The propagation paths of fluid-driven fractures in layered and faulted rocks

Agust Gudmundsson

Department of Earth Sciences, Queen's Building, Royal Holloway University of London, Egham TW20 0EX, UK (rock.fractures@googlemail.com)

Abstract

Fractures that form when fluid pressure ruptures the rock are referred to as fluid-driven fractures or hydrofractures. These include most dikes, inclined sheets, and sills, but also many mineral veins and joints, as well as human-made hydraulic fractures. While considerable field and theoretical work has focused on the geometry and arrest of hydrofractures, how they select their propagation paths, particularly in layered and faulted rocks, has received less attention. Here I propose that of all the possible paths that a given hydrofracture may follow, it selects the path of least (minimum) action as determined by Hamilton's principle. This means that the selected path is the one along which the energy transformed (released) multiplied by the time taken for the propagation is a minimum. Hydrofractures advance their tips/fronts in steps, with a time lag between the fracture front and the fluid front. In the present framework, each step is then controlled by Hamilton's principle. The results suggest that when the hosting rock body is regarded as homogeneous, isotropic and non-fractured, hydrofracture paths are everywhere perpendicular to the trajectories of the minimum compressive (maximum tensile) principal stress σ_3 and follow the trajectories of the maximum principal compressive stress σ_1. When applied to layered and faulted rock body, the results indicate that hydrofracture paths may follow existing faults for a while, depending primarily on (1) the dip of the fault (steep faults are the most likely to be used by vertically propagating hydrofractures), and (2) the tensile strength across the fault as compared with the tensile strength of the host rock along a path following the direction of σ_1. The results suggest that hydrofractures may use faults as parts of their paths primarily if the fault is steeply dipping and with close to zero tensile strength.

Keywords: Dikes, inclined sheets, mineral veins, hydraulic fractures, faults, Hamilton's principle of least action.

1. Introduction

What are the mechanical conditions that allow rock fractures to initiate and, subsequently, propagate? And, further, what factors control eventual rock-fracture arrest? More specifically, what physical conditions determine how fractures select their paths? To answer these and related questions about rock fractures, we first need to be clear about the main mechanical types of rock fractures. They are only two, namely shear fractures and extension fractures. Shear fractures form by shear stresses and include all faults and some rock fractures classified as joints. Extension fractures include all tension fractures and fluid-driven fractures. The latter comprise almost all dikes, inclined sheets, sills, and human-made hydraulic fractures, all as well as many mineral veins and joints; in short, all fractures where the fluid pressure is high enough to rupture the rock and form and to drive or propagate the resulting fracture. Here the focus is on fluid-driven rock fractures, which are also referred to as hydrofractures. These are thus extension fractures that can be modelled as mode I (opening-mode) cracks.

The above are mostly standard definitions from structural geology/tectonics and rock physics, and have been widely used for decades (Griggs and Handin, 1960; Dennis, 1972; Jaeger and Cook, 1979; Segall, 1984; Pollard and Aydin, 1988; Twiss and Moores, 1992). These and other definitions on rock fractures and crustal fluids used in the paper are mostly the same as used in modern textbooks (Pollard and Fletcher, 2005; Schultz, 2019). The main divergence in the definition of terms used in the paper concerns the concept of overpressure or driving pressure of hydrofractures. In the hydrocarbon industry, overpressure commonly denotes fluid pressure in the crust in excess of hydrostatic pressure and that definition of the term is sometimes used in structural geology (Bons et al., 2012). This definition is well established in the hydrocarbon industry for pore-fluid pressure in rock layers and hydrocarbon reservoirs. But it is not a very helpful definition when dealing with hydrofactures because then overpressure cannot be related to the conditions for rock rupture and hydrofracture propagation, nor to their eventual dimensions, using results from fracture mechanics. For analysing rock rupture and hydrofracture propagation, overpressure must refer to the normal stress on the fracture which, for hydrofractures as defined here, is normally the minimum principal compressive stress, σ_3 – and that is the definition of overpressure used here. In the paper, all definitions from rock-fracture mechanics and fluid mechanics follow those given by Gudmundsson (2011, 2020) all of which are widely used in earth sciences. Additionally, fault slips in a zone of high fluid pressure, as are very common, are not regarded as being related to hydrofracture/extension fracture propagation but rather as being the result of shear failure and thus modelled as mode II or mode III (or, possibly, mixed-mode) cracks.

In the paper, I explore and explain the physics of fluid-driven fracture (hydrofracture) paths. The focus is on magma-driven fractures, particularly on dikes,

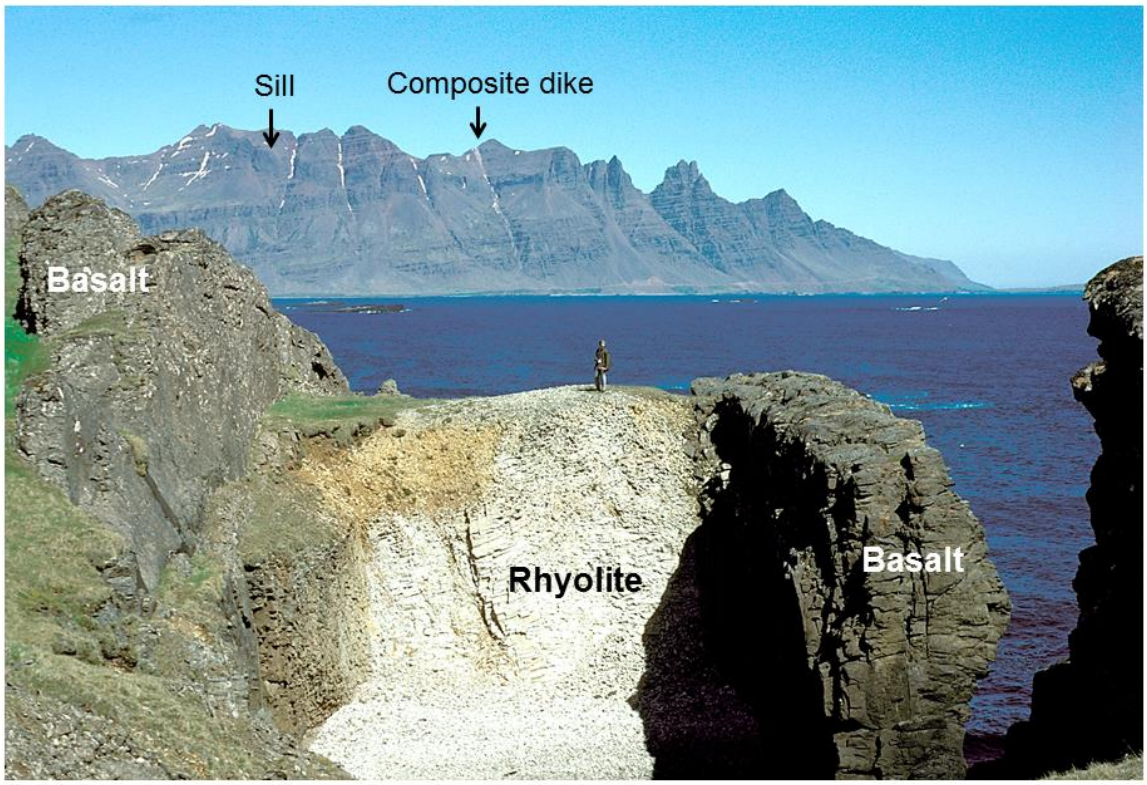

Fig. 1. Composite dike in East Iceland. View north, the thickness of the rhyolite part is 13 m, that of the western basalt part 7.5 m, and that of the eastern basalt part 5 m, bringing the total thickness to 25.5 m. The dyke can be traced laterally for about 14 km, its strike changing from N20°E here to N14°E towards its northern end. The dike extends to the top of the 700 m high mountain on the other side of the fjord (seen here), but the basaltic parts drop out in the middle part of the mountain so that at the top the dike is purely acid and 35 m thick. A 120-m-thick multiple sill dissects the dike, and is thus younger than the dike.

for the simple reason that numerous well-exposed dikes have been studied in the field where they are seen as having propagated through heterogeneous and layered (anisotropic) crustal segments (Geshi et al., 2010, 2012; Galindo and Gudmundsson, 2012; Geshi and Neri, 2014; Drymoni et al., 2020, 2021; Gudmundsson, 2020). Thus, many dike paths have been studied in detail vertically over hundreds of metres and laterally for kilometres and traced for tens, and occasionally, hundreds of kilometres. In addition, the propagation of many dikes has been inferred from the migration of dike-induced earthquake swarms and surface deformation in volcanoes during unrest periods (Peltier et al., 2005; Grandin et al., 2011; Gudmundsson et al., 2014; Agustsdottir et al., 2016).

Much research has also been done on human-made hydraulic fractures for over 70 years in the hydrocarbon industry. These are also fluid-driven fractures and their studies, including seismic monitoring (Shapiro, 2018), cores taken through the seismic volumes, focusing on the fracture types generated (Gale et al., 2019, 2021),

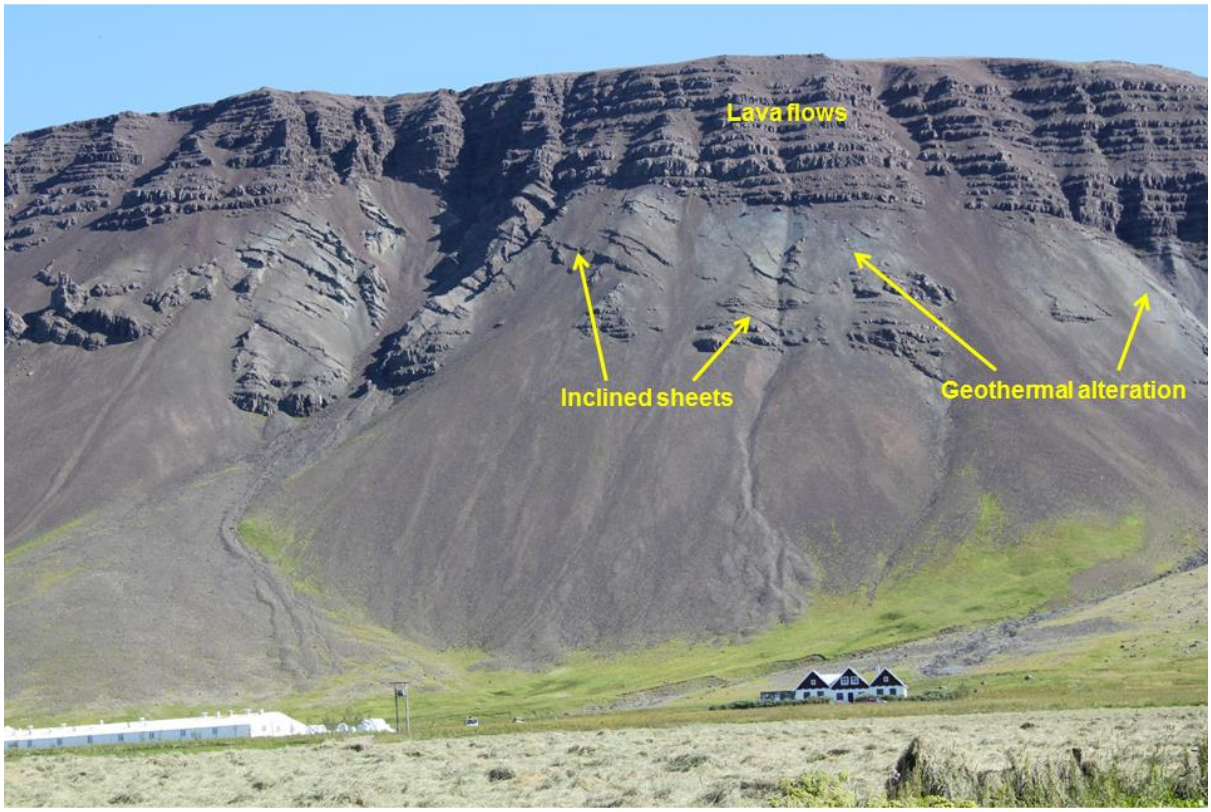

Fig. 2. Inclined sheets and (above) lava flows in a fossil central volcano in West Iceland. Also seen is alteration due to circulation of geothermal water while the central volcano was active some 2.8 Ma ago. The thicknesses of most of the inclined sheets are 0.5-1 m.

and their propagation paths and geometries (Howard and Fast, 1970; Warpinski and Teufel, 1987; Warpinski et al., 1993; Fast et al., 1994; Mahrer, 1999; Fisher and Warpinski, 2011; Davis et al., 2012; Flewelling et al., 2013; Fall et al., 2015), that provide additional information on such fractures. Mineral-filled extension (opening-mode) fractures, known as mineral veins (Philipp, 2008, 2012; Bons et al., 2012), also throw light on the propagation of fluid-driven fractures, on a much smaller scale than for dikes (Hillis, 2003; Cobbold and Rodrigues, 2007). The role of chemistry (including crystal growth and dissolution rates) in the development of mineral veins has received much attention in recent years (Laubach et al., 2019). Here, however, the focus is on the mechanical aspects of mineral-filled fracture initiation and propagation. While some mineral-filled fractures are shear fractures, field studies suggest that close to 80% of mineral veins, even inside fault zones, are extension fractures (Gudmundsson et al., 2001, 2002; Philipp, 2008, 2012). Most mineral veins may thus be regarded as extension (opening-mode) fractures whose attitude (strike and dip) was perpendicular to the local orientation of σ_3 at the time of vein formation. But very many veins are confined to single layers (stratabound/layerbound), that is, arrested in vertical sections, and the mechanics of

Fig. 3 Large mineral veins (of calcite) in limestone in Southwest England. The measuring tape is 8 m long.

arrest and its implication for the vein-propagation-path selection are not fully understood (Gudmundsson, 2011).

In conventional hydraulic fracturing in the oil industry, the likely generalised propagation path of a fracture injected laterally into a single mechanical unit/layer has often been forecasted based on current stress data and other information (Valko and Economides, 1995; Yew and Weng, 2014; Shapiro, 2018). In fact, small-scale hydraulic fracturing is used to determine the orientation of σ_3 in hydraulic fracturing stress measurements (Amadei and Stephansson, 1997; Zoback, 2007; Zang and Stephansson, 2014). In unconventional, primarily vertical, propagation of hydraulic fractures, such as used in the extraction of gas from shales, the forecasted paths are much less reliable, with the fractures commonly becoming deflected laterally along contacts to form water sills or deflected into faults or otherwise following non-predicted paths (Warpinski and Teufel, 1987: Fisher and Warpinski, 2011; Fisher, 2014; Davis et al., 2012). As for natural hydrofractures, the generalised attitude (strike and dip) of dikes has for a long time been known to be crudely perpendicular to the regional direction of σ_3 at the time of dike emplacement (Anderson, 1942). Yet, despite the high-density instrumentation (seismic and geodetic networks) in many volcanoes and volcanic zones today, no detailed dike path in an active volcano has

Fig. 4 Stratabound (layerbound) calcite veins in limestone in Southwest England. Mineral veins in sedimentary basins such as here in the Bristol Channel are commonly arrested at contacts between mechanically dissimilar layers. Here the veins are arrested at contacts between comparatively stiff limestone layers and comparatively compliant (at the time of vein formation) shale layers. (Cf. Philipp, 2012).

ever been successfully forecasted (Gudmundsson, 2020). Because a feeder-dike must reach the surface to erupt, and since volcanoes and volcanic zones are layered and faulted, the propagation path up through the layers on the way to the surface is of the greatest interest. Vertical and inclined propagation paths of hydrofractures through layered and faulted rocks are thus the main focus of this paper.

More specifically, the principal aim of the paper is to answer the question: How does a fluid-driven rock fracture (a hydrofracture) select its path when propagating up through a layered and faulted crustal segment? As is explained in the paper, each hydrofracture has theoretically and infinite number of potential propagation paths to select from. So the question is: which path does it select and why? Here, for the first time, Hamilton's principle of least action is applied to hydrofractures with a view of forecasting their likely paths through layered rocks. Additionally, using energy considerations, I provide a quantitative analysis of the potential of existing faults to act as parts of hydrofracture paths. For these purposes, the paper begins with a review of relevant field observations of hydrofracture paths. This is followed by a brief discussion of the conditions for hydrofracture initiation. The physics of

Fig. 5. Dense network of mineral veins (of various minerals) in the Husavik-Flatey Fault, a transform fault zone in North Iceland. About 80% of the veins are extension fractures. The veins are here exposed at a depth of about 1500 m below the original surface of the lava pile which the fault dissects (Gudmundsson et al., 2001, 2002).

hydrofracture-path selection in layered rocks, the main theme of the paper, is then discussed in considerable detail, including the potential effects of faults to provide parts of the paths. The paper also includes a discussion of the implications for the theory presented here for the understanding of the physics of hydrofracture propagation with application to dikes reaching the surface to supply magma to volcanic eruptions.

2. Field observations

Observations of natural fluid-driven fractures (hydrofractures) in the field are primarily of sheet intrusions and mineral veins. In addition, there are detailed data on human-made hydraulic fractures used in the hydrocarbon and geothermal industries. This section provides a short overview of these aspects of fluid-driven fractures, with a focus on the natural fractures. Much more detailed field descriptions of natural hydrofractures are provided by Segall (1984), Pollard and Aydin (1988), Hillis (2003), Cobbold and Rodrigues (2007), Philipp (2008, 2012),

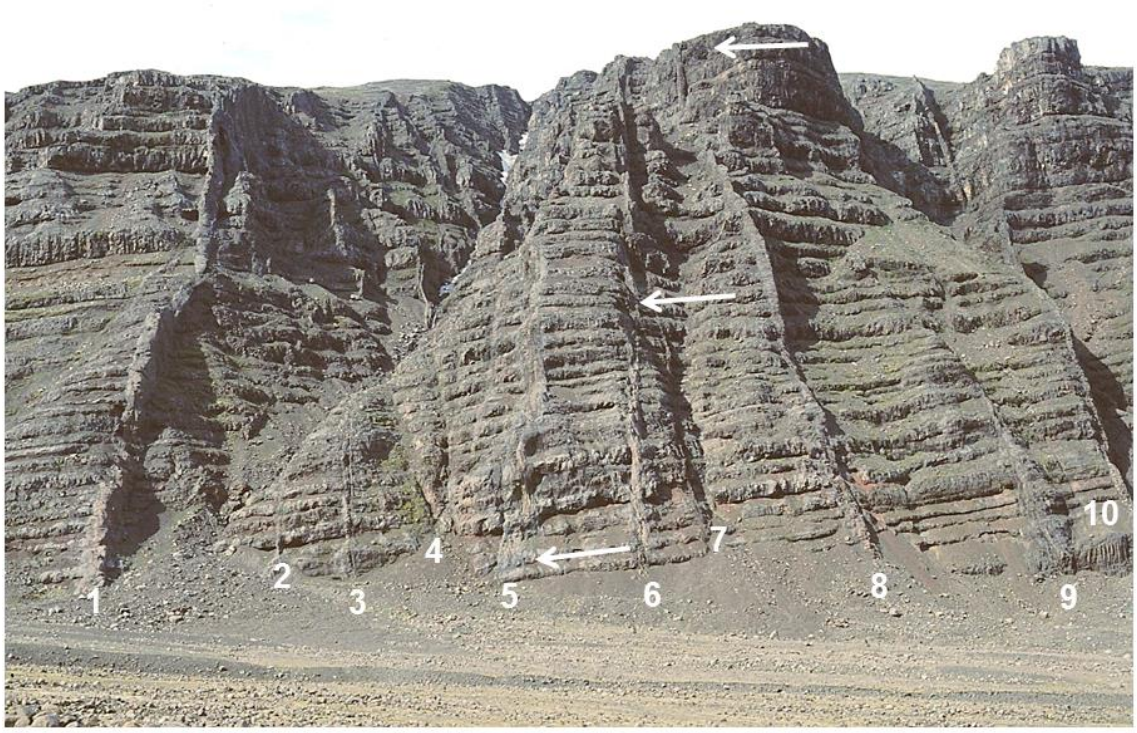

Fig. 6. Regional basaltic dikes dissecting a pile of mostly basaltic lava flows in Southeast Iceland. The dikes dissect the lava flows at right angles. The absence of vertical displacements parallel to the dikes suggests that the dikes are extension fractures. Ten dikes are indicated, and the subhorizontal arrows indicate how the dip of the lava flows increases with depth in the crust. The floor of the valley where the photo is taken is at about 2000 m depth below the initial top of the lava pile.

