A review of shear-wave splitting in the compliant crack-critical anisotropic Earth

Stuart Crampin a,b,*, Sheila Peacock c

a Shear-Wave Analysis Group, School of GeoSciences, University of Edinburgh, Grant Institute, West Mains Road, Edinburgh EH9 3JW, Scotland, UK
b Edinburgh Anisotropy Project, British Geological Survey, Murchison House, West Mains Road, Edinburgh EH9 3LA, Scotland, UK
c School of Earth Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Received 1 March 2004; accepted 17 May 2004
Available online 20 July 2004

Abstract

Shear-wave splitting due to stress-aligned anisotropy is widely observed in the Earth’s crust and upper mantle. The anisotropy is the result of stress-aligned fluid-saturated grain-boundary cracks and pore throats in almost all crustal rocks, and we suggest by stress-aligned grain-boundary films of liquid melt in the uppermost 400 km of the mantle. The evolution of such fluid-saturated microcracks under changing conditions can be modelled by anisotropic poro-elasticity (APE). Numerical modelling with APE approximately matches a huge range of phenomena, including the evolution of shear-wave splitting during earthquake preparation, and enhanced oil recovery operations. APE assumes, and recent observations of shear-wave splitting confirm, that the fluid-saturated cracks in the crust and (probably) upper mantle are so closely spaced that the cracked rocks are highly compliant critical systems with self-organised criticality. Several observations of shear-wave splitting show temporal variation displaying extreme sensitivity to small stress changes, confirming the crack-critical system. Criticality has severe implications for many Solid Earth applications, including the repeatability of seismic determinations of fluid flow regimes in time-lapse monitoring of hydrocarbon production. Analysis of anisotropy-induced shear-wave splitting is thus providing otherwise unobtainable information about deformation of the inaccessible deep interior of the Earth.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Crack-critical systems; Earth’s crust and upper mantle; Seismic anisotropy; Seismic birefringence; Shear-wave splitting

1. Introduction

In 1981, one of us published in Wave Motion [1] a review of theoretical aspects of wave propagation in anisotropic and cracked elastic-media pertaining particularly to the assumed anisotropy of the Earth’s crust and upper-mantle. This present review reports the progress of over 20 years study of shear-wave splitting in an anisotropic Earth. Stress-aligned anisotropy-induced seismic shear-wave splitting (seismic birefringence or double refraction), first positively identified in 1980 in both the Earth’s crust [2] and upper mantle [3], has now been widely observed and...
investigated. There are several thousand papers on the theory and observation of shear-wave splitting in the Earth, and shear-wave splitting is an anticipated phenomenon in almost all crustal and upper-mantle rocks. It has been inferred that the crust and probably the uppermost 400 km of the mantle are critical systems of closely spaced fluid-saturated microcracks with great sensitivity to changing conditions. This means that changes in shear-wave splitting directly monitor changes in rock mass deformation. We suggest that the ability to monitor low-level deformation with shear-wave splitting is probably the most fundamental advance in solid Earth geophysics for several decades [4–7]. Thus monitoring shear-wave propagation in anisotropic rocks is opening a new window into our understanding of rock deformation and the evolution of crustal and mantle properties.

1.1. Background of seismic anisotropy in geoscience

The first confirmed observations of seismic anisotropy in the Earth were measurements of azimuthal velocity variations of $P$-waves refracted in oceanic basins recognised by Hess in 1964 [8]. Hess interpreted these observations in terms of wave propagation through olivine crystals, a common mantle material, aligned by the movement of the mantle away from spreading centres. $P$-waves are sensitive to a wide variety of phenomena and identifiable observations of $P$-wave velocity anisotropy require accurate measurements of travel times over a wide range of directions in a uniform homogeneous rock mass. Except for oceanic basins, large blocks of uniform homogeneous rock are very uncommon in the Earth and there have been few studies where $P$-wave anisotropy has been reliably identified.

The next observations of seismic anisotropy, in 1966, were measurements of systematic anomalies in the particle motion of seismic surface waves [9], particularly the Second Rayleigh and Second Love Modes, which may become wholly or partially coupled in layered anisotropic solids. These modes have most of their energy (and most of the coupling) in the top few km of the upper mantle, immediately below the Mohorovičić discontinuity at around 33 km depth [10]. Theoretical modelling [10–12] showed that the observed particle motions and cross-mode coupling could be caused by as little as 4 km of 4% shear-wave velocity anisotropy at the top of the mantle. At that time, the cause of the anisotropy was not specified. A severe problem in analysing and interpreting such surface-wave anisotropy is the difficulty in determining the depth of the anisotropy, and there have again been limited further studies.

Wave propagation through plane-layered anisotropic models can be calculated by setting up $6 \times 6$ propagator matrix equations [13] matching the properties across interfaces. A major advance was to always calculating propagation in the $x$-plane with the $z$-plane horizontal [11], so that differences in direction, or changes of anisotropic alignment, could be accommodated merely by rotating the anisotropic matrices. This great convenience is obvious now, but previous use of Christoffel transformations made it less obvious at the time [11].

