The Alliance of Laboratories in Europe for Education, Research and Technology

ALERT Doctoral School 2019

The legacy of Ioannis Vardoulakis to Geomechanics

Editors:

Jean Sulem

Cino Viggiani

Editorial

The ALERT Doctoral School 2019 on "The legacy of Ioannis Vardoulakis to Geomechanics" will take place in Aussois, from October 3rd to 4th, 2019. The School has been organized by Prof. Jean Sulem (Ecole des Ponts Paris Tech) and Prof. Cino Viggiani (Université Grenobles-Alpes). I sincerely thank the organizers and all the contributors to this book for their effort!

Professor Ioannis Vardoulakis has been a prolific researcher, whose interests include many aspects of the geomechanics but also physics, mathematics... It is therefore hardly possible to cover all his contribution to science. The Doctoral school 2019 will focus on the granular materials, and will show how theoretical concepts can impact engineering applications. I am therefore convinced that this school will be beneficial to the ALERT community.

The organizers of the school decided to highlight the legacy of Professor Ioannis Vardoulakis with three topics: the mechanics of granular materials (both experimental approach and modelling), the modelling of strain localization in geomaterials and the hydromechanical couplings. The school will also cover some geotechnical applications, like the landslide mechanics and petroleum geomechanics. It has to be pointed out that the school will run over two days (instead of 2.5 days in the previous years). We expect therefore that every participant will attend the school up to Friday evening.

As usual, the pdf file of the book can be downloaded for free from the website of ALERT Geomaterials (http://alertgeomaterials.eu/publications/) after the school.

On behalf of the ALERT Board of Directors I wish all participants a successful ALERT Doctoral School 2019!

Frédéric Collin
Director of ALERT Geomaterials
University of Liege

Contents

The legacy of Ioannis Vardoulakis to Geomechanics – Foreword

Jean Sulem[(1)] and Gioacchino (Cino) Viggiani[(2)]

(1) Laboratoire Navier-CERMES, Ecole des Ponts ParisTech, IFSTTAR, CNRS, Université Paris-Est, Marne-la-Vallée
(2) Univ. Grenoble Alpes, CNRS, Grenoble INP, 3SR, F-38000 Grenoble, France

Professor Ioannis Vardoulakis was not only a highly recognized scientist. He was also a passionate and talented educator; all his former students and colleagues remember his tremendous energy, enthusiasm and generosity in teaching and mentoring young scientists. He was among the pioneering members of the ALERT Geomaterials network who established the ALERT Doctoral School, as he strongly believed in the need of international networks for educating the engineers and scientists of today and tomorrow. In his lecture notes and courses, Ioannis Vardoulakis had a special ability to present at the same time fundamentals and advanced notions (for "good" students, as he used to say). For example, he could introduce in a very natural and physical way, complex concepts of higher order continuum theories when teaching a class on basic strength of materials. Although Ioannis Vardoulakis had a strong background in theoretical mechanics, the way he used to communicate was that of an engineer – using simple and precise sketches drawn on the blackboard. As a researcher, Ioannis Vardoulakis had a special curiosity and taste for challenging scientific questions. Among them, one of the first topics that he addressed was related to slope stability or consolidation process in a softening ground. His large scientific culture gave him the ability to link various scientific fields and to propose novel and inspiring ideas. This is how he early found out that exploring advanced mathematical concepts of uniqueness and bifurcation could help in solving practical engineering issues, as for example borehole stability problems as encountered in petroleum industry. His scientific approach was based on first exploring the physical processes and identifying the dominant mechanisms, then imagining 'simple' conceptual experiments, writing down the governing equations in the simplest possible way, and then developing more advanced models by progressively relaxing some assumptions and hierarchically increasing the level of complexity. In doing so, Ioannis Vardoulakis showed interest in, and contributed to, the whole chain of research from fundamentals to applications. Ioannis Vardoulakis has produced milestone papers in various fields of geomechanics. Although it is not possible to cover in a single ALERT doctoral school all the topics he addressed, some emblematic subjects have been selected for this school. The chapters in this volume have been prepared to highlight the legacy of Ioannis Vardoulakis contributions to the current developments. Already in the mid-seventies, Ioannis Vardoulakis published highly cited papers on strain localization analysis

including discussions on fundamental mathematical issues of uniqueness and stability and experimental observations in element tests on sand. He recognized the key role of the microstructure of the material in strain localization phenomena and the necessity to look at the small scale physics in order to enhance the constitutive description of the material and introduce the necessary mathematical tools for regularizing the underlying governing equations. In his milestone paper coauthored with Hans-Bernd Mühlhaus [Müh87], Ioannis Vardoulakis showed that the framework of Cosserat continuum theory was appropriate to describe shear banding in granular materials; advanced experimental testing and Discrete Element Method simulations confirmed later significant grains rotations inside a shear band. This topic is revisited in the chapter by ***Ioannis Stefanou and Eleni Gerolymatou****, which presents the basics of bifurcation theory and higher order continuum regularization techniques. In the mid-eighties, Ioannis Vardoulakis published the first images of the progressive development of a shear band in a sand sample using x-rays [Var85]. Since then, advanced 3D imaging techniques and full-field measurements in experimental geomechanics have permitted an in-depth analysis of the micro-mechanisms involved in strain localization for various types of geomaterials. This topic is addressed in the chapter by* ***Edward Andò****. Another research area of Ioannis Vardoulakis was the mechanics of granular materials in relation with internal erosion processes and hydrodynamic instabilities. Modelling issues in the mechanics of granular materials are addressed in the chapter by* ***Itai Einav, Benjy Marks, and Pierre Rognon*** *with special attention the evolution of the structure of a grain assembly due to grain mixing/segregation and grain crushing, whereas hydromechanics issues are addressed in the chapter by* ***Holger Steeb****. For geotechnical applications, Ioannis Vardoulakis emphasized the key role of thermo-hydro-mechanical (THM) couplings in strain localization and stability analyses. One can mention another milestone paper on THM couplings in landslide mechanics with application to the Vajont case study [Var02]. This topic is addressed in the chapter by* ***Alexander M. (Sasha) Puzrin*** *with special attention on slip surface growth and post-failure evolution. Although the scientific contributions of Ioannis Vardoulakis might appear as being too sophisticated, or too "academic", for practical engineers, it turns out that his work on bifurcation phenomena in geomechanics [Var95] has attracted the interest of petroleum engineers for dealing with difficulties encountered in operations such as borehole stability and sand production. This is the topic of the chapter by* ***Euripides Papamichos and Panos Papanastasiou****.*

We would like to thank all the contributors to this volume and hope that the papers collected herein will provide a good overview of the impact that the work by Ioannis Vardoulakis has on the Geomechanics of today and tomorrow.

Jean Sulem
Cino Viggiani

References

[Müh87] Hans Bernd Mühlhaus, Ioannis Vardoulakis. The thickness of shear bands in granular materials, *Géotechnique* 37(3):271–283, 1987.

[Var85] Ioannis Vardoulakis, B. Graf, B. Calibration of constitutive models for granular materials using data from biaxial experiments. *Géotechnique*. 35, 299-317, 1985.

[Var95] Ioannis Vardoulakis, Jean Sulem. *Bifurcation Analysis in Geomechanics*. CRC Press — ISBN 0203697030, 1995.

[Var02] Ioannis Vardoulakis. Dynamic thermos-poro-mechanical analysis of catastrophic landslides, *Géotechnique*, 52(3), 157-171, 2002.

Mechanics of granular materials I. Experiments on grains

Edward Andò

Laboratoire 3SR, Université Grenoble Alpes and CNRS. F-38041 Grenoble, France

This chapter discusses current experimental capabilities in measuring the mechanical behaviour of granular materials. A short introduction is made regarding fundamental concepts in experimental mechanics. Thereafter more advanced measurement techniques for capturing granular behaviour are introduced and discussed. This chapter closes with a look at future techniques which may disrupt the field bringing even more information into the light.

1 Experimental Geomechanics – Opening

The primary objective of experimental geomechanics is to capture the behaviour of geomaterials as they respond to an applied stress or strain path. Traditional soil mechanics unit-testing aims to study the stress-strain response of supposedly homogeneous specimens in order to measure some inherent "constitutive behaviour" – as opposed to structural tests with explicitly induced inhomogeneity conditions such as a penetration test or a trapdoor experiment. In classic unit testing forces/displacements are applied on the boundaries of carefully-prepared specimens and the resulting displacements/forces are also measured on the boundaries (or on average). The mechanics that can be captured from such testing requires the assumption of homogeneity in order to interpret boundary measurements and can end up averaging/smoothing out inhomogeneous behaviour.

This chapter will stay with traditional unit testing, but will give an idea about how full-field measurements can be used to capture:

- first of all descriptors of granular materials such as their shape and size as well as their arrangement
- then the measurement of granular kinematics as a sample responds to applied stresses or strains

2 Crash course in measurement theory

2.1 Objectives of making measurements

The objective of experiments on the mechanics of granular materials is to characterise the deformation processes – a description of a state or a change of state.

An understanding of the quality and limitations of the measurements deployed in an experiment is essential. Without descending into the fundamental limitations of measurement at the particle physics scale, the types of measurements used in conventional soil mechanics are not perfect, and in order to describe reality properly the error that measurements can make is a fundamental part of the description. The accounting for experimental error is unfortunately not as prelevant in experimental mechanics in our field as it could be. For the usual effort of performing experiments to develop or calibrate mathematical models of behaviour, the uncertainty in the measurement (as well as the inter-sample variability) is an essential piece of information for the modelling effort.

2.2 Classical sources of uncertainty

Measurements are fundamentally uncertain, and the characterisation of the level of uncertainty is necessarily basic good practice in experimental mechanics. The level of uncertainty of a measurement is the sum of a number of different sources of error or noise.

As an initial example let us take the measurement of a displacement. Figure 1, shows two potential devices for measuring displacements – the body of both are attached to some reference point and movement of the probe at the bottom is measured. The internal mechanisms of both devices is very different – on the left there is a mechanical (dial) indicator, which essentially *multiplies* the possibly small displacement of the probe as a large displacement of the dial on the face, which is then read by the experimentalist. On the right, an LVDT ("linear variable differential transformer") changes a source voltage into output voltage by some linear function proportional to the displacement of the probe.

In both systems a further readout system is required to record the measurement. In the case of the mechanical system this is the experimentalist's eye and notebook, whereas in the electronic case this is most usually a digital voltage meter – and ADC (Analogue to Digial Converter) which records the analogue voltage over a given range, say [0, 5 V] and describes fractions of this range as a digital value (an integer running from 0 to a pre-defined maximum value). Both readout systems have limitations which resemble each other more than one might think – the possible values that can be read out depend on the level of *discretisation* that is made of the relatively continuous (we will return to this) measurement range. In the case of the mechanical system, the eye can tell only so many different values between the marks on the face, with fundamentally limits the smallest displacement that can be measured *even though the*

Example measurements of displacements

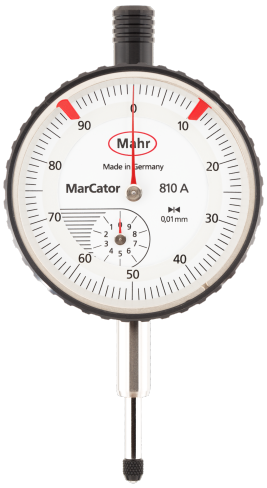
(mechanical) dial indicator	(electronic) LVDT
	

Figure 1: Examples of tools to measure displacements, an analogue mechanical indicator. `https://en.wikipedia.org/wiki/Indicator_(distance_amplifying_instrument)` (left) and an electronic LVDT. `https://bditest.com/product/sensors/displacement/lvdt-displacement-sensor/` (right).

hand may be making smaller but undetectable movements. In the same way, the fact that the [0, 5 V] range is explicity discretised into a number of integers also limits the smallest displacement that can be measured. For example a 16-bit sensor can record values from [0, 65535] meaning that the [0, 5V] range is discretised into $\frac{(5-0)}{(65535-0)} = 0.000\,076\,\mathrm{V}$ increments. As before, although the LVDT might be sensitive to smaller voltage changes, there is this strong limitation in the readout system. In both cases there is a minimum *sensitivity* below which a change cannot be measured.

Furthermore, in both systems in a static situation one might notice fluctuations of the measured value when repeated with time. This can be due to the precision of the estimate of the position by eye, but often with electrical measurements there are electrical fluctuations (the driving voltage in the example above) which when uncontrolled can scatter the measurement. This scatter can be characterised and in a number of cases reduced with understand of the measurement system. Common sources of this sort of noise in electrical systems are $50\,\mathrm{Hz}$ temporal fluctuations from the grid supply, and inherent variations in the ADC used to read the signal. If this scatter follows a statistical law that can be identified – Gaussian distributions often appear – this knowledge can be used to give an uncertainty to the measured value, as well as means to find the

true value by repeated measurements – which are usually taken to be independent. In the case of a Gaussian distribution, averaging N samples will reduce the noise in the measurement by $\sqrt{N}$.

Furthermore if the measurement in question requires some kind of accumulation, or counting of discrete events – such as counting sheep, or more commonly in experimental mechanics photons, further complications can arise. In the case of photons, a continuous flux of photons can only be detected by detecting individual, discrete photons. For a given integration time, if very few discrete photons are captured, the measured flux will be very erroneous, and any noise in the measurement will be very damaging. This sort of sampling bias can often be modelled as a Poisson process and therefore be expected to follow a Poisson distribution.

2.3 Sources of uncertainty in full-field measurements

Full-field measurements can vaguely be defined by measurements where the distribution of a variable is measured in many different points. Just as the "point" measurements discussed above are often mediated by a change of voltage/resistance which is linked to the change in the desired measurement, in the case of full-field measurements, the properties sought are often related to some light emission/reflection which can be captured with a camera (may exceptions exist, such as ultrasonic tomography). Full-field measurements suffer from all the above problems, but the fact that many measurements are repeated in space mean some additional sources of error that are difficult to consider for "point" measurements above. This is essentially that the so-called "point" measurements sometimes are an uncontrolled local average of some physical value that is being measured and are therefore not punctual. A clear example of this could be a temperature measurement that represents an average along the sensing tip, and even around the tip. In the context of full field measurements, a similar concept applies: each one of the measured values is subject to a discretisation error, has a minimum sensitivty and a time-varying noise (all as above) but furthermore there can be some lack of independence of nearby measurements. This effect is not an uncorrelated random noise as might be expected for individual measurements, but is rather strongly correlated in space. This effect can have different sources – in the case of optical methods like taking photos with a camera it can be related to the optics (bad focus means blurry images where there is a correlated noise), or in the readout system, where values can "bleed" onto nearby ones.

2.4 Rapid introduction to x-ray tomography

Without going into too much detail on techniques, x-ray tomography is an optical full-field measurement technique that takes advantage of the penetrating character of x-ray radiation to make an indirect measurement of the x-ray attenuation coefficient in a 3D volume.

