detectable earthquakes seems quite real.

While the evidence for larger-than-lab fracture energies is clear, it is not yet known if this significantly impedes propagation. The nominal energy release rate G is proportional to $\Delta P^2 l / M$. For $\Delta P = 1$ MPa, laboratory values of G_c are reached for $l \approx 1$ m; thus a factor of 1000 increase might be necessary to significantly affect the propagation of 1-km dikes. Delaney et al (1986) counted close to 100 dike-parallel joints for some kilometer-scale dikes on the Colorado Plateau; assuming that each consumed the energy of a laboratory fracture, this implies a factor of 100 increase in G_c . An estimate of the fracture energy implied by the seismicity associated with the July 1978 intrusion at Krafla, subject to large error, is 100 to 1000 times laboratory values. Hints of a significant effect also come from monitoring artificial hydrofractures, where measured pumping pressures are significantly higher than predicted for a given flux. This is frequently attributed to larger-than-lab fracture energies, although this inference is subject to uncertainty as well (Shlyapobersky & Chudnovsky 1992).

Inelastic deformation at dike tips interferes with propagation because stored elastic strain energy that could have gone into widening the tip is instead consumed in distributed deformation of the host rock. In extreme cases this deformation might lead to dike arrest. The dike in Figure 13 stops shortly beyond the

point at which slip along a zone of fractures decreases the thickness abruptly; enhanced freezing in the thinned region may have played a role. Similar processes may occur in the rift zones of Kilauea, where some regions of high seismicity apparently act as consistent barriers to dike propagation (Klein et al 1987). The transition from tensile crack propagation to shear failure of the host rock also plays an important role in the growth of some laccoliths from sills (Johnson & Pollard 1973).

“Anomalous” Earthquakes During Dike Propagation

In addition to the “normal,” “volcano-tectonic,” or “short-period” earthquakes mentioned above, several types of seismicity are frequently associated with volcanism generally, and perhaps with dikes in particular. These include harmonic or volcanic tremor, “long-period” earthquakes, and nondouble-couple earthquakes; each has been observed in association with hydrofractures as well (Bame & Fehler 1986, Majer & Doe 1986, Talebi & Cornet 1987). Harmonic tremor accompanies essentially all eruptions and most known dike intrusions; it consists of sustained ground motion over a narrow frequency interval near 1 Hz. Long-period earthquakes have an amplitude envelope similar to normal earthquakes, but the frequency content of tremor. The similar spectra, and the fact that they frequently occur together in time and space, suggest that long-period events and tremor share a common source; in fact, it is often suggested that tremor is simply a swarm of long-period events (Koyanagi et al 1987).

One candidate mechanism for producing “anomalous” earthquakes is the dynamic propagation of tensile cracks. Unlike shear failure, tensile failure generates compressional P-wave arrivals everywhere. Such focal mechanisms have sometimes been observed during hydrofracture growth in laboratory tests (Majer & Doe 1986). Foulger (1988) concluded that numerous nondouble-couple magnitude -2 to 1 earthquakes near the Hengill geothermal region, Iceland, resulted from rapid tensile crack growth due to rock cooling. Three magnitude 5 to 6 events at Mammoth Lakes in 1980 are candidate nondouble-couple earthquakes, which are sometimes attributed to dike propagation (Julian & Sipkin 1985). However, this interpretation remains controversial among both seismologists and volcanologists, in part because of the difficulty in getting cracks filled with viscous fluid to propagate rapidly on such a scale. Sammis & Julian (1987) show that, under certain conditions, a fracture at the margin of a magma reservoir can propagate dynamically to the requisite size even if magma does not follow the crack tip. However, this mechanism requires an initially dry crack and an exceedingly large magma pressure so that absolute tensions are reached outside the reservoir; it does not work if the seed crack is filled with fluid at the reservoir pressure.

Several mechanisms have been proposed for tremor and long-period events;

(a)



[see Chouet (1992) for a recent review]. Most models consider the peaked spectrum to result from normal-mode vibrations of a fluid-filled cavity; cracks are the favored geometry when known to be present or when the spectra suggest more than one dominant length scale (Ferrazini et al 1990, Chouet et al 1994). Because the important point is the resonance, numerous processes could act as the trigger; suggestions include “jerky” dike propagation (Aki et al 1977), bubble nucleation and collapse in the magma (Chouet 1992) or associated

(b)



Figure 13 Dike termination in shear zone on horizontal pavement on Colorado Plateau. (a) Dike (dotted outline) with bleached zone of fractures extending a short distance to upper left of tip (near lens cap) (photo courtesy of P Delaney). In (b), magma front (arrow) extends slightly beyond the intersection of the fault zone with the dike; crack followed by magma continues beyond top of photo. A small amount of magmatic material has infiltrated the fracture zone. (Lens cap has been moved to the opposite side of the dike.)

hydrothermal systems (Leet 1988), and nearby fault slip (Lahr et al 1994). The peaked spectrum is indicative of a weakly damped source, which in turn implies a large impedance contrast between the fluid and the rock, requiring at least a small fraction of bubbles in the magma. Large impedance contrasts are hard to explain at mantle depths, where hypothetical volatiles (presumably CO_2) are not compressible. For this reason, kinematic models of tremor deeper than 20 km beneath Hawaii have relied on repeated triggering at about the observed tremor frequency (Aki & Koyanagi 1981).