The breakthrough in studies of anisotropy in the Earth came when theoretical studies [1,13–16] demonstrated that stress-aligned shear-wave splitting was the most diagnostic phenomenon for recognising anisotropy in cracked rocks, and such stress-aligned shear-wave splitting was commonly observed in the Earth [2,3,17,18]. Shear-wave splitting, analogous to the birefringence of light in crystals, occurs when shear waves enter a system of aligned cracks. The waves are split into two approximately orthogonal polarisations which travel at different velocities and write characteristic easily recognisable signatures into three-component seismograms. Fig. 1 shows a schematic illustration of the phenomenon. The cracks tend to be aligned perpendicular to the direction of minimum compressional stress, just like fractures induced by the injection of high-pressure fluid in the oil industry. Since below the critical depth, where the increasing vertical stress equals the minimum horizontal stress (usually between 500 and 1000 m), the minimum stress is horizontal and cracks tend to be vertical striking parallel to the maximum horizontal stress. This means that orientations of the polarisations generally provide information about the anisotropic symmetry and stress direction, whereas time-delays between the split shear-waves provide information about the average crack density along the ray paths. Such stress-aligned crack-induced shear-wave splitting is now widely observed in the Earth’s crust [18–21] and upper mantle [22]. Shear-wave splitting is recognised in seismograms in igneous and metamorphic rocks above small earthquakes [19,23–25], and in seismic reflection surveys and vertical seismic profiles in sedimentary hydrocarbon reservoirs [17–21].
Fig. 1. Schematic illustration of shear-wave splitting in distributions of stress-aligned fluid-saturated parallel vertical microcracks aligned normal to the direction of minimum horizontal stress, $\sigma_h$. For nearly vertical propagation the polarisation of the faster split shear-wave is parallel to the strike of the cracks, parallel to the direction of maximum horizontal stress, $\sigma_H$. Such parallel vertical crack orientations are typically found below the critical depth (usually between 500 and 1000 m), where the vertical stress, $\sigma_V$, becomes greater than $\sigma_h$.

Note that such stress-aligned anisotropic symmetry is close, but not usually identical, to hexagonal symmetry (transverse isotropy) with a horizontal axis of symmetry. As a result, it is sometimes, misleadingly, referred to as TIH-anisotropy. In practice, there is always a distribution of orientations leading (usually) to greater or smaller orthorhombic perturbation of the hexagonal TIH-anisotropy.

1.2. Other causes of anisotropy in crustal rocks

In contrast to TIH-anisotropy, TIV-anisotropy is transverse isotropy with a vertical axis of symmetry, where the shear-waves split into strictly SH- and SV-polarisations with no azimuthal variations [26]. Such symmetry is characteristic of finely layered horizontal sedimentary strata, where the anisotropy is caused by the interactions of reflections and transmissions through thin isotropic layers. TIV-anisotropy is also characteristic of many shales and clays, where the anisotropy is caused by horizontal platelets of mica and other minerals. Such TIV-anisotropy has vertical and horizontal velocities that may differ by up to 40% which causes severe problems in seismic processing, and reconciling seismically derived depths with true depths at boreholes in the hydrocarbon industry. These technical problems can usually be accommodated by processing [27], although, like almost all examples of azimuthal anisotropy, the behaviour of shear-wave splitting in in situ rocks with TIV anisotropy has only recently been effectively calibrated [28].

Thus, TIV-anisotropy is comparatively well understood and probably has few surprises left. However, almost the only real geophysical information it carries is that the rock was laid down in some sort of sedimentary process in some sort of fluid, and that gravity is vertical. In general TIV-anisotropy is not stress-sensitive, and we shall not refer to TIV-anisotropy again.

Note that alignment of grains and intergranular pores not related to present-day stress also occurs in metamorphic rocks, particularly in slates, in which grains have been aligned by re-crystallisation under stress. This alignment can give rise to very high anisotropy [29,30] at arbitrary orientations determined by the tectonic history of the rock. Where ray paths cross zones of such strongly anisotropic rocks, shear-wave splitting due to aligned cracks elsewhere on the ray path may be severely disturbed [31].
1.3. What will be discussed

We shall discuss observations of stressed-aligned shear-wave splitting in the Earth’s crust. It can be shown that stress-aligned crack-induced shear-wave splitting is controlled by the incipient low-level deformation of rock before fracturing takes place (pre-fracturing deformation) [32,33]. The effect of stress on microcrack geometry, which shear-wave splitting images directly, depends on the rock properties, and in a complex heterogeneous Earth, may be difficult to deduce from a few determinations of the effective anisotropy. Potentially more important are time-lapse observations, where seismograms repeated with the same or similar source-to-receiver geometry can be differenced to show the effects of very small changes to crack geometry between recording sessions. Similar source-to-receiver geometry means similar ray paths, so that differences can often be interpreted directly in terms of the effects of the changing stress on the microcrack geometry along the ray paths.

Such time-lapse observations in the Earth have shown extraordinary sensitivity to near negligible changes. The reasons for this sensitivity is thought to be that the fluid-saturated cracks are critical systems verging on fracture-criticality with all the sensitivity, universality, and calculability that criticality implies [4,5,7,34–36]. This criticality means that there are spatial and temporal heterogeneities so the detailed behaviour and properties of much of the crust and upper mantle cannot be extrapolated from place to place or from time to time. The validity of any particular seismic measurement degrades with time so that many common procedures, such as averaging, are only possible in strictly limited circumstances. Consequently, the results of many conventional geophysical measurements, for example, reservoir characterisation by the oil industry, may be correct over much more limited temporal and spatial resolution than was previously thought.

2. Basic theory and observation of shear-wave splitting in cracked rocks

Garbin and Knopoff [37] were the first to calculate the behaviour of wave propagation through aligned (parallel) cracks. They did not recognise that aligned cracks would make a rock effectively anisotropic, and their analysis shows shear-wave polarisations varying smoothly with direction. Crampin [14] used the calculations of Garbin and Knopoff to calculate anisotropic elastic constants for wave propagation through parallel cracks. Hudson [16] in a more sophisticated analysis determined the anisotropic elastic moduli for both velocity variations and attenuation through distributions of parallel cracks. Hudson confirmed the elastic constants of Crampin [14].