(Gudmundsson (2011, 2020), Geshi et al. (2010, 2012), Bons et al. (2012), Galindo and Gudmundsson (2012), Kusumoto et al. (2013), Kusumoto and Gudmundsson (2014), Fall et al. (2015), Tibaldi (2015), Gale et al. (2019), Laubach et al. (2019).

The largest exposures of natural hydrofractures are those of dikes. Some dikes have continuous vertical exposures of hundreds of metres (Geshi et al., 2010, 2012) and can be traced laterally for many kilometres (Anderson, 1942; Gudmundsson, 1983, 2020; Pollard and Fletcher, 2005; Geshi and Neri, 2024) – although the lateral exposures may not be continuous – and the same applies to sills (Fig. 1). The exposures of inclined (cone) sheets, by contrast, both in vertical and in horizontal sections, are normally much more limited (Fig. 2), for the simple reason that these are much smaller structures than regional dikes (Galindo and Gudmundson, 2012; Kusumoto et al., 2013; Tibaldi, 2015; Gudmundsson, 2020).

The dimensions of mineral veins are mostly of the order of tens of centimetres to a few metres (Philipp 2008, 2012; Gudmundsson, 2011; Bons, 2012), but some reach many metres or more in length or strike-dimension (Fig. 3). Their height or dip-

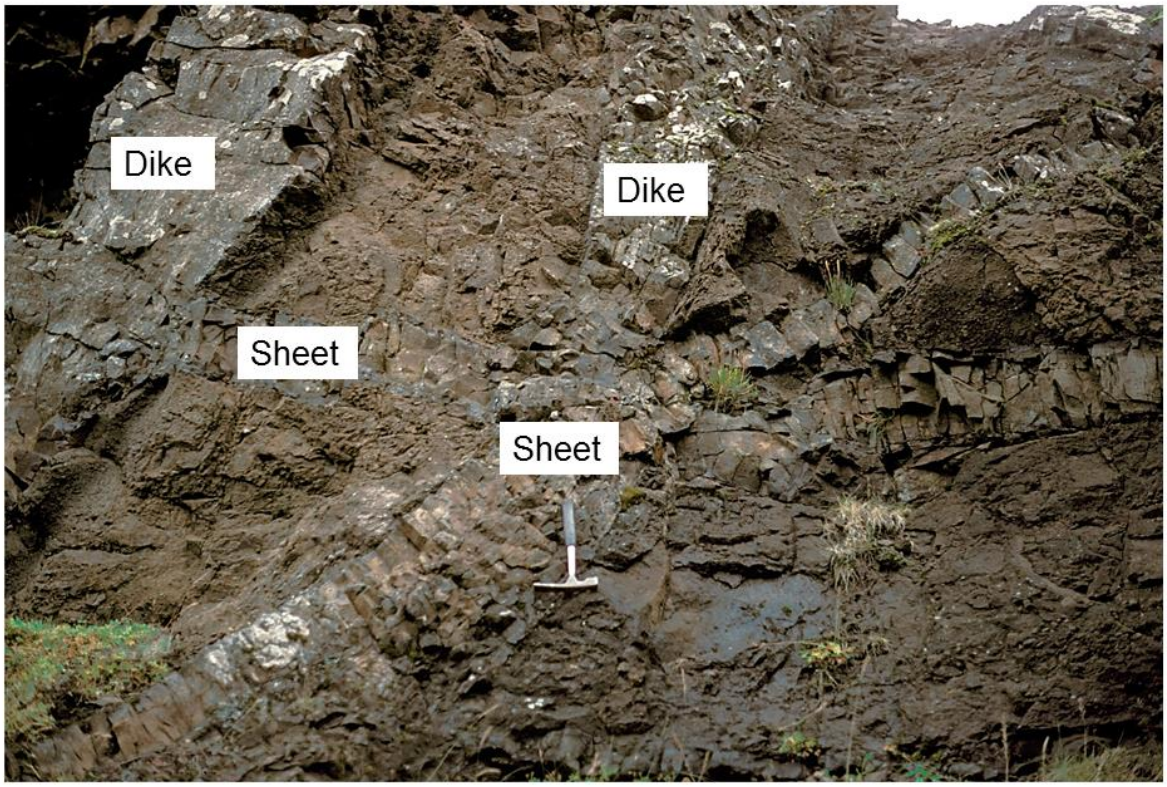

Fig. 7. Cross-cutting basaltic dikes and inclined sheets in South Iceland. The relationships indicate that there is no displacement parallel with any of the sheet intrusions, which, thereby, should be interpreted as extension fractures (and modelled as mode I cracks). The host rock is lake sediment. The length of the hammer is about 30 cm.

dimension can also reach several metres, but more commonly, in layered rocks such as sedimentary piles, the dip dimensions are tens of centimetres to a few metres, with the veins commonly being stratabound (Fig. 4). In contrasts to dikes (and sills and inclined sheets), mineral veins commonly form dense networks where many of the hydrofractures were active, that is, transporting fluids, at the same time (Fig. 5).

For understanding hydrofracture propagation paths, and for modelling purposes, it is of fundamental importance to determine to which mechanical type of fracture most hydrofractures belong. Cross-cutting relationships (Fig. 5) as well as orientation of fibres in the veins (Philipp, 2008, 2012) indicate that most mineral veins are extension fractures, even inside fault zones (Fig. 5). For dikes and inclined sheets, the type of fracture is also easiest to determine in the field from cross-cutting relationships; either between dikes and sheets, or between these intrusions and the layers that they dissect (mostly lava flows or pyroclastic layers).

On a regional scale, the cross-cutting relationships between dikes and the lava flows (and pyroclastic layers) that they dissect provides ample evidence that most dikes are

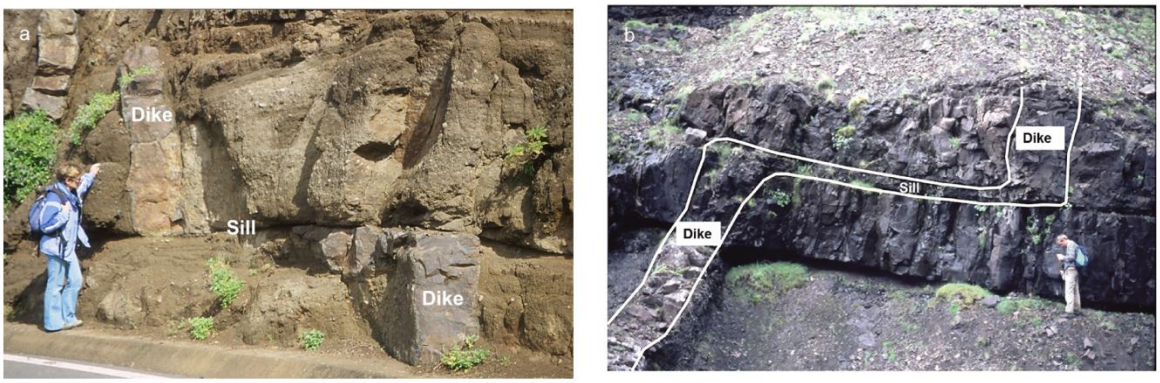

Fig. 8. Zig-zag geometry of dikes. (a) Basaltic dike becomes deflected into a thin sill along a contact between pyroclastic layers in Tenerife (Canary Islands). The lower dike segment is 0.43 m thick, strikes N12° E and dips 87° W, whereas the upper segment is 0.46 m thick, strikes N1°E and dips 89°W. Thus, the dike strike changes by about 11° on crossing the contact, whereas the dike thickness remains similar. At the contact itself, however, the dike changes into a sill with a thickness of as little as 0.02 m (2 cm) for a lateral distance of about 3.4 m. (b) Dike changing into a sill along part of its path in West Iceland. The change occurs at the contact between mechanically dissimilar rocks, the contact itself being composed of comparatively soft scoria (along which the sill is deflected), whereas the layer above (the one on top of the present sill) is stiff basaltic lava flow. The horizontal length of the sill is about 8 m. The vertical dike segments are about 0.8 m thick.

extension (opening-mode) fractures. Such cross-cutting relationships in vertical cross-sections demonstrate that the dikes do not occupy dip-slip faults (Fig. 6). When the dikes are followed along their length (the strike-dimension), similar cross-cutting relationships show that the dikes do not occupy strike-slip faults. On a local scale, cross-cutting relationships and the absence of slickensides further demonstrate that most dikes and inclined sheets are extension fractures. In some cases the cross-cutting relationship includes several dikes and inclined sheets, in which cases the evidence for their being extension fractures is very clear (Fig. 7).

Many regional dike paths in vertical sections are comparatively straight (Fig. 6). This applies particularly to dikes propagating up through a basaltic lava pile where all the mechanical properties of the flows are similar and there are no thick layers of compliant (soft) scoria, soil, or pyroclastics in-between the lava flows. When such layers exist, or when the lava flows are of widely different mechanical properties and the pile contains very stiff sills as well as lava flows, then the dike paths tend to become much more irregular. In particular, in a pile of layers with contrasting mechanical properties – layers with widely different Young's moduli – the propagation paths of dikes and inclined sheets (as seen in vertical sections) commonly show abrupt changes in attitude, resulting in a zig-zag geometry (Fig. 8). Where the attitude of the path changes abruptly - such as where a vertical dike for a while deflects into a sill and then back into a dike – the thickness normally also

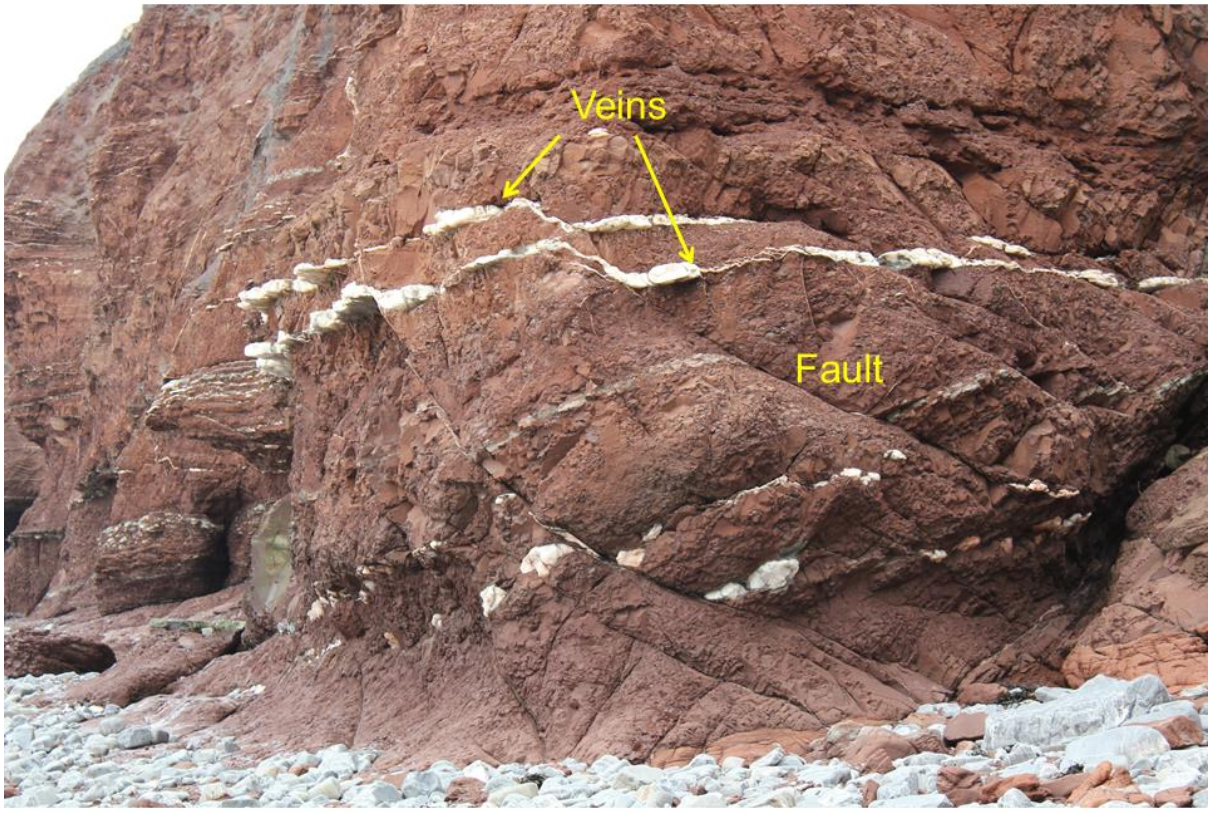

Fig. 9. Gypsum veins deflect into a curving (listric) normal fault for parts of their paths in the Bristol Channel in Southwest England. The veins used the fault as paths presumably because the tensile strength of the fault at the time of vein formation was zero (assuming that the fault had recently slipped or formed). The veins are notably thinner in the parts of the paths that are along the fault because they are there no longer perpendicular to the minimum principal compressive stress σ_3 but rather to a normal stress σ_n on the fault plane (and also because their paths within the faults are very short). The host rock is red mudstone (cf. Philipp, 2008).

changes. In particular, the sill-part of the path may be much thinner than the nearby dike parts (Fig. 8).

Abrupt changes in the propagation-path attitude are also common in mineral veins. This happens, in particular, when the veins meet existing discontinuities with little or no tensile strength, such as active or recently active faults (Fig. 9). As is the case for dikes, the veins are commonly much thinner in the part of the propagation path that uses the fault.

Dikes are generally segmented (Fig. 10a). The segmentation is partly because the dikes propagate as individual 'fingers' and at different rates. For feeder-dikes the first finger that reaches the surface initiates the eruption (Fig. 10b). Segmentation is in fact a universal feature of rock fractures; thus, mineral veins are commonly seen as segmented, even when their exposed lengths are of the order of metres or less (Fig. 11).

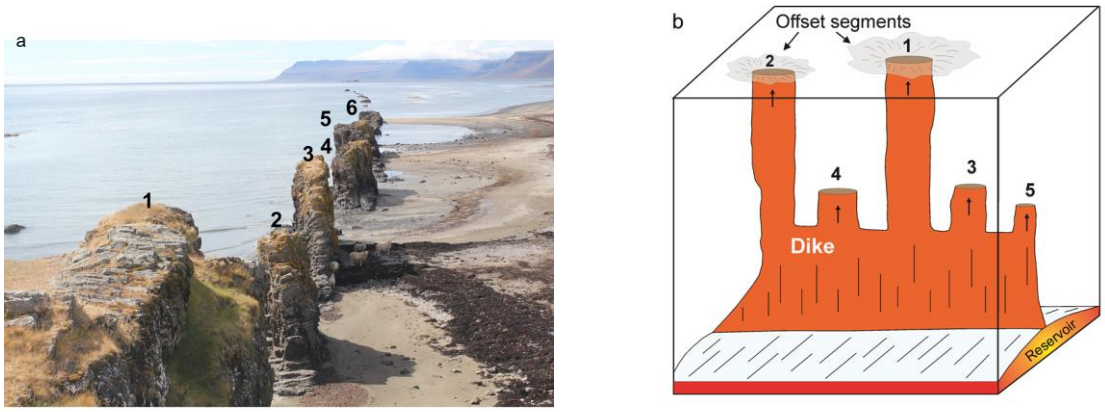

Fig. 10. Segmented dikes. (a) Six segments of a 5-m-thick basaltic regional dike in Northwest Iceland. This dike is more resistant to erosion than the host rock, and therefore stands as segmented ridge above the surroundings. (b) Schematic illustration of a propagating segmented dike. The first dike 'finger' to reach the surface initiates the resulting fissure eruption. The first segments at the surface are commonly short, tens of metres, but they commonly propagate laterally at the surface and may eventually link up into larger fissure segments (Cf. Gudmundsson, 2020).

Most hydrofracture segments become vertically arrested and commonly stratabound (Fig. 12). The fracture tip is most commonly arrested at a contact between mechanically dissimilar layers, or just below the contact between such layers. Mechanically dissimilar layers include, for example, limestone (stiff) and shale (often compliant during vein formation) layers in sedimentary basins. Many mineral veins are seen arrested at contacts between limestone and shale layers (Fig. 12a, b). Most dike segments that are seen to end vertically do not reach the surface but rather become arrested. While the arrested tips are commonly exactly at the contact between mechanically dissimilar layers there are also cases where the dikes become arrested just below the contact of layers with contrasting mechanical properties – when the layer above the contact is much stiffer (has a much higher Young's modulus) than the layer below the contact and hosting the top-part of the dike (Fig. 12c, d; cf. Forbes Inskip et al., 2020).

Fig. 11. Segmented mineral veins in vertical and lateral sections in the Bristol Channel in Southwest England. (a) Segmented calcite veins in a vertical section though (mostly) shale layers and (at the top) limestone layers. The length of the measuring tape is 1 m. (b) Segmented calcite vein, with an *en echelon* arrangement, in a lateral section in limestone.