Here I briefly address possible roles of rock fracture in generating “anomalous” earthquakes. Consider first how jerky dike propagation might be, given a heterogeneous distribution of laboratory-type (interlocking grain) fracture resistance at the tip. If it is assumed that the fluid in the tip cavity is highly compressible, and thus is able to maintain pressure as the cavity expands, and that the fracture toughness drops rapidly from laboratory values of K_c to zero,

one may show that the jump in crack length (including elastodynamic effects) is given dimensionally by

$$\frac{\Delta R}{R} \sim \frac{K_c}{\Delta P \sqrt{l}}, \quad (20)$$

where R is the tip cavity length and ΔR is the jump in length. Because for large dikes the nominal fracture toughness is considerably less than $\Delta P l^{1/2}$, the jump in length is considerably less than the tip cavity length, and, given that dynamic propagation velocities are close to elastic wave speeds, the crack tip probably stops before finding out if the magma is compressible. For R less than some tens of meters, Equation (20) suggests that crack jumps of meters are plausible. Whether the resulting seismic radiation would be detectable remains to be determined.

Of course, different types of seismicity are possible if different types of crack-tip deformation occur. Because the fault zone in Figure 13 cuts only one dike margin, it clearly happened while the magma was fluid; perhaps it generated a long-period event. In fact, some long-period events have a high-frequency onset, suggesting that dike resonance was triggered by shear failure of the host rock (the “hybrid events” of Lahr et al 1994). One possibility is that the difference between short-period and long-period earthquakes is simply the distance separating the causative fault and the dike. The gap in both long- and short-period earthquakes at 20 km depth beneath Hawaii (Klein et al 1987, Koyanagi et al 1987) suggests that such an association is possible. What determines the critical separation has yet to be established.

Cracks in the Magma Source Region

Our understanding of dike initiation is meager, owing to the inaccessibility of magma source regions and a paucity of experimental data on the rheology of partially molten rock. The chemistry of mid-ocean ridge basalts indicates that much of the melt produced deep within the melting region ascends without reequilibrating chemically with shallower mantle (Kelemen et al 1995). It has been suggested that rapid ascent through fractures could allow such disequilibrium to occur. A limited number of experiments indicate that, when stressed, melt pockets in deforming partial melts are preferentially aligned with their long axes parallel to the maximum compression direction. The underlying mechanism for this alignment is unknown, however, and as such experiments are generally carried out at very large differential stresses and with an uncontrolled melt pressure, more experiments are desirable.

Stevenson (1989) showed that if the viscosity of the host matrix decreases with increasing melt fraction, then regions of high melt fraction are also regions of low melt pressure. For likely parameters this gives rise to an instability in which melt flows into regions where it has already accumulated. Such a

process may be a necessary step in the formation of cracks in a partial melt. Analogous behavior occurs once the melt fraction locally becomes so large that those regions start looking more like cracks in a homogeneous medium and less like low-viscosity regions in a heterogeneous medium. That is, the cracks possess an internal fluid pressure that is greater than the least compressive stress (so they dilate) but less than the regional pore fluid pressure (so they draw fluid). Sleep (1988) examined the dilation of veins in a viscous matrix, but did not consider how the veins lengthened. Without specifying a propagation criterion, it is not clear when it becomes appropriate to view deformation of the host rock during crack growth as elastic. To the extent that propagation rates are limited by influx of magma, the large difference in viscosity between basalts and rhyolites implies that elastic deformation takes over at a much earlier stage in the growth of basalt dikes; this could make initiation of basalt dikes easier.

The processes discussed by Stevenson and Sleep require that the melt pressure exceed the least compressive stress in the host matrix. If this criterion is met, there may be several unexplored mechanisms by which cracks could grow. Stress concentrations at the tips of small existing cracks, or equivalently along the appropriate grain-grain boundaries in regions of high melt fraction, increase the deviatoric stress and decrease the pressure (average stress) in the solid. Because the melting temperature of peridotite decreases with decreasing pressure, melting will occur preferentially in these regions; this is essentially a form of subcritical crack growth.

The important issue of the melt pressure relative to the matrix stresses may not yet be resolved. In most studies of mantle melt migration, it has been concluded that, except within the compacting layer of order 100 m at the base of the melting region, the melt pressure is less than the average stress in the solid by only about 1 kPa, owing to surface tension. This is negligible compared to the expected differential stress in the solid. However, in a more complete thermodynamic treatment, Fowler (1990) argues that compaction occurs at a reduced rate throughout the entire melting column, and that melt pressures are likely to be well below the least compressive stress everywhere except within the freezing boundary layer at the top. In this case fractures would not be expected except in this freezing region; if correct, then some other mechanism is required to generate the melt channels implied by the geochemical data. One possibility is the "reactive infiltration instability" advocated by Kelemen et al (1995). In this scenario, the pressure dependence of phase relations causes melts ascending by porous flow to preferentially dissolve pyroxene, producing high-porosity, nonreactive dunite channels that capture most of the flow.

There is abundant field evidence that melt-filled fractures grow in magma source regions. In peridotite massifs, these veins or dikes range from

undeformed to strongly deformed and have consistent orientations relative to deformation fabrics, indicating that fracture and large-scale flow of the host rock were contemporaneous (Nicolas & Jackson 1982, Nicolas 1986). Frequently, the more extensive dikes have centimeter-scale dunite margins where orthopyroxene has been selectively dissolved, indicating that the host rock was partially molten during intrusion (Kelemen et al 1992). Ceuleneer & Rabinowicz (1992) describe small gabbro dikes in the Oman ophiolite that apparently root within melt-impregnated pods of peridotite. Textural relations suggest that some aplite dikes in granites also intruded while the host was partially molten (Hibbard & Watters 1985).