Many other authors have made modifications and ‘improvements’ to the theory, but Hudson [16] is still the basic source for elastic constants of distributions of parallel cracks in effective media calculations. (Effective media are homogeneous solids with the same elastic properties as the cracked rock, that can be used as an approximation in mathematical modelling when the cracks are much smaller than the seismic wavelength.) The attenuation in Hudson [16] was derived from scattering. The effects of both primary and secondary scattering from neighbouring microcracks are minor and probably not important in many Earth structures.

These early developments were reviewed in Wave Motion in 1981 [1]. The underlying reason for those developments was that observations of three-component seismograms had shown evidence of shear-wave splitting [2,3], and the only possible source of anisotropy that satisfies all observations is fluid-saturated stress-aligned microcracks [19,38,39]. Almost all suitable three-component seismograms display shear-wave splitting in almost all in situ rocks. Fig. 2 shows the examples of shear-wave splitting at two stations in Iceland [24], and Fig. 3a shows the shear-wave alignments as rose diagrams of shear-wave polarisations observed above small earthquakes in Iceland [25]. The dominant direction of the diagrams is approximately NE to SW and is the direction of maximum tectonic compressional stress, the strike of the cracks in Fig. 1.

2.1. The shear-wave window

An absolute requirement for useful observations of shear-waves at a free surface of an isotropic half-space is that the shear waves must be recorded within the shear-wave window defined by angles of incidence less than $\theta$.
The recognition that the distributions of fluid-saturated microcracks in the Earth are so closely spaced that they are critical systems verging on fracture-criticality and self-organised criticality [6,7,19], so that shear-wave
The key observations were that shear-wave splitting in in situ rocks varies from a minimum of about 1.5% shear-wave velocity anisotropy (SWVA) to a maximum of about 4.5% in uniform unfractured rock [19]. The crack density, \( \varepsilon \), defined by \( \varepsilon = \frac{N a^3}{v} \), where \( N \) is the number of cracks of radius \( a \) in volume \( v \), is approximately equal to a hundredth of the percentage shear-wave velocity anisotropy (SWVA/100) [19]. Thus the dimensionless crack distributions in Fig. 4 represent the range of observed SWVA. Observed shear-wave splitting suggests that most rocks are close to fracture-criticality at about 5% SWVA broadly agreeing with Fig. 4 where at 5% SWVA cracks are so closely spaced that shear-strength is lost and fracturing and earthquakes can occur. It can be shown that the theoretical percolation threshold for

Fig. 3. (a) Normalised equal-area rose diagrams superimposed on equal-area polar plots (out to 45°) of the polarisations of the faster split shear-wave observed above small earthquakes at the Iceland Seismic Network (red triangles) in years 1996–2000. Glaciers and ice caps are shown in white. (b) Rose diagrams and polar plots for the Tjörnes Fracture Zone outlined in (a). Green petals are polarisations for earthquakes in years 1996–2000, and red petals are polarisations in 2001, when stations BRE, FLA, and HED were installed. Earthquake locations are shown for 2000 and 2001. Note rose diagrams in (a) and (b) are normalised for each station, and green and red rose diagrams have different normalisations (after [7]).

Fig. 4. Cross-sections of uniform distributions of parallel stress-aligned fluid-saturated microcracks representing shear-wave velocity anisotropy in percent, crack density \( \varepsilon \), and crack radius \( a \) (after [19]).
distributions of stress-aligned parallel cracks is about \( \varepsilon = 0.055 \), or 5.5% SWVA [33], for a Poisson’s ratio of 0.25. This suggests that if in situ rock is disturbed in any way, it is the geometry of the fluid-saturated microcracks that responds to the disturbance. A remarkable feature of the 1.5–4.5% range of SWVA is that it is observed, as expected from APE, to be independent of rock type (igneous, sedimentary, or metamorphic with the exception noted above), porosity, or tectonics. It appears to be a universal feature of almost all rocks. The only recognised exception is that areas of high heat-flow such as Iceland may have larger percentages of SWVA [24,25]. This is believed to be not due to higher crack density (implying a higher fracture-criticality limit) but to a higher temperature reducing the elastic modules of the crack-filling fluid, and hence increasing the shear-wave velocity anisotropy for a given crack density. Theoretically for given crack densities, levels of shear-wave velocity anisotropy are very sensitive to the elastic properties of the rock matrix and the pore fluid [41].

(2) The evolution of closely spaced compliant fluid-saturated grain-boundary cracks and flat pores and pore throats is so pervasive throughout most rocks that the behaviour can be modelled by anisotropic poro-elasticity (APE) [32,33]. The APE mechanism for low-level deformation is fluid migration by flow or diffusion along pressure gradients between neighbouring microcracks at different orientations to the stress-field. APE-modelling is fully 3D, but 3D behaviour is difficult to illustrate in 2D diagrams, and Fig. 5 shows the schematic illustration of the effects of such stress changes on a distribution of initially randomly oriented vertical fluid-saturated cracks. Note that the APE mechanism involves micro-scale (<2 mm, say) fluid movement around grain boundaries and does not involve any large-scale movement of fluids. Consequently, the response of the microcracks to changing conditions is almost immediate [33].