Studies of human-made hydraulic fractures yield similar results as to propagation and arrest as those for the natural hydrofractures. Propagating hydraulic fractures, used in the hydrocarbon and geothermal industries to increase the permeability of reservoirs, generate earthquake swarms (Shapiro, 2018). The swarms are mostly of microseismicity and their propagation paths can be monitored (Fisher and Warpinski, 2011; Davis et al., 2012; Flewelling et al., 2013; Fisher, 2014). It should be noted that part of the microseismicity associated with hydrofracture propagation in general (including dikes) is attributable to fault slip in the process zone ahead of the propagating fracture tip, and part to fault slip in the walls of the fracture (Gudmundsson, 2011, 2020; Geshi and Neri, 2014; Agustsdottir et al., 2016; cf. Shapiro, 2018). Close to the surface, fractures may be induced by a propagating dike

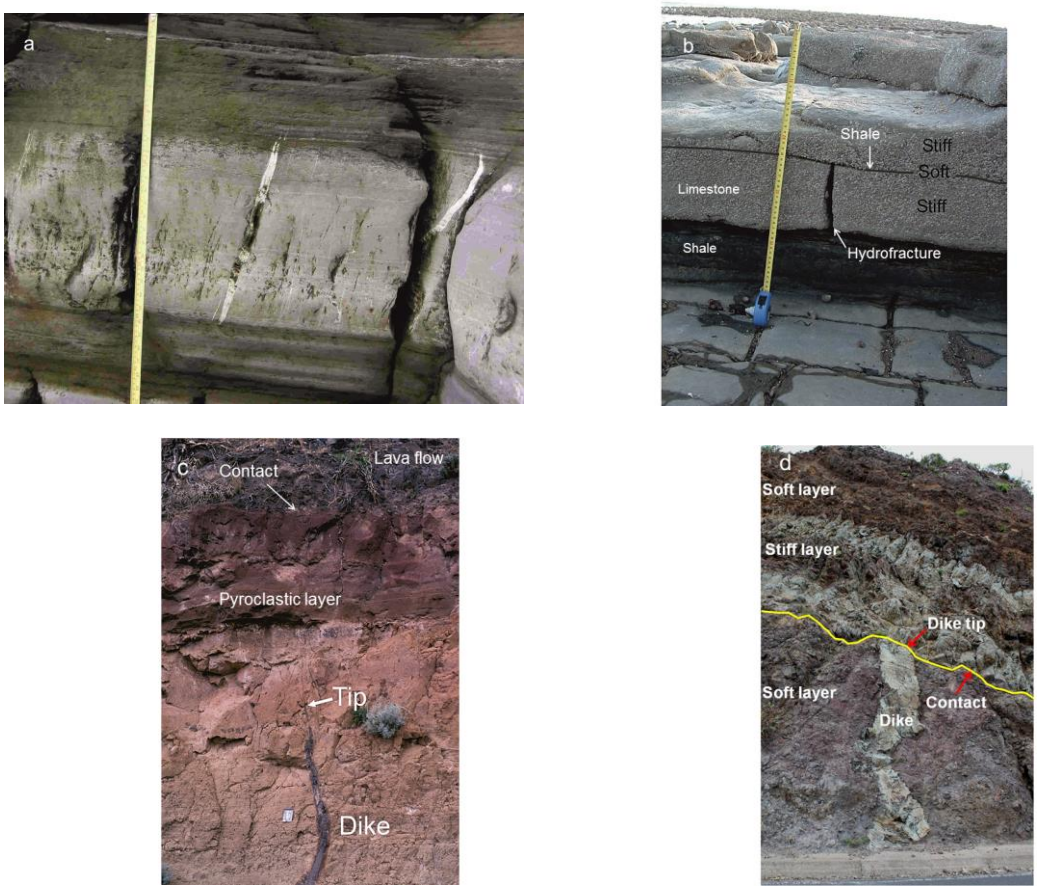

Fig. 12 Arrested hydrofractures, mineral veins and dikes. (a) Calcite vein arrested at contacts between shale and limestone layers. (b) Hydrofracture arrested, and forming a T-shaped fracture, at the contact between a very thin shale layer (indicated) and limestone layers. Both the fractures in a and b are in the Bristol Channel of Southwest England. (c) Arrested dike at some distance below the contact between a basaltic lava flow and a pyroclastic layer. The thickness of the basaltic dike gradually decreases from 25 cm at the bottom of the figure to 2 cm at the dike tip. Next to the dike the host rock is baked. (d). A vertically propagating basaltic dike becomes arrested on meeting with a gently dipping stiff intrusive sheet, marked as a stiff layer. View north-northeast, the maximum thickness of the dike is about 0.8 m. Both the basaltic dikes in c and d are in Tenerife (Canary Islands).

even if the dike itself, eventually, does not reach the surface (Al Shehri and Gudmundsson, 2018; Bazargan and Gudmundsson, 2019; Gudmundsson, 2020).

In conventional hydraulic fracturing the fracture that is injected laterally into a layer from a vertical well is supposed to be arrested at the top and bottom of that layer. The aim is thus to confine the hydraulic fracture to the 'target layer', the layer or unit containing oil or gas. The fracture should thus not propagate much into the layers above and below the target layer (cf. Howard, and Fast, 1970: Valko and Economides, 1995; Yew and Weng, 2014; Shapiro, 2018). The generalised propagation path of such lateral hydraulic fractures can often be forecasted based on information such as the mechanical properties of the target layer in relation to those of the adjacent

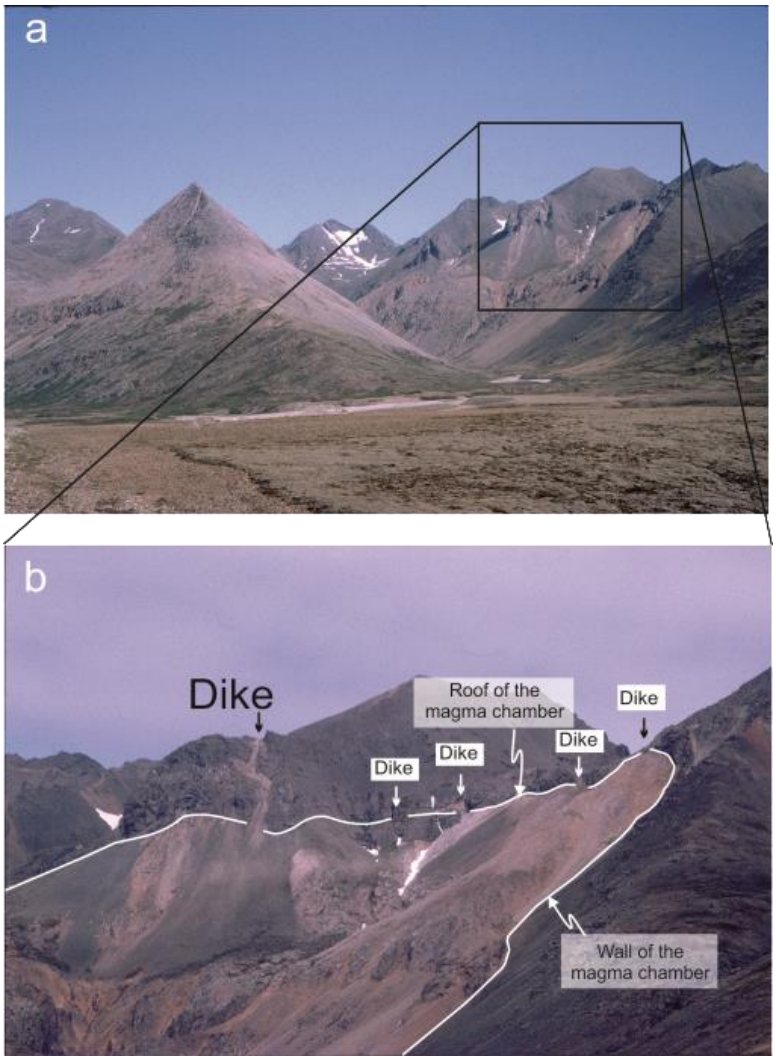

Fig. 13. (a) Exceptionally well-exposed fossil magma chambers, like this one here, commonly show part of the roof where many (here felsic) dikes can be seen cutting the roof. (b) In addition to the roof, parts of the wall of the fossil chamber (now a felsic pluton) are seen and indicated. The pluton is made of granophyre and is hosted by a pile of basaltic lava flows in Southeast Iceland. The floor of the valley (seen in a) is at 2000 m below the initial top of the lava pile (cf. Gudmundsson, 2020).

layers and that of the current stress fields. In unconventional hydraulic fracturing, developed in the past decades, the fractures are injected vertically into the rock layers from a horizontal well (Wu, 2017; Shapiro, 2018). This technique has been extensively used for getting gas out of shales (gas shales) and is now well developed. The (mostly) vertical hydraulic fractures are mechanically very similar to vertical dikes. In particular, the dip-dimension or height of some hydraulic fractures exceed 1 km (Davis et al., 2012), although most are much shorter (Fisher and Warpinski, 2011; Fisher, 2014), and may thus reach dip-dimensions similar to those of many radial dikes and inclined sheets injected from shallow magma chambers (Gudmundsson, 2020).

Studies of thousands of hydraulic fractures show that their propagation paths are commonly complex and similar to those of dikes. The microseismic studies suggest that vertical hydraulic fractures commonly deflect into discontinuities such as faults and contacts, particularly at shallow depths (Fisher and Warpinski, 2011; Flewelling et al., 2013). While deflection of vertical hydraulic fractures into horizontal water-sills can occur at any depth in the sedimentary basins, such a deflection is particularly common at crustal depths of less than 700-800 m (Fisher, 2014). Hydraulic fractures have also been observed to deflect into faults, using the faults as part of their propagation paths (Davis et al., 2012; Lacazette and Geiser, 2013), as discussed further below.

Following hydraulic fracture experiments, many such fractures have been studied in the subsurface in excavated sections. These studies show that the geometric features of hydraulic fractures are similar to those observed of dikes and, on a much smaller scale, mineral veins. Some hydraulic fractures, or fracture segments, are seen to be arrested at contacts between mechanically dissimilar layers, while others become deflected into water-sills (e.g., Fisher and Warpinski, 2011; Fisher, 2014). Clearly, the three principal mechanisms of the arrest and deflection of extension (opening-mode) fractures (mode I cracks), namely Cook-Gordon delamination, stress barriers, and elastic mismatch (Gudmundsson, 2011, 2020), all operate on hydraulic fractures in the same way as they operate on mineral veins and dikes.

3. Fracture initiation

When the following condition is satisfied then a fluid-filled source or reservoir will rupture and inject a hydrofracture:

$$p_l + p_e = \sigma_3 + T_0 \qquad (1)$$

Here, p_l is the magmatic pressure in the source or reservoir when the reservoir is in mechanical equilibrium with the host rock (neither expanding nor contracting) –

that is, before any excess pressure p_e is generated in the reservoir. For equilibrium, p_l must be equal in magnitude to the lithostatic stress or overburden pressure at the source/reservoir boundary. p_e is the excess fluid pressure in the source at the time of rupture and hydrofractur initiation, and σ_3 is the minimum compressive (maximum tensile) principal stress, and T_0 the local in situ tensile strength, at the rupture site. Compressive stress is considered positive and tensile stress negative. The excess fluid pressure is the pressure in excess of the lithostatic stress. Lithostatic refers to an isotropic (hydrostatic or spherical) state of stress – meaning that all the principal stresses are equal - where the stress magnitude increases proportionally with depth in the crust (or lithosphere). The rate of increase of the lithostatic stress magnitude with depth is determined by the density of the crustal/lithospheric rocks (Gudmundsson, 2011). When the fluid-filled source is in lithostatic or mechanical equilibrium with its host rock, then $p_e = 0$ (there is no excess pressure) and the state of stress in the roof (and elsewhere at the boundary) of the source is so that $\sigma_1 = \sigma_2 = \sigma_3 = p_l$. This means that all the principal stresses are equal (isotropic state of stress) and equal to the lithostatic stress or overburden pressure which, because of lithostatic equilibrium, must also be equal to the fluid pressure p_l inside the source. Then the pressure in the source entirely balances the stresses on it, implying that the source is neither expanding nor contracting (shrinking). For example, most active magma chambers are presumably in the state of lithostatic or mechanical equilibrium when there is neither inflation nor deflation occurring in the associated volcano (Gudmundsson, 2020). Eq. (1) is sometimes given as $p_t = \sigma_3 + T_0$ where p_t is the total fluid pressure in the source at the time of rupture (equal to $p_l + p_e$ in Eq. (1)) - an equation that can easily be derived from Griffith's theory of fracture (Jaeger and Cook, 1979).

The conditions of Eq. (1) can be reached either through increasing p_e by addition fluid to the source, by decreasing σ_3 through regional extension, such as at divergent plate boundaries, or both. These conditions are well known and already implied in Anderson's (1942) models on dike formation (cf. Section 6.2). When the source receives an additional fluid, then the source will no longer be in lithostatic/mechanical equilibrium with the host rock. This is well established in the case of shallow magma chambers that on receiving new magma enter a period of unrest. Additional fluid makes the excess pressure positive ($p_e > 0$) and increasingly so as more new fluid is received by the source. Tensile stress concentration around the source due to positive p_e also reduces σ_3 (so that the state of stress is no longer lithostatic). For a shallow magma chamber, the increase in p_e would result in inflation. Eventually, if the condition of Eq. (1) is reached, either as fluid is added to the source or the source is subject to regional extension, the source boundary (the roof for a sill-like source as assumed here) ruptures and a hydrofracture (a dike or an inclined sheet in the case of a magma chamber) is injected (cf. Gudmundsson, 2020).

As indicated, we assume here that the source/reservoir is 'sill-like', that is, penny-shaped or an oblate ellipsoid. We also assume that the lateral dimensions of the source are large in relation to the opening of the hydrofracture. The local stress conditions around a well, a drill hole of a circular horizontal cross-section, and the fact that its lateral dimension in relation to the opening of the hydraulic fracture is not as large as typically for the sources of natural hydrofactures, mean that the conditions of Eq. (1) do not apply as well to the initiation of human-made hydraulic fractures (cf. Gudmundsson, 2011). Equation (1) implies that when the excess pressure p_e in the source decreases, so that $p_e \rightarrow 0$, the opening of the fracture at the contact with the source/reservoir closes. This applies generally to water-driven fractures (hydrothermal/geothermal and human-made), and to fractures driven by

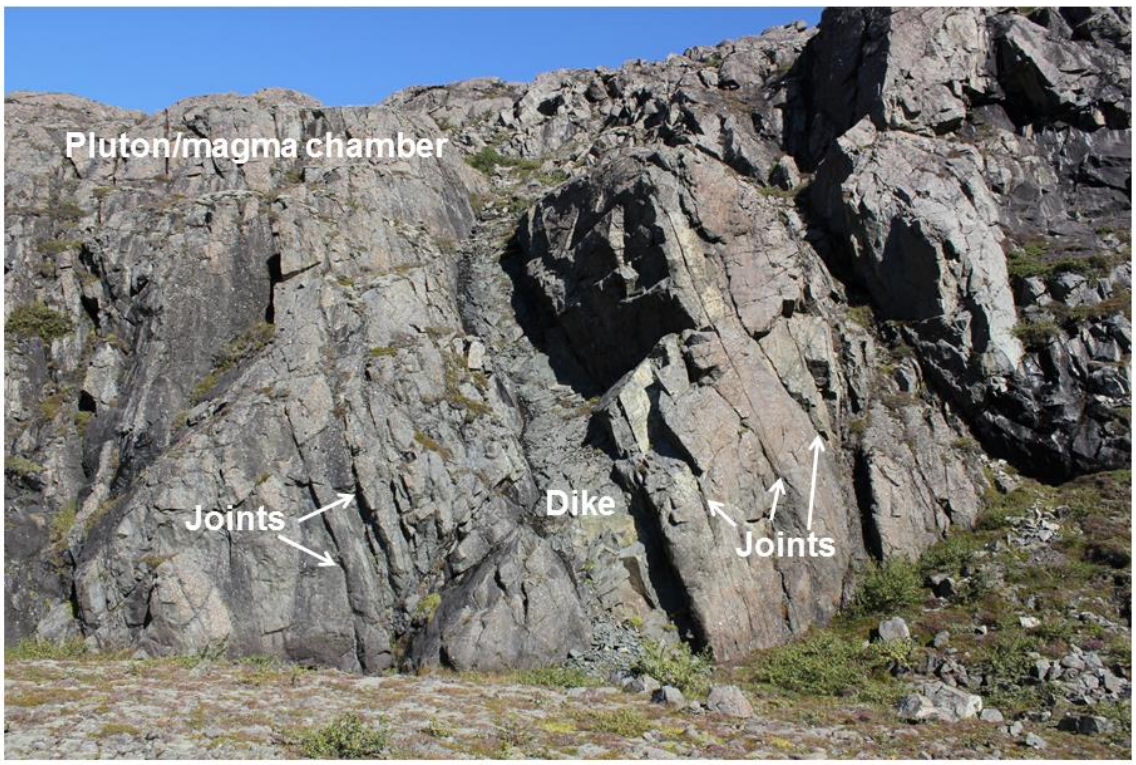

Fig. 14. Dike, about 2 m thick, using existing cooling (columnar) joints as a part of its path through a gabbro body. The exposure is a part of the outermost part of a fossil magma chamber (now a gabbro pluton) and located some 2000 m below original top of the associated central volcano. Dikes and inclined sheets commonly use favourably oriented joints as parts of their paths. The joints are extension fractures.

gas and low-viscosity magmas, but less so to fractures driven open by high-viscosity magmas. It is, in fact, well known that many acid dikes, particularly those that become arrested, that is, non-feeders, stay open at their contacts with the source chamber when their propagation paths come to an end (Fig. 13).

When using fracture-mechanics theory, the condition for hydrofracture initiation at its source is normally formulated in terms of toughness, which is a measure of the material (here, rock) resistance to fracture. The two closely related toughness measures used are material toughness and fracture toughness, which may be briefly defined as follows (cf. Gudmundsson, 2011). Material toughness is a measure of the energy absorbed in a material per unit area of fracture or crack in that material and is also referred to as the critical strain energy release rate. Its unit is given as J m^{-2}, energy (work) per unit area, or as Nm^{-1}, that is, as force per unit length of crack. When the latter unit is used, material toughness is sometimes referred to as crack extension force. The energy release rate of a material is normally denoted by G and its critical value, the material toughness, by G_c. Material toughness G_c is thus the critical energy release rate needed for a crack to propagate.

Fracture toughness, the critical stress intensity factor for a fracture to propagate, has the unit of N $m^{-3/2}$ or Pa $m^{1/2}$ (normally given in MPa $m^{1/2}$), and is denoted by the symbol K_c. Fracture toughness increases with increasing volume of material at the fracture tip that deforms plastically, through microcracking, or both. The energy release rate and stress intensity, and thus material toughness and fracture toughness, are all related but provide different measures of the resistance to fracture of the material and have, as we have seen, different units. The energy release rate G measures the energy per unit area of crack extension whereas the stress intensity factor K measures the magnitude (intensity) of the stress field close to the crack tip.