Fundamentally, the first order physical model that describes the interaction of x-rays with matter is the well-known Beer-Lambert Law:

$$I = I_0 e^{-\mu\rho x} \tag{1}$$

This states that a photon beam with flux (*e.g.*, in photons/sec) I_0 travelling through a given material is reduced by en exponential function of the path length of the beam x and the material properties (x-ray attenuation and density) $\mu\rho$ into an output flux I. $\mu\rho$ depends on material properties as well as the energy of the photons.

X-ray tomography aims to measure a 3D field of $\mu\rho$. This cannot be done directly, so the fact that x-rays travel through matter is exploited. If a beam of initial flux I_0 interacts with a specimen and is recorded I after the interaction with the object, this is called a radiography, or a radiographic projection, since the (unknown) x-ray attenuation field is *projected* (integrated) in the direction of the beam. During a tomographic scan, the (unknown) x-ray attenuation field is projected in a number of different directions and the recorded.

An inverse problem can then be solved (traditionally by a technique called "filtered back-projection") in order to obtain the 3D field of $\mu\rho$ which will be referred to as the x-ray attenuation field – this process is called "reconstruction". In most x-ray tomography setups, $\mu\rho$ cannot be reconstructed quantitatively because x-ray detectors to measure the x-ray flux after interaction with the specimen I/I_0. This immediately means that the reconstructed field is a somehow averaged $\mu\rho$ over all the frequencies that were emmitted and detected. The fact that in reality μ depends on frequency creates well-known non-phisical errors ("artefacts") in the reconstructed field known as "beam-hardening" for which there are some mitigation strategies which will not be discussed further.

Depending on the physical size of the pixels that are used, and the optical zoom factors, each 3D pixel ("voxel") in the reconstructed field represents a the x-ray attenuation coefficient in a small physical volume. In most scanners used in experimental mechanics, these little volumes are cubes, and their size – which is purely geometrical – is expressed as the length of one of the sides of the cube, the "pixel size" in micrometers/pixel. It is important to note at this stage that given this discretisation of *space*, that a voxel situated on the edge between two materials will (discussions about noise left aside) have a volume-averaged value of the x-rayh attenuation coefficients present in the voxel. This is known as the "partial volume effect" and is often (and in my opinion incorrectly) considered a source of noise – it is simply the consequence of the discretisation of space.

Armed with this knowledge, we can now conclude that with penetration-based volumetric tomography provides:

- A 3D volume measurement of a non-quantitative x-ray attenutation field
- Every measurement "point" (small subvolume of a fixed size called a "voxel") has a certain amount of random noise
- Due to optical effects there can be noise with spatial correlations, such as blur

- There are other sources of noise commonly referred to as artefacts, which are closely tied to the measurement technique. Ring (not discussed so far) and Beam Hardening are the most important types

The above list of deviations from reality gives an important first-order approach to define the quality of a tomographic measurement:

- Signal-to-Noise Ratio (the desired measurement – *e.g.,* the x-ray attenuation coefficient of the grain and pore phases – compared to the level noise)
- Level of blur (expressed in mm)

This in turn leads to a convienent way to characterise a given scanner for a given sample:

- Signal-to-Noise ratio
- Spatial resolution (mm)
- Temporal resolution (scanning time in minutes)

On a typical lab-based scanning system these three quantities can be traded off, and so when designing an experiment it is important to understand the sensitivity of a full-field measurement to these three quantities.

3 3D static measurements of granular media

Geological granular media typically consist of silicate or carbonate particles in water, oil or air. The ability to measure a 3D field of x-ray attenuation coefficient is very convenient for characterising these materials since there is a significant difference in density (and therefore an expected difference in x-ray attenuation coefficient).

This chapter will focus on discrete measurements, where grains can easily be detected in the reconstructed x-ray tomography volumes. Adopting the language from the previous section, the identification of particles clearly requires the spatial resolution to be far below (for argument sake let's say 10 times smaller when measured in mm) than the size of the particles. Furthermore, in order for particles to be distinguished from pores, the signal to noise ratio must be sufficiently high.

It is important to note at this stage, that the requirement for high spatial resolution is only needed for a *discrete* handling of the particles, and that much important experimental work has been done at lower spatial resolution – the foundational work of Jacques Desrues immediately comes to mind. Since density is a key variable in the behaviour of granular media, a local density can easily be measured with x-ray tomography, even if grains cannot be seen. In the work mentioned above, this permitted the key discovery that the concept of a critical state density is particularly true locally within shear bands compared to globally in the specimen.

3.1 Identifying particles

Returning to discrete measurements, if the signal to noise ratio is very high, one would expect to be able to classify each voxel in the reconstructed image by simply applying a *threshold value*. Figure 2 shows the sort of image which is nowadays easily obtained of a sample of sand within an experiment. Grains can easily be distinguished by eye from the air in the pores and the material of the confining vessel.

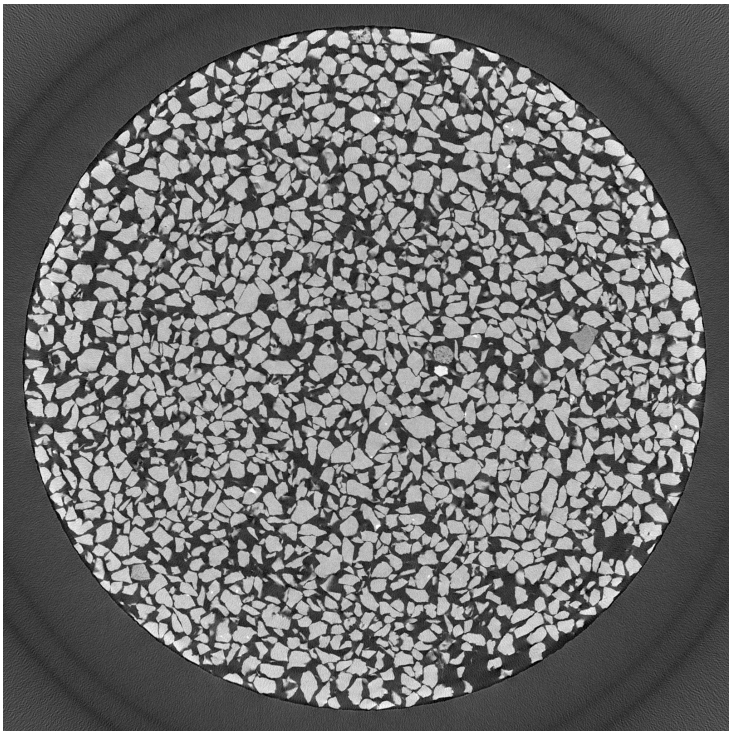


Figure 2: Horizontal slice of a reconstructed x-ray attenuation volume. As is conventional high values are white and low values dark. Material is Hostun sand (D_{50} = 330 µm) and pixel size is $10\,\mu\mathrm{m/px}$. The sample is housed in a PEEK oedometer, experiment performed by Max Wiebicke

At the very first order, the solid phase can be separated from voids by establishing a threshold grey value. Figure 3 shows the distribution of reconstructed x-ray attenuation coefficients for a crop in Figure 2 which contains only air and grains (*i.e.,* the material of the cell is excluded). It is clear that the values in the image corresponding to the air (≈12000) and the solid grains (≈ 48000) are well, separated, which means that the signal-to noise ratio is relatively low. If a Gaussian noise model were to be used (which is reasonable in this case), the signal-to-noise ratio can be calculated as the difference in greyvalues of the peaks, compared to standard deviation (also has

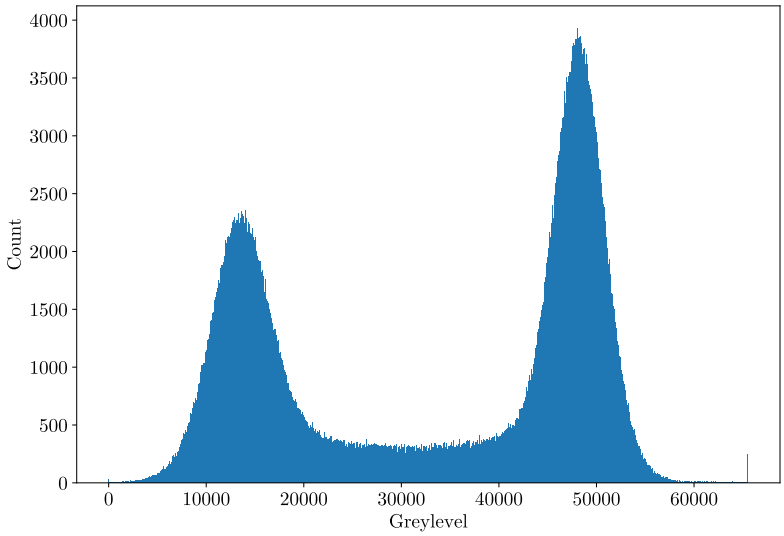


Figure 3: Histogram of grey values of the central part of Figure 2

units of greyvalues) of the noise, which can be obtained by fitting either peak. The standard deviation of of both peaks is expected to be the same, if the solid material is "pure" – if the solid material presents gaussian variations of x-ray attenuation, this will obviously increase the standard deviation of this phase.

The partial-volume effect is clearly visible as a raised value between the peaks, which is not present below the lower peak or above the higher peak. Its relative significance is related to the small size of grains in the image – or more correctly to the large numbers of interface voxels found in the image.

Starting from such a 3D measurement and from the appearance of the distribution of grey values, it is clear that a greyscale threshold will work appropriately. The way in which this can be set is twofold: either by modelling the grey value distribution in the reconstructed images (*e.g.*, finding the minimum of the intersection between the two Gaussian fits of each peak), or in such a way to obtain the total solid volume measured for example at the end of a test. It is important to note that this thresholding operation represents a significant coarsening of the image – more advanced techniques for the representation of the solid phase such as level-sets can avoid this coarsening, but are much more complicated to use for simple analysis.

The solid phase must then be separated into individual particles. This is a challenging problem called segmentation. The classic way that this is solved is using a "water-

shed algorithm", which tends to work quite well for relatively convex particles with point contacts and not too wide particle size distribution. Very briefly, at their base, watershed algorithms are geometrical, especially the "morphological" variant which will be discussed here. Given the binary map of the grain phase to be split into individual particles, the distance of each grain-phase voxel to the nearest pore voxel is found – this is expressed as a "distance map" of the grain phase. What is expected is that voxels on the edge of each grain will have a low value, and that this value will increase steadily towards the centre of each particle. In fact, the centres of particles are defined directly (as "markers" for the watershed) from the distance map – the 3D high points in the distance map are considered the centres of grains – in order to allow some non-sphericity very close high points are merged. These high points are numbered, and nearby voxels are associated to each number based on a flooding of the inverted distance map.

Higher fidelity segmentation is an active field of research, since it is such a critical step – especially for non-spherical grains. Regardless of the technique used, the result is that every voxel belonging to the grain-phase is given a unique number or label, as can be clearly seen in Figure 4.

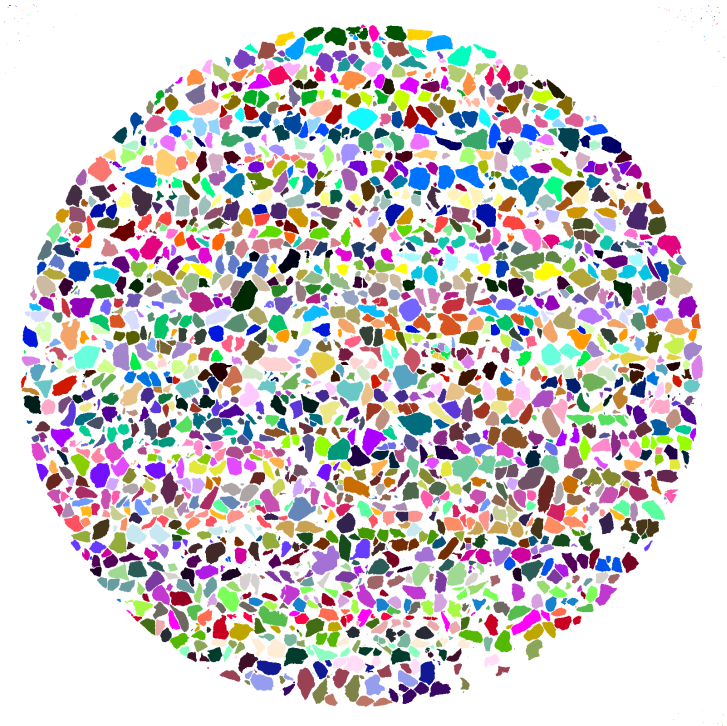


Figure 4: Slice through a labelled image, with different (scattered) colours indicating differently numbered grains

A labelled 3D image is the ideal starting point for all grain-based measurements: for example each grain can be individually interrogated as a discrete and contiguous cloud of voxels. This allows, for example, for a grain-size distribution to be obtained, on the condition of having a well-defined "size" to measure, with a relative error in the order of the pixel size.

3.2 Characterising particles

Particle Volumes Starting from a correctly labelled image, and armed with the fact that sand grains, at the large scale are on the whole rather convex, a simple counting of the number of voxels making up each labelled grain gives an accurate and rather unbiased and stable (with respect to a change of resolution) measurement of volume. This measurement can be rendered a little more sensitive to partially-filled voxels on the edges by weighting voxel densities proportional to their grey value (since the reconstructed grey value is roughly proportional to density).

The extension of this sort of measurement to the reproduction of sieve sizes is not particularly challenging, the main question being how close to simulating the actual sieving procedure one wants to go. Some examples can be found in [ADAR$^+$14, KAP$^+$17].

Particle Positions The centre of mass of a label is simply calculated as the mean position of the voxel-cloud making up the grain – again, weighting by the grey value can improve sensitivity. Outside of the extraordinary case of cubic particles aligned with the coordinate axes, averaging means that the accuracy of this measurement is far better than the pixel size, around $\frac{1}{20}$th of the pixel size for a grain measuring roughly 15 pixels in diameter according to [And13].

Grain orientation can be defined in a number of different ways. Some typical methods used in the literature are:

- *Feret* diameters (or caliper lengths) where voxelised objects' sizes are probed with digital calipers (or more correctly plates) at a number of different orientations. The shortest orientation (for lentil-like particles) or the longest orientation (for rice-like particles) can be unique and thus useful for detecting the orientation of such particles
- *The Moment of Inertia Tensor* and its eigenvectors are also a common way to measure the orientation of a particle's long and short axes – although these are not always aligned with the axes in a symmetric shape. This method has the significant advantage with respect to Feret of not having to scan a number of discrete angles which can limit sensitivity, as well as taking advantage of the full body of the 3D data available
- *Geometrical fits* of particle shapes can be good ways of detecting their orientation – 3D ellipsoid fits are often used – however for good results some *a-priori*

information about the shape of the particle is needed, which is unfortunately not typically available for natural particles

Generally speaking, the quality of the definition of a particle's orientation depends both on the method used to measure the orientation, and on the shape of the grain. It is clear that in the extreme case of a perfect sphere, its orientation is not defined, and for particles very close to spheres, making an accurate and stable measurement will be difficult. It can be safely said that the less spherical the shape of a grain, the better the measurement of orientation will be, regardless of the method used to measure it. The overall picture is not so simple, however, since there can be some features of a grain's shape which are unmistakable and others where the shape can induce errors: for the examples given above, only the long axis of rice and the short axis of lentils are representative and stable when measured in a number of different states.