![Fig. 5. Schematic illustration of APE-modelling the evolution of crack aspect-ratios in an initially random distribution of vertical cracks (solid lines) for four values of increasing maximum horizontal differential stress, normalised to the critical value at which cracks first begin to close. Below the critical stress there is zero SWVA. Pore-fluid mass is preserved and aspect-ratios are chosen to give a porosity \( \phi \) = 5%. Please see [6] for detailed interpretation (after [34]).](image-url)
It can be shown [32,33] that APE prediction of the evolution of an initially random distribution of cracks when a deviatoric stress is applied is independent of porosity. Fig. 5 indicates that there is a minimum of about 1.5% shear-wave velocity anisotropy (Table 1) independent of porosity. This minimum is observed in the Earth (Fig. 4) with almost identical parameters in a vast range of sedimentary, igneous, and metamorphic rocks of different porosity, geology, and tectonic history [17–21]. APE is tightly constrained with no free parameters yet appears to approximately match a range of some 15–20 different phenomena and millions of individual observations [5,6,35]. Such uniformity and calculability were initially inexplicable until APE-modelling was developed.

The underlying reason for the universality is that the fluid-saturated microcrack distributions are so closely spaced that they are critical systems possessing self-organised criticality [4–6,36]. Table 1 summarises the match of APE-modelling to observations. The success of APE-modelling shear-wave anisotropy implies that all fluid-saturated microcracked rocks, of whatever origin or rock-type, can be modelled from a rock containing an initial distribution of randomly oriented fluid-saturated cracks with crack densities close to fracture-criticality. (Note that below a few metres depth, in situ cracks are subject to increasingly high pressures and temperatures in corrosive environments, so that direct physical calibration of APE is difficult.)

It is interesting to note that it is only shear-wave splitting that has this comparatively direct relationship with aligned crack distributions and anisotropy. Seismic P-waves are sensitive to so many phenomena that the effects of anisotropic velocity variations are difficult to observe and very difficult to isolate. Other electrical and

<table>
<thead>
<tr>
<th>Table 1: Summary of match of APE-modelling to observations (references in [5])</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static effects</td>
</tr>
<tr>
<td><strong>Field observations of SWVA</strong> (below 500–1000 m depth)</td>
</tr>
<tr>
<td>(1) SWVA in all rocks independent of porosity and geology</td>
</tr>
<tr>
<td>(2) Minimum SWVA in intact rock: observed ≈1.5%; APE-modelled ≈1.0%</td>
</tr>
<tr>
<td>(3) Maximum SWVA in intact rock: observed ≈4.5%; APE-modelled ≈5.5%</td>
</tr>
<tr>
<td>(4) Narrow range of observed crack density: 0.015 ≤ ϵ ≤ 0.045</td>
</tr>
<tr>
<td>(5) Proximity of intact rocks to fracture-criticality at crack density ϵ ≈ 0.055 (≈5.5% SWVA), where fracture-criticality can be associated with the percolation threshold</td>
</tr>
<tr>
<td>Other field observations:</td>
</tr>
<tr>
<td>(6) Fracture-criticality limit specifies crack distributions with a range of dimensions of about nine-orders of magnitude</td>
</tr>
<tr>
<td>(7) π/2 shear-wave polarisation changes (90°-dips) in over-pressured reservoirs</td>
</tr>
<tr>
<td>Dynamic effects</td>
</tr>
<tr>
<td>Temporal changes in SWVA during production procedures</td>
</tr>
<tr>
<td>(8) Changes between measurements made before and after hydraulic pumping tests</td>
</tr>
<tr>
<td>(9) Changes between measurements made before and after both high-pressure and low-pressure CO₂-injections in a carbonate reservoir</td>
</tr>
<tr>
<td>Temporal changes in SWTD before earthquakes</td>
</tr>
<tr>
<td>(10) Variations of time-delays (with hindsight) before some dozen earthquakes worldwide with magnitudes from M = 1.7–6</td>
</tr>
<tr>
<td>(11) Successful forecast of time and magnitude of an M = 5 earthquake in SW Iceland</td>
</tr>
<tr>
<td>Temporal changes in SWTD before volcanic eruption</td>
</tr>
<tr>
<td>(12) Variations in SWTD (over 5 months before 30 September 1996, Vatnajökull eruption, Iceland, at a range of azimuths and distances up to 240 km)</td>
</tr>
<tr>
<td>Variations of shear-wave properties in laboratory experiments</td>
</tr>
<tr>
<td>(13) Ultrasonic variations of SWVA and permeability in uniaxial stress cell</td>
</tr>
<tr>
<td>(14) Ultrasonic variations of (isotropic) shear-wave velocities to changes in confining pressure and pore-fluid pressure for oil-, water-, and gas- (dry) saturations in sandstone cores in stress cells</td>
</tr>
<tr>
<td>(15) Ultrasonic variations of velocity and attenuation from sonic (transducers) to seismic (resonant bar) frequencies</td>
</tr>
</tbody>
</table>

* Anisotropic poro-elasticity.
* Shear-wave velocity-anisotropy.
* Shear-wave time-delays.
electro-magnetic techniques are in principle sensitive to crack-induced anisotropy, but these techniques involve potential analysis that cannot resolve small-scale details of crack distributions. Consequently, seismic waves are the only direct way to image the critical distributions of microcracks in the interior of the Earth, and seismic interpretation of shear-wave splitting appears to be the only source of direct information about in situ anisotropy, and the pre-fracturing deformation. Critical systems and their implications for conventional geophysics are discussed more fully in Section 6.