In terms of toughness, the conditions for hydrofracture initiation from its source (Eq. 1) can be presented in various forms, for example as (cf. Gudmundsson, 2011; Kusumoto et al., 2013):

$$p_e = \frac{K_{Ic}}{(\pi a)^{1/2}} \tag{2}$$

where p_e is the fluid excess pressure in the source at the location of rupture and hydrofracture initiation, K_{Ic} is the fracture toughness for an extension (mode I) fracture, and a is the dip dimension (height) of the hydrofracture. When the material toughness G_{Ic} for a mode I crack is used instead of the fracture toughness, then the condition for hydrofracture initiation becomes:

$$p_e = \left(\frac{EG_{Ic}}{\pi(1-\nu^2)a} \right)^{1/2} \tag{3}$$

where E is Young's modulus and ν is Poisson's ratio of the rock. Eq. (3) is for plane-strain conditions, which are most suitable when dealing with hydrofractures that are hosted by crustal segments or layers/units where all the dimensions are similar, or the thickness (dip dimension) is somewhat greater than the other dimensions. When

the crustal segments/layers/units have large lateral dimensions in relation to their thickness, then plane-stress conditions apply, in which case the factor $(1-\nu^2)$ is omitted.

Equations (2) and (3) indicate that the excess fluid pressure in the source needed for hydrofracture initiation is inversely proportional to the dip-dimension of the 'flaw' from which the fracture initiates (due to stress concentrations). On the scale of dikes injected from shallow magma chambers (Fig. 13), the flaws are mostly joints, and in igneous rocks primarily columnar (cooling) joints (Fig. 14). The dimensions of such joints are from tens of centimetres to ten metres or more, which is thus the minimum size of a in these equations.

When the hydrofracture begins to propagate up into the roof and away from the contact with its source (Figs. 13 and 14), buoyancy contributes to the driving pressure, which then becomes the overpressure p_o defined as follows:

$$p_o = p_e + (\rho_r - \rho_f)gh + \sigma_d \qquad (4)$$

Here, ρ_r is the average host-rock density, ρ_f is the average fluid density, g is the acceleration due to gravity, h is the dip-dimension of the dike, and σ_d is the differential stress (the difference between the maximum and the minimum principal stress) in the host rock where the hydrofracture – such as a dike or a mineral vein -- is studied in the field. The buoyancy term is because of the difference in density between the host rock and the hydrofracture fluid. The buoyancy term can be positive (average fluid density less than average rock density up to the point of interest along the hydrofracture), zero or neutral (average rock and fluid density the same), or negative (average rock density less than average fluid density).

For groundwater and geothermal water, the buoyancy term is always positive and reaches the order of mega-pascals once the hydrofracture dip-dimension or height attains 100 m or more. Notice that few individual mineral veins reach this height (Figs 3 and 4), but the fluid path does so as a part of a network of interconnected hydrofractures or veins (Fig. 5). For acid and intermediate magmas, the buoyancy is mostly positive, but may be zero or somewhat negative in the very shallow crust – the uppermost 1-2 km – for intermediate magmas (typical densities of andesite, for example, are 2450-2500 kg m^{-3} and of shallow crustal layers 2300-2500 kg m^{-3}, Gudmundsson, 2020). Mafic magmas, however, have common densities of 2650-2800 kg m^{-3} and so the buoyancy is almost always negative in the uppermost part of the crust for such magmas – but zero and then positive as deeper crustal levels. Since the density difference is generally much smaller – positive or negative – between magmas and crustal rocks than between water and crustal rocks, the buoyancy pressure effects reach the order of mega-pascals only when the height or dip-dimension of the propagating dike reaches 1000 m or more.

Because it is not only fluid excess pressure p_e in the source but rather fluid overpressure p_o where buoyancy comes into play (Eq. 4) that drives hydrofractures, the conditions for hydrofracture propagation (once the fracture has initiated) can be formulated in terms of fracture mechanics criteria as follows:

$$p_o = \frac{K_{Ic}}{(\pi a)^{1/2}} \tag{5}$$

$$p_o = \left(\frac{EG_{Ic}}{\pi(1-\nu^2)a} \right)^{1/2} \tag{6}$$

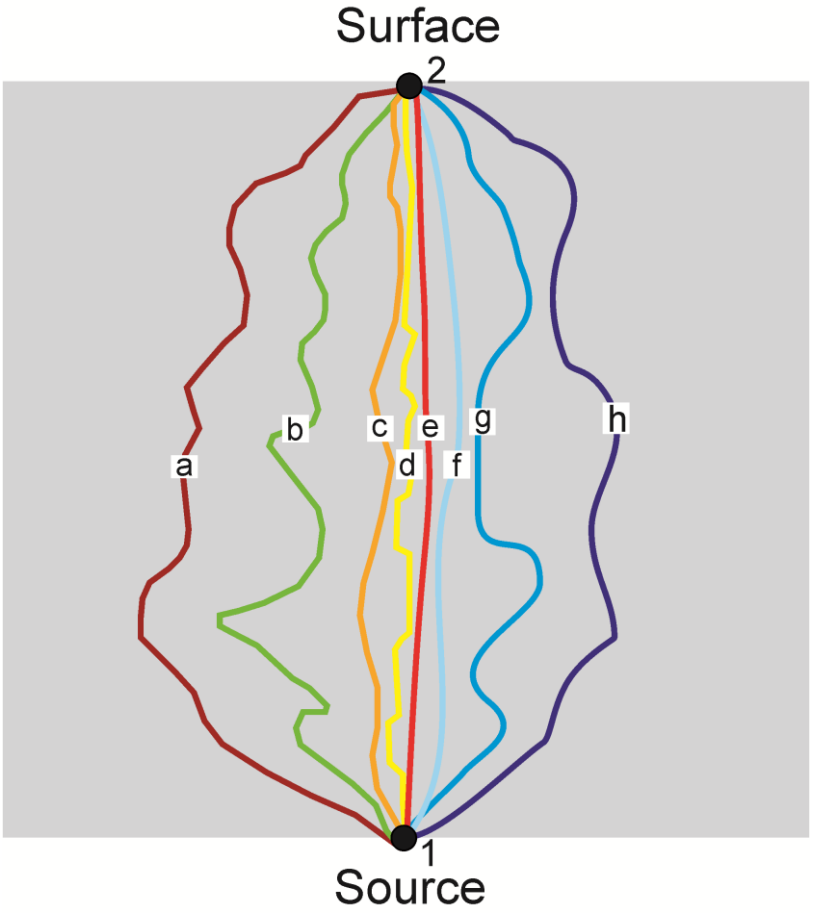

Fig. 15. Hydrofractures initiated at a source can, theoretically, choose among an infinite number of paths to reach the surface (or its point of arrest within the crust). Here the point of initiation is denoted by 1 and the surface point (applicable to a feeder-dike, for example) by 2. Possible hydrofracture paths – only 8 are shown here – are denoted by a-h. Hamilton's principle of least action implies that the hydrofracture selects the path along which the time integral of the difference between the kinetic and potential energies is stationary (is an extremum), and most commonly a minimum, relative to all other possible paths with the same initiation and arrest/surface points.

where Eq. (6) is for plane-strain conditions. When plane-stress conditions apply, then the term $(1-\nu^2)$ in Eq. (6) is omitted. All the symbols have been defined above.

Equations (1) to (6) define the conditions for the initiation and propagation of hydrofractures in terms of overpressure, tensile strength, and toughness. These equations, however, were initially derived for homogeneous and isotropic materials. While they can be applied to heterogeneous and anisotropic materials, they do not, as such, allow us to understand and forecast the commonly complex paths of hydrofractures (e.g. Figs. 5-12).

4. Theory of fracture-path selection

4.1 Hamilton's principle

If a hydrofracture has sufficient energy (Section 5) to begin its vertical or inclined propagation into the commonly layered and heterogeneous host rock from its place of initiation at the source, the fracture can choose among many potential propagation paths. In a crustal segment composed of such rock there is, theoretically, an infinite number of paths that the hydrofracture could select from its initiation to its vertical end (Fig. 15). The vertical end can be either within the crust (an arrested fracture) or, in the case of dikes and inclined sheets that feed eruptions (but more rarely mineral veins), the Earth's surface. Current theoretical understanding does not make it possible to forecast with any reliability the likely detailed path of a hydrofracture – or, for that matter, any outcrop-scale (or larger) rock fracture – through heterogeneous and layered (or, on a finer scale, laminated) rocks, particularly when the contacts and layers have (as is common) widely different mechanical properties (Gudmundsson, 2011, 2020). This is a very unfortunate situation because rock-fracture propagation controls earthquakes, much crustal fluid transport, landslides (lateral collapses), calderas (vertical collapse), formation and development of all types of plate boundaries, as well as most volcanic eruptions.

Here I propose that the eventual path selected by a hydrofracture is the one of least action as determined by Hamilton's principle. Briefly, the principle, as used here, states that the hydrofracture selects the path along which the time integral of the difference between the kinetic and potential energies is stationary (is an extremum) relative to all other possible paths with the same initiation and end points. For most processes to which the Hamilton's principle applies the extremum turns out to be a minimum. Most hydrofractures propagate slowly (commonly at 0.01-1 m s^{-1}) so that the kinetic energy is primarily associated with seismic waves of the induced earthquakes. By contrast, the potential energy is the strain energy stored in the crustal segment or volcano/volcanic zone/sedimentary basin plus the elastic energy supplied by the forces acting on the segment/volcano/basin during fracture propagation. As is explained below, when the kinetic energy is omitted, that is, when

the forces associated with the hydrofracture propagation are conservative and there are no constraints, Hamilton's principle of least action reduces to the principle of minimum potential energy. This is a well-known principle in solid mechanics and was postulated as a basis for understanding dike propagation by Gudmundsson (1986). We come to this point in a later section, but first provide an overview of Hamilton's principle.

In its simplest form, Hamilton's principle is given by:

$$\delta S = \delta \int_{t_1}^{t_2} L dt = \delta \int_{t_1}^{t_2} (T - V) dt = 0 \qquad (7)$$

Here S is the action, L the Lagrangian, t_1 and t_2 are two specified and arbitrary chosen times during the evolution of the system, and δ is the variational symbol which simply denotes a small change. Furthermore, T is the kinetic energy and V is the potential energy, to that the Lagrangian is equal to the difference between the kinetic energy and the potential energy, that is:

$$L = T - V \qquad (8)$$

Hamilton's principle states that the actual path chosen by the system in moving from time t_1 to t_2 is such that the variation of the action δS is zero. This means that actual path taken is the one for which the action integral (Eq. 7) is an extremum and normally minimised along the chosen path. The dimensions of action are energy × time (or linear momentum × distance) and the unit is J s (joule-second). Thus, based on the Hamilton's principle as used here, the selected path of a hydrofracture (Fig. 15) is normally the one along which the energy transformed multiplied by the time taken for the fracture propagation is the least (is a minimum). Alternatively, this conclusion may be stated so that the chosen path is the one along which the difference between the time-averages of the kinetic and potential energies is as small as possible (cf. Theobald, 1966).

Equation (7) applies to a conservative system, that is, a system where the work done by a force is independent of path (reversible) and equal to the difference between the final and initial values of the potential energy (potential energy function). The associated force field can then be expressed as a gradient of the potential. When all the forces acting on a system are conservative, it follows that the system itself is also conservative. In many systems in solid mechanics, such as when applied to the brittle crust, the external force system or loads are such that the body force and the surface stresses are independent of the solid-body deformation (Fung and Tong, 2001). Such systems cannot involve friction. Friction is obviously important when dealing with faulting or, in general, fractures modelled as mode II or mode III, but much less so for fractures modelled as mode I, such as hydrofractures.

4.2 Hamilton's principle for crustal segments

As presented by Eq. (7), Hamilton's principle of least action applies to discrete systems composed of (normally very many) particles. During hydrofracture propagation the system is the rock hosting the fracture, the crustal segment, and thus a continuous system and, normally to a first approximation, an elastic system.

One difference between continuous and discreet systems relates to the degrees of freedom. Degrees of freedom is the minimum number of independent coordinates required to specify completely the position of each and every part of the system - the configuration of the system – which must also be compatible with any constraints on the system. Configuration is here the position, at a given moment, of all the particles (for a discrete system) or all the material points (for a continuous system; Reddy, 2002). The degrees of freedom are thus the number of independent parameters needed to define the system's configuration. For a discrete system of N particles without constraints there are 3N degrees of freedom. N may be very large, but is always finite in a discrete system, whereas for continuous system, N is infinite, that is, the degrees of freedom of a continuous system is infinite.

Another difference between continuous and discreet systems relates to the potential energy *V*. In a discrete mechanical system the potential energy *V* is only a function of the external forces (force field), such as gravity. For a continuous elastic system there is, in addition to the potential energy of the external forces (loading), an internal strain energy - due to internal forces – which, for hydrofractures, concentrates in the elastic rock before the hydrofracture propagation initiates. For example, during unrest and inflation in a volcano, strain energy concentrates in the rocks around and above the magma chamber, including the volcano itself, before a dike/sheet is injected. The generalised external forces Q_i, when conservative and thus derivable from potential energy *V,* are defined by:

$$Q_i = -\frac{\partial V}{\partial q_i} \tag{9}$$

where q_i denotes generalised coordinates.

With these terms clarified, the Hamilton's principle of least action for an elastic solid body (a crustal segment) may be given as follows (Tauchert, 1981; Bedford, 1985; Reddy, 2002):

$$\delta S = \delta \int_{t_1}^{t_2} (T - V - U)\,dt = 0 \tag{10}$$

where S is the action, T is the kinetic energy, V is the potential energy – here due to generalised external forces acting on the elastic rock body (Eq. 9) - and U is the strain energy in the body. In this formulation the external forces or loads that act on the rock body are assumed independent of the elastic displacements that they generate – thus, to be conservative (Eq. 9) - as is common in elastic deformation (Tauchert, 1981; Fung and Tong, 2001) and that there are no constrains.

Together, the strain energy stored in the rock body and the potential energy attributable to the external generalised forces acting on the body are known as the total potential energy, denoted by Π. Thus, we have:

$$\Pi = V + U \tag{11}$$

Then the Lagrangian (Eq. 8) becomes:

$$L = T - \Pi \tag{12}$$

in which case Hamilton's principle (Eq. 10) becomes:

$$\delta \mathrm{S} = \delta \int_{t_1}^{t_2} (T - \Pi) dt = 0 \tag{13}$$

4.3 Minimum potential energy principle

While there is commonly earthquake activity associated with hydrofracture propagation, particularly for large-scale hydraulic fracturing and the injection of dikes, the rate of hydrofracture propagation is slow in comparison with that of seismic ruptures. Thus, for modelling hydrofractures we normally use static elastic moduli (such as Young's modulus) while we use the dynamic moduli when modelling earthquake rupture (Gudmundsson, 2011). If we assume that the kinetic energy can be regarded as effectively zero when a hydrofracture is propagating, so that T = 0, and the system is conservative and elastic, then it can be shown (Richards, 1977; Tauchert, 1981; Reddy, 2002) that for such a body in equilibrium, the total potential energy is a minimum, that is:

$$\delta(V + U) = \delta \Pi = 0 \tag{14}$$

where V is the potential energy due to the generalised external forces, U is the strain energy due to the internal forces, and Π is the total potential energy.

Equation (14) represents the principle of minimum potential energy. One way of expressing the principle in words is as follows. Of all the possible displacement fields or configurations of an elastic body (linear elastic or nonlinear) that satisfy the

external and internal loads and the constraints, the actual displacements are those that make the total potential energy of the body a minimum. Alternatively, the principle can be stated thus. For an elastic body to be in stable equilibrium, it is necessary and sufficient that the total potential energy of the body be a minimum.

The principle of minimum potential energy was already suggested as a basic mechanical framework for forecasting dike paths by Gudmundsson (1986). The present analysis is an extension of that general framework into a much more detailed one that includes all hydrofractures and the results of numerical modelling of their likely paths. Before we turn to the models of hydrofracture paths, we first give a brief overview of the energy aspects of hydrofracture propagation.

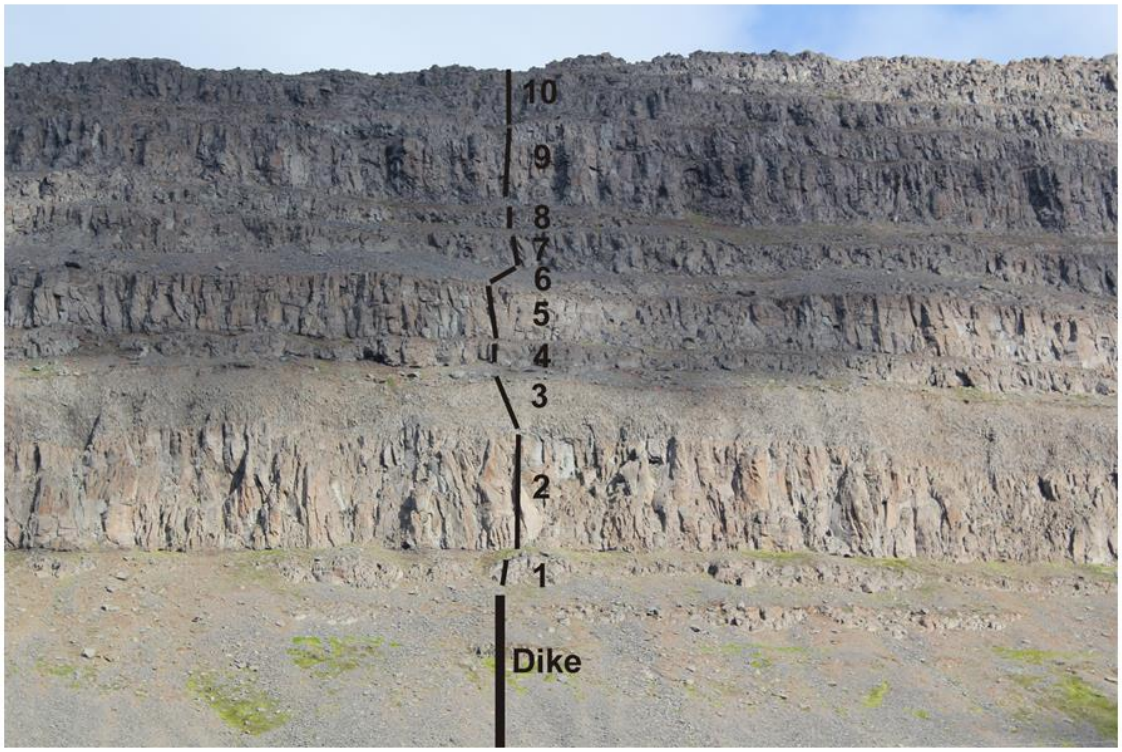

Fig. 16. Dikes and other hydrofractures propagate in steps. When the fracture propagates through layers with widely different mechanical properties and sharp contacts, each step is likely to be similar in length (here height) to the thickness of the mechanical layer through which the dike is propagating. This is indicated schematically here where the potential steps for the further vertical propagation of a dike through a lava pile (in Northwest Iceland) is indicated by the numbers 1 to 10. In the basaltic lava flows, the propagation steps would tend to follow existing columnar joints (Fig. 14). While the steps are discrete, the resulting dike-fracture is normally physically continuous in that the segments/steps are in physical contact. If, with time and burial, the mechanical layers become 'welded together' so as to form thicker units (Fig. 18), each composed of many lava flows, the steps/segments become vertically more extended (longer in the dip-dimension direction).