It must be said that the evolution of measured particle orientation is not necessarily stable for natural sand grains, where for small changes in orientation, non-unique axes can "snap", see [AHV+12]. In this case an initial orientation can be updated using incremental rotations measured with image correlation. The identification of a unique or representative axis to orient natural particles is not necessarily obvious but can be studied on high-resolution images.

Grain surface/topology is a challenging measurement because of the fractality of natural grain surfaces [SB89, MT90, SVK+02], which means that the measurement of the surface area, for example, depends on the scale of measurement (*i.e.*, the spatial resolution in this case). This makes a direct measurement very problematic (by any means), however *a-priori* knowledge of the fractal length of the surfaces of the material in question can be of use for interpreting a measurement made at a given scale. Furthermore, starting from a 3D image where the grey values have already been thresholded to define the volume of a grain, a significant amount of precious greyscale information for reconstructing the surface of the grain is lost.

Inter-particle distance measurement for inter-particle contact detection. The series of steps described above to obtain labelled 3D volumes representing each individual particle are relatively standard, however they introduce a tacit assumption that damages the detection of contacts: the grey value threshold. Although this seemingly innocuous step retrieves the correct solid volume, in a voxel partially filled with edges of two close-but-not-touching grains there is no obvious way to assign solid pixels to each particle and see whether the grains are really in contact[1]; the single voxel may appear dense enough to seem solid, and therefore the obvious conclusion to draw is: since there is a solid bridge between two particles, they must be touching. Based on this simple reasoning, the number of contacts that a particle has, can be shown to be systematically over-estimated. Some remedial measures being developed in the ongoing PhD work of Max Wiebicke[2] allow the *over-detection distance* to be reduced but

[1] in the case of natural materials where there is no *a-priori* model for the shape which would significantly improve the accuracy of the interparticle distance measurement

[2] working between TU-Dresden and Laboratoire 3SR in Grenoble

never eliminated when dealing with images of irregular natural particles [WAHV17]. Furthermore, given that one deals with an over-detection *distance*, a doubling of the spatial resolution only yields a halving of this over-detection distance, a situation very unlike the measurement of particle positions which improves with a very strong function of the spatial resolution (the actual numbers depend on the shape of the object).

Inter-particle contact orientation Also of interest in the PhD work of M. Wiebicke is – given correctly-identified interparticle-contact points – the measurement of the local orientation of the contact plane. This is of key mechanical importance for finding the orientation of the force which can be transmitted with no tangential components (and thus the cone of possible forces given a constant friction) as well as the contact-sliding direction. This poses a serious metrological problem: this key mechanical variable – a contact – does not actually exist! Current attempts to reconstruct the orientation of the contact plane from the "watershed line" that separates contiguous parts of the solid phase have large errors (in the order of 10 degrees) and *do not benefit at all* from increases of spatial resolution, since fractal surfaces in contact in reality have very few touching points. The future directions to explore to overcome this *impasse* must involve the use of the local surface topology to make two independent measurements of surface orientation to be compared and combined.

4 3D kinematic measurements of granular media

When a series of images has been acquired where particles rearrange (without breaking) following the trajectory of every particle is of clear interest.

Again, individual grain labels are a very convenient way of defining grains in the context of grain tracking from a 3D image of a given state to find their corresponding positions in a 3D image of a different state.

In time-series data where particles are rearranging, all imaged states can be labelled, and labelling can be made consistent across time using techniques such as ID-Track [AHV$^+$12] or equivalent [SFA$^+$12, AJDAR16]. Changes in particle position therefore naturally provide a measurement of 3D particle displacements. This technique is however very sensitive to segmentation errors (*i.e.*, if a particle is correctly identified in one scan, and split in two in the next scan). However if a reliable labelled image is obtained for two states, and labels can be made consistent between states, then displacements can be measured as a change of centre of mass of each particle. Rotations can also be defined by follwing the rotation of the basis offered by the eigenvalues of the Moment of Inertia tensor discussed above (on the condition that these axes are quite unique).

Displacements and rotations can also be measured using only one labelled image (typically the initial one) by using the greyscale data in the context of Digital Image Correlation (DIC), as first done by [HBD$^+$10]. This general family of techniques aims to find an optimal transformation to minimise the difference between a reference and

a deformed scan. Generally speaking, if transformations are not too large these techniques tends to be quite robust. With image correlation, particles are tracked from one image to the next in order to measure their change in postition, the same order of accuracy can be obtained (see [PLB10]).

However, in order to be able to use the "particle characterisation" tools detailed above to monitor the *evolution* of each particle's properties (the coordination number for instance), subsequent labelled images are highly desirable. To this end, 4D-segmentation techniques are of clear interest for this sort of data.

5 Discussion and future directions

There is no doubt that for micro-mechanical work of real engineering relevance, it is essential to go beyond the particle – kinematics and descriptions of shape – and focus on the way in which forces are transmitted through a rearranging granular assembly. To this end, the above discussion presents a bleak state of affairs for the "direct" measurement of granular entities attached to inter-particle contacts, which is particularly bad news for the experimental measurement of variables depending on these quantities; a perfect example of this is the Fabric tensor mentioned in the introduction.

However, the ease with which particles can be identified and followed during a mechanical test can be exploited to the experimentalist's advantage: a stunning recent example [KBB+17] shows how given an initial 3D image of a granular assembly, particle positions and orientations can be updated with degraded tomographies, with as few as 1% of the number of projections required to reconstruct a whole image. Here we propose two possible directions for improving inter-particle contact measurements which make use of the very high quality of measurements that can be made of particle kinematics.

5.1 Using particle registration to increase spatial resolution

[KBB+17] show the extraordinary attractiveness (from an optimisation standpoint) of rigid particle transformations through space, and suggest a potential method by which the spatial resolution/field-of-view trade-off can be improved in order to increase the quality of the measurement of granular Fabric for example. The idea, which is at the very heart of the PhD project of O. Okubadejo[3] would be to improve particle tracking in a framework in which segmentation errors and tracking errors can be reduced iteratively taking into account the whole series of 3D images acquired throughout a test. A successful tracking of each grain would then mean that the experimentalist has a (relatively large) number of *registered* greyscale 3D images of the same particle. At the very first order this could be used to obtain a significantly de-noised image of the grain, but could subsequently yield a super-resolution image of each grain (especially from views from different positions and angles). Super-resolution works in the particular case where the real spatial resolution is better than the geometrical voxel

[3] working in Grenoble between Laboratoire 3SR and GIPSA-Lab

size, meaning that each voxel in this case can be subdivided and a higher resolution image can be reconstructed on this new basis meaningfully. This means that the over-detection distance for inter-particle contact detection is reduced, and the quality of the map of the surface of each grain is improved, allowing a better contact plane to be estimated in the technique proposed above.

5.2 Using particle kinematics to deduce local conditions

Another piece of information that can be exploited to improve measurements of contact properties from a series of images starts again from a high quality tracking of particles; however in this case we will interest ourselves with particle rearrangements. The precision that can be obtained for the displacements and rotations of each grain is at least an order of magnitude greater than the over-detection distance for contacts: this leaves some room to scrutinise granular displacements for signs of reorganisation, which in turn give indications regarding the balance of forces, or presence of contacts. This discussion of course works increasingly well where small increments of rearrangement separate each image in the sequence, improving the chances of catching individual reorganisation events. When studying a neighbourhood of particles where some rearrangement is measured between two states, out-of balance forces must have existed at some point during the interval studied. This causes a particle to move with respect to its neighbours while losing, shifting along, or gaining contacts. This in itself is not a profound discovery, however a rearrangement event means that a number of pieces of information can be associated with this interval: for example to bound the uncertainly regarding apparent contact points (subject to the over-detection distance error developed above), as well as to set limits (given a simple friction model) on the forces at play over these contacts. This sort of approach really starts to make sense when deeply integrated with grain-scale modelling able to take each individual grain's shape into account in the style of [KAVA18].

5.3 Far in the future: breakage

A challenge with a significant jump in complexity is the handling of grain breakage, which first of all introduces the problem of non-persistent grains in the tracking of particles, and thus the correct measurement of the kinematics of the fragments of each grain that breaks. The fact that particle sizes reduce while the spatial resolution of the images stays constant also causes problems to the measurement of fabric, even given the potential developments outlined above.

Currently, studies allowing full quantification of breakage are severely limited in the number of particles [CA12, ZWC$^+$15]. Studies with a larger number of particles suffer from the above-mentioned problems of resolution, but are improving in the direction of full fragment tracking [CA14, AAGK$^+$15, OKA$^+$17, KAP$^+$17].

5.4 Conclusion

In conclusion, it appears that for natural sand grains the first two of the three points of micro-mechanical characterisation set out in [CCL97][4], can be safely obtained for grains, however the presence and the properties of grain-to-grain contacts are a significant measurement challenge. This paper has shown that the essential grain-scale ingredients (based on inter-particle contacts) clearly set out in [RTR04] or [FD11b, FD11a] are not easily and directly accessible from 3D images issued from experiments carried out *in-situ* inside an x-ray scanner. Two potential paths to follow for the improvement of these measurements have been outlined, both of which are based around a deeper and more comprehensive use of the fact that the 3D data is *time-resolved*.

References

[AAGK+15] R Alikarami, E Andò, M Gkiousas-Kapnisis, A Torabi, and G Viggiani. Strain localisation and grain breakage in sand under shearing at high mean stress: insights from in situ x-ray tomography. *Acta Geotechnica*, 10(1):15–30, 2015.

[ADAR+14] Khalid A Alshibli, Andrew M Druckrey, Riyadh I Al-Raoush, Taylor Weiskittel, and Nickolay V Lavrik. Quantifying morphology of sands using 3d imaging. *Journal of Materials in Civil Engineering*, 27(10):04014275, 2014.

[AHV+12] E. Andò, S. A. Hall, G. Viggiani, J. Desrues, and P. Bésuelle. Grain-scale experimental investigation of localised deformation in sand: a discrete particle tracking approach. *Acta Geotechnica*, 7(1):1–13, 2012.

[AJDAR16] Khalid A Alshibli, Maha F Jarrar, Andrew M Druckrey, and Riyadh I Al-Raoush. Influence of particle morphology on 3d kinematic behavior and strain localization of sheared sand. *Journal of Geotechnical and Geoenvironmental Engineering*, 143(2):04016097, 2016.

[And13] E. Andò. *Experimental investigation of micro-structural changes in deforming granular media using x-ray tomography*. PhD thesis, Université de Grenoble, 2013.

[CA12] MB Cil and KA Alshibli. 3d assessment of fracture of sand particles using discrete element method. *Géotechnique Letters*, 2(3):161–166, 2012.

[CA14] MB Cil and KA Alshibli. 3d evolution of sand fracture under 1d compression. *Géotechnique*, 64(5):351, 2014.

[CCL97] F. Calvetti, G. Combe, and J. Lanier. Experimental micromechanical analysis of a 2d granular material: relation between structure evolution

[4] in short: 1. Description of grain and contact positions; 2. Description of the displacements, rotations and evolution of contacts; 3. Description of intergranular forces

and loading path. *Mechanics of Cohesive-Frictional Materials*, 2:121–163, 1997.

[FD11a] Pengcheng Fu and Yannis F. Dafalias. Fabric evolution within shear bands of granular materials and its relation to critical state theory. *International Journal for Numerical and Analytical Methods in Geomechanics*, 35(18):1918–1948, 2011.

[FD11b] Pengcheng Fu and Yannis F. Dafalias. Study of anisotropic shear strength of granular materials using dem simulation. *International Journal for Numerical and Analytical Methods in Geomechanics*, 35(10):1098–1126, 2011.

[HBD+10] S.A. Hall, M. Bornert, J. Desrues, Y. Pannier, N. Lenoir, G. Viggiani, and P. Bésuelle. Discrete and continuum experimental study of localised deformation in hostun sand under triaxial compression using x-ray μct and 3d digital image correlation. *Géotechnique*, 60(5):315–322, 2010.

[KAP+17] Z Karatza, E Andò, S-A Papanicolopulos, JY Ooi, and G Viggiani. Evolution of deformation and breakage in sand studied using x-ray tomography. *Géotechnique*, pages 1–11, 2017.

[KAVA18] Reid Kawamoto, Edward Andò, Gioacchino Viggiani, and José E Andrade. All you need is shape: predicting shear banding in sand with ls-dem. *Journal of the Mechanics and Physics of Solids*, 2018. In press.

[KBB+17] MH Khalili, Sébastien Brisard, Michel Bornert, Patrick Aimedieu, J-M Pereira, and J-N Roux. Discrete digital projections correlation: a reconstruction-free method to quantify local kinematics in granular media by x-ray tomography. *Experimental Mechanics*, pages 1–12, 2017.

[MT90] A Majumdar and CL Tien. Fractal characterization and simulation of rough surfaces. *Wear*, 136(2):313–327, 1990.

[OKA+17] Olumide Okubadejo, Zeynep Karatza, Edward Andò, Laurent Bonnaud, Mauro Dalla Mura, and Gioacchino Viggiani. Identification and tracking of particles undergoing progressive breakage under stress with 3d+ t image analysis. In *Proceedings ICTMS 2017, Lund*, 2017.

[PLB10] Y. Pannier, N. Lenoir, and M. Bornert. Discrete volumetric digital image correlation for the investigation of granular type media at microscale: accuracy assessment. *EPJ Web of Conferences*, 6:35003, 2010.

[RTR04] F Radjai, H Troadec, and S Roux. Key features of granular plasticity. *Granular materials: fundamentals and applications*, pages 157–184, 2004.

[SB89] Charles G Sammis and Ronald L Biegel. Fractals, fault-gouge, and friction. *Pure and Applied Geophysics*, 131(1):255–271, 1989.

[SFA+12] Mohammad Saadatfar, Nicolas Francois, Alon Arad, Mahyar Madadi, Ron Cruikshank, Mehdi Alizadeh, Adrian Sheppard, Andrew Kingston, Ajay Limay, Tim Senden, et al. 3d mapping of deformation in an unconsolidated sand: A micro mechanical study. In *SEG Technical Program Expanded Abstracts 2012*, pages 1–6. Society of Exploration Geophysicists, 2012.

[SVK+02] Frank Stallmach, Corina Vogt, Jörg Kärger, Katrin Helbig, and Franz Jacobs. Fractal geometry of surface areas of sand grains probed by pulsed field gradient nmr. *Physical Review Letters*, 88(10):105505, 2002.

[WAHV17] Max Wiebicke, Edward Andò, Ivo Herle, and Cino Viggiani. On the metrology of interparticle contacts in sand from x-ray tomography images. *Measurement Science and Technology*, 2017.