3. The significance of changing crack aspect-ratios

Fig. 5 shows that the major effect of small increases of stress in the APE-model is to increase the aspect-ratio of cracks opening perpendicular to the direction of minimum compressional stress. Hudson’s analysis [15,16] shows that increasing crack aspect-ratios of fluid-saturated cracks changes the time-delays between split shear-waves in a specific range of azimuths and incidence angles [35]. Interpreting field observations before a large ($M_s = 6$) earthquake, Peacock et al. [42–44] first suggested that the observed temporal changes of time-delays in a range of stress-oriented directions could be explained by the accumulation of stress before an impending large earthquake enlarging crack aspect-ratios. The effects of stress on aspect-ratios was later confirmed with APE-modelling the evolution of fluid-saturated cracks as in Fig. 5 [32,33].

Changes in shear-wave splitting have only been recognised (with hindsight) before about a dozen earthquakes world-wide ranging in magnitude from $M_s = 1.7$–6 [25,35,45]. Observations of temporal changes at Station BJA in SW Iceland (Fig. 6) indicating increasing stress were used to stress-forecast in real time the time and magnitude of the $M_s = 5$ earthquake of 13 November 1998 3 days before the earthquake occurred [46]. Although shear-wave splitting provides no direct information about the impending epicentre, knowledge that a large earthquake is approaching allows other precursors to be interpreted more realistically. In the case of the successful stress-forecast, continuing
seismic activity on a nearby fault allowed the fault break to be identified before the earthquake occurred [46]. (Note that the explanation for the enormous scatter in shear-wave time-delays in Fig. 6 is given in Section 5.3.)

Such forecasting using small earthquakes as the source of shear-waves requires persistent seismicity, adequate recording and processing networks, and nearby large earthquakes [35]. Such circumstances are rare, and even if found, sufficiently persistent seismicity is unreliable. The largest recent earthquake in SW Iceland, the $M_s = 6.6$ earthquake of June 2000, was not forecast because the source seismicity was quiescent for the initial 2 months of the seven-month build up of stress. With no signals, the increase was not recognised, and the earthquake was not stress-forecast [25].

4. Significance of pore space in porous rocks: squirt flow, the source of frequency-dependent attenuation

Thomsen [47] recognised that velocity variations would be modified by pore-space where distributions of aligned cracks occur in media also containing grain-scale ‘equant’ pores. These so-called “dual porosity” media are common as oil reservoir rocks. He modelled velocity anisotropy in effective media with aligned cracks and ‘equant’ pore-space. Modifications to velocity anisotropy were comparatively minor but, in some circumstances, equant pore-space can make significant modifications to seismic attenuation [48].

One of the effects of pore space is to modify attenuation by squirt flow, where fluid is pumped in and out of thin pores and pore throats as seismic waves propagate through porous media. First identified for isotropic (random) crack distributions [49], squirt flow was analysed for distributions of aligned cracks by Pointer et al. [48]. The effects can be significant but are controlled by the ratio of wavelength to crack size. For sedimentary rocks, squirt flow is only likely to be measurable at the comparatively high frequencies of exploration seismics.

However, squirt flow could provide valuable information. A major difficulty in exploration seismics is estimating the size of the fractures causing the shear-wave splitting. This is because the cube of the crack size trades off against number density in the crack density, which controls velocity variations in media containing aligned cracks. In industrial seismology shear-wave splitting is frequently interpreted in terms of macro-fractures enabling oil flow and aiding hydrocarbon production. Although recognising that large fractures may affect shear-wave splitting [50,51], we suggest that stress-aligned fluid-saturated microcracks are the major source of shear-wave splitting [52]. One way to distinguish crack size is by measuring and interpreting frequency-dependent seismic attenuation, where squirt flow is (probably) the dominant mechanism. The basic equations have been developed [53] and modelling data has been attempted [54]. However, correct interpretation requires very well-recorded observations which are difficult to obtain, and there are no definitive results at the time of going to press.

5. Significance of high pore-fluid pressures and 90°-flips in shear-wave polarisations

In APE-modelling of the evolution of fluid-saturated cracks under changes of stress, when pore-fluid pressures are low, as is assumed for Fig. 1, the aligned cracks tend to be perpendicular to the direction of minimum compressional stress, and the faster split shear wave is polarised parallel to the direction of maximum horizontal stress. However, when pore-fluid pressures increase to critical levels, the effective stress-field realigns, microcrack distributions are modified, and the faster split shear-wave does a 90°-flip and becomes parallel to the direction of minimum horizontal stress [33]. Such 90°-flips are very distinctive phenomena, and faster split shear waves with polarisation perpendicular to the direction of maximum compressive stress (stress-perpendicular polarisations) have been observed in both exploration and earthquake seismology.

5.1. 90°-flips in exploration seismology

Stress-perpendicular polarisations caused by high pore-fluid pressures were first recognised in a vertical-seismic profile (VSP) in a reservoir in the Caucasus Oil Field [55]. The most definitive observation of 90°-flips has been
in fluid-injection experiments in Vacuum Field, New Mexico by the Reservoir Characterisation Project of the Colorado School of Mines [56]. There were both high-pressure and low-pressure CO₂-injections and the effects were recorded by time-lapse seismic reflection surveys. The principal seismic effects were variations in the travel times of the split shear-waves. Angerer et al. [50] modelled the behaviour by setting up a suitable initial fractured model and simulating the injections by inserting the appropriate CO₂-presures into the APE-modelling procedure.

The effects on the shear-waves of both high- and low-pressure injections were modelled almost perfectly. Only the high-pressure results are shown in Fig. 7. Before injection (Fig. 7a), the faster split shear-wave is polarised in the S1-direction. After the high-pressure injection (Fig. 7b), the faster split shear-wave has done a 90°-flip and is parallel to the previous S2-direction. These results are the best in situ calibration of APE-modelling to date.