5. Energy in hydrofracture propagation

The host rock needs energy input for a hydrofracture to initiate and propagate. The energy used to create new surfaces is referred to as surface energy (Anderson, 2005; Gudmundsson, 2011). At an atomic level, surface energy is needed to rupture the solid so that two atomic planes in the solid become separated from each other to a distance where there are no longer any interacting forces between the planes. At the scale of hydrofractures observed in the field, the rupture is rarely at an atomic level, but rather at the level of existing pores, joints, and other flaws in the host rock. For a hydrofracture the surface energy W_s is energy that must be added to the host rock for the fracture to initiate and propagate. Since the energy must be added to the system (the hosting rock body), thermodynamically W_s is regarded as positive.

For a hydrofracture to initiate and propagate the total energy U_t of the hosting rock body or crustal segment must be large enough to overcome the surface energy W_s. The total energy is here composed of two parts (Sanford, 2003; Anderson, 2005), that is:

$$U_t = \Pi + W_s \tag{15}$$

where, as before, Π is the total potential energy of the hosting body. From Eq. (11) it is known that Π derives from two sources, namely the strain energy U and the potential energy V. Here the potential energy V is due to the generalised forces Q_i (Eq. 9), which include body forces such as gravity as well as surface forces/stresses. The external forces contribute to the overpressure of the hydrofracture (Eq. 4). Here the strain energy U is attributable to the internal forces between the material points in the deformed hosting rock body. The strain energy is stored in the body because of changes in the relative location of its material points, that is, changes in its internal configuration as the body deforms and the material points become displaced. This happens before the hydrofracture initiates, so that the strain energy is available to drive the fracture propagation, provided certain conditions are satisfied. Perhaps the best relevant example of large-scale strain-energy storage is during unrest and inflation of a volcano or volcanic field just before magma-chamber rupture and dike injection (Gudmundsson, 2020).

The total strain energy U in a rock body or crustal segment is obtained through multiplying stress × strain × volume and has the unit of joule (J). Strain energy per unit volume U_o is given by:

$$U_0 = \int_V \frac{\sigma_{ij}\varepsilon_{ij}}{2} dV_v \tag{16}$$

Here, σ is stress and ε is strain (the subscripts *ij* indicate the 9 components of the stress and strain tensors), and dV_v is the unit volume of the strained body or crustal segment. (I use the subscript v in V_v for volume to distinguish it from the letter V which denotes potential energy.) Dropping the subscripts *ij* to simplify the notation, and using Hooke's law, $\sigma = E\varepsilon$, where E is Young's modulus, Eq. (16) can be rewritten in terms total strain energy U and strains or stresses for the total volume V_v as follows:

$$U = \frac{\sigma\varepsilon V_v}{2} = \frac{E\varepsilon^2 V_v}{2} = \frac{\sigma^2}{2E} V_v \qquad (17)$$

Thus, the expansion or inflation of the source of a hydrofacture (such as that of a magma-chamber) prior to hydrofacture initiation and propagation results in strain energy being stored in the hosting rock body. From Eq. (17) this strain energy can be calculated either using strain or stress together with Young's modulus for the entire volume of the strained rock body. Once a hydrofracture is initiated and begins to propagate, the strain energy (Eq. 17) is partly used to form the two new fracture surfaces (that is, used as surface energy), and partly for microcracking and plastic deformation in the process zone at the tip of the propagating hydrofracture.

When the hydrofacture has initiated it can begin to propagate away from its source, but does so only if the total energy U_t in Eq. (15) remains constant or decreases during each hydrofracture-front advancement. More specifically, for equilibrium conditions the hydrofracture propagates if $U_t = k$, where k is a constant. During hydrofracture propagation, new surface area dA must be generated. It then follows from Eq. (15) and the condition $U_t = k$ that:

$$\frac{dU_t}{dA} = \frac{d\Pi}{dA} + \frac{dW_s}{dA} = 0 \qquad (18)$$

Thus, from Eq. (18) we have:

$$-\frac{d\Pi}{dA} = \frac{dW_s}{dA} \qquad (19)$$

Equation (19) shows that the decrease in total potential energy Π during hydrofracture propagation is equal to the increase in surface energy W_s, namely that when the hydrofracture propagates stored potential energy in the host rock is released and transformed into surface energy. The rate at which this release or

transformation occurs, known as energy release rate and denoted by *G*, is (from Eq. 19) given by:

$$G = -\frac{d\Pi}{dA} \tag{20}$$

Here, G may be regarded as the energy available to drive the hydrofracture propagation. This means that the hydrofracture propagates only if the energy release rate *G* reaches the critical value on the right-hand side of Eq. (19):

$$G_c = \frac{dW_s}{dA} \tag{21}$$

where the critical value G_c is, as discussed earlier, known as material toughness of the hosting rock body (Anderson, 2005; Gudmundsson, 2011).

6. Hydrofracture paths

6.1 General

Hamilton's principle (Eq. 7) normally means the 'path' along which a system 'moves' and reflects changes in its configuration rather than an actual movement of the system as a whole. Each point on the path or curve along which the system moves through time corresponds to one configuration of the system – one arrangement of the particles (for a concrete system) or the material points (for a continuous system). Here, however, we extend the principle so as to refer to actual propagation paths of hydrofractures in space and time, in its reduced version as the principle of minimum potential energy.

This extension is particularly appropriate when dealing with hydrofracture propagation, because hydrofractures advance in steps (Fig. 16). The steps are partly a consequence of the time lag between the fracture front and the fluid front at any particular instant, especially when the fluid is magma. When the overpressure at the tip of the hydrofracture reaches the condition for propagation (Eqs. 5 and 6) the upper end (the tip) of the hydrofracture advances very quickly (at about the velocity of S-waves, kilometres per second) for a certain distance and then stops (becomes temporarily arrested). The fluid viscosity, particularly if the fluid is magma, does not allow it to flow as quickly as the hydrofracture tip propagates so that, following each fracture-tip advance (step), the fracture front will, for a while, be essentially empty of the fracture-driving fluid (Fig. 17). Thus, the fluid front has to move into the fracture

front, fill it, and build up a pressure so high that the hydrofracture tip can advance again (Eqs. 5 and 6). This process, filling the fracture-front and building up the overpressure for further rupture takes time, hence the time lag between the fracture front and the fluid front.

Hydrofractures therefore propagate in steps, each of which may be regarded as following Hamilton's principle of least action (Eq. 13) or, if the kinetic energy associated with earthquakes is omitted, the principle of minimum potential energy (Eq. 14). Following each fracture-segment advancement, an approximate mechanical equilibrium with the surrounding host-rock body/layer is reached. However, when the fluid continues to flow to the fracture-front and builds up the overpressure again

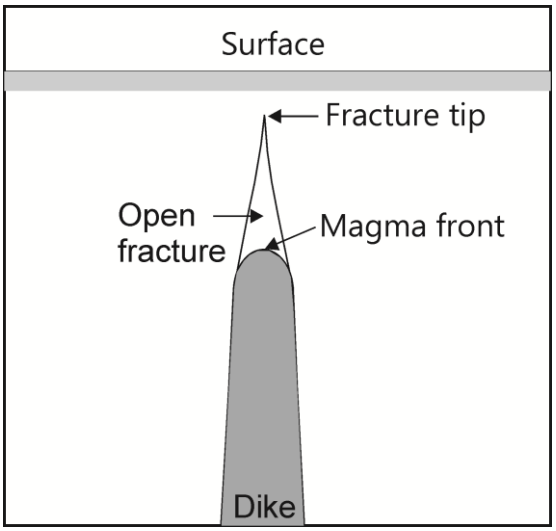

Fig. 17. During hydrofracture propagation (and particularly during the propagation of magma-driven fractures) the fluid front normally lags behind the fracture front or tip. At intervals during the propagation there will thus be an open fracture front ahead of the fluid front. This conclusion is supported by field experiments and theories on human-made hydraulic fractures (Davis et al., 2012; Flewelling et al., 2013; Yew, 2013). Later on the fracture-driving fluid front 'catches up' with the fracture front.

the equilibrium gradually becomes unstable at and close to the fracture front until the rupture and a new fracture-front advancement occurs.

The size of a typical vertical fracture-front advancement during hydrofracture propagation depends much the mechanical layering of the host rock. When the host rock is a pile of lava flows and pyroclastic or sedimentary/soil layers, the vertical advancement (the steps) may correspond to the thicknesses of the layers ahead of the fracture-front (Figs 4, 6, 14, 16 and 18). This applies particularly to a young pile, such as is most common in active volcanic areas and sedimentary basins, with sharp mechanical discontinuities at the contacts between layers. When the pile becomes older, there is commonly a gradual reduction in mechanical contrast between layers (partly due to secondary mineralisation and general compaction). Also, the contacts

between layers in volcanic fields may be partly welded together. In both these cases many layers may then function as single mechanical layers/units during hydrofracture propagation (Figs 6 and 18). The fracture-front, even for dikes, however, cannot normally propagate through very thick layers (tens or hundreds of metres) in a single step.

For each hydrofracture advance there is, theoretically, an infinite number of possible paths (Fig. 18), and also for the overall path of the hydrofracture (Fig. 15). The actual or true path taken from point the point of hydrofracture initiation at time t_1 to the endpoint or tip of the fracture, to which it reaches at time t_2, is the one along which the action S – or if kinetic energy is omitted, the potential energy Π - is minimised.

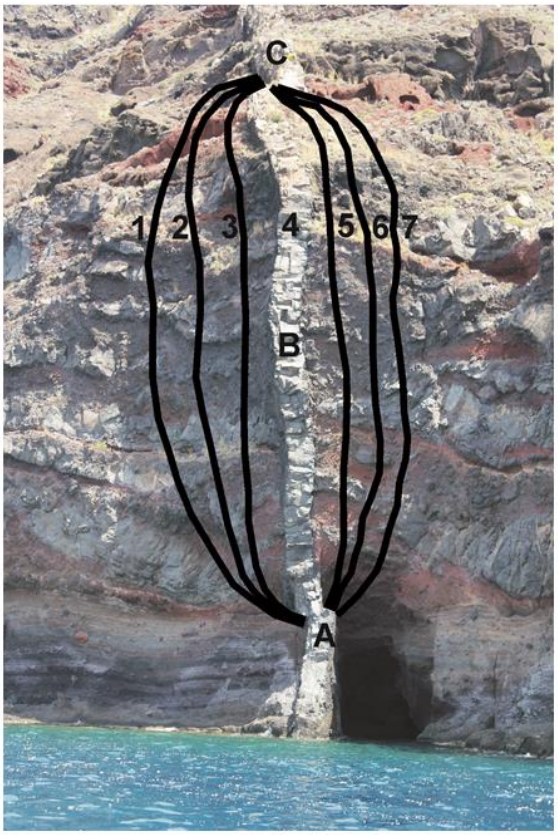

Fig. 18. For each dike-fracture step (Fig. 16) there is, theoretically, an infinite number of possible paths (Fig. 15). Some of the potential paths that the propagating tip of the dike seen here could have followed from segment A to segment C - so as to form segment B - are indicated by the numbers 1 to 7. The actual path taken to form segment B is the one number 4 (the dike is about 1.5 m thick). Here many lava flows and scoria layers are 'welded together' so as to function as single mechanical layers of thickness similar to the heights or dip-dimensions of the individual dike segments, such as segment B.

6.2 Application to dike propagation

Dikes constitute the largest hydrofractures and have been studied intensively, both as solidified sheet-like structures in the field and as propagating magma-driven fractures during unrest periods and volcanic eruptions (Rivalta et al., 2015;

Gudmundsson, 2020). Here the focus is on the application of the theoretical framework to dikes, but all the principal conclusions as regards mechanics apply to other types of hydrofractures (inclined sheets, sills, mineral veins, many joints, and hydraulic fractures) as well.

Hamilton's principle indicates that dikes seek the path that minimises the action, that is, the used energy × time. We have seen that a dike-fracture will propagate if the energy release rate reaches the critical value of the material toughness given by Eq. (21). In addition to rupturing the rock, a dike-fracture has a certain opening – for solidified dikes, thickness. The final thickness of a dike may be about 10% less than the opening of the dike-fracture; yet the thickness of solidified dikes may be taken as a good measure of their opening displacements. Additionally, even if not as accurate as the direct field measurements (Geshi et al., 2010, 2012; Kavanagh and Sparks, 2011; Galindo and Gudmundsson, 2012; Geshi and Neri, 2014), geodetic data give indications of the opening-displacements of present-day dikes, particularly feeder-dikes (Rubin and Pollard, 1988).

To open up the dike-fracture, work has to be done against a force, namely the normal force that acts on the dike-fracture walls. The amount of work needed varies positively with the magnitude of the normal force. Stress and pressure are defined as force per unit area. It follows that when the magmatic overpressure first breaks the rock and then pushes the walls aside to form the dike-fracture then more work is needed, for a given opening displacement, when the push is against a compressive stress/force per unit area that is high rather than low. Much more energy (work) is therefore required of the dike-fracture to open up (the fracture walls to be displaced) against the maximum σ_1 or the intermediate σ_2 principal compressive stresses than against the minimum compressive (maximum tensile) principal stress σ_3. Based on Hamilton's principle of least action (Eq. 10) or, if the kinetic energy is zero, the principle of minimum potential energy (Eq. 14), dike-fracture propagation should thus be along a path that is perpendicular to the direction of σ_3 and, therefore, coincides with the trajectories (directions) of σ_1. The time constraints in Hamilton's principle would furthermore suggest that the dike would tend to follow the shortest path that is compatible with the other constraints.

Homogeneous, isotropic host rock

Let us first look at the simplest case, that is, dike propagation from a shallow chamber located in a homogeneous and isotropic crustal segment. The least action/minimum principles indicate that the dike path should everywhere be perpendicular to σ_3 and thus follow the trajectories of σ_1 for the shortest distance between t_1 and t_2 (Eq. 10). That the path is straight follows from the arrangement of the σ_1-trajectories (Fig. 19) and is also a well-known result from the calculus of variations (Washizu, 1975; Cassel, 2013), demonstrating the fact that a straight line is the shortest distance between given points. Thus, provided the step-like average rate

of dike propagation (Fig. 16) is approximately constant, the straight path also minimises the time needed for the propagation among a family of nearby curved or somewhat irregular paths (Figs. 15 and 18) with the same endpoints, t_1 and t_2.

Of the many straight and potential paths from the source chamber to the surface, the path selected is the one from t_1 and parallel with σ_1 to t_2 (Fig. 19). Her point t_1, which basically determines where the dike initiates, is the point or location of highest tensile stress concentration at the margin of the chamber, namely where the conditions of Eq. (1) is first satisfied. As for point t_2, it is a point on the tip line for an arrested dike and on the volcanic fissure for a feeder-dike. Using Eqs (6) and (21),

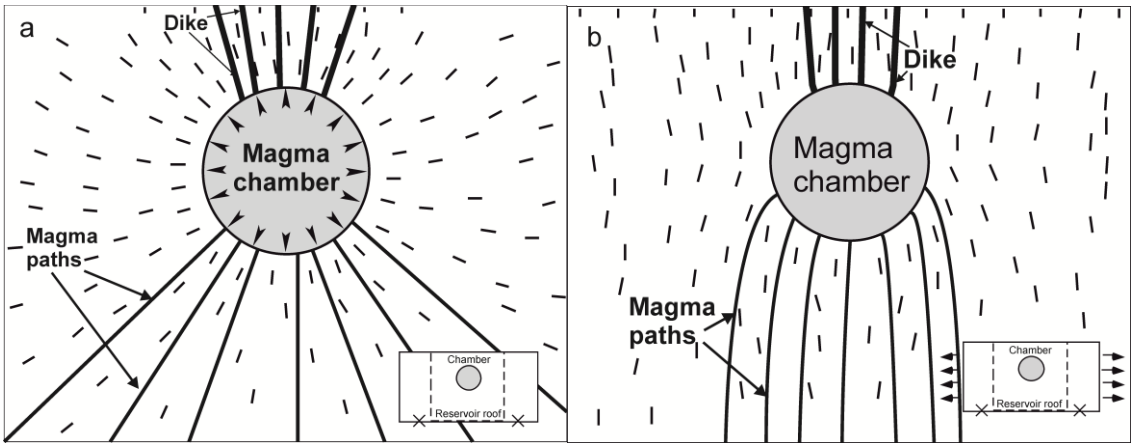

Fig. 19. Theoretical dike-propagation paths in a homogeneous, isotropic crustal segment. The crustal segment has a uniform Young's modulus of 40 GPa and a Poisson's ratio of 0.25, both values being appropriate generalised average static values for the crust in Iceland (Gudmundsson, 1988). The model is fastened at the corners (indicated by crosses) and was made using the boundary-element program BEASY (www.beasy.com). The boundary-element method, with application to BEASY, is described by Brebbia and Dominguez (1992) and on the homepage www.beasy.com. Several potential paths from the roof of the shallow chamber (of a circular, vertical cross section) to the surface are indicated. Also indicated are some potential magma paths from the source reservoir to the floor (lower margin) of the chamber. The ticks show the trajectories of σ_1, the likely dike paths (and magma paths) being parallel with these. When the loading of the chamber is (a) internal magmatic excess pressure of 5 MPa (see the inset), the potential dike-paths are more spread, more fan-shaped, than in (b) where the only loading is external tensile stress of 5 MPa (see the inset). More specifically, the loading conditions in (a) could be reached when the chamber receives new magma from the deeper source. By contrast, the loading conditions in (b) could be reached when a chamber initially in lithostatic/mechanical equilibrium with the host rock becomes subject to tensile stress, such as would be common at divergent plate boundaries.

the dike-fracture propagation becomes arrested at t_2 somewhere inside the crust when:

$$G_I = \frac{p_o^2(1-\nu^2)\pi a}{E} < G_c = \frac{dW_s}{dA} \qquad (22)$$

Here, G_I is the plane-strain energy release rate, p_o is the magmatic overpressure in the dike, a is half the height (dip dimension) of the dike, E is Young's modulus of the host rock, G_c is the material toughness of the host rock, and W_s is the surface energy

needed to form the dike-fracture as it extends so as to form the new surface area *A*. Thus, if the energy release rate during step-like dike-tip propagation (Fig. 16) is less than the material toughness, the dike-fracture propagation stops or becomes arrested. Presumably, however, most dikes injected into an approximately homogeneous, isotropic crustal segment would reach the surface – become feeders (cf. Galindo and Gudmundsson, 2012).