[ZWC+15] B Zhao, J Wang, MR Coop, G Viggiani, and M Jiang. An investigation of single sand particle fracture using x-ray micro-tomography. *Géotechnique*, 65(8):625–641, 2015.

Landslide mechanics and growth of slip surfaces

Alexander M. Puzrin

Institute for Geotechnical Engineering, ETH Zurich

Quantification of the phenomenon of growing slip surfaces in soils is critical for understanding progressive and catastrophic failure in geotechnical problems. A large variety of landslide mechanisms is affected by these phenomena. The chapter introduces simple criteria for catastrophic and progressive slip surface propagation using energy balance and process zone approaches and provides their experimental validation. Progressive slip surface propagation is then used together with the landslide pressure calculations to identify the landslide influence zone and to predict evolution of the compression zone in constrained terrestrial landslides. For submarine landslides, catastrophic slip surface propagation allows explaining their large dimensions, while progressive propagation helps to understand the mechanisms of their post-failure evolution and the resulting geomorphological patterns.

1 Introduction

Progressive failure in landslides has been traditionally associated with long-term stability of slopes in overconsolidated clays [Ske64, Bje67, Bur77, Var02], although the short-term stability can be affected as well. Progressive failure in normally consolidated clays has also been documented [*e.g.,* Ber89]. Many theories developed to explain the progressive failure phenomenon are concerned with definition of shear resistance of soil and assume that the slip surface appears instantaneously along the entire length of the landslide.

Experimental studies of strain localization pioneered in clays by [Mor67] and in sands by [Var81] demonstrated formation of the slip surfaces (shear bands) with thickness of about 200 and 20 mean grain diameters, respectively. To describe strain localization, [Rud75] proposed a bifurcation condition obtained by combining the constitutive relationship with equilibrium and compatibility conditions on assumed localized shear band. The great advantage of the [Rud75] approach is its ability to account for elasto-plastic soil behavior and to provide the shear band orientation and

the displacement discontinuity. These solutions are applicable to granular soils, and for a long time this approach dominated the literature related to quasi-static strain localization analysis in soils [review by Var95]. The next step would be to model its gradual propagation observed in both normally and overconsolidated clays. [Ske64] and [Bje67] suggested that fracture mechanics concepts might throw light on progressive failure. Palmer and Rice [Pal73], proposed an approach for analysis of the growth of localized shear bands in the progressive failure of overconsolidated clay. An important advantage of this approach is that the shear band *evolution* is treated as a true physical process. This is achieved by a proper mechanical consideration of the end (process) zone of the shear band and by utilization of an energy based criterion for its propagation. In reality, this evolution may have stable and unstable stages, so that an important difference between the progressive and catastrophic types of failure can be distinguished.

This chapter introduces criteria for progressive and catastrophic growth of slip surfaces and demonstrates how these criteria can be used to improve our understanding of certain mechanisms of terrestrial and submarine landslides.

2 Criteria for the slip surface growth

The problem formulated below is closely related to the problem of a long slope inclined by angle α to the horizontal, into which a step of height h is cut, causing the shear band of the length l to propagate upward from the base of the cut, parallel to the slope surface (figure 1a). This problem was analyzed approximately by [Pal73] using the J-integral with the assumption that h and ω are small in comparison to l. Here ω is the size of the end zone near the tip, beyond which the shear resistance τ is essentially equal to residual shear strength τ_r.

While applicable to overconsolidated clays, the problem illustrated by figure 1a is not really meaningful for sands and normally consolidated clays, because the stresses developed in the sliding part of the slope are tensile and cannot be sustained by soils with no tensile strength. In order to address these limitations, [Puz05] considered a related problem (figure 1b) where all the stresses are compressive.

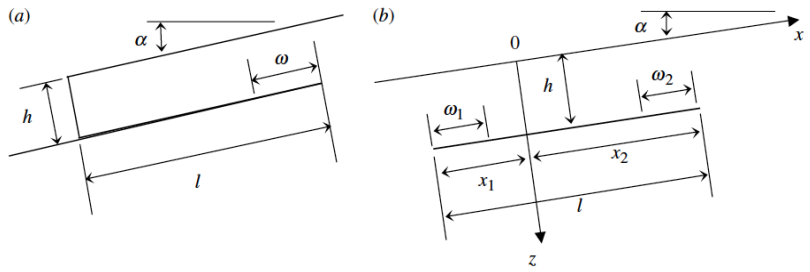


Figure 1: Shear band propagation caused by: a) a cut in the slope; b) discontinuity parallel to the slope surface.

Following [Puz05], we consider a thin linear discontinuity that is parallel to the slope and its length l is sufficiently larger than both its depth h within the slope and the length ω of its end zones (figure 1b). Apart from the two small end zones, the shear resistance τ along this discontinuity drops to its residual value τ_r. At the very tips of the shear band as well as at any point outside the band, the shear resistance is equal to its peak value τ_p. We are interested in conditions, under which the initial shear band will grow parallel to the slope surface. This will allow for a proper identification of catastrophic and progressive types of soil failure.

2.1 Theoretical approaches

Extension of the [Pal73] approach to a wider variety of soils requires for non-elastic soil properties (e.g., hardening plasticity, strain softening, zero tensile strength, active and passive failure modes, etc.) to be taken into account. This section demonstrates how the energy balance and process zone approaches can be applied to the simple problem of the shallow shear band propagation in an infinite slope built of such a soil.

Energy balance approach

The energy balance approach requires that the energy surplus produced in the body by incremental propagation of the shear band should exceed the work required for this incremental propagation. Mathematically this can be expressed as the following inequality:

$$\Delta W_e - \Delta W_i - \Delta D_l > \Delta D_w \qquad (1)$$

where ΔW_e is external work made in our case by gravitational forces on downslope movements of the layer above the shear band; ΔW_i is the internal work made by the normal stress acting parallel to the slope surface on deformations of the layer caused by changes in these stresses; ΔD_l is the plastic work dissipated on the shear band, which is required to overcome the residual shear resistance along the band; ΔD_w is the plastic work dissipated in the shear band during its propagation, which is required to overcome the shear resistance in excess of residual in the end zones of the band.

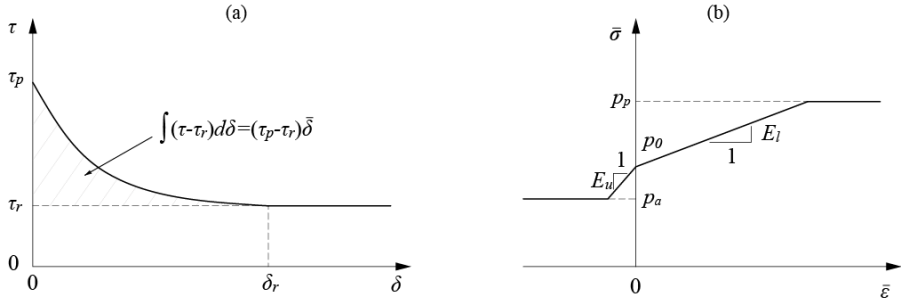


Figure 2: Constitutive behaviour: (a) in the shear band; (b) in the sliding layer.

[Puz05] used the energy balance approach to demonstrate that in an infinite slope (figure 1b) with inclination angle α, where:

- the slip surface is strain softening (figure 2a), with undrained shear strength decreasing from peak τ_p to residual τ_r over the relative displacement δ_r ,
- the sliding layer is elasto-plastic (figure 2b) with different plane strain deformation moduli (E_l and E_u) and passive/active failure criteria (p_p and p_a) in loading and unloading, respectively,

the shear band will propagate catastrophically (*i.e.*, under existing external forces) at the depth h parallel to the slope surface once its initial length l_0 exceeds the critical value of:

$$l_0 \geq l_{cr} = \left(1 + \sqrt{\frac{E_l}{E_u}}\right)\frac{l_u}{r} \tag{2}$$

where

$$l_u = \sqrt{\frac{2E_u\bar{\delta}h}{\tau_p - \tau_r}};\ r = \frac{\tau_g - \tau_r}{\tau_p - \tau_r};\quad \tau_g = \gamma' h\sin\alpha \tag{3}$$

are the characteristic length, shear stress ratio and gravitational shear stresses, respectively,

$$\bar{\delta} = \frac{\int_0^{\delta_r}(\tau - \tau_r)d\delta}{\tau_p - \tau_r} \tag{4}$$

is the characteristic displacement, proportional to the area below the softening curve, representing the energy dissipated in the process zone of the shear band. This parameter is critical for the analysis of the SBP: the larger is $\bar{\delta}$ – the more work has to be done to propagate the band. Because in the energy balance approach the length of the end zone is neglected, the above equations apply equally to different shapes of the degradation curve.

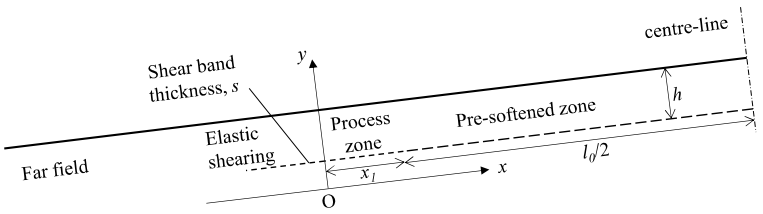


Figure 3: Identification of shear zones in the process zone approach.

For a finite length of the end zone, however, the critical length of the shear band will be affected by the shape of the degradation curve. [Zha15] analyzed such a case using a process zone approach for a simplified symmetric problem in figure 3. If there were no discontinuity, the slope down to the depth h would be subjected to an initial internal lateral force per unit thickness (P_g) caused by earth pressure, and an initial shear stress (τ_g) caused by the soil weight and slope angle. Note, all 'forces' referred to in the following are forces per unit (out of plane) thickness. Conceptually,

the problem is explored by fixing a length, l_0, of initial failure zone, where the strength has been reduced to its residual value, τ_r, along its entire length, flanked by the material where the softening has not yet been completed. Then the slope is brought to catastrophic failure by increasing the gravity loading such as might occur due to sedimentation or increase in slope angle caused by diapirism.

The initial failure zone within the slope will disturb the stress and strain fields in its vicinity for any level of gravity loading, τ_g, that exceeds the residual shear strength, τ_r, in this zone. As the relative magnitude of τ_g increases (relative to the peak and residual strengths outside the initial failure zone, τ_p and τ_r), first a region of elastic shear and then the process zone will start to grow adjacent to the initial failure zone, with plastic deformations causing partial softening. As it does so, the internal lateral force P will increase above P_g within the elastic and process zones and the leading half of the pre-softened zone. In general, the x coordinate at the end of the process zone (start of the pre-softened zone) is taken as x_1, which approaches a value ω at the critical condition of catastrophic failure.

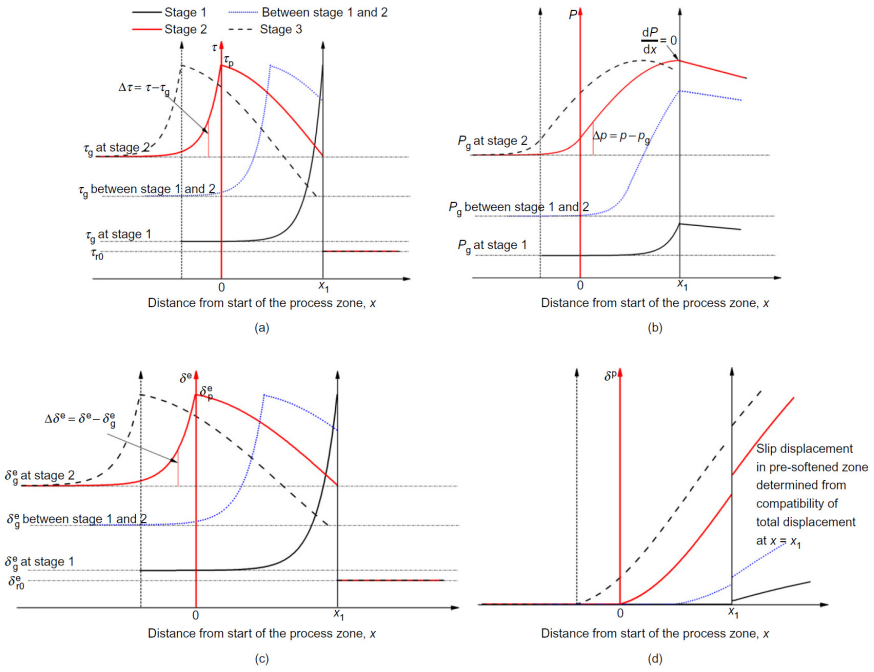


Figure 4: Profiles of key variables at different stages of failure: (a) shear stress; (b) lateral force; (c) elastic shear displacement within shear band; (d) plastic shear displacement within shear band.

Figure 4 shows profiles of key variables (shear stress τ, P, δ^e and δ^p) at three stages of gradually increasing gravity shear stress, τ_g: Stage 1, where the peak strength had just been mobilised due to elastic deformations in the shear band immediately adjacent to the pre-softened zone; Stage 2, at the onset of catastrophic failure, where the lateral force, P, in the material above the shear band has reached a maximum value, P_{max}, at the interface between the pre-softened and process zones; and Stage 3, just after catastrophic failure where additional displacement has caused the shear strength of the shear band in the process zone to degrade below τ_g and towards the residual value at the interface with the pre-softened zone. For convenience with later algebra, the origin, O, of the Cartesian coordinates system in figure 3 has been placed at the interface between the process zone and the region of purely elastic deformations, and thus moves as the gravity shear stress increases.

According to the process zone approach [Zha15], in case of the linear decay of shear strength in figure 2a, the shear band will propagate catastrophically once its initial length l_0 exceeds the critical value of:

$$l_0 \geq l_{cr} = 2(1-r)\frac{l_u}{r} \tag{5}$$

In case of the exponential decay, catastrophic propagation of the shear band will take place once its initial length l_0 exceeds the critical value of:

$$l_0 \geq l_{cr} = 2\sqrt{1-r(1-\ln r)}\frac{l_u}{r} \tag{6}$$

In summary, three stages of progressive failure of a clay slope with a softenable shear band may be distinguished (figure 5): 1) purely elastic shearing; 2) development of the process zone; 3) catastrophic failure. The linear and exponential decay criteria, as well as the energy balance criterion PG_1 (equation (2)) are plotted for comparison. For detailed discussion of these results see [Zha15].

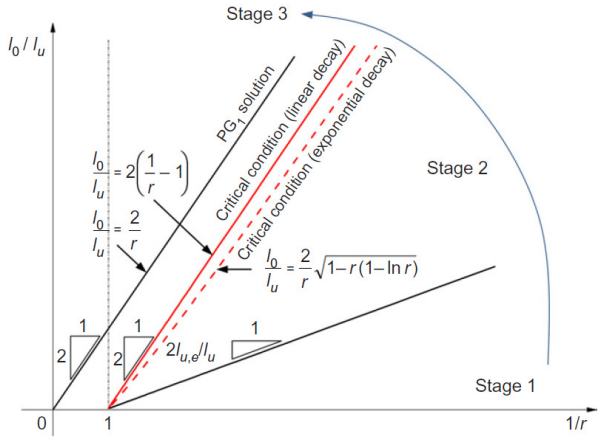


Figure 5: Distinguishing features of the three failure stages in l_0/l_u vs $1/r$ plane.