5.2. 90°-flips above large seismically active faults

Fig. 3b shows rose diagrams for the shear-wave polarisations at seismic stations in northern Iceland. As in Fig. 3a, rose diagrams at most stations are approximately NE to SW, imaging the direction of maximum horizontal stress. Three stations on the Húsavík-Flatey Fault, however, show fault-parallel rose diagrams approximately orthogonal to the polarisations elsewhere. The Húsavík-Flatey Fault is a major transform fault of the Mid-Atlantic Ridge which runs onshore in northern Iceland. Similar polarisation anomalies have also been observed in two places above another major fault, the San Andreas Fault in California [42,57]. Modelling the effects with APE, Crampin et al. [58,59] shows that the most likely explanation is that the critically high pore-fluid pressures on all seismically active fault planes cause shear-wave polarisations perpendicular to the direction of maximum compressive stress. If normally pressured paths result in positive time-delays, highly pressured paths cause stress-perpendicular first polarisations and hence negative time-delays. At major faults such as the Húsavík-Flatey and the San Andreas, which extend through a large part of the crust, the normally pressured segment of the ray path above the fault is too short to reverse 90°-flips in the fault zone and the stress-perpendicular 90°-flips are observed at the surface.

It has also been suggested [60] that the seemingly stress-perpendicular shear-wave polarisations, observed at the surface above very shallow earthquakes in João Câmaras, northeast Brazil [61], are also pressure-induced 90°-flips.

5.3. 90°-flips above small seismically active faults and scatter of shear-wave time-delays

Almost all measurements of time-delays between split shear-waves above small earthquakes display a large ±80% scatter, as in Fig. 6. In contrast, time-delays in shear-wave splitting in reflection profiles and VSPs in exploration seismics away from seismically active faults show negligible scatter, so the scatter appears to be an earthquake-related phenomenon. This scatter has been impossible to explain by conventional geophysics [24].

Section 5.2 suggests that there are critically high pore-fluid pressures on large seismically active faults in order to relieve friction and overcome asperities. We suggest [59] that such pore-fluid pressures exist on all seismically active faults so that stress-perpendicular polarisations must be expected in the vicinity of all earthquake faults. Measures of shear-wave splitting above earthquakes are typically from small earthquakes rupturing small faults, or small segments of larger faults. Consequently, the high pore-fluid pressures surrounding faults are limited in extent, so that stress-perpendicular polarisations only occur in a limited zone around the source.

In most circumstances, the remainder of the ray path to the surface recorder through normally pressured rock is longer than the critically high-pressured path so that polarisations observed at the free surface are parallel to the regional stress direction. Each earthquake releases stress and changes the stress field, and it has been shown that small temporal variations in the ratio of normally pressured path to high-pressured path, causing 90°-flips along varying lengths of the path, can easily cause the ±80% scatter in observed time-delays [59]. Thus what appears to be ±80% scatter is not an indication of measurement inaccuracies. It is actually a consequence of the sensitivity of the critical system to small changes of stress and pressure following every earthquake.
Although it has long been recognised that seismically active faults must be pervaded by high pore-fluid pressures in order to permit slip without frictional heating [62–65], in the past there has been no way to demonstrate their presence. One of the major objectives of the current SAFOD (San Andreas Fault Observatory at Depth) plans to drill into the San Andreas Fault is to directly measure pore-fluid pressures on an active fault [66]. Thus the ±80% scatter in shear-wave time-delays above small earthquakes may be a valuable demonstration that all seismically active faults are permeated by critically high pore-fluid pressures.

6. Critical systems of stress-aligned fluid-saturated cracks

The reason that the almost parameterless APE-modelling can calculate and predict the response of stress-aligned fluid-saturated microcracked rock to changing conditions in an extraordinarily complicated heterogeneous Earth is because the cracks are so closely spaced (see Fig. 4) that the cracked rock is a critical system [6,35,67]. Systems verging on criticality maintain dynamic interactive processes that evolve locally until they reach criticality, when all members influence all other members [67–69]. Equilibrium thermo-dynamics is the classical critical system, and enthusiasts argue that almost all complex systems in nature are critical systems [70]. Such systems are said to possess self-organised criticality (SOC), and evolve in response to some disturbance (increasing stress in the case of the Earth) until criticality is reached. At fracture-criticality, when fractures are so pervasive that shear-strength disappears, the energy is released (by fracturing and earthquakes), the system relaxes to below criticality, and the evolution continues from the new position. The behaviour of critical systems near criticality is closer to that of other critical systems than it is to the underlying sub-critical physics [5,36]. This universality is believed to be the reason APE-modelling matches such a wide range of phenomena (Table 1).

The best-known evidence for SOC in geophysics is the (empirical) self-similarity of the Gutenberg–Richter relationship, where the log of the cumulative number of earthquakes is self-similar to the earthquake magnitude for several orders of magnitude. The APE mechanism for the evolution of aligned fluid-saturated crack distributions provides the sub-critical physics of the Gutenberg–Richter relationship, which had previously not been identified.

The uniformity and coherence of shear-wave splitting parameters throughout the crust and upper-mantle and the limited range of implied crack densities suggest that all fluid-saturated rocks have similar underlying, initially random, crack distributions with similar crack densities as illustrated schematically in Fig. 4 [19].

As mentioned above, the interior of the Earth is difficult to access, and APE criticality is difficult to demonstrate in situ. In the APE-modelling of the CO2-injections [50] in Fig. 7, and in changes before earthquakes [25,35,46], the criticality is implied. The only direct observations of extreme compliance are from a prototype stress-monitoring site (SMS) [71] measuring shear-waves in cross-hole seismics between boreholes, adjacent and parallel to the Húsavík-Flatey transform fault of the Mid-Atlantic Ridge in northern Iceland [7]. A calibration test measuring travel times horizontally at 500 m depth between boreholes 315 m-apart happened to coincide with a 4-day swarm of small earthquakes with energy equivalent to a single $M \approx 3.5$ earthquake on a parallel transform fault 70 km NNW [7].