Currently, it is difficult to forecast exactly where at the boundary of the magma chamber the rupture leading to dike injection will occur. Analytical and numerical

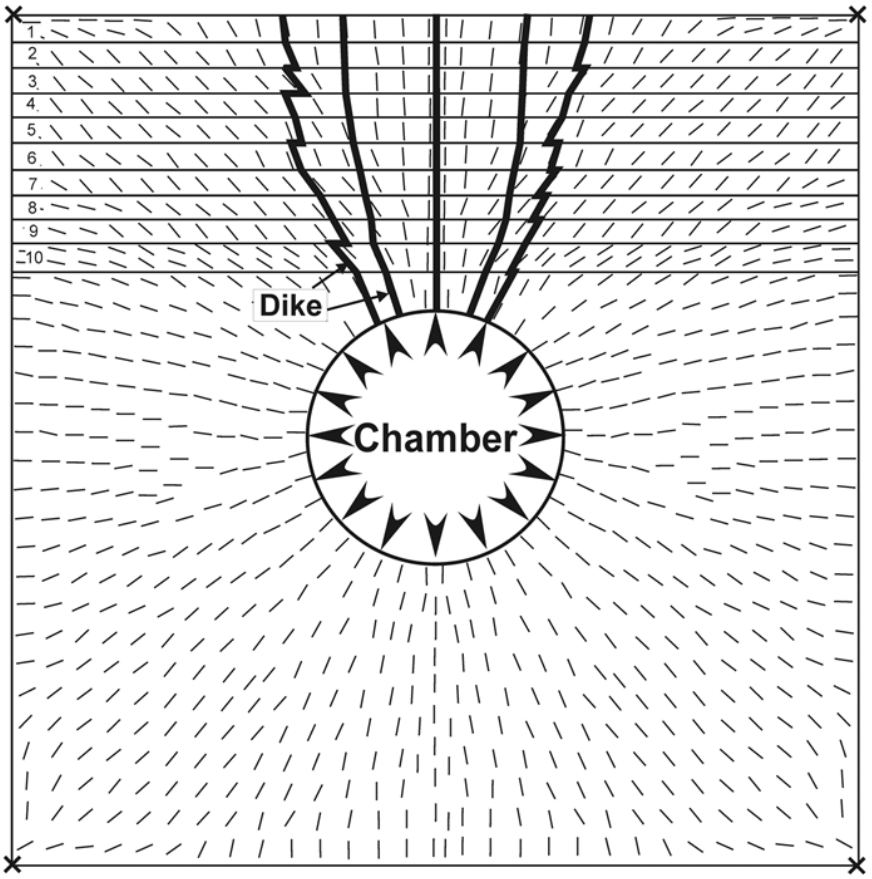

Fig. 20. Theoretical dike-propagation paths in a layered crustal segment. The numerical model was made using the finite-element software ANSYS (www.ansys.com). The finite-element method is described by Logan (2002). There are 10 layers above the unit hosting the chamber. Top layer (layer 1) has a Young's modulus of 10 GPa, which then gradually increases with depth by 2 GPa for each layer so that layer 10 has a Young's modulus of 28 GPa and the layer hosting the chamber a Young's modulus of 30 GPa. All the layers have a Poisson's ratio of 0.25. The range in values for Young's modulus and Poisson's ratios of rocks in general is provided by Gudmundsson (2011). The only loading in the model is magmatic excess pressure of 10 MPa in the chamber. The model is fastened at the corners (indicated by crosses). This crustal segment is thus approaching homogenisation as regards mechanical layering, as some segments do when they become older. The three potential dike paths above the central part of the chamber roof are comparatively smooth, whereas the two outermost two paths show greater variation in geometry and overall length from the source to the surface. Hamilton's principle implies that the paths of least action would be somewhere above the central part of the chamber.

models can be used to infer regions of highest stress concentration at the chamber boundary during unrest periods, based on the mechanical properties of the host rock and the geometry and properties of the chamber (Gudmundsson, 2020). Additionally, when the chamber ruptures and a dike initiates, the associated earthquakes provide a rough guide to the approximate location and timing of the rupture, and thus to the location of t_1 (Hobe et al., 2022). Once the location of t_1 is determined, then the principles above make it theoretically possible to forecast the likely propagation path to the point t_2.

Heterogeneous, anisotropic rock

The above assumptions, namely that the host rock is homogeneous and isotropic, are very common in geodetic studies in volcanology but generally not warranted. Stratovolcanoes are composed of strata, of layers, with widely different mechanical properties (and the same applies to sedimentary basins and thus to the host rock of hydrofractures in general). Basaltic edifices (shield volcanoes) are composed of layers where the mechanical properties are more uniform than in stratovolcanoes, but the Young's modulus can still vary by orders of magnitude between compliant scoria or soil layers and stiff lava flows (or sills) in basaltic edifices. The rocks of real volcanoes and volcanic zones are also heterogeneous (meaning: material properties change with position in the body), but here the focus is on the effects anisotropy (meaning: material properties have different values in different directions at a given point within the body), particularly in connection with layering, on the dike-propagation paths.

Mechanical layering can have strong effects on the dike-path geometry and the conditions for dike-path arrest. We have already discussed the effects of layering on dike arrest, so here the focus is on the path geometry. Layering gives rise to local stresses that involve changes in the orientation of the trajectories of σ_1 (and, consequently, the orientations of σ_2 and σ_3). Hamilton's principle implies that the dike-fracture seeks to be everywhere parallel with σ_1 and perpendicular to σ_3 so as to minimise the energy used during the propagation. The dike-fracture also seeks to minimise the time needed to propagate from t_1 to t_2. The propagation velocity of dikes varies somewhat, primarily between about 0.01 and 1 m s^{-1} (Hobe et al., 2022). For typical rates of dike propagation, say 0.1 m s^{-1}, it follows that minimising the duration implies minimising the dike-path length, for the given constraints.

Of all the possible dike-paths following the trajectories of σ_1 the shortest one between t_1 and t_2 would normally be selected. To see what this implies, we can consider a simple two-dimensional model where the stiffnesses (Young's moduli) of the layers gradually increases with depth (as is common) in the volcanic zone/volcano (Fig. 20). Five possible paths of feeder-dikes (out of numerous potential paths) are indicated. The shortest path, and the one likely to be selected in this case (but also depending on the tensile stress concentration around the chamber), is one in the centre. The trajectories of σ_1 indicate that many of the paths would reach the surface, particularly above the central part of the magma chamber. That is largely because

there is little contrast in the mechanical properties (stiffness) of the layers – as would be common if the chamber were located in a basaltic edifice.

The 30-layer model (Fig. 21) provides more details as to possible dike paths. The shallow magma chamber is located in a single thick layer or unit (with a Young's modulus of 40 GPa) and subject to 10 MPa excess pressure – which is the only loading in the model. The roof above the chamber is mostly composed of 30 layers of alternating Young's moduli of 1 GPa and 100 GPa. The contrast in stiffness between the layers is thus by two orders of magnitude, as is common in many volcanic fields

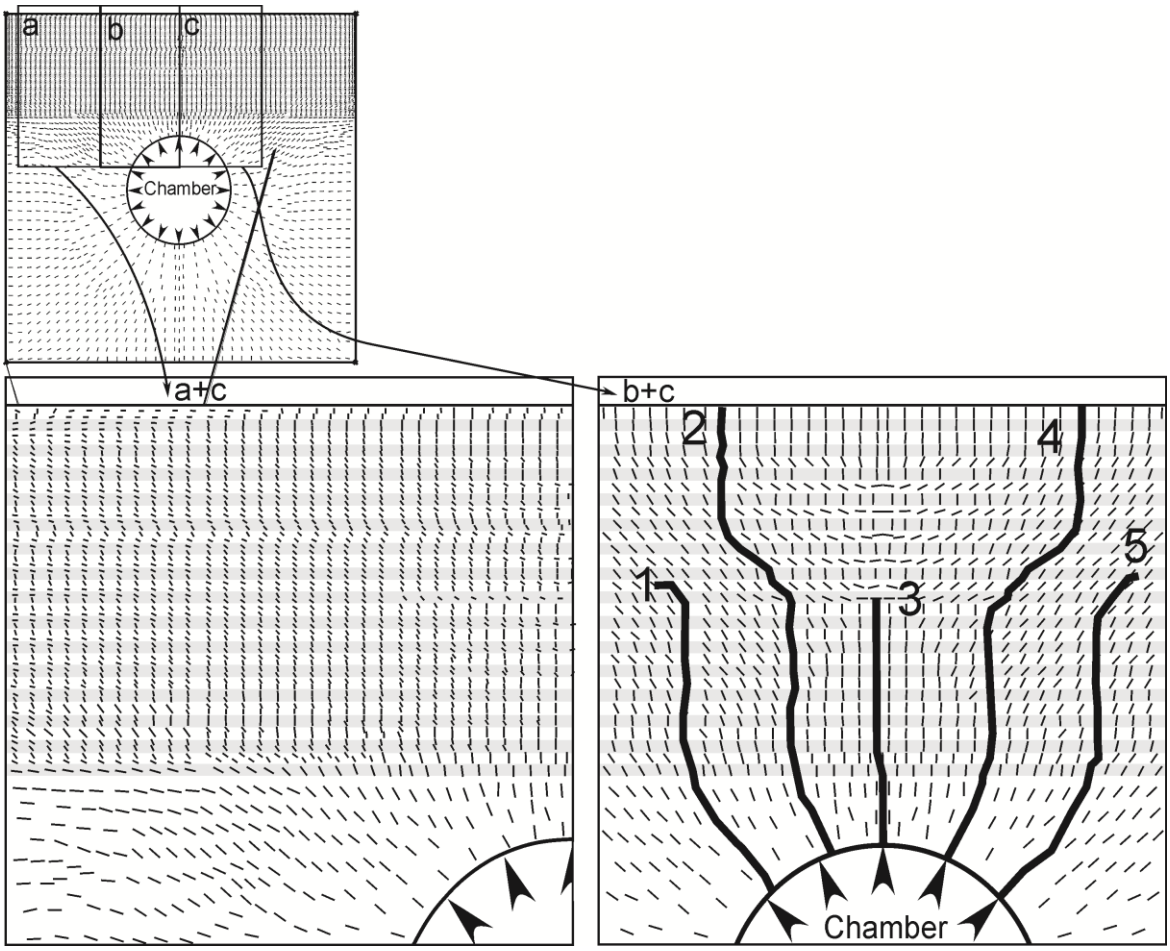

Fig. 21. Some potential paths of dikes and inclined sheets injected from a magma chamber of circular cross-section into a roof composed of 30 layers of alternating Young's moduli of 1 GPa and 100 GPa. Internal magmatic excess pressure of 10 MPa is the only loading. The numerical model is fastened at the corners (indicated by crosses in the top illustration). The model was made using the finite-element software ANSYS (www.ansys.com; cf. Logan, 2002). The top illustration provides and overview of the entire model. The lower left illustration is a close-up of part of the model, showing the σ_1- trajectories (ticks) in the individual layers above a part of the magma chamber. In the lower right figure some potential dike paths, following the σ_1- trajectories, are shown. Only five potential paths are indicated, two of which may possibly reach the surface.

and in stratovolcanoes in particular. The absolute stiffness values, however, in many stratovolcanoes could be somewhat different. Thus, the stiff layers might have a

Young's modulus of 50 GPa and the thin layers of 0.5 GPa; but the factor to be explored here is the effect of stiffness contrast on potential dike paths, rather than the absolute stiffnesses (which are, however, all within the general range of Young's moduli of volcanic rocks – Gudmundsson, 2020).

The thicknesses of the layers in the model (Fig. 21) depends on the depth of the magma chamber. If, for example, the magma chamber had the top part of its roof at 1.5-2 km, as is common at divergent plate boundaries such as in Iceland (Gudmundsson, 2020), the thickness of each of the 30 layer would be between 50 m and 67 m. Many layers of pillow lava, hyaloclastite, and sediments, for example, reach these thicknesses, whereas most ordinary lava flows would be thinner (5-20 m for basalt being common). The layering in the model may thus be regarded as an approximation to the layering exposed in Iceland and many volcanic rift zones, but also appropriate for many crustal segments at convergent plate boundaries.

For this local stress field, the geometry and eventual fate of the dike path depends much on the location of the point of magma-chamber rupture, that is, the point (line, curve at the surface of the chamber) of origin t_1 of the dike. Within the thick unit hosting the chamber, some of injected intrusions in the marginal parts of the roof would be inclined sheets (Fig. 2). Once a propagation path enters the 30 layers, however, the path geometry becomes more complex, particularly in the upper half of the pile of layers. This is in harmony with many dike paths observed in the field (Fig. 8). Also, many of the dikes would be likely to become arrested (Fig. 12c, d). Of the five theoretical dike paths indicated here, three would be expected to become arrested, while two have a chance of reaching the surface to supply magma to an eruption. In particular, many dikes would be expected to become arrested in the upper central part of the pile. The arrest is there attributable to abrupt 90° flip in the orientation of σ_1 in many of the layers in that part. However, in the uppermost 5-6 layer, that is, close to the surface, the orientation of σ_1 again flips by 90°, back to vertical as in the lower layers and would, therefore, be favourable to dike propagation.

Despite being comparatively simple, these numerical models illustrate that the least action or minimum potential energy principles can be used as a means of forecasting the likely dike-propagation paths during unrest periods with dike injection. Such forecasts, even if they are rather elementary at this stage, are of great theoretical and practical importance because they allow us to predict the likely paths. Predictions of this kind would include, say, the direction in which a dike is likely to propagate within a stratovolcano, perhaps many days before the actual propagation is completed. Furthermore, the predicted dike-paths can sometimes be compared with the actual paths, as determined by induced earthquake swarms, almost in real time. The latter, however, requires a dense seismic network that is operational at the time of magma-chamber rupture and dike injection. Some volcanoes and volcanic zones have comparatively dense permanent seismic networks, such as exist for the volcanic

zones of Iceland (e.g., Jakobsdottir, 2008; Hobe et al., 2022). These can be used to determine the dike propagation (Gudmundsson et al., 2014), while more details can be obtained through very dense, transportable networks (Agustsdottir et al., 2016). During dike propagation, the earthquakes are generated mainly in the process zone ahead of the propagating dike tip but also in the walls of the dike due to the magmatic pressure (Gudmundsson, 2020).

These forecasts (Figs. 19-21), however, consider only the effects of layering on the dike paths. In addition to layering, all volcanoes and volcanic zones contain numerous fractures of varying sizes. The cooling or columnar joints are used to form

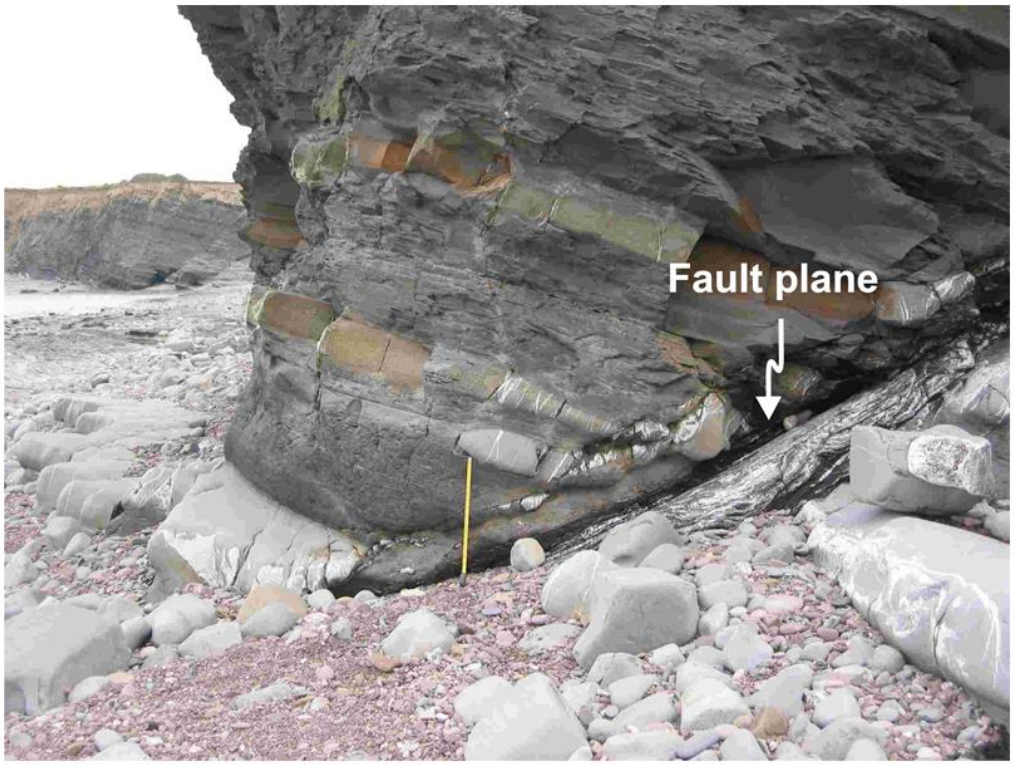

Fig. 22. Mineral veins along the plane (indicated) of a normal fault in the Bristol Channel of Southwest England. The fault dissects layers of limestone and shale. The veins are of calcite (cf. Philipp, 2008).

the paths (Gudmundsson, 1986), but these are generally evenly distributed in in the pile and therefore do not encourage the dike to deviate from the path parallel with σ_1. Field observations show that some dikes use faults as parts of their paths, and those parts would normally not be parallel with σ_1 – yet, as will be shown below, the parts of dike paths that may follow faults would still be in agreement with Hamilton's principle.

6.3 Effects of faults on hydrofracture paths

For a constant tensile strength T_o (Eq. 1) hydrofracture paths would be expected to follow the trajectories of σ_1. All rocks, however, contain fractures, most commonly joints, which are weaknesses used by many hydrofractures (Fig. 14). Many joints are roughly uniformly distributed in the hosting layers, such as columnar joints in lava flows, and would normally not result in significant deviation of a hydrofracture path from the direction of σ_1. Some hydrofractures, however, use faults for short (Fig. 9) or long (Fig. 22) parts of their paths. In fact, faults commonly contain networks of mineral veins, most of which are extension fractures (Fig. 5). To understand the conditions that favour hydrofractures using faults as parts of their paths, we again focus on dikes (Fig. 23).