Many dikes within peridotites are perpendicular to stretching lineations, but smaller melt segregations may be parallel to shear planes oblique to the foliation. The former are interpreted as tensile fractures that formed when the melt pressure exceeded the least compressive stress by some tensile strength; the latter may be shear fractures that formed when the excess melt pressure was less, but still large enough to dilate the resulting fractures (Shaw 1980). Davidson et al (1994) describe considerably larger conjugate melt-filled shear fractures that formed in partially molten crustal rocks. Dick et al (1993) use geochemical evidence to argue that large-scale shear zones in mantle peridotites are preferred paths for melt migration. While the evidence for magma-filled fractures in partial melts is beyond dispute, there is a substantial gap between the field observations and our understanding of fracture initiation and growth in these settings.

DIKE INTRUSION AND TECTONICS

The interaction between dike propagation and tectonics is varied; here I provide only a flavor of what one might consider. The influence of the ambient stress on dike orientation has already been mentioned. Tectonics can also determine the direction of propagation within the dike plane. Figure 14 shows a stress state appropriate for a region undergoing horizontal extension by faulting in the upper crust and ductile flow at depth (Hickman 1991). Note that within large portions of the lithosphere, rock is sufficiently strong that gradients in the tectonic stress can significantly exceed $\Delta\rho g$. That is, changes in dS_v/dz due to density changes (*arrow*) are barely perceptible, relative to changes in dS_h/dz due to the brittle-ductile transition or rheological contrasts at the crust-mantle boundary. Thus, tectonic extension may dramatically increase both the depth of effective neutral buoyancy and the tendency for lateral rather than vertical propagation at this depth. For the horizontal stress indicated, a dike would require an elastic thickness of approximately 100 m in order to erupt (Equation 1 with $l \sim 20$ km, $\Delta P \sim 50$ MPa, and $M = 20$ GPa) and would probably be hundreds, if not thousands, of kilometers long. In this sense extension makes the eruption of individual dikes harder, not easier; in addition, an ambient stress

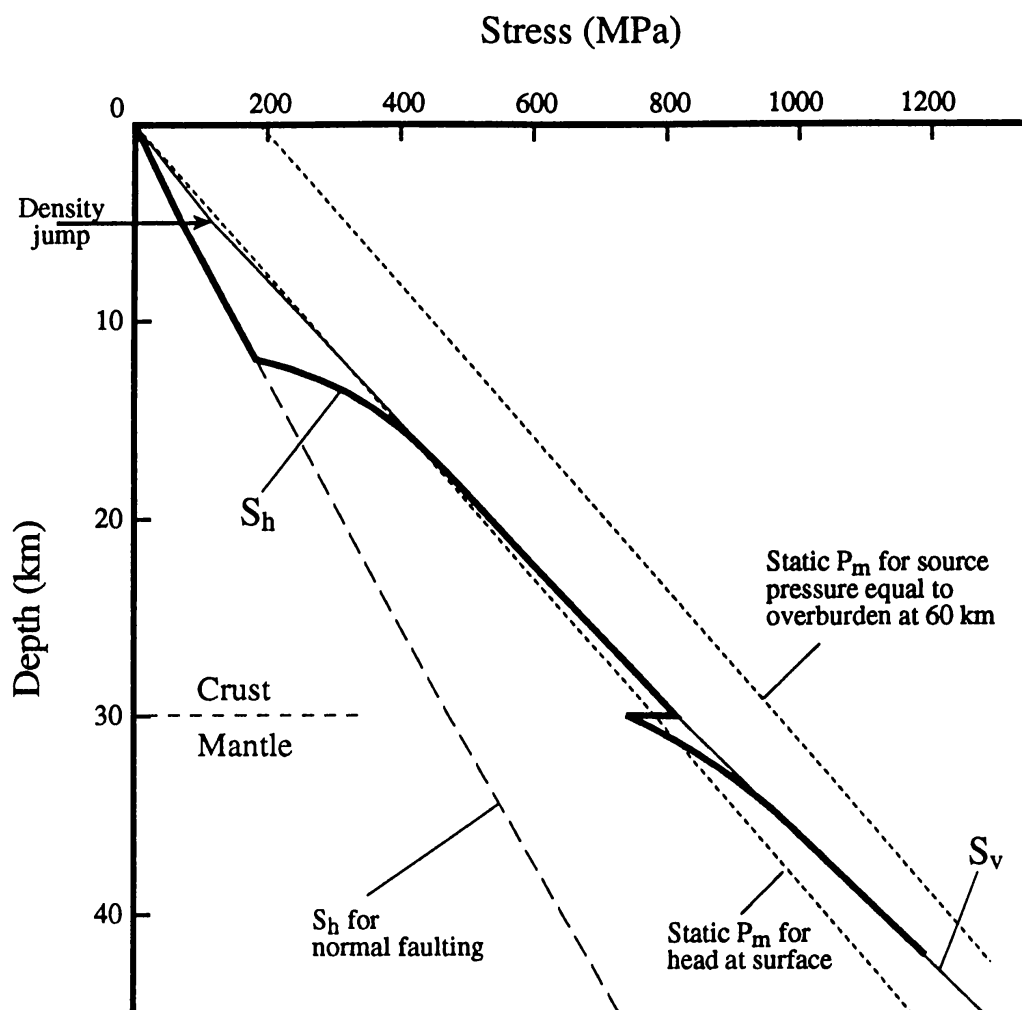


Figure 14 Hypothetical stress state, drawn to scale, for a region undergoing horizontal extension. S_v , vertical stress; S_h horizontal stress. Host rock density $\rho_r = 2.3 \times 10^3 \text{ kg m}^{-3}$ above 5 km depth, $2.8 \times 10^3 \text{ kg m}^{-3}$ below 5 km depth, $3.1 \times 10^3 \text{ kg m}^{-3}$ within mantle. Reference magma pressure curves P_m are for magma density $\rho_m = 2.6 \times 10^3 \text{ kg m}^{-3}$. It is assumed that normal faults dipping at 60° are in a state of incipient slip above the brittle-ductile transition at 12 km depth, and that pore pressure is hydrostatic. Within the middle to upper crust and at the crust-mantle boundary, gradients in tectonic stress may be much more significant than magma buoyancy in determining dike propagation direction.

state near failure increases the chance that ascent will be inhibited by induced normal faulting.