2.2 Experimental validation

In order to obtain the experimental evidence of the SBP in conditions close to observed submarine landslides, a testing setup has been built at the ETH Zurich [Puz16]. A 4 cm thick layer of saturated kaolinite clay has been consolidated in a 2 m long and 25 cm wide chute with steel bottom and glass walls (figure 6a). The slope was then inclined by 10° and a miniature 'bulldozer' applied a gradually increasing static force at the top of the sliding layer, which was measured using a set of three load-cells. Displacements of the 'bulldozer' were measured by a laser displacement sensor. This experiment provided for the first time direct measurements of the evolving length of the shear band. The measurements were performed using a swept wavelength interferometry fiber optic strain sensing technology by installing two optical fibers parallel to the slope in the sliding layer with the help of plate micro anchors, spaced at 1 cm distance from each other, resulting in 400 individual high precision strain gages along the slope. The cables were pre-tensioned and the compressive strain distribution in the sliding layer could be measured for different positions of the 'bulldozer' (figure 6b).

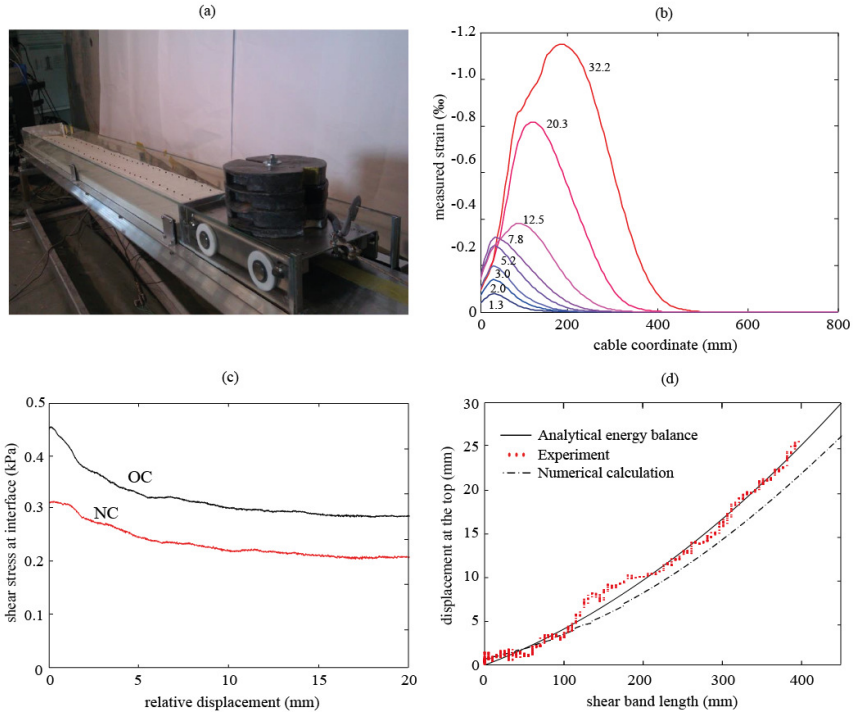


Figure 6: Shear band propagation in a kaolinite clay slope: (a) experimental setup; (b) evolution of measured compressive strains in the sliding layer for different levels of applied force (in N); (c) strain softening in the interface tests; (d) evolution of the shear band length.

Closer to the "bulldozer", reliable strain measurements can be taken only until the compressive strain does not exceed the pre-tensioning. Fortunately, in order to capture the SBP propagation the relevant measurements are those further away from the "bulldozer". Indeed, because the layer only gets compressed where the shear band has already propagated, its length is determined by the length of the strain profile in figure 6b, identified with the unprecedented spatial resolution of 1 cm.

Analytical and numerical investigations of test conditions required soil parameters, which were obtained from independent interface shear, ring shear, oedometer and chute stiffness tests. The interface tests are here of particular interest, showing softening behaviour at the contact between the sliding layer and the steel of the chute for both normally and lightly overconsolidated (OCR=3) clays (Figure 6c). These softening curves have been used directly as a constitutive relationship for the interface behavior in ABAQUS. Numerical simulation of the evolution of the length of the shear band propagating progressively at the bottom of an elastic layer (inclined by $10°$) is shown by a dashed line in Figure 6d as a function of the 'bulldozer' displacements.

The same test has been simulated analytically using the energy balance approach, where the characteristic displacement $\bar{\delta}$ was calculated from the area below softening curves in the figure 6c. The length L_{sb} of the propagating shear band is given as a function of the stress ratio *r* defined from equation (3) and of the pressure p_b, applied by the „bulldozer" at the top of the slope:

$$L_{sb} = -\frac{p_b - p_0}{\tau_p - \tau_r}\frac{h}{r} + \frac{1}{r}\sqrt{\frac{2E_l h\bar{\delta}}{\tau_p - \tau_r}} \tag{7}$$

where p_0 is the initial lateral pressure (at rest). The "bulldozer" displacement is given by

$$u_b = \int_0^{L_{sb}} \varepsilon_x(x)dx = \frac{1}{E_l}\left(p_b L_{sb} - (\tau_r - \tau_g)\frac{L_{sb}^2}{2h}\right) \tag{8}$$

and eliminating p_b between equations (7) and
(8) gives the parabolic relationship between u_b and L_{sb} shown as the solid line in figure 6d. Analytical and numerical predictions in figure 6d match within 10%, but more importantly, they both fit reasonably well the measured shear band length (data points in figure 6d).

This example provides both a direct experimental evidence of the SBP in a slope built of saturated, normally to lightly overconsolidated clay and a confirmation that energy balance approach is capable of its quantification.

2.3 Progressive vs catastrophic failure

The above experiment simulates progressive SBP, *i.e.*, under gradually increasing external forces (in this case the weight of the 'bulldozer', figure 7a). This was achieved intentionally, by choosing the $10°$ inclination of the chute, which was not sufficient for the gravitational shear stress τ_g to exceed the residual shear strength τ_r

(*i.e.*, $r < 0$). Under such conditions propagation of the shear band can be terminated at any moment by keeping the bulldozer load constant, allowing for more accurate measurements of the shear band length.

By increasing inclination of the chute and making the gravitational shear stress larger than the residual shear strength (*i.e.*, $0 < r < 1$), the catastrophic SBP could be recorded (figure 7b). After exceeding the critical length, the catastrophic SBP took place under practically constant 'bulldozer' force and was around 30 times faster than the progressive SBP, which, as mentioned above, was fully controlled by the rate of increase of the 'bulldozer' force. The measured velocity of the catastrophic shear band propagation was, as expected, of the order of the shear wave velocity.

In summary, the shear band propagation phenomenon allows for the following types of failure to be identified. In *progressive failure*, propagation of the shear band is stable in the sense that it requires work of the additional external forces. In *catastrophic failure*, propagation of the shear band is unstable, and takes place under existing external forces. The term *delayed failure*, often used in the geotechnical, is in fact a particular case of catastrophic failure, in which propagation of the shear band is unstable, but delayed in time owing to, for example, local pore water suctions caused by dilation within the shear band, or viscoelastic creep of soil and bulk pore water diffusion.

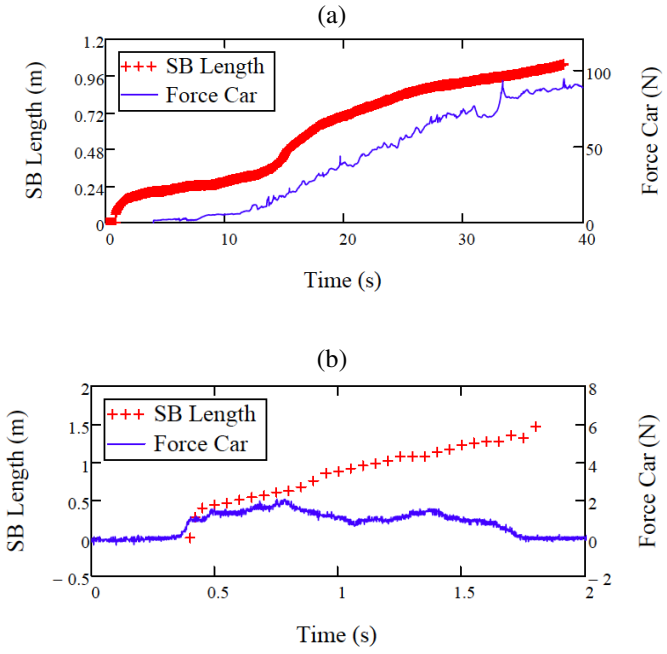


Figure 7: Evolution of the applied force and the shear band length in time for (a) progressive and (b) catastrophic failure.

3 Progressive failure in constrained terrestrial landslides

Understanding of failure mechanisms of creeping terrestrial landslides is of critical importance for assessment and mitigation of their hazard. In general, when a creeping landslide is constrained by an obstacle, either natural (a milder slope or a rock outcrop) or artificial (a retaining wall), it slows down, creating an impression of being stabilized. Sometimes, however, the stabilization phase is followed by acceleration, caused by the soil failure at the obstacle. At what pressure will the soil fail? What causes the failure in the compression zone above the obstacle? Is it safe to excavate below the failed zone? These and other questions can be answered by considering combination of the kinematic failure mechanisms in the sliding layer with progressive growth of slip surfaces underneath the sliding layer.

3.1 Landslide pressure

The problem of the limiting landslide pressure on an obstacle was first formulated in 1944 by Robert Haefeli of ETH Zurich, who recognized that the kinematics of the problem does not allow for classical active and passive earth pressure theories to be applied (figure 8). He derived an approximate solution using a limit equilibrium approach with a number of rather arbitrary assumptions and simplifications. Since then, the Haefeli solution has been widely applied for the design and analysis of landslide retaining structures. [Fri17] revisited this old landslide pressure problem (figure 9) by means of a rigorous upper- and lower-bound limit analysis and derived the exact landslide pressure solution for a planar landslide with a weak slip surface parallel to the slope ($\alpha = \theta$) and a vertical retaining structure in cohesionless soil:

$$K_{lhw} = \frac{\sigma_x}{\gamma z} = \frac{2E_{lhw}}{\gamma H^2} = \frac{\cos^4 \alpha}{\cos^2 \varphi'} \cdot \left(1 + \sqrt{1 - \cos^2 \varphi' (1 + \tan^2 \alpha)}\right)^2 \qquad (9)$$

The landslide pressures from this solution increase with the strength of the sliding layer and are significantly higher than the active, but much lower than the passive, earth pressures. Of even higher practical importance, however, is that due to their oversimplifying assumptions, the widely used approximate solutions appear to get close to the exact solution only over a very narrow range of slope and friction angles (figure 10a). It appears that for mildly inclined weak slip surfaces and high strengths of the sliding layer, analysis and design of retaining structures based on well-known approximate solutions can become dramatically unsafe.

Effect of the ground water on the total normalized landslide force were demonstrated by [Fri17] in figure 10b. The lower and upper bound solutions are no longer identical and, therefore, they do not represent the exact solution. However, it can be concluded that the presence of water will always weaken the sliding body and the force needed to fail the soil in the vicinity of a constraining obstacle will always be smaller than the one for a landslide in dry condition.

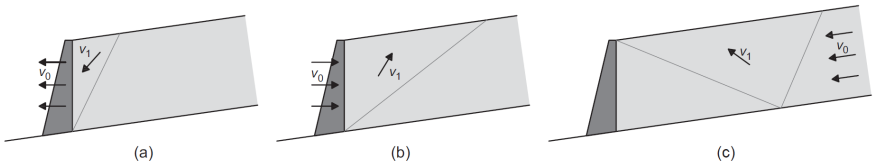


Figure 8: Kinematics of different limiting earth pressure situations: (a) active; (b) passive; (c) the landslide case.

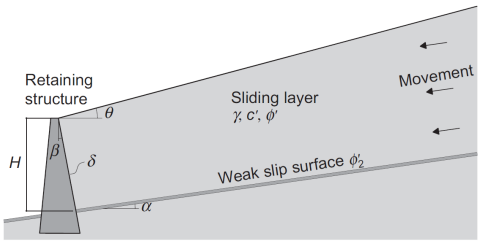


Figure 9: Formulation of the problem of landslide pressures.

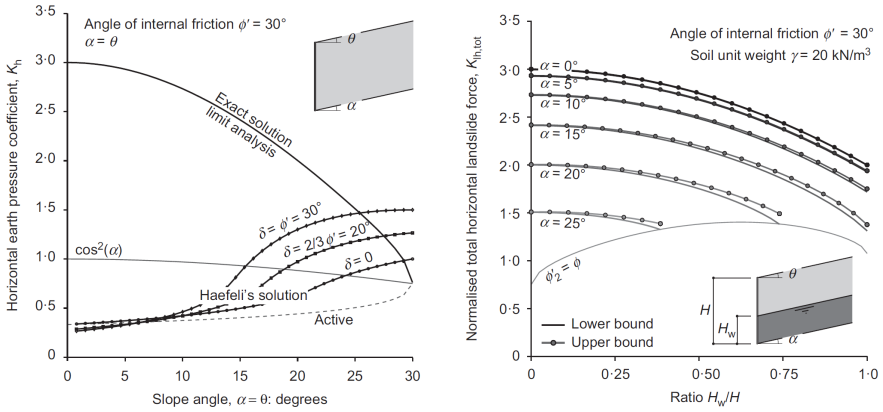


Figure 10: Landslide pressure coefficients: (a) comparison between the Haefeli and exact solutions; (b) effects of the ground water.

3.2 Landslide influence zone

Continuing urbanization brings construction into landslide prone areas. Since 1940s it has been understood that earth pressures acting on the structures within the sliding body may significantly differ from those in the stable slopes with rigorous solutions for landslide pressures presented by [Fri17]. Design codes reflect this knowledge and recommend application of landslide pressures for construction in permanent landslides. It is less clear, however, what kind of design pressures should be used outside of the sliding body, in the proximity of the existing landslides? This question is relevant for both permanent and one-time events and of particular practical interest is the area just below the boundary of the high-pressure landslide compression zone. Intuitively, there should be a transition zone, within which high landslide pressures subside to the regular (at rest) earth pressures in the stable slope. How large is this "landslide influence zone" and what are the lateral earth pressures acting within it (figure 11)?

Understanding the mechanism of the downhill pressure transfer within the stable part of the slope below the landslide is critical for finding the answers to these questions. Progressive propagation of a slip surface into the stable part of the slope (figure 6) has been investigated experimentally and analytically in Section 2.2 of this Chapter. Once the steeper part of the slope fails, the lateral earth pressure in the failed zone remains constant and is equal to the landslide pressure from equation (9). The slip surface stops growing and its length is determined from equation (7) with p_b equal to the landslide pressure. The earth pressure at the tip of the slip surface can be calculated, but not below that, because equation (7) has been derived using the energy balance approach, i.e., neglecting the process zone.