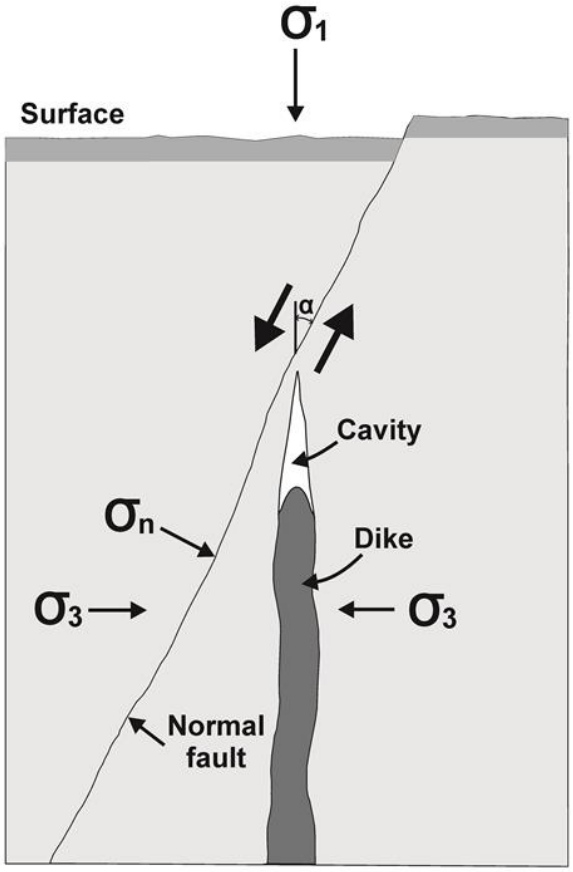

Fig. 23. When a propagating hydrofracture (here a dike) meets a fault, such as the normal fault seen here, the hydrofracture may enter the fault and use it as a part of its path (Figs 9 and 24). Here we explore the situation where a vertical dike meets a steeply dipping normal fault, but the analytical results are easily generalised to other types of faults and hydrofractures. As before, σ_1 and σ_3 are the maximum and minimum principal stresses, respectively, σ_n is the normal stress on the fault plane, and α is the acute angle between σ_1 and the fault plane. Dike propagates in steps (Figs 16 and 17) and during each step the fracture front propagates much faster than the fluid front, generating a temporary (air or gas-filled) cavity devoid of magma at the dike tip, a cavity which the magma subsequently flows into.

Faults, by definition, are shear fractures and the fault plane is oblique to σ_1 (and the other principal stresses) at the time of fault formation or slip. The tensile strength across many active or recently active faults is likely to be essentially zero. Less energy may thus be needed for a dike or a dike-segment to use the fault, even though the segment is not perpendicular to σ_3 but rather to the normal stress on the fault plane, σ_n, which, for an active fault is always higher than σ_3. Apart from ring-dikes, all of which occupy faults, dikes rarely follow faults. Some, however, do so for a part of their paths, and here we analyse under what conditions they would be likely to do so.

When a dike segment follows a path that is perpendicular to σ_n the effective overpressure available to drive the dike-fracture open becomes less than the one given by Eq. (4), which assumes the dike-segment to be perpendicular to σ_3. The normal stress σ_n on a fault plane (Fig. 23) is given by:

$$\sigma_n = \frac{\sigma_1 + \sigma_3}{2} - \frac{\sigma_1 - \sigma_3}{2}\cos 2\alpha \qquad (23)$$

where σ_1 and σ_3 are the maximum and minimum principal compressive stresses, respectively, and α is the angle between the fault plane and the direction of σ_1. The difference between the normal stress σ_n and the minimum principal compressive stress σ_3 is given by:

$$\sigma_n - \sigma_3 = \frac{\sigma_d}{2}(1 - \cos 2\alpha) \qquad (24)$$

where $\sigma_d = (\sigma_1 - \sigma_3)$, as defined in Eq. (4).

Using Eqs. (1) and (4) in combination with Eq. (24), the condition that a dike is likely to inject an existing fault and use it as a part of its path becomes:

$$\sigma_n - \sigma_3 = \frac{\sigma_d}{2}(1 - \cos 2\alpha) \le \Delta T_0 \qquad (25)$$

where

$$\Delta T_0 = T_0^{\sigma_3} - T_0^{\sigma_n} \qquad (26)$$

is the difference between the tensile strength along a potential dike path that is perpendicular to σ_3 (and parallel with σ_1) and a dike path that is perpendicular to σ_n. If tensile strength along the path perpendicular to σ_n is zero, which is probably often the case for an active or recently active fault, then we have:

$$\Delta T_0 = T_0 \tag{27}$$

Equation (4) gives the dike/hydrofracture overpressure assuming that the dike-segment opens up against σ_3. As indicated above, the overpressure is less when the dike-fracture opening is against σ_n. Denoting the overpressure against σ_n by $p_o^{\sigma_n}$ and using Eqs. (4) and (25), we get:

$$p_o^{\sigma_n} = p_e + (\rho_r - \rho_m)gh + \frac{\sigma_d}{2}(1 + \cos 2\alpha) \tag{28}$$

As long as α is positive, in which case $\sigma_n \neq \sigma_3$, then $P_o^{\sigma_n} < p_o$, that is, the overpressure in Eq. (28) is less than that in Eq. (4).

The energies needed to form a dike-segment in a direction perpendicular to σ_n and σ_3 can now be compared. Work, which is force × displacement, is positive if the work is in the direction of the force, but negative if the work is in a direction opposite to the force. The work W needed to form a dike-segment of final volume is ΔV_v against the normal stress σ_n is:

$$W = \Delta V_v \sigma_n \tag{29}$$

For same-volume segment opened against σ_3, the work is:

$$W = \Delta V_v \sigma_3 \tag{30}$$

The force or overpressure that opens the dike-segment varies linearly with the opening displacement so that the corresponding elastic energy must be multiplied by ½. The total energy $U_t^{\sigma_n}$ needed to form the dike-fracture, that is, rupture the rock and then open the rupture/fracture up against σ_n is therefore:

$$U_t^{\sigma_n} = W_s + \frac{\sigma_n \Delta V_v}{2} \tag{31}$$

where W_s is the surface energy (Eq. 15). Similarly, the total energy needed to form and open the dike-segment against σ_3 is:

$$U_t^{\sigma_3} = W_s + \frac{\sigma_3 \Delta V_v}{2} \tag{32}$$

When the surface energy of rupture, W_s, is constant for the host rock of a given dike-segment then, because $\sigma_3 < \sigma_n$, it follows on comparing Eqs (31) and (32) that less

energy is needed to form a segment perpendicular to σ_3 than to σ_n. Thus, as expected, a dike (and any hydrofracture) tends to follow the path perpendicular to σ_3, namely the path parallel with σ_1.

If the tensile strength across a recently active fault is zero while being several mega-pascals along the potential dike-path following σ_1 at a given locality, then Eqs. (25) to (27) suggest that less energy, or less action, for a dike segment of a given length is needed if the dike follows the recently active fault than if the dike forms its own path (and thus rupturing the rock) along σ_1. More specifically, a dike may follow an existing fault along part of its path (Fig. 24) if the difference $\sigma_n - \sigma_3$ (Eq. 25) is less than the difference in tensile strength (Eq. 26) between the rock along σ_1 and along

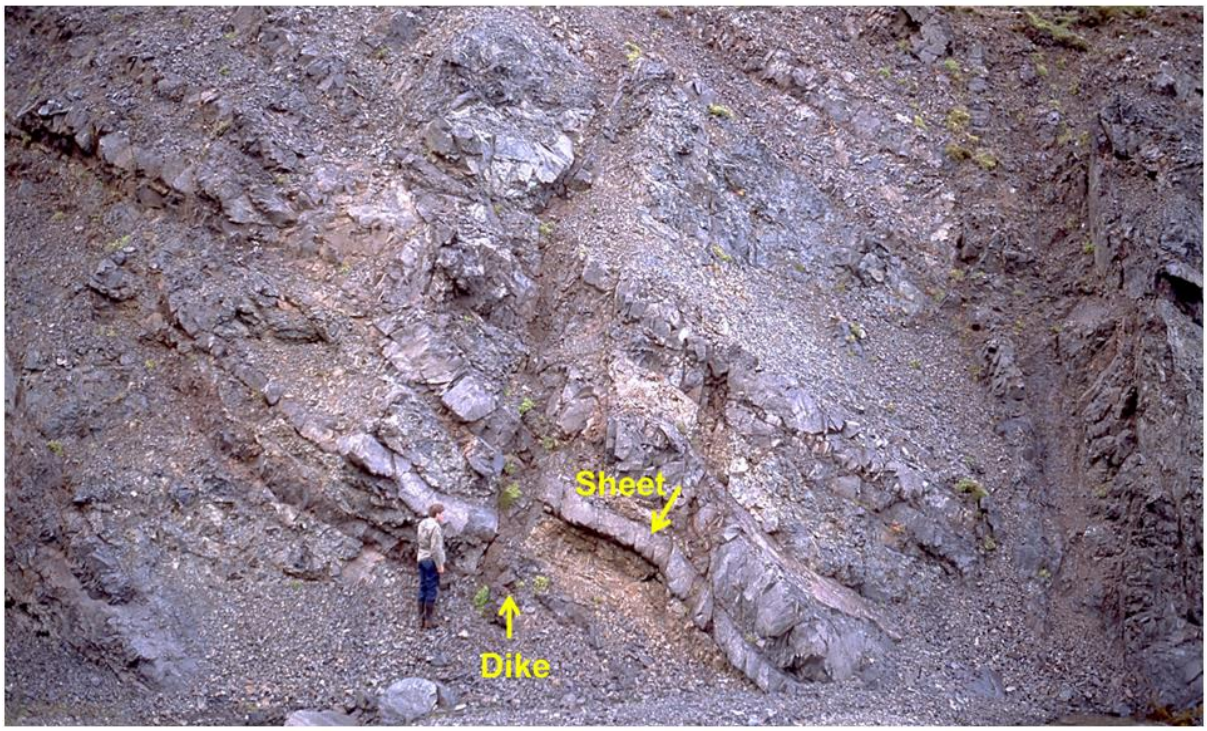

Fig. 24. Some ordinary dikes, such as the basaltic dike here (indicated), follow faults for a while along their paths, particularly steeply dipping normal faults. The fault dissects a swarm of gently dipping inclined sheets (one sheet is indicated). The structures seen here are a part of fossil central volcano in West Iceland.

the fault. The stress difference $\sigma_n - \sigma_3$ depends much on the angle α between the fault plane and σ_1 (Eq. 24); the smaller the angle α, the more likely the dike is to enter the fault plane on meeting the plane (Figs. 23 and 24). In particular, for normal faults, as are most common in volcanic zones with dike injection, small α means

steeply dipping faults, which are the most likely to be used as parts of dike paths. In some volcanic areas, however, vertical strike-slip faults are common, some of which may be used as parts of dike paths, particularly close to the surface (Hobe et al., 2022).

7. Discussion

One of the important unsolved problems in solid-earth geosciences is to provide a theoretical framework that allows us to make reliable forecasts for rock-fracture propagation paths through layered and faulted rocks. The reason why the solution of this problem is so important is that brittle deformation, which dominates in the upper part of the crust, is mostly through fracture initiation and propagation and most crustal segments are composed of layered and faulted rocks. Thus, fracture propagation controls earthquakes, landslides (lateral collapses), calderas (vertical collapse), and the formation and development of all types of plate boundaries. Furthermore, fractures to a large degree determine fluid transport in the Earth's crust, including flow of groundwater, geothermal water, and hydrocarbons. Additionally, most volcanic eruptions are supplied with magma through magma-driven fractures, primarily inclined sheets and dikes.

There has been much research on hydrofracture propagation over many decades, particularly in the hydrocarbon industry (Valko and Economides, 1995; Yew and Weng, 2014; Shapiro, 2018). For more than 70 years, conventional hydraulic fractures have been propagated from vertical wells laterally for distances of up to 1 km or more with the aim of increasing the permeability (for oil and gas) of the target layer. In recent decades a technique has been developed whereby hydraulic fractures are injected vertically from horizontal wells so as to obtain gas from tight shale layers (Wu, 2017; Shaphiro, 2018). Some of the vertical hydraulic fractures exceed 1 km in dip-dimension (height) and are thus with dip-dimensions similar to those of many radial dikes and inclined sheets.

In conventional human-made hydraulic fracturing, the likely propagation path of a fracture injected laterally so as to stay in the same mechanical unit/layer can often be forecasted with reasonable reliability. The forecasts are based on current stress data and other mechanical information and made easier since the fracture is (supposed to be) confined to a single mechanical unit (Valko and Economides, 1995; Yew and Weng, 2014; Shapiro, 2018). In unconventional vertical propagation of hydraulic fractures, such as used in the extraction of gas from shale, the fractures propagate through layered and faulted rocks. Then the forecasted paths are much less reliable, with the fractures commonly becoming deflected laterally along contacts to form water sills (Fisher and Warpinski, 2011; Fisher, 2014) or deflected into faults (Davis et al., 2012).

More specifically, the propagation paths of thousands of vertically injected hydraulic fractures have been studied through microseismicity. Many hydraulic fractures have been observed to deflect into subhorizontal contacts, while some have deflected into faults and use them as parts of their paths (Fisher and Warpinski, 2011; Davis et al., 2012; Flewelling et al., 2013; Lacazette and Geiser, 2013). Deflection of vertical hydraulic fractures into contacts to form water sills is particularly common at crustal depth of less than 700-800 m (Fisher, 2014). The general conditions for hydrofracture arrest have also been studied; Forbes Inskip et al. (2020) provide new numerical models on these conditions and a review of much of the literature.

As for natural hydrofractures such as igneous sheet-like intrusions, the details of no new paths of dikes (excluding multiple injections) or inclined sheets have ever been successfully forecasted (Gudmundsson, 2020). By this I mean that it has not been possible to forecast whether or not a new dike path (a dike injected during an unrest period) will become arrested, or reach the surface – and in the latter case, where and when it would reach the surface and how large the resulting volcanic fissure is likely to be. This is despite the fact that many dike (and some inclined sheet and sill) propagation paths during unrest periods in active volcanoes/volcanic zones have been well monitored using seismic methods. The results have indicated the overall propagation paths as well as crude estimates of the rate of dike propagation. Thus dike propagation in Iceland for horizontal distances (strike-dimensions) of many tens of kilometres have been recorded (Gudmundsson, 1995; Gudmundsson et al., 2014; Agustsdottir et al., 2016). Similarly, many dike-propagation events in the Manda Hararo-Dabbahu spreading centre in Africa from 2005-2010 were recorded seismically (Grandin et al., 2011). Dike-propagation events in the volcano Etna in Italy have also been recorded – some primarily from surface deformation of very shallow dikes (Falsaperla and Neri, 2015) and in Kilauea in Hawaii (Rutherford and Gardner, 2000).

Most of the above events have been interpreted as primarily related to lateral dike propagation. But some of the recorded dike-injections have been primarily through vertical propagation. These include some dike-propagation events in Piton de la Fournaise in Reunion, France, and, most recently, the dike-propagation events that eventually reached the surface to supply magma to the 2021 Fagradalsfjall eruption in Iceland (Hobe et al., 2022). Both the lateral and vertical dike-propagation events vary widely as regards rates, but are most commonly in the range of 0.01-1 m s^{-1}.

Arrested dikes and inclined sheets and the surface deformation they induce have been widely studied. Recent publications on these topics, summarising much of the previous research, are by Dzurisin (2006), Segall (2010), Kavanagh and Sparks (2011), Tibaldi (2015), Al Shehri and Gudmundsson (2018), by Bazargan and Gudmundsson (2019, 2020), and by Drymoni et al. (2020). These and related publications demonstrate that most observed segments of dikes and inclined sheets – and of hydrofactures in general (Gudmundsson and Brenner, 2002: Forbes Inskip et al., 2020) – become arrested at or just below contacts between mechanically

dissimilar rocks; in particular at contacts where the layer above the contact is considerably stiffer (with higher Young's modulus) than the layer below the contact. Hydraulic fractures are also commonly observed to arrest at contacts between dissimilar layers (Fisher and Warpinski, 2011; Fisher, 2014). Furthermore, that the fracture arrest can normally be attributed to one or more of the following factors: Cook-Gordon delamination, stress barrier, and elastic mismatch (Gudmundsson, 2011). There have also been many studies on the percentage of dikes occupying faults (Gudmundsson, 1983), as well as of dike-fault interactions (Rubin and Pollard, 1988; Rubin, 1995; Tibaldi, 2015; Drymoni et al., 2021).

While the above works have focused on various aspects of hydrofracture emplacement, arrest, and induced earthquakes and surface deformation, the physics of formation of hydrofracture paths has received comparatively little attention. There have been discussions in the fracture mechanics literature on dynamic fracture propagation (Freund, 1998: Ravi-Chandar, 2004; Shulka, 2006), but the results apply primarily to propagation in homogeneous, isotropic materials, are not based on Hamilton's principle, and are more appropriate for earthquake rupture as regards velocities (even if these works focus much on mode I cracks) than the comparatively slow propagation of fluid-driven fractures in layered rock bodies.

In the present paper a new theoretical framework is provided for forecast the likely propagation paths of hydrofractures propagating through layered and faulted rocks. At this stage, the framework does not apply to shear fractures such as faults. With further development and proper adjustment of the boundary conditions, I think it will be possible to include shear-fracture propagation paths in this theoretical framework in the future. The proposal is based on the suggestion that hydrofracture paths are controlled by Hamilton's principle of least action.

Based on Hamilton's principle, it is suggested that of the theoretically infinite number of possible paths that a given hydrofracture may follow, the entire fracture (and each fracture segment) selects the path of least (minimum) action. More specifically, the path chosen is the one along which the variation of the action is zero. Action has the dimensions of energy × time and the unit of Js (joule-second). For an elastic solid, such as a host-rock body, action (S) is the kinetic energy (T) minus the potential energy due to generalised external forces (V) and minus the strain energy due to internal forces (U), or $S = T - V -$ U. The right-hand side of this equation is also referred to as the Lagrangian (the difference between the kinetic and the total potential energy). Thus, in the theoretical framework proposed here, which applies to conservative systems, the path taken by a hydrofracture is normally the one along which the energy transformed (released) multiplied by the time taken for the propagation is the least (is a minimum).

For the normally slow-propagating hydrofracture – with typical rates for dikes of 0.01-1 m s^{-1} in comparison with seismic ruptures whose rates are or the order of km s^{-1} - the kinetic energy may with some justification be assumed zero. On that

assumption, Hamilton's principle of least action reduces to the principle of minimum potential energy (or least work). This latter principle states that of all the possible displacement fields (configurations) of an elastic body that satisfy the constraint conditions and the external and internal forces, the actual (true) displacements are those that make the total potential energy of the body a minimum. This principle applies only to (linear and non-linear) elastic bodies, and was postulated as a general principle for dike propagation by Gudmundsson (1986).

A quantitative comparison between the paths forecasted here and the actual observed paths must, at this stage, be of a general (statistical) nature. Neither the present method nor other methods developed over the years have so far been used to make a reliable prediction as to the details of dike-propagation path during actual volcano unrest and dike injection. Such a prediction is the next step – the next development – of the theories and observations presented here and should, at first, be applied to volcanoes/volcanic zones where the layering and faulting is well known, such as from exposures in caldera walls, cliffs, and drill holes. Unconventional (vertical) hydraulic fracturing, where drilling provides information on the mechanical layering, can also be used to test the present theories of hydrofracture propagation paths. While rigorous testing of the forecasts has thus not be possible so far, the general path predictions are supported by direct field observations, both as regards the paths of dikes (Figs 1, 6, 8, 13, 14 and 18) and the paths of mineral veins (Figs 4, 5, 9, 11, 12).