On the length scale of the dike, intrusion raises the dike-perpendicular stress to near the magma pressure [see Pollard & Segall (1987) for the crack-scale variation of this and other stresses]. Thus eruption could occur if the “stress deficit” produced by extension in Figure 14 were filled by many small dikes, rather than a single large dike. This seems to have been the case during the 1975–1984 activity at Krafla, which progressed from primarily subsurface intrusions prior

to 1980, to primarily eruptive fissures thereafter (Björnsson 1985). Apparently the early dikes could be intruded at low enough pressure that the magma head was below the surface, while later dikes required pressures sufficient to erupt in order to intrude. The inference of an increasing magma pressure is supported by the general increase in reservoir elevation from intrusion to intrusion. Analogous behavior might be required for eruption in other settings. The computed bending stresses beneath some ocean island volcanoes are so large that a dike within the lower elastic lithosphere would be driven downward, regardless of the magma buoyancy. The fact that magma erupts in these settings suggests that enough dikes have intruded to raise the horizontal stress significantly.

In the absence of mechanisms for relaxing the stresses due to intrusion, repeated intrusion can raise the dike-perpendicular stress until it is no longer the least compressive stress, causing future sheet intrusions to take the form of sills or dikes of a different orientation. Thus a wide range of magmatic and tectonic styles is possible, depending upon the relative rates of magma supply and tectonic extension. Radial dike swarms, such as at Spanish Peaks, result when the intrusion rate outpaces the ability of tectonic extension to renew the preintrusion stress orientations. That is, after one dike intrudes, the preferred location for the next is elsewhere. At the other extreme, linear rift zones develop when extension continually renews the orientation of the least compressive stress, such as in Iceland as a result of plate spreading. Mauna Loa, with numerous radial dikes outside its two primary rift zones, may be an intermediate case (Rubin 1990).

Parsons et al (1992) suggest that if a dike crosses the crust-mantle boundary in a region of horizontal extension, the large ΔP within the uppermost mantle (Figure 14) increases the horizontal compression within the lower crust. Because S_h is already close to S_v in this region, following intrusion S_h could exceed S_v even if the magma pressure were less than S_v . This could cause subsequent ascending dikes to turn into sills in the lower crust; such sills could be the cause of the horizontal seismic reflectors prevalent in the lower crust of extending regions. It remains to be established whether this mechanism can operate effectively in three dimensions, however. A dike that encounters the stress state of Figure 14 will preferentially spread laterally within the uppermost mantle; determining if S_h exceeds S_v in this case would require two-dimensional elastic calculations. Gudmundsson (1986) similarly suggests that dike intrusion across rheological contrasts might cause subsequent dikes to turn into sills; however, he relies on variations in elastic stiffness between layered basalt flows and breccias to cause the stiffer basalt layers to bear most of the increased horizontal compression due to the dike.

Because intrusion raises the least compressive stress to the level of the magma pressure, in the long term one expects to see decreased normal fault activity in regions of frequent dike intrusion (notwithstanding the local increase in fault

slip beyond the dike perimeter). Bursik & Sieh (1989) documented how the extension of $\sim 1 \text{ mm yr}^{-1}$ along the Sierra Nevada range front fault system has been superseded by dike intrusion in the Mono Craters area over the last 40 ka. Parsons & Thompson (1991) give numerous other examples where, in regions undergoing horizontal extension, active faulting and topographic relief have been suppressed by magmatic activity and presumed dike intrusion.

At times it may be too limiting to view dike intrusion and regional extension as independent processes. This is not a concern along mid-ocean ridges, where extension may lead to tectonic thinning and exposure of mantle peridotites during periods of low magma supply. In oceanic islands such as Hawaii, however, intrusion is probably an important part of the force balance that leads to widening of the rift zone. More generally, the integrated strength of the lithosphere is represented by the area between the S_h and S_v curves in Figure 14. If the force available to drive extension (resulting from local uplift, for example) is less than this strength, then continental breakup or rift propagation may occur only if localized dike intrusion "weakens" the lithosphere (raises the horizontal stress) sufficiently.

THE FUTURE

While our understanding of dike propagation has increased significantly over the past decade, our ability to incorporate this knowledge in large-scale geochemical or tectonic pictures is still quite limited. I close with a list of some areas where advances in the near future could facilitate this process. Beginning with near source processes, there has been considerable speculation on how cracks initiate in a partial melt, but more experiments are crucial, particularly ones in which the melt pressure is accurately characterized. Once such cracks form we need to understand better how they tap melt from the source region; models of this process need to address crack growth, magma buoyancy, and the viscous-elastic transition in the host rock. Once dikes grow to be fairly large, we need to understand better how they generate both normal and long-period earthquakes. Although earthquakes occur to depths of 60 km beneath Hawaii, currently the relation of those in the mantle to magma transport can only be guessed at.

Finally, we need to understand what processes may stop dikes beneath the surface, such as depletion of the source region, magma freezing, sill formation, and inelastic deformation of the host rock, and whether any of these processes may lead to the growth of shallow magma chambers. Incorporating thermal processes into more complete models of dike propagation is a difficult task and this seems to be one area where field observations have yet to be exploited to the extent that they might. Once these (and other) processes are better understood, we can more fully address such issues as how fracture transport through the source and lithosphere affects magma chemistry, how tapping of source regions

by dikes influences volcanic episodicity, how hundreds of cubic kilometers of magma can erupt in a single event, how long-lived volcanoes form, how granites ascend, how dike propagation in the crust might influence the spacing of oceanic transforms, and more.