To overcome these limitations, [Puz19] used the process zone approach (figure 3), similar to the one of [Zha15], considering lateral pressure distributions below the shear band, in the process zone and the zone of elastic shearing. This allowed for estimating both the length of the "landslide influence zone" (figure 12a) and the elevated earth pressures acting on the buildings embedded within this zone (figure 12b). This becomes particularly important when a building in this area has to be replaced by a new one, raising questions about the magnitude of earth pressures that should be used in design of both the excavation and the structure.

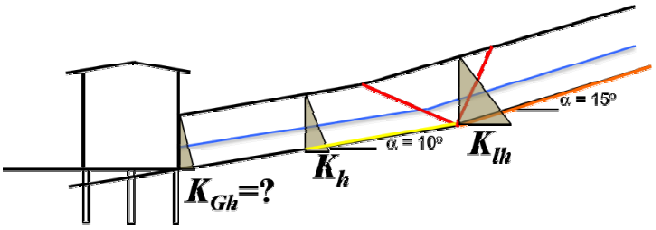


Figure 11: Progressive propagation of the slip surface into a stable part of the slope below the failure in the upper unstable part of the slope .

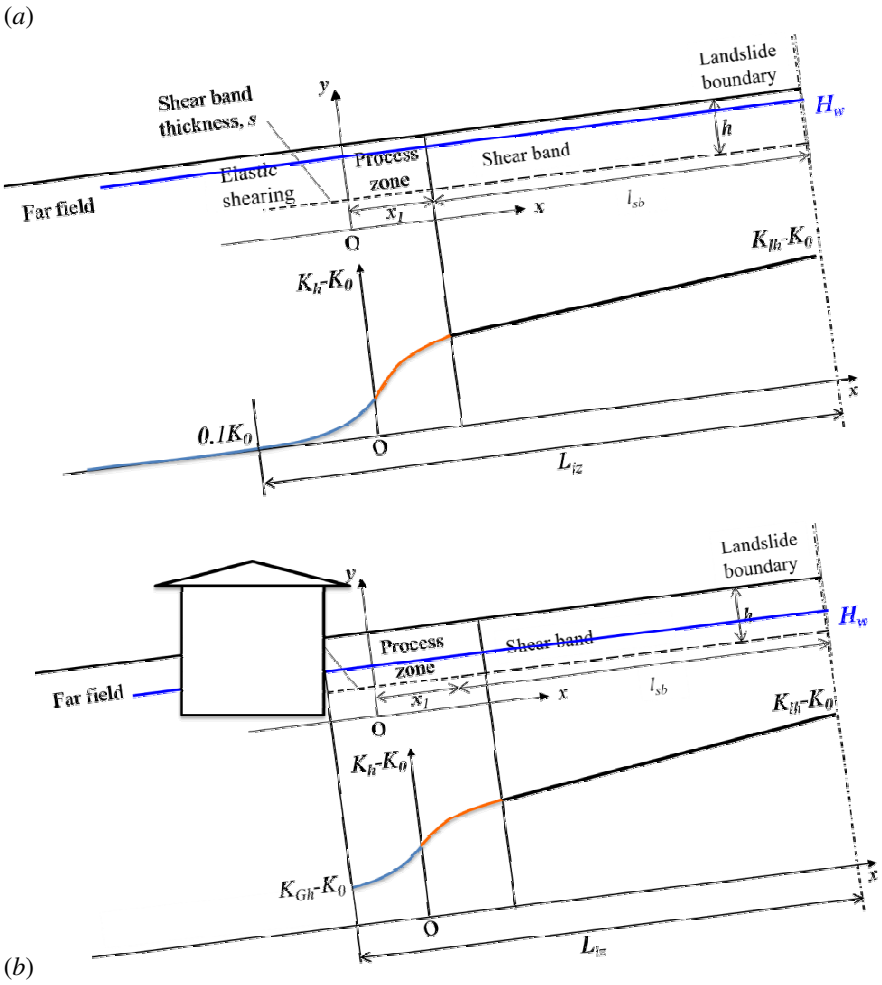


Figure 12: Process zone approach for estimating (a) the length of the "landslide influence zone"; (b) earth pressures acting on the buildings within this zone.

3.3 Evolution of the landslide compression zone

The ski resort town of St Moritz, Switzerland, is partially constructed on a large creeping landslide (figure 13a), which has been causing damage to buildings and infrastructure. At the town centre, the landslide is constrained by a rock outcrop, creating a compression zone in the sliding mass. After decades of gradual slowing down, in the beginning of 1990s the landslide started to accelerate (figure 13b), in spite of the fact that the average yearly precipitation remained fairly constant (figure 13b). [Puz11] demonstrated that a constrained creeping landslide experiences the

progressive propagation of the slip surface, driven by the visco-elastic deformation of the sliding layer (figures 14a and 14b). This results in significant earth pressure increase in the compression zone, until the landslide pressure described by equation (9) is reached, with subsequent failure and visco-plastic yielding of soil in this zone (figure 14b).

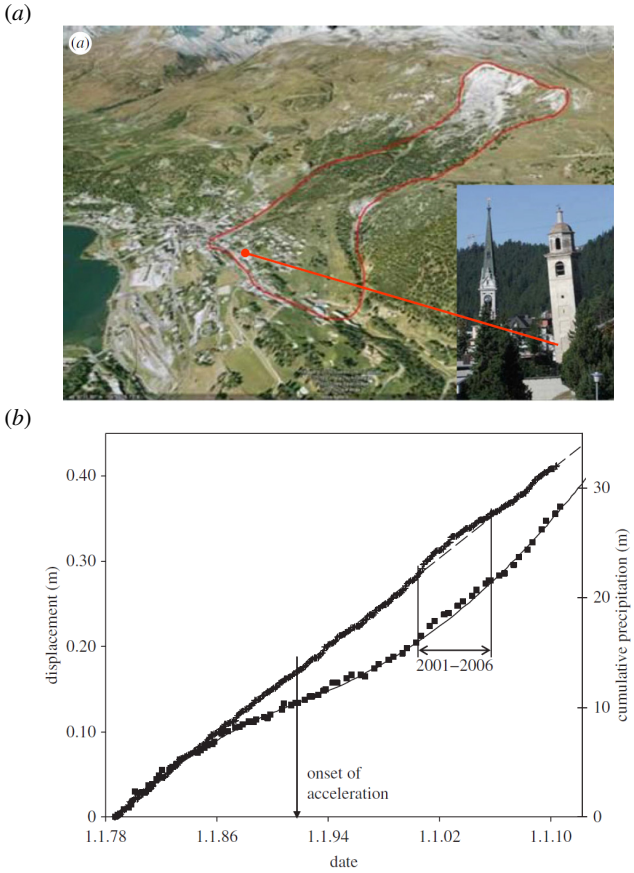


Figure 13: St Moritz landslide: (a) the outline; (b) cumulative precipitation and evolution of displacements within the compression zone.

This basic physical mechanism, relying on extensive laboratory and field tests and long-term displacement monitoring, explains the paradox of the St Moritz landslide acceleration (solid line in figure 13b). Although the model predicts that the landslide could eventually slow down [Puz11], reaching certain constant velocity, its displacements may become excessive for some buildings, requiring an early warning system and further stabilization of the historic Leaning Tower (inset in figure 13a). In general, by predicting the onset of yielding, the model can provide an important timeframe for stabilization of constrained landslides.

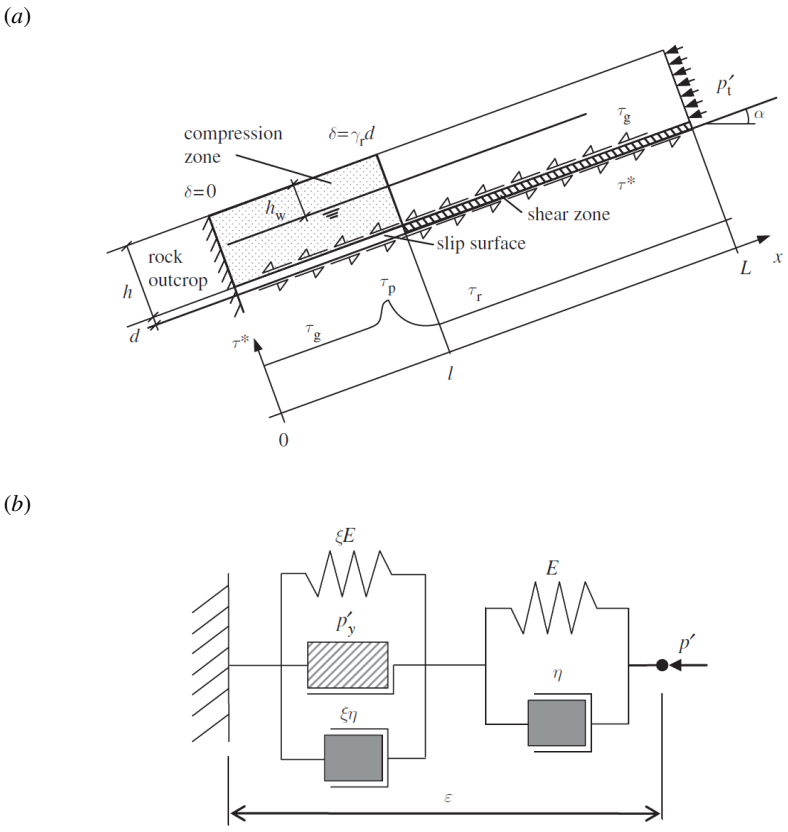


Figure 14: Conceptual model: (a) of the shear band propagation; (b) of the visco-elastic-plastic behaviour of the soil in the compression zone.

4 Catastrophic failure in gigantic submarine landslides

A Geographical Informational System (GIS)-based deterministic and probabilistic slope stability analysis (PSSA) for offshore developments requires slope stability calculations to be repeated millions of times, thus excluding possible use of finite element analysis and relying mainly on analytical limiting equilibrium (LE) slope

failure criteria. In strain softening soils, however, LE approach assumes that the failure takes place simultaneously along the entire sliding surface where the shear stress exceeds the peak shear strength. It cannot explain the failure of the parts of the slope, where the shear stress is lower than the peak shear strength, frequently observed in gigantic submarine landslides. There is a clear need in an alternative approach, which would allow for overcoming this limitation while maintaining the simplicity of the LE analysis.

A potential candidate to fill this gap is the shear band propagation (SBP) approach [Pal73; Puz05; Zha15], which in contrast to the LE provides criteria for an initial slip surface, in which the shear stress exceeds the peak shear strength, to propagate into portions of the slope, where the shear stress is lower than the peak shear strength but exceeds the residual strength. This helps to explain dimensions and evolution of some terrestrial and submarine landslides in the nature.

4.1 Geomorphological patterns

Besides its inability to explain the gigantic dimensions of the observed submarine landslides, LE approach also cannot distinguish between different landslide failure modes, such as slab failures, spreadings, ploughings and runouts (figure 15).

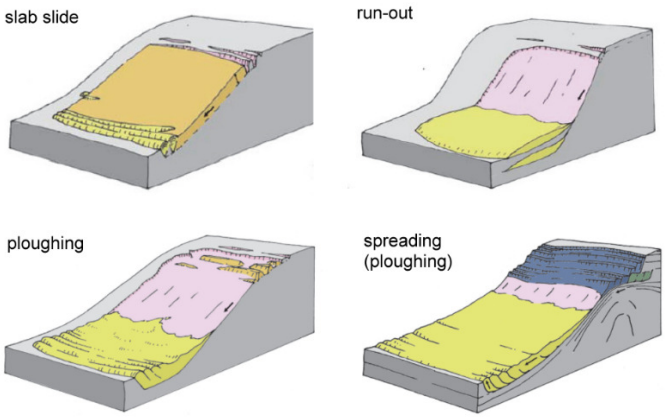


Figure 15: Geomorphological patterns of submarine landslides

4.2 Shear band propagation in a non-linear slope

The existing SBP approaches have been developed for a case of an infinite or semi-infinite slope. In submarine settings, however, slopes can be typically characterised by monotonically decreasing functions (figure 16), with many slopes exhibiting an

'S-shaped' Gaussian profile, formed as a result, for example, of tectonic regime, a drape-covered scarp or a pile of debris from a previous landslide. In all these cases, it is logical to assume that the initial weak zone would appear parallel to the slope in its steepest part, where the combined action of the gravity and seismic loads could overcome the degraded peak shear resistance of the soil.

For this 2D slope geometry, [Puz16] proposed to treat the submarine landslide evolution as a continuous sequence of catastrophic and progressive SBP mechanisms (figure 16). The SBP mechanism is capable of explaining the failure evolution from a relatively short initial shear band, triggered (e.g. by an earthquake) in the steepest part of the slope, where gravitational and seismic forces exceed the peak shear strength (shear stress ratio $r > 1$, primary failure in figure 16). If this initial shear band becomes sufficiently long, it can propagate catastrophically parallel to the slope surface, into those parts of the slope where gravitational and seismic forces exceed the residual shear strength ($0 < r < 1$), triggering a slab failure (secondary failure in figure 16). Once the slab fails in active or passive mode at its ends, this causes changes in the seabed level, driving progressive propagation of the shear band into those parts of the slope where gravitational and seismic forces are smaller than the residual shear strength ($r < 0$) and triggering spreadings, ploughings and run-outs (tertiary failure in figure 16).

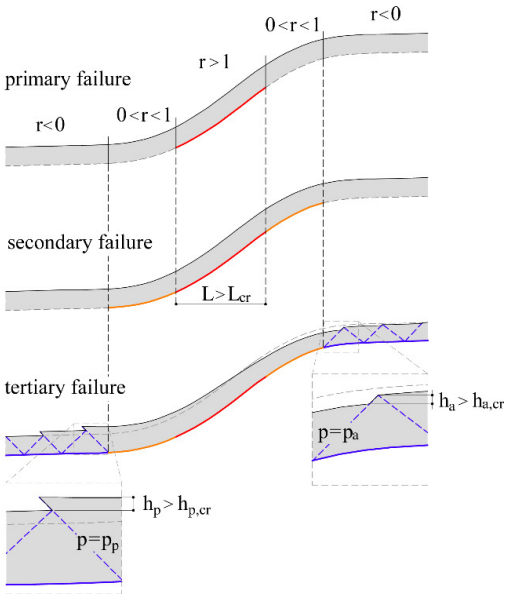


Figure 16: Evolution of submarine landslides.