Fig. 8 shows the travel-time variations over 14 days, accurate to ±0.02 ms of P-waves, SH- and SV-waves, and shear-wave anisotropy SH–SV. The seismic variations correlate with the low-level seismicity, but also correlate with anomalies in N–S and E–W global positioning system (GPS) measurements and with a 1 m 5-day decrease.
Fig. 8. Variations at the SMStES SMS from August 8–24, 2001: (a) P-wave travel times in ms; (b) travel times of SV-waves (green crosses) and SH-waves (blue crosses) in ms; (c) Time-delay (SV-SH) in ms; (d) GPS displacements around Húsavík in mm, north–south (blue circles) and east–west (red crosses); (e) pressure at 33 m depth in water well on Flatey Island in bars showing ocean tides and anomalous ∼1 m drop in water level; (d) twelve-hourly histogram of seismicity within 100 km of SMStES, Húsavík (after [7]). For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.
in level of water in a well immediately over the fault. Direction of the seismic-wave propagation is parallel to the fault and is a symmetry direction so splitting into $SH$- and $SV$-waves is expected. The SMS recorded data during four brief sessions, and we were fortunate to record these anomalies as continuous well-level measurements indicated that the 1 m decrease in well-level was the only such decrease in 15-months of available data. $M = 3.5$ is comparatively small earthquake with a fault dimension of a few tens of metres. Thus the variations in seismic properties show great sensitivity at a distance several hundred times the equivalent source dimensions. Although not yet fully interpreted, the anomalies in the figure are internally consistent. Such sensitivity would not be expected in conventional geophysics, and is a direct indication of deformation in a critical system of fluid-saturated cracks.

The high accuracy of the seismic measurement ($\pm 0.20 \text{ ms}$) is due to highly repeatable DOV source transmitting a 200 ms sweep of frequencies up to 200 Hz (a vibroseis ‘chirp’) three or four times a minute over 315 m in a very quiet environment at 500 m depth with borehole stability and low ambient noise [7]. The repeatability allows 100-fold stacking and the exceptional accuracy for field measurements.

7. Discussion

Criticality has major implications for conventional geophysical investigations. Taken from Paper [6] where they are discussed in more detail, Table 2 lists some of the effects. Critical systems have great sensitivity where minor remote disturbances can have major effects on local phenomena. This is the butterfly’s wings sensitivity of theoretical chaos. This extreme compliance means that the detailed behaviour of rock varies from place to place and from time to time, with the implication that any particular observation degrades with time. These variations may be in response to known disturbances, as in the seismic activity in Fig. 8, but may also be caused by the effects of solid earth and ocean tides, movements of magma (Item 11 in table), and by remote seismic activity. Since the earthquake magnitude scale is exponential, a unit increase in magnitude implies approximately an order of magnitude more energy. Consequently, extrapolation from Fig. 8, suggests that a single SMS would be able to monitor changes induced by $M = 5$ earthquakes up to (conservatively) 200 km; $M = 6$, up to 700 km; $M = 7$, up to 2000 km; and $M = 8$, earthquakes to probably worldwide.

This implies that the Earth is continually ‘flexing’ in response to typically unattributable local or distant perturbations. This may have serious consequences. For example, oil companies are currently investing billions of dollars in seismic time-lapse surveys of producing reservoirs, where (usually) three-component reflection surveys over extensive areas are repeated with similar source-to-receiver geometries in order to monitor the movement of fluids in producing reservoirs. If there are temporal variations in the seismic properties of the reservoir that are not associated with oil extraction (Items b1 and b2 in Table 2), then the basis of the technique is undermined. Note however, that the seismic travel times in Fig. 8 that display the effects of criticality are measured to exceptionally high accuracy ($\pm 0.20 \mu s$ with wavelengths of order of $\sim 6 \text{ cm}$). In contrast if resolution is low, many of the difficulties associated with criticality may be ignored (although criticality as an explanation of the behaviour of shear-wave splitting would still be required [35]).

Despite these various advances outlined, above, there are still at least two serious postulates on which the SOC hypothesis depends, that have yet to be proved.

1. Shear-wave splitting in the crust is caused by microcracks. We consider that the overall observations of compliance in Table 1 indicates stress-aligned microcracks, grain-boundary cracks, flat pores, and pore throats as the source of the splitting (see discussion at the end of Section 4, above). As pointed out in Section 4, there is ambiguity between micro- and macro-cracks as the cause of velocity anisotropy. This controversy might be resolved by better measurements of seismic attenuation. Macrofractures are nevertheless merely the end point of the process of crack growth and might be expected to mirror the properties of alignment, in particular, of the progenitor microcracks. However, macrofractures lead to seismic reflections and refractions and cause shear-wave splitting only when they are numerous enough and then they lead to significant attenuation and deterioration of the slower split shear-wave [51]. In addition, microcracks are sensitive to the contemporary
Table 2
Some practical implications of critical crack systems with SOC for fluid–rock interactions within the Earth’s crust (references in [4])