ACKNOWLEDGMENTS

The ideas presented in this paper were developed over many years through discussions with many others, particularly Paul Delaney, David Pollard, Jim Rice, and Norm Sleep. I thank Dominique Gillard for critically reviewing the manuscript. Research carried out in the process of writing the paper was supported by the National Science Foundation under grant EAR-9219824 and by NASA under grant NAGW-3589.

Any *Annual Review* chapter, as well as any article cited in an *Annual Review* chapter, may be purchased from the Annual Reviews Preprints and Reprints service.
1-800-347-8007; 415-259-5017; email: arpr@class.org

Literature Cited

- Aki K, Fehler M, Das S. 1977. Source mechanism of volcanic tremor: fluid-driven crack models and their application to the 1963 Kilauea eruption. *J. Volcanol. Geotherm. Res.* 2:259–87
- Aki K, Koyanagi R. 1981. Deep volcanic tremor and magma ascent mechanism under Kilauea, Hawaii. *J. Geophys. Res.* 86:7095–109
- Anderson EM. 1936. The dynamics of the formation of cone-sheets, ring-dykes and cauldron-subsidences. *Proc. R. Soc. Edinburgh* 56:128–57
- Atkinson BK, ed. 1987. *Fracture Mechanics of Rock*. San Diego: Academic. 534 pp.
- Atkinson BK, Meredith PG. 1987a. The theory of subcritical crack growth with applications to minerals and rocks. See Atkinson, 1987, pp. 111–66
- Atkinson BK, Meredith PG. 1987b. Experimental fracture mechanics data for rocks and minerals. See Atkinson 1987, pp. 477–525
- Baer G. 1991. Mechanisms of dike propagation in layered rocks and in massive, porous sedimentary rocks. *J. Geophys. Res.* 96:11, 911–29
- Bame D, Fehler M. 1986. Observations of long period earthquakes accompanying hydraulic fracturing. *Geophys. Res. Lett.* 13:149–52
- Barenblatt GI. 1962. The mathematical theory of equilibrium cracks in brittle fracture. *Adv. Appl. Mech.* 7:55–129
- Batchelor GK. 1967. *An Introduction to Fluid Dynamics*. Cambridge: Cambridge Univ. Press. 615 pp.
- Bieniawski ZT. 1984. *Rock Mechanics Design in Mining and Tunneling*. Boston: Balkema. 272 pp.
- Bird RB, Stewart WE, Lightfoot EN. 1960. *Transport Phenomena*. New York: Wiley. 780 pp.
- Björnsson A. 1985. Dynamics of crustal rifting in Iceland. *J. Geophys. Res.* 90:10,151–62
- Bruce PM, Huppert HE. 1990. Solidification and melting along dikes by the laminar flow of basaltic magma. In *Magma Transport and Storage*, ed. MP Ryan, pp. 87–101. Chichester, England: Wiley. 420 pp.
- Bursik M, Sieh K. 1989. Range front faulting and volcanism in the Mono basin, eastern California. *J. Geophys. Res.* 94:15,587–609
- Carslaw HS, Jaeger JC. 1959. *Conduction of Heat in Solids*. Oxford: Clarendon. 510 pp. 2nd ed.
- Ceuleneer G, Rabinowicz M. 1992. Mantle flow and melt migration beneath oceanic ridges: models derived from observations in ophiolites. In *Mantle Flow and Melt Migration at Mid-Ocean Ridges*, ed. JP Morgan, DK Blackman, JM Sinton, *Am. Geophys. Union Geophys. Monogr.* 71, pp. 123–54. Washington, DC: Am. Geophys. Union. 361 pp.
- Chouet B. 1992. A seismic model for the source of long-period events and harmonic tremor. In *Volcanic Seismology, IAVCEI Proc. Volcanology 3*, ed. P Gasparini, R Scarpa, K Aki, pp. 212–22. Berlin/New York: Springer-Verlag. 572 pp.
- Chouet BA, Page RA, Stephens CD, Lahr JC,