Although the energy balance calculations using equilibrium in curvilinear coordinates are rather cumbersome [Puz15], their outcome is surprisingly simple. It follows that the one-dimensional (1D) SBP criterion in equation (2) for an infinite

slope also gives a reasonably conservative estimate for a 2D topography. The only difference is that the 2D criterion uses an average shear stress ratio calculated based on the average gravitational shear stress, which in turn is calculated using the average slope of the initial shear band (figure 17):

$$l_0 \geq l_{cr} = \left(1 + \sqrt{\frac{E_l}{E_u}}\right)\frac{l_u}{\bar{r}} \tag{10}$$

where

$$\bar{r} = \frac{\bar{\tau}_g - \tau_r}{\tau_p - \tau_r}; \quad \bar{\tau}_g = \gamma' h \sin\bar{\alpha} \tag{11}$$

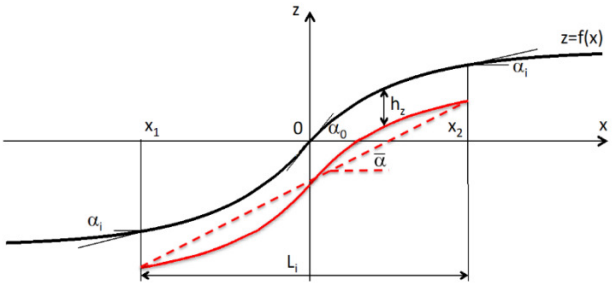


Figure 17: 2D criterion of the SBP in non-linear slopes.

4.3 Post-failure evolution

[Bus19] applied the energy balance kinematic method of plasticity theory to the large deformation problem of initiation and propagation of the spreading and ploughing failure outside a failed slab in submarine sediments (figure 18). The models account for the phenomenon of the progressive propagation of a slope parallel slip surface, which is also quantified using the energy balance approach. In contrast to existing approximate analytical and numerical solutions, the proposed approach provides a theoretical basis for spreading and ploughing criteria as well as the comprehensive dynamic solution of the problem of post-failure landslide evolution. Incremental integration of the derived analytical expressions for kinetic energy in time allows for modelling recurrent initiation of new kinematic failure mechanisms with their subsequent large-scale deformation. Treating the failed slab as well as the spreading and ploughing mechanisms as one composite dynamically evolving mass movement allows for the final post-failure geomorphology of the failed slope to be predicted using basic mechanical principles (figure 19).

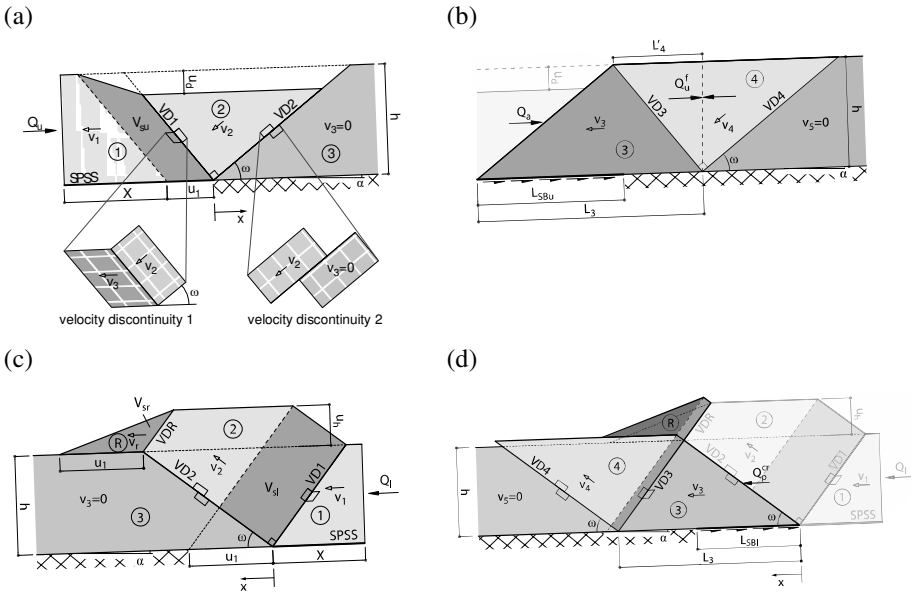


Figure 18: Kinematic mechanisms: (a) initial active failure; (b) spreading; (c) initial passive failure; (d) ploughing.

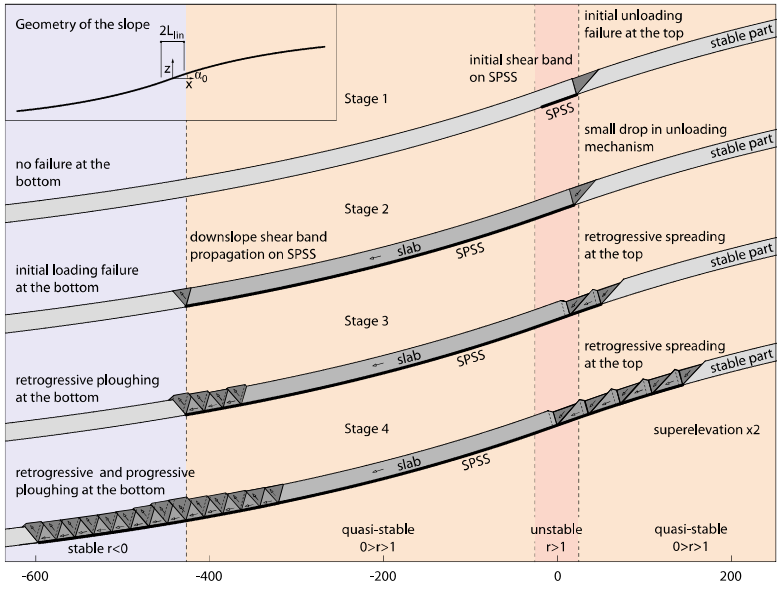


Figure 19: Submarine landslide evolution: numerical integration of analytical kinematic energy balance expressions.

While similar results have been obtained by [Zha19] using Large-Deformation Finite Element (LDFE) (figure 20) and by [Sto19] using Coupled Eulerian Lagrangian (CEL) analysis (figure 21), the [Bus19] approach has significant computational advantages for the GIS based probabilistic stability analysis: it uses analytical criteria to identify spreading, ploughing and run-out and is based on numerical integration of analytical expressions for kinetic energy.

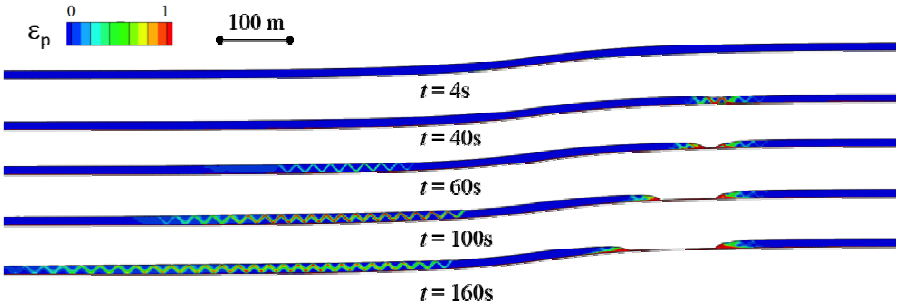


Figure 20: LDFE simulation of ploughing and spreading (plastic strains).

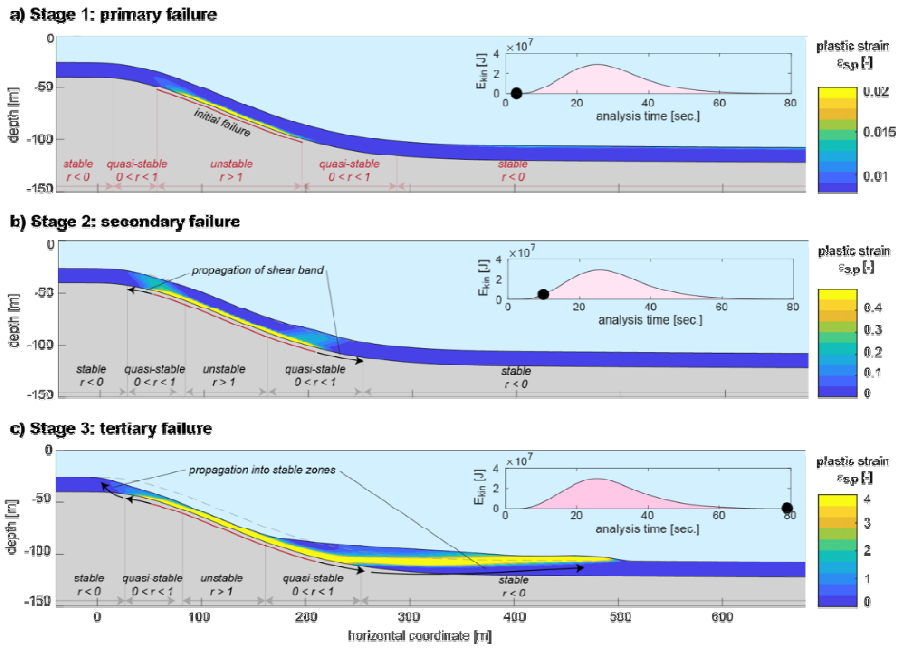


Figure 21: CEL simulation of a run-out (plastic strains and kinetic energy).

5 Conclusions

Quantifying the role of the growth of the slip surfaces in landslide mechanisms does improves our understanding of these important geo-hazards, leading to some interesting practical applications. For constrained terrestrial landslides, it has helped to identify the landslide influence zone and to predict evolution of the compression zone in a number of projects in Switzerland. For submarine landslides, [Rus15] incorporated the SBP approach, into GIS-based deterministic and probabilistic stability analysis and applied it to the area of the Caspian Sea floor designated for a large new oil and gas development. The resulting annual probability of failure predicted by the SBP approach (figure 22) was an order of magnitude higher than the one predicted by the LE approach, approaching the observed historical landslide frequencies and contributing to a more realistic landslide risk assessment. The participants of the ALERT doctoral school are encouraged to think about further potential applications.

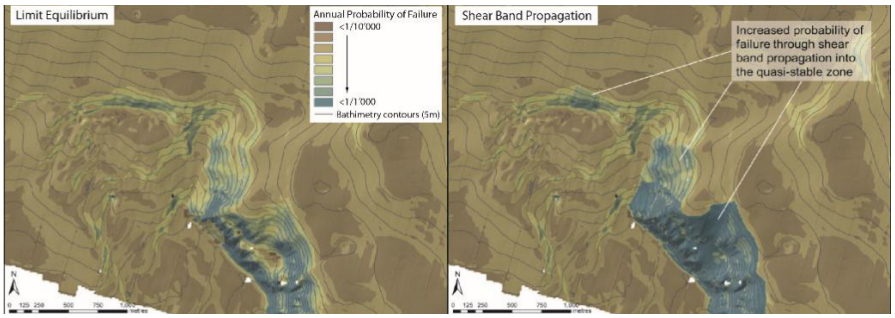


Figure 22: Annual probabilities of failure predicted using: LE (left) and SBP (right) approaches [Rus15], with permission of OTC.

Acknowledgements

The author would like to thank Leonid Germanovich, Mark Randolph, Andreas Schmid, Balz Friedli, Christian Buss, Andreas Stoecklin, Wangcheng Zhang, David Rushton, Tom Gray, Andy Hill and many others with whom he had a pleasure of collaborating on this research.

References

[Ber89] Bernander, S., Gustas, H. and Olofsson, J. Improved model for progressive failure analysis of slope stability. In: *Proceedings of the International Conference on Soil Mechanics and Foundation Engineering, 3, Balkema, Rotterdam, Neth., 1539-1542*, 1989.

[Bje67] Bjerrum, L. Progressive failure in slopes of overconsolidated plastic clay and clay shales. *ASCE Journal of the Soil Mechanics and Foundation Engineering Division, 93(SM5):1-49,* 1967.

[Bur77] Burland, J. B., Longworth, T. I. and Moore, J. F. A. Study of ground movement and progressive failure caused by a deep excavation in Oxford clay. *Geotechnique, 27(4):557-591*, 1977.

[Bus19] Buss Ch., Friedli B., and Puzrin A. M. Kinematic energy balance approach to submarine landslide evolution. *Canadian Geotechnical Journal (in print),* 2019.

[Fri17] Friedli, B., Hauswirth, D., Puzrin, A. M. Lateral earth pressures in constrained landslides. *Geotechnique, 67: 890-905*, 2017.

[Mor67] Morgenstern, N.R. and Tschalenko, J.S. Microscopic structures in kaolin subjected to direct shear. *Geotechnigue, 17(4):309-328*, 1967.

[Pal73] Palmer, C. and Rice, J.R. The growth of slip surfaces in the progressive failure of overconsolidated clay. *Proceedings of the Royal Society of London, A 332: 527-548*, 1973

[Puz05] Puzrin, A.M. and Germanovich, L.N. The Growth of Shear Bands in the Catastrophic Failure of Soils, *Proceedings of the Royal Society of London: Mathematical, Physical and Engineering Sciences, 461(2056):786-795*, 2005.

[Puz11] Puzrin, A.M. and Schmid, A. Progressive failure of a constrained creeping landslide, *Proceedings of the Royal Society of London A, 467(2133): 2444-2461*, 2011.

[Puz15] Puzrin, A. M., Gray, T. and Hill, A. J. Significance of the true non-linear slope geometry for catastrophic failure in submarine landslides, *Proceedings of the Royal Society of London A, Vol. 471(2175): Article No. 2014077*, 2015

[Puz16] Puzrin, A. M., Germanovich, L. N. & Friedli, B. Shear band propagation analysis of submarine slope stability. *Geotechnique, 66(3): 188-201*, 2016.

[Puz19] Puzrin, A.M. Landslide influence zone *(in preparation),* 2019

[Rud75] Rudnicki, J.W. and Rice, J.R. Conditions for the localization of deformation in pressure-sensitive dilatant materials. *Journal of the Mechanics and Physics of Solids, 23: 371-394*, 1975.

[Rus15] David Rushton, Thomas Gray, Alexander M. Puzrin, and Andy Hill. GIS-Based probabilistic slope stability assessment using shear band propagation. *Proceedings of the Annual Offshore Technology Conference, Red Hook, NY: Curran*, 2015.

[Ske64] Skempton, A.W. Long-term stability of clay slopes. Geotechnique, *14: 77-102*, 1964.

[Sto19] Andreas Stoecklin. *Written communication.* 2019

[Var81] Vardoulakis, I., Graf B. and Gudehus G. Trap-door problem with dry sand: A statical approach based upon model test kinematics. *Int. J. Num. Anal. Meth. Geomechanics, 5: 57-78*, 1981.

[Var95] Vardoulakis, I. and Sulem, J. *Bifurcation analysis in geomechanics.* Blackie Academic & Professional, an imprint of Chapman & Hall, Glasgow, U.K., 1995.

[Var02] Vardoulakis, I. 2002. Dynamic thermo-poro-mechanical analysis of catastrophic landslides. *Geotechnique, 52(3):157-171*, 2002.

[Zha15] Zhang, W., Wang, D., Randolph, M. F. & Puzrin, A. M. Catastrophic failure in planar landslides with a fully softened weak zone. *Geotechnique, 65(9):755-769*, 2015.

[Zha19] Zhang, W., Randolph, M., Puzrin, A.M. and Wang, D. (2019). Transition from shear band propagation to global slab failure in submarine landslides. *Canadian Geotechnical Journal, 56(4): 554-569*, 2019.