8. Conclusions

The main conclusions of the paper may be summarised as follows.

- The principal aim of the paper is to explain, and provide a theoretical framework to forecast, the paths of fluid-driven fracture, that is, hydrofractures when propagating through layered, heterogeneous, and faulted crustal segments. The term hydrofracture includes all dikes, inclined sheets, sills, most mineral veins and many joints, as well as human-made hydraulic fractures. For the first time, Hamilton's principle of least action is here applied to hydrofractures with a view of forecasting their likely paths through layered rocks. Additionally, using energy considerations a quantitative estimate is made of the potential of existing faults acting as parts of hydrofracture paths.
- It is proposed that of the theoretically infinite number of possible paths that a given hydrofracture may follow, it selects the path of least (minimum) action as determined by Hamilton's principle. The path chosen is thus the one along which the variation of the action is zero, which means that the selected path is the one where the energy transformed (released) multiplied by the time taken for the propagation is the least (is a minimum).

- For many hydrofractures the kinetic energy may be taken as zero, in which case Hamilton's principle of least action reduces to the principle of minimum potential energy. This latter principle states that of all the possible displacement fields (configurations) of an elastic body, such as a rock body or a crustal segment, that satisfy the constraint conditions and the external and internal forces, the actual (true) displacements are those that make the total potential energy of the body a minimum.
- Hydrofractures advance their tips/fronts in steps, with a time lag between the fracture front and the fluid front. Each vertical step may be similar in length to the thickness of the layers, such as lava flows (for dikes) or shale/limestone layers (for hydrofractures in sedimentary basins) through which the hydrofracture propagates.
- In the present theoretical framework, each propagation step of a hydrofracture is controlled by Hamilton's principle or, in case the kinetic energy is omitted, by the principle of minimum potential energy. The energy needed to advance the dike-fracture is the surface energy, whereas the energy transformed or released during the hydrofracture propagation is part of the total potential energy (potential energy due to external forces and strain energy due to internal forces).
- When applied to dikes, some of the main conclusions of the new framework are as follows. When a crustal segment is regarded as homogeneous, isotropic and non-fractured dike-propagation paths are everywhere perpendicular to the trajectories of the minimum compressive (maximum tensile) principal stress σ_3 and thus follow the trajectories of the maximum principal compressive stress σ_1.
- For a faulted and/or a layered (anisotropic) crustal segment or rock body, the dike-propagation paths can locally follow existing faults for a while, and thus, in that part, be oblique to the trajectories of the maximum principal compressive stress σ_1 (and oblique to σ_3). Whether the dike uses a fault as a part of its path depends primarily on (1) the dip of the fault (steep faults are the most likely to be used), and (2) the tensile strength across the fault as compared with the tensile strength of the host rock along a path following the direction of σ_1. The results suggest that dikes use faults as parts of their paths primarily if the fault is a comparatively steeply dipping, preferably a normal fault or currently in an extensional regime, and with close to zero tensile strength.

Acknowledgements. The results presented here are based on work over many years that has been supported by several funding agencies such as the Icelandic Science Foundation, the Research Council of Norway, the European Commission, and the Natural Environment Research Council of the United Kingdom. I thank the journal reviewers for very helpful comments.

References

Agustsdottir, T., Woods, J., Greenfield, T., Green, R.G., White, R.S., Winder, T., Brandsdottir, B., Steinthorsson, S., Soosalu, H., 2016. Strike-slip faulting during the 2014 Bardarbunga-Holuhraun dike intrusion, central Iceland. *Geophysical Research Letters*, **43**, 1495-1503, doi: 10.1002/2015GL067423.

Al Shehri, A., Gudmundsson, A., 2018. Modelling of surface stresses and fracturing during dyke emplacement: Application to the 2009 episode at Harrat Lunayyir, Saudi Arabia. *Journal of Volcanology and Geothermal Research*, **356**, 278-303.

Amadei, B., Stephansson, O., 1997. *Rock Stress and its Measurements*. London: Chapman and Hall.

Anderson, E.M., 1942. *The Dynamics of Faulting and Dyke Formation with Application to Britain*. London: Oliver and Boyd.

Anderson, T. L., 2005. *Fracture Mechanics: Fundamentals and Applications*, 3rd ed. London: Taylor & Francis.

Bazargan, M., Gudmundsson, A., 2019. Dike-induced stresses and displacements in layered volcanic zones. *Journal of Volcanology and Geothermal Research*, **384**, 189-205.

Bazargan, M., Gudmundsson, A., 2020. Stresses and displacements in layered rocks induced by inclined (cone) sheets. *Journal of Volcanology and Geothermal Research*, **401**, doi.org/10.1016/j.jvolgeores.2020.106965.

Bedford, A., 1985. *Hamilton's Principle in Continuum Mechanics*. London: Pitman Publishing.

Bons, P.D., Elburg, M.A., Gomez-Rivas, E., 2012. A review of the formation of tectonic veins and their microstructures. *Journal of Structural Geology*, 43, 33-62.

Brebbia, C.A., Dominguez, J., 1992. *The Boundary-Element Analysis System User Guide*. Boston, Mass: Beasy International.

Cassel, K.W., 2013. *Variational Methods with Application in Science and Engineering*. Cambridge: Cambridge University Press.

Cobbold, P.R., Rodrigues, N., 2007. Seepage forces, important factors in the formation of horizontal hydraulic fractures and bedding-parallel fibrous veins ('beef' and 'cone-in-cone'). *Geofluids*, **7**, 313-322**.**

Davis, R.J., Mathias, S.A., Moss, J., Hustoft, S., Newport, L., 2012. Hydraulic fractures: how far can they go? *Marine and Petroleum Geology*, **37**, 1-6.

Dennis, J.G., 1972. *Structural Geology*. New York: Ronald Press.

Drymoni, K., Browning, J., Gudmundsson, A., 2020. Dyke-arrest scenarios in extensional regimes: Insights from field observations and numerical models, Santorini, Greece. *Journal of Volcanology and Geothermal Research*, **396**, doi: 10.1016/j.jvolgeores.2020.106854.

Drymoni, K., Browning, J., Gudmundsson, A., 2021. Volcanotectonic interactions between inclined sheets, dykes, and faults at the Santorini Volcano, Greece. *Journal of Volcanology and Geothermal Research*, **416**, doi.org/10.1016/j.jvolgeores.2021.107294

Dzurisin, D., 2006. *Volcano Deformation: New Geodetic Monitoring Techniques*. Berlin: Springer Verlag.

Fall, A., Eichhubl, P., Bodnar, R.J., Laubach, S.E., Davis, J.S., 2015. Natural hydraulic fracturing of tight-gas sandstone reservoirs, Piceance Basin, Colorado. *Geological Society of America Bulletin*, 127, 61-75.

Falsaperla, S., Neri, M., 2015. Seismic footprints of shallow dyke propagation at Etna, Italy. *Scientific Reports* **5**, 11908; doi: 10.1038/srep11908.

Fast, R. E., Murer A. S., Timmer R. S., 1994. Description and analysis of cored hydraulic fractures - Lost Hills field, Kern County, California. *SPE Production & Facilities,* 9, 107–114.

Fisher, K., 2014. *Hydraulic fracture growth: real data*. Presentation given at GTW-AAPG/STGS Eagle Ford plus Adjacent Plays and Extensions Workshop, San Antonio, Texas, February 24-26, 2014.

Fisher, K., Warpinski, N., 2011. *Hydraulic fracture-height growth: real data*. Society of Petroleum Engineers Annual Technical Conference and Exhibition, SPE 145949.

Flewelling, S.A., Tymchak, M.P., Warpinski, N., 2013. Hydraulic fracture height limits and fault interactions in tight oil and gas formations. *Geophysical Research Letters*, **40**, 3602-3606.

Forbes Inskip, N.D., Browning, J., Meredith, P.G., Gudmundsson, A., 2020. Conditions for fracture arrest in layered rock sequences. *Results in Geophysical Sciences*, **1-4,** doi.org/10.1016/j.ringps.2020.100001

Freund, L.B., 1998. *Dynamic Fracture Mechanics*. Cambridge: Cambridge University Press.

Fung, Y. C., Tong, P., 2001. *Classical and Computational Solid Mechanics*. Singapore: World Scientific Publishing.

Gale, J.F.W., Elliott, S.J., Li, J.Z, Laubach, S.E., 2019. Natural fracture characterization in the Wolfcamp Formation at the Hydraulic Fracture Test Site (HFTS), Midland basin, Texas. *SPE/AAPG/SEG Unconventional Resources Technology Conference*, URTEC-644. doi: 10.15530/urtec-2019-644.

Gale, J.F.W., Elliott, S.J., Rysak, B.G., Ginn, C.L., Zhang, N., Myers, R.D., Laubach, S.E., 2021. Fracture description of the HFTS-2 slant core, Delaware Basin, West Texas. In: *SPE/AAPG/SEG Unconventional Resources Technology Conference*. OnePetro. doi.org/10.15530/urtec-2021-5175

Galindo, I., Gudmundsson, A., 2012. Basaltic feeder dykes in rift zones: geometry, emplacement, and effusion rates. *Natural Hazards and Earth System Sciences,* **12,** 3683–3700.

Geshi, N., Kusumoto, S., Gudmundsson, A., 2010. Geometric difference between non-feeder and feeder dikes. *Geology*, **38**, 195–198.

Geshi, N., Kusumoto, S., Gudmundsson, A., 2012. Effects of mechanical layering of host rocks on dike growth and arrest. *Journal of Volcanology and Geothermal Research*, **223-224**, 74-82.

Geshi, N., Neri, M., 2014. Dynamic feeder dyke systems in basaltic volcanoes: the exceptional example of the 1809 Etna eruption (Italy). *Frontiers in Earth Science*, **2**, doi: 10.3389/feart.2014.00013.

Grandin, R., Jacques, E., Nercessian, A. et al., 2011. Seismicity during lateral dike propagation: Insights from new data in the recent Manda Hararo–Dabbahu

rifting episode (Afar, Ethiopia). *Geochemistry, Geophysics, Geosystems*, **12**, doi:10.1029/2010GC003434.

Griggs, D.T., Handin, J., 1960. Observations on fracture and an hypothesis of earthquakes. In: *Rock Deformation*. Geological Society of America Memoir 79, 39-104.

Gudmundsson, A. 1983. Form and dimensions of dykes in eastern Iceland. *Tectonophysics*, **95**, 295-307.

Gudmundsson, A. 1986. Formation of dykes, feeder-dykes and the intrusion of dykes from magma chambers. *Bulletin of Volcanology*, **47**, 537-550.

Gudmundsson, A. 1988. Effect of tensile stress concentration around magma chambers on intrusion and extrusion frequencies. *Journal of Volcanology and Geothermal Research,* 35, 179-194.

Gudmundsson, A., 1995. The geometry and growth of dykes. In: *Physics and Chemistry of Dykes* (edited by Baer, G. & Heimann, A.). Rotterdam: Balkema, 23-34.

Gudmundsson, A., 2011. *Rock Fractures in Geological Processes*. Cambridge: Cambridge University Press.

Gudmundsson, A., 2020. *Volcanotectonics: Understanding the Structure, Deformation and Dynamics of Volcanoes*. Cambridge: Cambridge University Press.

Gudmundsson, A., Brenner, S.L., 2001. How hydrofractures become arrested. *Terra Nova*, **13**, 456-462.

Gudmundsson, A., Berg, S.S., Lyslo, K.B., Skurtveit, E., 2001. Fracture networks and fluid transport in active fault zones. *Journal of Structural Geology*, **23,** 343-353.

Gudmundsson, A., Fjeldskaar, I., Brenner, S., 2002. Propagation pathways and fluid transport of hydrofractures in jointed and layered rocks in geothermal fields. *Journal of Volcanology and Geothermal Research*, 116, 257-278.

Gudmundsson, A., Lecoeur, N., Mohajeri, N., Thordarson, T., 2014. Dike emplacement at Bardarbunga, Iceland, induces unusual stress changes, caldera deformation, and earthquakes. *Bulletin of Volcanology*, **76,** doi: 10.1007/s00445-014-0869-8.

Hillis, R.R., 2003. Pore pressure/stress coupling and its mplications for rock failure: In: *Subsurface Sediment Mobilization* (Van Rensbergen, P., Hillis, R.R., Maltman, A.J., and Morley, C.K., eds). Geological Society of London Special Publication, **216**, 359-368.

Howard, G. C., Fast, C. R., 1970. *Hydraulic Fracturing*. New York: Society of Petroleum Engineers of AIME.

Hobe, A., Bazargan, M., Selek, B., Tryggvason, A., SIL Seismological Group, Gudmundsson, A., 2022. Quantifying the volcanotectonic parameters of the 2021 Fagradalsfjall eruption, Iceland. *Nature Communications* (in review).

Jaeger, J.C., Cook, N.G.W., 1979. *Fundamentals of Rock Mechanics, 3rd ed*. London, Chapman and Hall.

Jakobsdottir, S., 2008. Seismicity in Iceland: 1994-2007. *Jokull. The Icelandic Journal of Earth Sciences*, 58, 75-100.

Kavanagh, J.L., Sparks, R.S.J., 2011. Insights of dyke emplacement mechanics from detailed 3D dyke thickness datasets. *Journal of the Geological Society of London,* **168**, 965-978.

Kusumoto, S., Geshi, N., Gudmundsson, A., 2013. Inverse modeling for estimating fluid-overpressure distributions and stress intensity factors from arbitrary open-fracture geometry. *Journal of Structural Geology*, **46**, 92-98.

Kusumoto, S., Gudmundsson, A., 2014. Displacement and stress fields around rock fractures opened by irregular overpressure variations. *Frontiers in Earth Science*, **2**, doi: 10.3389/feart.2014.00007.

Lacazette, A., Geiser, P., 2013. Comment on Davis et al., 2012 - Hydraulic fractures: how far will they go? *Marine and Petroleum Geology*, **43**, 517-519.

Laubach, S.E., Lander, R.H., Criscenti, L.J., Anovitz, L.M., Urai, J.L., Pollyea, R.M. et al., 2019. The role of chemistry in fracture pattern development and opportunities to advance interpretations of geological materials. *Reviews of Geophysics*, 57, 1065-1111.

Logan, D.L., 2002. *A First Course in the Finite Element Method*. Pacific Grove, California: Brooks/Cole.

Mahrer, K.D., 1999. A review and perspective on far-field hydraulic fracture geometry studies. *Journal of Petroleum Science and Engineering*, 24, 13-28.

Peltier, A., Ferrazzini, V., Staudacher, T., Bachelery, P., 2005. Imaging the dynamics of dyke propagation prior to the 2000-2003 flank eruptions at Piton de la Fournaise, Reunion Islands. *Geophysical Research Letters*, **32**, doi:10.1029/2005GL023720.

Philipp, S.L., 2008. Geometry and formation of gypsum veins in mudstones at Watchet, Somerset, SW England. *Geological Magazine*, **145**, 831-844.

Philipp, S.L., 2012. Fluid overpressure estimates from the aspect ratios of mineral veins. *Tectonophysics*, **581**, 35-47.

Pollard, D.D., Aydin, A., 1988. Progress in understanding jointing over the past century. *Geological Society of America Bulletin,* 100, 1181-1204.

Pollard, D.D., Fletcher, R.C., 2005. *Fundamentals of Structural Geology*. Cambridge: Cambridge University Press.

Ravi-Chandar, K., 2004. *Dynamic Fracture*. Amsterdam: Elsevier.

Reddy, J.N., 2002. *Energy Principles and Variational Methods in Applied Mechanics*, 2nd ed. Hoboken, New Jersey: Wiley.

Richards, T.H., 1977. *Energy Methods in Stress Analysis*. Chichester: Ellis Horwood.

Rivalta, E., Taisne, B., Bunger, A.P., Katz, R.F., 2015. A review of mechanical models of dike propagation: schools of thought, results and future directions. *Tectonophysics*, **638**, 1-42.

Rubin, A.M., 1995. Propagation of magma-filled cracks. *Annual Reviews of Earth and Planetary Sciences*, **23**, 287-336.

Rubin, A.M., Pollard, D.D., 1988. Dike-induced faulting in rift zones of Iceland and Afar. *Geology*, **16**, 413-417.

Rutherford, M.J., Gardner, J.E., 2000. Rates of magma ascent. In: *Encyclopedia of Volcanoes* (edited by Sigurdsson, H. et al.). Academic Press, 207-217.

Sanford, R.J., 2003. *Principles of Fracture Mechanics*. Upper Saddle River, New Jersey: Prentice-Hall.

Schultz, R.A., 2019. *Geologic Fracture Mechanics*. Cambridge: Cambridge University Press.

Segall, P., 1984. Formation and growth of extensional fracture sets. *Geological Society of America Bulletin,* 95, 454-462.

Segall, P., 2010. *Earthquake and Volcano Deformation*. Princeton: Princeton University Press.

Shapiro, S.A., 2018. *Fluid-induced Seismicity*. Cambridge: Cambridge University Press.

Shulka, A (ed.), 2006. *Dynamic Fracture Mechanics*. London: World Scientific.

Tauchert, T.R., 1981. *Energy Principles in Structural Mechanics*. Malabar (Florida): Krieger.

Theobald, D.W. 1966. *The Concept of Energy*. London: Spon.

Tibaldi, A., 2015. Structure of volcano plumbing systems: A review of multi-parametric effects. *Journal of Volcanology and Geothermal Research*, **298,** 85–135.

Twiss, R.J., Moores, E.M., 1992. *Structural Geology*. New York: W.H. Freeman.

Valko, P., Economides, M.J., 1995. *Hydraulic Fracture Mechanics*. New York: Wiley.

Warpinski, N. R., and Teufel, L. W., 1987. Influence of geologic discontinuities on hydraulic fracture propagation. *Journal of Petroleum Technology*, 39, 209–220.

Warpinski, N. R., Lorenz, J. C., Branagan, P. T., Myal, F. R., and Gall, B. L., 1993. Examination of a cored hydraulic fracture in a deep gas well. *SPE Production & Facilities,* 8*,* 150-158.

Washizu, K., 1975. *Variational Methods in Elasticity and Plasticity*. Amsterdam: Elsevier.

Wu, Y.S. (editor), 2017. *Hydraulic Fracture Modeling*. Houston: Gulf Publishing.

Yew, C.H., Weng, X., 2014. *Mechanics of Hydraulic Fracturing,* 2nd ed. Houston: Gulf Publishing.

Zang, A., Stephansson, O., 2014. *Stress Field of the Earth's Crust*. Berlin: Springer Verlag.

Zoback, M.D., 2007. *Reservoir Geomechanics*. Cambridge: Cambridge University Press.