- Power JA. 1994. Precursory swarms of long-period events at Redoubt Volcano (1989–1990), Alaska: their origin and use as a forecasting tool. *J. Volcanol. Geotherm. Res.* 62:95–135
- Clemens JD, Mawer CK. 1992. Granitic magma transport by fracture propagation. *Tectonophysics* 204:339–60
- Cotterell B, Rice JR. 1980. Slightly curved or kinked cracks. *Int. J. Fracture* 16:155–69
- Crawford G, Stevenson D. 1988. Gas-driven water volcanism and the resurfacing of Europa. *Icarus* 73:66–79
- Davidson C, Schmid SM, Hollister LS. 1994. Role of melt during deformation in the deep crust. *Terra Nova* 6:133–42
- Davies JH, Stevenson DJ. 1992. Physical model of source region of subduction zone volcanics. *J. Geophys. Res.* 97:2037–70
- Decker RW. 1987. Dynamics of Hawaiian volcanoes: an overview. See Decker et al 1987, pp. 1019–186
- Decker RW, Wright TL, Stauffer PH, eds. 1987. *Volcanism in Hawaii*. US Geol. Surv. Prof. Pap. 1350. 1667 pp.
- Delaney PT. 1987. Heat transfer during emplacement and cooling of mafic dykes. See Halls & Fahrig 1987, pp. 31–46
- Delaney PT, Pollard DD. 1981. Deformation of host rocks and flow of magma during growth of Minette dikes and breccia-bearing intrusions near Shiprock, New Mexico. *US Geol. Surv. Prof. Pap.* 1202. 61 pp.
- Delaney PT, Pollard DD. 1982. Solidification of basaltic magma during flow in a dike. *Am. J. Sci.* 282:856–85
- Delaney PT, Pollard DD, Ziony JI, McKee EH. 1986. Field relations between dikes and joints: emplacement processes and paleostress analysis. *J. Geophys. Res.* 91:4920–38
- Dick HJB, Keleman PB, Hirth G. 1993. Deformation and melt flow in the oceanic lower crust and mantle. *Eos, Trans. Am. Geophys. Union* (Spring Meet. Suppl.) 74(16):284 (Abstr.)
- Einarsson P, Brandsdottir B. 1980. Seismological evidence for lateral magma intrusion during the July 1978 deflation of the Krafla volcano in NE-Iceland. *J. Geophys.* 47:160–65
- Emmerman SH, Turcotte DL, Spence DA. 1986. Transport of magma and hydrothermal solutions by laminar and turbulent fluid fracture. *Phys. Earth Planet. Inter.* 36:276–84
- Ernst RE, Baragar WRA. 1992. Evidence from magnetic fabric for the flow pattern of magma in the Mackenzie giant radiating dyke swarm. *Nature* 356:511–12
- Ferrazzini V, Chouet B, Fehler M, Aki K. 1990. Quantitative analysis of long-period events recorded during hydrofracture experiments at Fenton Hill, New Mexico. *J. Geophys. Res.* 95:21,871–84
- Foulger GR. 1988. Hengill triple junction, SW Iceland 2. Anomalous earthquake focal mechanisms and implications for process within the geothermal reservoir and at accretionary plate boundaries. *J. Geophys. Res.* 93:13,507–23
- Fowler A. 1990. A compaction model for melt transport in the earth's asthenosphere. Part II: applications. In *Magma Transport and Storage*, ed. M Ryan, pp. 15–32. New York: Wiley. 420 pp.
- Fujii N, Uyeda S. 1974. Thermal instabilities during flow of magma in volcanic conduits. *J. Geophys. Res.* 79:3367–69
- Gartner AE. 1986. Geometry, emplacement history, petrography, and chemistry of a basaltic intrusive complex, San Rafael and Capitol Reef areas, Utah. *US Geol. Surv. Open File Rep.* 86-81. 112 pp.
- Griffith AA. 1920. The phenomena of rupture and flow in solids. *Philos. Trans. R. Soc. London* 221:163–97
- Gudmundsson A. 1986. Formation of crustal magma chambers in Iceland. *Geology* 14:154–66
- Gudmundsson A. 1990. Dyke emplacement at divergent plate boundaries. In *Mafic Dykes and Emplacement Mechanisms*, ed. AJ Parker, PC Rickwood, DH Tucker, pp. 47–62. Rotterdam: Balkema. 541 pp.
- Haimson B, Fairhurst C. 1969. Hydraulic fracturing in porous-permeable materials. *J. Petrol. Technol.* 21:811–17
- Halls HC, Fahrig WH, eds. 1987. *Mafic Dyke Swarms*. Geol. Assoc. Can. Spec. Pap. 34. 503 pp.
- Hargraves RB, Johnson D, Chan CY. 1991. Distribution anisotropy: the cause of AMS in igneous rocks? *Geophys. Res. Lett.* 18:2193–96
- Heimpel M, Olson P. 1994. Buoyancy-driven fracture and magma transport through the lithosphere: models and experiments. In *Magmatic Systems*, ed. M Ryan, pp. 223–40. San Diego: Academic. 401 pp.
- Hibbard MJ, Watters RJ. 1985. Fracturing and diking in incompletely crystallized granitic plutons. *Lithos* 18:1–12
- Hickman SH. 1991. Stress in the lithosphere and the strength of active faults. *US Natl. Rep. Int. Union Geol. Geophys. 1987–1990, Rev. Geophys.* 29:759–75
- Hill DP. 1977. A model for earthquake swarms. *J. Geophys. Res.* 82:1347–52
- Hutton DHW. 1992. Granite sheeted complexes: evidence for the dyking ascent mechanism. *Trans. R. Soc. Edinburgh: Earth Sci.* 83:377–82
- Ingraffea AR. 1987. Theory of crack initiation and propagation in rock. See Atkinson 1987, pp. 71–110
- Johnson AM, Pollard DD. 1973. Mechanics of growth of some laccolithic intrusions in the