(a) General implications
1. Fluid-saturated crack distributions are highly compliant and crack geometry responds to small nearly negligible changes of stress, pressure, and physical properties of the pore fluids
2. Since fluid–rock properties vary with time, and vary from place to place, measured fluid–rock properties are only strictly valid at the place and time they are measured. Hence, the need for measurements with single-well imaging if accurate specifications are required
3. Since fluid–rock interactions have a dominant effect on almost all physical and chemical behaviour within the crust and mantle (see a2, above), these various effects apply to a large range of geophysical phenomena, particularly those associated with any deformation, including almost all processes during hydrocarbon recovery
4. Behaviour of stress-aligned fluid-saturated crack distributions appears to be remarkably uniform (within certain limits) even in very heterogeneous structures
5. Pre-fracturing deformation of any given fluid-rock configuration can be monitored by observations of shear-wave splitting (a5, above)
6. Pre-fracturing deformation can be modelled by anisotropic poro-elasticity (APE)
7. Response of fluid–rock systems to known changes can be calculated by APE
8. Response to calculated changes (a6, above) can be monitored by shear-wave splitting (a5, above), and the response controlled by feedback by adjusting changes to optimise the response

(b) Specific implications
Implications for hydrocarbon exploration and production
1. Reservoir properties may change from place to place
2. Reservoir properties may change with time, even without production processes
3. Relevant properties need to be measured at the place and time they are needed
4. Response to known changes can be calculated and predicted
5. Response of a reservoir can be controlled, in the sense of a8, above
6. Possibility of long-range and long-time correlations across and between reservoirs
7. There is a limit to the temporal and spatial resolution of any particular measurement

(c) Implications for earthquake geophysics
1. Deterministic prediction of time, magnitude, and place of large earthquakes is likely to be impossible
2. With sufficient source seismicity or appropriate cross-hole SMS observations, times and magnitudes of future large earthquakes can be stress-forecast. Other information may then indicate location
3. In presence of sufficient source seismicity, or appropriate cross-hole SMS observations, times of future volcanic eruptions can be stress-forecast
4. There is the possibility of long-range and long-time correlations between earthquakes

(d) Implications for rock physics
1. Much of the behaviour in stress-cells in the rock-physics laboratory can be modelled and predicted by APE

stress-field, whereas macrocracks have fixed orientations determined by a previous stress-field, and do not have a dynamic response to changes of stress.

(2) Shear-wave splitting in the uppermost mantle is caused by films of hydrologised melt. Following Hess [8], most seismologists consider that the shear-wave splitting in the upper mantle is due to aligned crystals known as lattice-preferred orientation (LPO). In fact the typical parameters of shear-wave splitting in the mantle are remarkably similar to those of the crack-induced shear-wave splitting in the crust. In the mantle there is the same stress (or flow) aligned polarisations, and the same range of shear-wave velocity-anisotropy with a minimum of about 1% and a maximum of about 5% compared to the 1.5–4.5% range observed in the crust (Table 1). Petrology suggests there is up to 0.1 wt.% water in the upper 400 km of the mantle [72] which promotes melting. When crystals melt, they first melt along grain boundaries, and it has been shown experimentally [73] that in the presence of triaxial stress the melt appears as films of liquid melt along grain boundaries for the appropriate dihedral angles. Such films of melt, like the fluid-saturated microcracks in the crust, would be aligned perpendicular to the direction of minimum stress, exactly analogous to the fluid-saturated grain-boundary cracks in the crust. In the mantle the fluid is water-induced melt rather than the water-based salt solution saturating cracks in the crust. Note that crack-induced shear-wave splitting for thin cracks is almost independent of porosity and aspect-ratio
and the water-induced melt expected in the upper mantle will yield aspect-ratios sufficient to eliminate tangential shear coupling across cracks and induce shear-wave splitting.

Note that observations of shear-wave splitting do not give any direct information about the location of the anisotropy along the ray path. However it is generally consistent with other observations that mantle anisotropy is confined to the uppermost 400 km, the approximate depth of the olivine-spinel phase change at 415–445 km [74].

Note also that polarisations of the faster split shear-wave in the mantle are generally found to be aligned with the presumed direction of flow of the heavily viscous mantle material. If mantle shear-wave splitting is oriented by stress-aligned films of hydrologised melt as suggested above, the polarisations will be perpendicular to the direction of minimum stress. Hence observations suggest that the direction of minimum stress is orthogonal to the direction of flow, implying that the mantle is pushed by spreading centres rather than pulled by subduction.

8. Conclusions

These various studies of shear-wave splitting in an anisotropic cracked Earth, including many hundreds of papers not referenced here, are probably the most comprehensive investigation of wave propagation in any anisotropic material. Stress-aligned shear-wave splitting is now expected in almost all rocks in the crust and the uppermost 400 km of the mantle. Our understanding of shear-wave splitting in the Earth’s crust, and its association with rock deformation, has advanced substantially in the last year or two, as the essential crack-critical nature of the crust has become better understood.

The implication is that the crust and uppermost 400 km of the mantle are critical systems of closely spaced fluid-saturated microcracks with great sensitivity to changing conditions. This ability to monitor low-level deformation, might well be the most fundamental advance in solid Earth geophysics for several decades, and is the driving force for the developments reviewed in this paper. Thus, it can be argued that wave-propagation in anisotropic rocks is likely to lead to major changes in geophysical practice by industrial as well as earthquake geophysicists. Shear-wave splitting is opening a window into a New Geophysics of the Solid Earth where many effects are calculable, sometimes predictable, and potentially even controllable [6].

Acknowledgements

We acknowledge with gratitude the vast contribution of all our colleagues in the ideas reported in this review. This work was partially supported by the European Commission SMSITES and PREPARED projects, contract numbers EVR1-CT1999-43082 and EVG1-CT2002-00073, respectively.

References