- Henry Mountains, Utah, I: field observations, Gilbert's model, physical properties and flow of the magma. *Tectonophysics* 18:261–309
- Julian BR, Sipkin SA. 1985. Earthquake processes in the Long Valley Caldera area, California. *J. Geophys. Res.* 90:11,155–69
- Karpin TL, Thurber CH. 1987. The relationship between earthquake swarms and magma transport: Kilauea Volcano, Hawaii. *Pure Appl. Geophys.* 125:971–91
- Kelemen PB, Dick HJB, Quick JE. 1992. Formation of harzburgite by pervasive melt/rock reaction in the upper mantle. *Nature* 358:635–41
- Kelemen PB, Whitehead JA, Aharonov E, Jordahl KA. 1995. Experiments on flow focussing in partially soluble porous media, with applications to melt extraction from the mantle. *J. Geophys. Res.* In press
- Kidd RGW. 1977. A model for the process of formation of the upper oceanic crust. *Geophys. J. R. Astron. Soc.* 50:149–83
- Klein F, Koyanagi RY, Nakata JS, Tanigawa WR. 1987. The seismicity of Kilauea's magma system. See Decker et al 1987, pp. 1019–186
- Knight MD, Walker GPL. 1988. Magma flow directions in dikes of the Koolau complex, Oahu, determined from magnetic fabric studies. *J. Geophys. Res.* 93:4301–20
- Koyanagi R, Chouet B, Aki K. 1987. Origin of volcanic tremor in Hawaii. See Decker et al 1987, pp. 1221–57
- Lahr JC, Chouet BA, Stephens CD, Power JA, Page RA. 1994. Earthquake classification, location, and error analysis in a volcanic environment: implications for the magmatic system of the 1989–1990 eruptions at Redoubt Volcano, Alaska. *J. Volcanol. Geotherm. Res.* 62:137–51
- Lawn BR, Wilshaw TR. 1975. *Fracture of Brittle Solids*. New York: Cambridge Univ. Press. 204 pp.
- Leet R. 1988. Saturated and subcooled hydrothermal boiling in groundwater flow channels as a source of harmonic tremor. *J. Geophys. Res.* 93:4835–49
- Linde A, Agustsson K, Sacks I, Stefansson R. 1993. Mechanism of the 1991 eruption of Hekla from continuous borehole strain monitoring. *Nature* 365:737–40
- Lister JR. 1990a. Buoyancy-driven fluid fracture: the effects of material toughness and of low-viscosity precursors. *J. Fluid Mech.* 210:263–80
- Lister JR. 1990b. Buoyancy-driven fluid fracture: similarity solutions for the horizontal and vertical propagation of fluid-filled cracks. *J. Fluid Mech.* 217:213–39
- Lister JR. 1994. The solidification of buoyancy-driven flow in a flexible-walled channel. Part 1. Constant-volume. *J. Fluid Mech.* 272:21–44
- Lister JR, Campbell IH, Kerr RC. 1991. The eruption of komatiites and picrites in preference to primitive basalts. *Earth Planet. Sci. Lett.* 105:343–52
- Lister JR, Kerr RC. 1991. Fluid-mechanical models of crack propagation and their application to magma transport in dykes. *J. Geophys. Res.* 96:10,049–77
- MacLeod CJ, Rothery DA. 1992. Ridge axial segmentation in the Oman ophiolite: evidence from along-strike variations in the sheeted dyke complex. In *Ophiolites and their Modern Oceanic Analogues*, ed. LM Parson, BJ Murton, P Browning, *Geol. Soc. Spec. Publ.* 60, pp. 39–63. London: Geol. Soc. 330 pp.
- Majer E, Doe T. 1986. Studying hydrofractures by high frequency seismic monitoring. *Int. J. Rock Mech. Min. Sci.* 23:185–99
- McClelland-Brown EA. 1981. Paleomagnetic estimates of temperatures reached in contact metamorphism. *Geology* 9:112–16
- McConnell RK. 1967. Dike propagation. In *Economic Geology in Massachusetts*, ed. OC Farquhar, pp. 97–104. Amherst: Univ. Mass. Grad. Sch.
- Muller O, Pollard DD. 1977. The stress state near Spanish Peaks, Colorado determined from a dike pattern. *Pure Appl. Geophys.* 115:69–86
- Murase T, McBirney AR. 1973. Properties of some common igneous rocks and their melts at high temperatures. *Geol. Soc. Am. Bull.* 84:3563–92
- Nakamura K, Jacob KH, Davies JN. 1977. Volcanoes as possible indicators of tectonic stress orientation—Aleutians and Alaska. *Pure Appl. Geophys.* 115:87–112
- Nicholson R, Pollard DD. 1985. Dilation and linkage of echelon cracks. *J. Struct. Geol.* 7:583–90
- Nicolas A. 1986. A melt extraction model based on structural studies in mantle peridotites. *J. Petrol.* 27:999–1022
- Nicolas A, Jackson M. 1982. High temperature dikes in peridotites: origin by hydraulic fracturing. *J. Petrol.* 23:568–82
- Okada Y, Yamamoto E. 1991. Dyke intrusion model for the 1989 seismovolcanic activity off Ito, central Japan. *J. Geophys. Res.* 96:10361–76
- Okamura AT, Dvorak JJ, Koyanagi RY, Tanigawa WR. 1988. Surface deformation during dike propagation. In *The Puu Oo Eruption of Kilauea Volcano, Hawaii: Episodes 1 Through 20*, ed. EW Wolfe, pp. 165–81. *US Geol. Surv. Prof. Pap.* 1463
- Olson J, Pollard DD. 1989. Inferring paleostresses from natural fracture patterns: a new method. *Geology* 17:345–48
- Parsons T, Sleep NH, Thompson GA. 1992. Host rock rheology controls on the emplacement of tabular intrusions: implications for

- underplating of extending crust. *Tectonics* 11:1348–56
- Parsons T, Thompson GA. 1991. The role of magma overpressure in suppressing earthquakes and topography: worldwide examples. *Science* 253:1300–2
- Petford N, Lister J, Kerr R. 1993. The ascent of felsic magmas in dykes. *Lithos* 32:161–68
- Phipps Morgan J. 1987. Melt migration within mid-ocean spreading centers. *Geophys. Res. Lett.* 14:1238–41
- Platten IM, Watterson J. 1987. Magma flow and crystallization in dyke fissures. See Halls & WH Fahrig 1987, pp. 65–73
- Pollard DD. 1987. Elementary fracture mechanics applied to the structural interpretation of dykes. See Halls & Fahrig 1987, pp. 5–24
- Pollard DD, Delaney PT, Duffield WA, Endo ET, Okamura AT. 1983. Surface deformation in volcanic rift zones. *Tectonophysics* 94:541–84
- Pollard DD, Muller OH, Dockstader DR. 1975. The form and growth of fingered sheet intrusions. *Geol. Soc. Am. Bull.* 86:351–63
- Pollard DD, Segall P. 1987. Theoretical displacements and stresses near fractures in rock: with applications to faults, joints, veins, dikes, and solution surfaces. See Atkinson 1987, pp. 277–349
- Pollard DD, Segall P, Delaney PT. 1982. Formation and interpretation of dilatant echelon cracks. *Geol. Soc. Am. Bull.* 93:1291–303
- Rice JR. 1968. Mathematical analysis in the mechanics of fracture. In *Fracture: An Advanced Treatise*, ed. H Liebowitz, pp. 191–311. San Diego: Academic
- Rubin AM. 1990. A comparison of rift-zone tectonics in Iceland and Hawaii. *Bull. Volcanol.* 52:302–19
- Rubin AM. 1992. Dike-induced faulting and graben subsidence in volcanic rift zones. *J. Geophys. Res.* 97:1839–58
- Rubin AM. 1993a. Dikes vs. diapirs in viscoelastic rock. *Earth Planet. Sci. Lett.* 119:641–59
- Rubin AM. 1993b. On the thermal viability of dikes leaving magma chambers. *Geophys. Res. Lett.* 20:257–60
- Rubin AM. 1993c. Tensile fracture of rock at high confining pressure: implications for dike propagation. *J. Geophys. Res.* 98:15,919–35
- Rubin AM. 1995. Getting granite dikes out of the source region. *J. Geophys. Res.* In press
- Rubin AM, Pollard DD. 1987. Origins of blade-like dikes in volcanic rift zones. See Decker et al 1987, pp. 1449–70
- Ryan M. 1987. Neutral buoyancy and the mechanical evolution of magmatic systems. In *Magmatic Processes: Physicochemical Principles*, ed. BO Mysen, pp. 259–87. Univ. Park, PA: Geochem. Soc. Spec. Publ. 1. 500 pp.
- Ryan M. 1993. Neutral buoyancy and the structure of mid-ocean ridge magma reservoirs. *J. Geophys. Res.* 98:22,321–38
- Sammis CG, Julian BR. 1987. Fracture instabilities accompanying dike intrusion. *J. Geophys. Res.* 92:3597–605
- Secor D, Pollard D. 1975. On the stability of open hydraulic fractures in the Earth's crust. *Geophys. Res. Lett.* 2:510–13
- Shaw HR. 1980. Fracture mechanisms of magma transport from the mantle to the surface. In *Physics of Magmatic Processes*, ed. RB Hargraves, pp. 201–64. Princeton, NJ: Princeton Univ. Press. 585 pp.
- Shlyapobersky J, Chudnovsky A. 1992. Fracture mechanics in hydraulic fracturing. *Proc. 33rd US Symp. Rock Mech.*, pp. 827–36. Rotterdam: Balkema
- Sigurdsson H, Sparks S. 1978. Lateral magma flow within rifted Icelandic crust. *Nature* 274:126–30
- Sleep NH. 1984. Tapping of magmas from ubiquitous mantle heterogeneities: an alternative to mantle plumes? *J. Geophys. Res.* 89:10,029–41
- Sleep NH. 1988. Tapping of melt by veins and dikes. *J. Geophys. Res.* 93:10,255–72
- Smith RP. 1987. Dyke emplacement at Spanish Peaks, Colorado. See Halls & Fahrig 1987, pp. 47–54
- Speight J, Skelhorn R, Sloan T, Knaap R. 1982. The dyke swarms of Scotland. In *Igneous Rocks of the British Isles*, ed. DS Sutherland, pp. 449–621. New York: Wiley
- Spence DA, Turcotte DL. 1985. Magma-driven propagation of cracks. *J. Geophys. Res.* 90:575–80
- Spence DA, Turcotte DL. 1990. Buoyancy-driven magma fracture: a mechanism for ascent through the lithosphere and the emplacement of diamonds. *J. Geophys. Res.* 95:5134–39
- Spera F. 1984. Carbon dioxide in petrogenesis III: role of volatiles in the ascent of alkaline magma with special reference to xenolith-bearing mafic lavas. *Contrib. Mineral. Petrol.* 88:217–32
- Stevenson DJ. 1982. Migration of fluid-filled cracks: applications to terrestrial and icy bodies. *Lunar Planet. Sci. Conf. 13th.* pp. 768–69 (Abstr.)
- Stevenson DJ. 1989. Spontaneous small-scale melt segregation in partial melts undergoing deformation. *Geophys. Res. Lett.* 16:1067–70
- Swanson PL. 1984. Tensile fracture resistance mechanisms in brittle polycrystals: an ultrasonics and in situ microscopy investigation. *J. Geophys. Res.* 92:8015–36
- Takada A. 1990. Experimental study on propagation of liquid-filled crack in gelatin: shape and velocity in hydrostatic stress condition. *J. Geophys. Res.* 95:8471–81

- Talebi S, Cornet FH. 1987. Analysis of the microseismicity induced by a fluid injection in a granitic rock mass. *Geophys. Res. Lett.* 14:227-30
- Turcotte DL, Schubert G. 1982. *Geodynamics*. New York: Wiley. 450 pp.
- Walker GPL. 1987. The dike complex of Koolau Volcano, Oahu: internal structure of a Hawaiian rift zone. See Decker et al 1987, pp. 961-93
- Weertman J. 1971. Theory of water-filled crevasses in glaciers applied to vertical magma transport beneath oceanic ridges. *J. Geophys. Res.* 76:1171-83