Strain localization in geomaterials and regularization: rate-dependency, higher order continuum theories and multi-physics

Ioannis Stefanou[1] **and Eleni Gerolymatou**[2]

[1] *Université Paris-Est, Laboratoire Navier, CNRS UMR 8205, IFSTTAR, ENPC, France*
[2] *Chalmers University of Technology, Sweden*

Strain localization is a central topic in geomechanics as it is often related to failure and other important physical phenomena and geological processes.
This chapter is addressed to graduate students and researchers interested in an introduction to strain localization analysis. We present fundamental notions that are frequently met in this topic, such as loss of uniqueness, bifurcation, stability, ill-posedness and mesh dependency.
We first show the inherent pathology of classical, Cauchy, rate-independent continua that leads to mesh sensitivity and we present methods for alleviating/regularizing this problem. These methods involve the use of theories that result in the introduction of characteristic time and length scales into the system. We focus mainly on rate-dependent constitutive laws, Micromorphic continua and multiphysics.
Regularization of strain localization is shown as general as possible using bifurcation and stability analysis and without prescribing exact constitutive relations. One-dimensional examples are then used to illustrate each regularization approach and show in a mathematically simple manner the main results.

1 Introduction

Strain localization is an important phenomenon in geomechanics. From a geometrical point of view, strain localization is related to the creation of (quasi-)periodic geometrical patterns as in figure 1. From an engineering point of view strain localization is related to failure. For instance, failure of a retaining wall happens through the localization of strain at the slip surface. At a larger scale, landslides or even earthquakes occur due to localized, intense shear deformation in a narrow zone of millimetric to centimetric scale.

Figure 1: Network of quasi-periodic compaction bands (see paragraph 2.3 for definition) at Valley of Fire, Nevada, USA (date of photo: 25/12/2015).

In this chapter we give the fundamental tools for studying a) the conditions for which strain localization takes place and b) its type. We focus mainly on deformation bands, which are frequently observed in nature and in engineering applications. However, the methodology that is developed can be applied for more complex patterns of strain localization as is for example the checkerboard pattern shown in figure 2, diffusion and/or run-away modes (e.g. [VS95, ND11, BSS17]).

The methodology we follow is based on bifurcation and stability theory of dynamical systems. This theory gives a unified, general and rigorous mathematical framework for studying strain localization in solids. It is worth emphasizing that, despite the various theoretical and mathematical complications related to constitutive modeling and multiphysics couplings (see [LCBD09, Bow09] for some good references in continuum mechanics and constitutive modeling), once the equations for the (dynamical) system are established, bifurcation (and stability) analysis is a standard methodology.

In Section 2 we give the necessary definitions of common terms that are often found in the literature when studying strain localization. We explain what is loss of uniqueness, bifurcation, stability, ill-posedness and mesh dependency and we emphasize their differences. We then focus on deformation bands, their various types and we discuss the necessary conditions for their triggering (onset of localization). Next we use the first Lyapunov method in order to derive qualitative estimations regarding the thickness of the deformation band and its evolution. Furthermore, we explain why in rate-independent Cauchy continua strain localization occurs on a mathematical plane

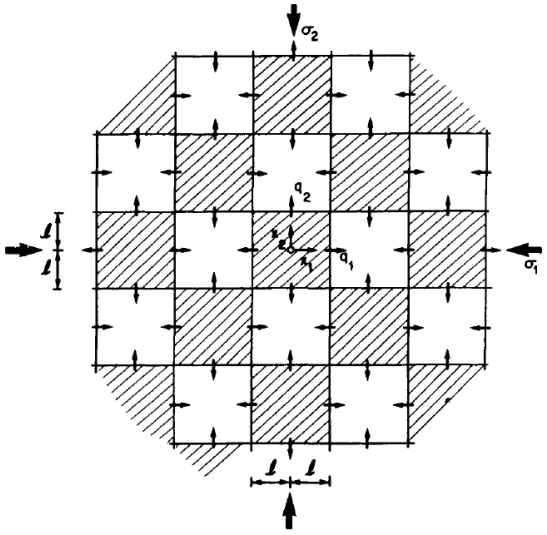


Figure 2: Checkerboard pattern of dilating and contacting cells in water-saturated granular medium (source: [VS95])

.

(deformation band of zero thickness), which is a mathematical and physical artifact when experimental evidence is taken into account. As a consequence we show why mesh dependency takes place when the finite element method is used for solving strain localization problems.

In order to overcome the aforementioned mathematical and numerical artifacts several approaches have been proposed in the literature for *regularization*. Here we explore the following three regularization techniques. First, we present rate-dependent Cauchy continua (Section 3), their ability to regularize the problem and their limitations. Scale analysis is performed showing the characteristic time scale that rate-dependent Cauchy materials introduce to the strain localization problem. In this way we highlight the physics and the interplay of viscous, inertia and rate-independent terms and we show when each one of them is dominant. Second, in Section 4, we show the general class of Micromorphic continua (e.g. strain gradient theory and Cosserat continuum [Ger73b, Ger73a, CCM01, Var09]), which regularize the problem by introducing characteristic lengths. Finally, going a step further from a pure mechanical description, we show in Section 5 the effect of multiphysics couplings that insert both characteristic lengths and time scales in the problem [Var96a, Var96b, Ben05]. It is needless to say that the literature in each one of this topics is huge and can be classified material-wise, application-wise and method-wise.

Here we use simple one-dimensional examples in order to present the mathematical developments in a simplified manner and to help understanding. These examples fol-

low after a general presentation of the regularization techniques in three-dimensions. Index notation is used throughout this chapter as it is easier for handling the simple case of Cauchy continua and of Micromorphic continua with or without multiphysics couplings.

This chapter is addressed to graduate students and researchers interested in strain localization analysis and can be read in any order. Readers that want a synthesis of the different regularization techniques are advised to follow the order of the sections. Readers that prefer a *"hands-on approach"* are advised to start from the 1D examples in each section and repeat the calculations.
After studying this chapter we hope that the reader will be able to:

- Understand fundamental notions related to bifurcation theory;
- Perform a bifurcation analysis using the first Lyapunov method and derive the conditions for strain localization under different constitutive assumptions and continua;
- Identify the dominant time and spatial scales in a class of problems;
- Draw qualitative conclusions regarding strain localization zone thickness and mesh dependency without cumbersome numerical analyses;
- Understand the added-value of Micromorphic continua such as the Cosserat and strain-gradient continua;
- Investigate the effect of multiphysics couplings on the localization of deformations.

Updated versions of this chapter can be found at: `http://coquake.eu/wp-content/uploads/2019/06/ALERT_2019.pdf`.

2 Strain localization

Strain localization is a phenomenon that is frequently met in (geo-)materials when strain is localized into narrow zones of increased deformation. Instability, loss of uniqueness, bifurcation, ill-posedness or even mesh dependency are terms that are frequently used in the literature (some times erroneously or as unwitting abuse of language) to describe this phenomenon from a mathematical point of view. Before studying strain localization in details and proposing various regularization strategies, it is worth define the meaning of each term.

2.1 Definitions

Loss of uniqueness, *bifurcation* and (*in*)*stability* are distinct, but related notions that are frequently used to describe the behavior of physical systems.

Loss of uniqueness and bifurcation are associated with the existence of one or several equilibrium states of a system. These equilibrium states are also called *steady-states*

of a system and can be periodic in time or time-independent. Of course, the system has to be non-linear in order to have several equilibria and not just only one. They can depend also on several parameters whose value can determine the existence and the number of this equilibrium states. In this case, these parameters are called *bifurcation parameters*. Examples of bifurcation parameters is the loading intensity, the constitutive parameters of a material, geometry etc.

Besides the existence of one or several steady states for given values of the bifurcation parameters, an equilibrium state may be *stable* or *unstable*. We say that an equilibrium (or steady state) is stable when it returns or stays close to this equilibrium after a small perturbation. An equilibrium is unstable when it is not stable. The notion of stability is well established and mathematically rigorously defined in the original work of Lyapunov [Lya92] in the end of 19^{th} century. Stability is directly connected with the *time evolution* of a system. This is in important point, because even if in common practice time is neglected (e.g. quasi-static conditions), the transition from a steady-state to another one happens in a certain time scale, which might be very short (sudden failure of a brittle material) or very slow (geological phenomena).

According to the above, stability and bifurcation (or loss of uniqueness) are two different notions. However, bifurcation points are commonly accompanied with stability change of the equilibrium states. This is illustrated in the following example. In figure 3 we present a simple mechanical system consisted of a rigid bar attached to a pivot point and a spring and loaded with a load P. We choose the applied load P as bifurcation parameter and we plot the angle θ at equilibrium as a function of P. The space (θ^*, P^*) is called (bifurcation) parameter space and the asterisk denotes equilibrium. Solid lines denote stable equilibrium states and dashed unstable ones. Obviously this system has several equilibria. For instance, for $P^* < -kl$, where l is the length of the bar, the system has two equilibrium states. The first one is when the bar is at vertical upward position $\theta^* = 0$ and it is stable. The second one is when the bar is vertical but downwards ($\theta^* = \pi$) and it is unstable. When $-kl < P^* < kl$ the system has three equilibrium positions, but all are unstable except the one corresponding to $\theta* = 0$. In this system there is no unique equilibrium and therefore the term loss of uniqueness is of no use. However, if one linearizes the system in the vicinity of the $\theta = 0$, then the equilibrium branch for $\theta^* = \pi$ disappears. In this case (see figure 3) the system has a unique equilibrium point for $\theta^* = 0$, which is lost at $P^* = kl$ (loss of uniqueness at the bifurcation point B). For rigorous mathematical definitions of bifurcation, loss of uniqueness and stability we refer to [Lya92, BN69, CCV04, BH91b, BH91a, SA16]. Another, important term that is common in the study of physical systems is *ill-posedness*. A mathematical system is said to be *well-posed* when:

- A solution exists;
- The solution is unique;
- The solution's behavior changes continuously with the initial conditions.

This definition is given by Hadamard [Had02]. Problems that are not well-posed in

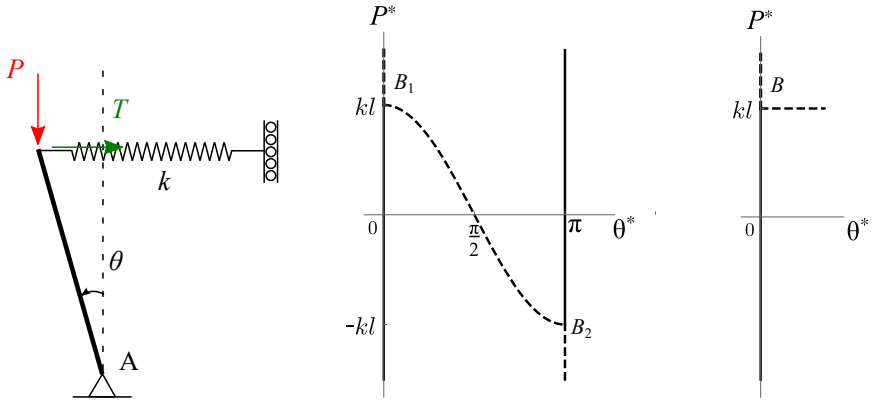


Figure 3: Left: Spring-rigid bar mechanical system. Middle: Bifurcation diagram of the spring-rigid bar system. B_1 and B_2 are bifurcation points. Right: Bifurcation diagram of the linearized spring-rigid bar system in vicinity of $\theta = 0$. The linearized system has only one bifurcation point, B. Solid lines denote stable equilibrium and dashed unstable ones.

the sense of Hadamard are termed ill-posed. Hadamard believed that problems that are physically important are both solvable and uniquely solvable. However, nowadays we know that the many among the most important modern problems are not uniquely solvable. Examples of very important ill-posed problems are found in all scientific disciplines. Strain localization, inverse problems (e.g. seismic inversion), earthquake nucleation, neural networks, population growth, weather and chaos are some problems involving ill-posed mathematical equations among many others. Even the very simple example given in figure 3 is ill-posed in the sense of Hadamard as multiple equilibria (solutions) exist.

Ill-posed problems are though difficult to solve and these difficulties appear in different forms depending on the application. For example, strain localization is connected with mesh dependency in finite elements analyses. Mesh dependency means that the stress-strain response of the system, as well as the strain localization thickness (when interested in deformation bands, see below) depend on the finite element discretization used for solving the problem and further refinement of the mesh does not assure convergence to a unique solution. In order to remedy mesh-dependency and other undesired and nonphysical phenomena that are frequently met in ill-posed problems, *regularization* is needed.

A mathematical problem is regularized either ad-hoc by changing the mathematical equations to alleviate undesired pathologies or by introducing more physics to the problem at hand. The mathematical problem might still remain ill-posed, but it will be free of nonphysical behaviors which are not confirmed by observations. Such an example is the formation of shear bands in rate-independent granular materials. As it

will be shown below, modeling with the classical Cauchy continuum leads to the formation of shear bands of zero thickness, which is contrary to experimental evidence. Experiments show (e.g. [AHV$^+$12]) that the shear band thickness is finite and equal to some grain particles in size. Mühlhaus and Vardoulakis [MV87] regularized this problem by introducing the missing integral lengths in the mathematical problem by resorting to Cosserat theory (see section 4).

2.2 Instability of homogeneous deformation

The general PDE's of the problem are:

$$\sigma_{ij,j} = \rho \ddot{u}_i, \tag{1}$$

where σ_{ij} is the Cauchy stress tensor, ρ is the density of the material, u_i represents the displacement at direction "i" and the double dot denotes the second time derivative (acceleration). The indices take values $1, 2, 3$ and Einstein summation convention is used herein. Suppose a homogeneous, homogeneously deformed solid that is in equilibrium:

$$\sigma^*_{ij,j} = 0. \tag{2}$$

We assume a perturbation, $\tilde{u}_i$ from the reference, homogeneous solution, u^*_i, such that $\tilde{u}_i = u_i - u^*_i$. $\tilde{u}_i = 0$ at the part of the boundary where displacements are applied and $\tilde{u}_{i,j} n_j = 0$ where tractions are applied. We consider the class of materials whose constitutive law can be written (linearized) as follows:

$$\tilde{\sigma}_{ij} = L_{ijkl} \tilde{\varepsilon}_{kl}, \tag{3}$$

where $\tilde{\sigma}_{ij} = \sigma_{ij} - \sigma^*_{ij}$, $\tilde{\varepsilon}_{ij} = \frac{1}{2}\left(\tilde{u}_{i,j} + \tilde{u}_{j,i}\right)$ and L_{ijkl} a fourth order tensor depending on the constitutive behavior of the material at the reference state, where the linearization is made. Injecting equation (3) into equation (1) we obtain:

$$L_{ijkl} \tilde{u}_{k,lj} = \rho \ddot{\tilde{u}}_i. \tag{4}$$

The above PDE is linear and can be solved by separation of variables (or Fourier transform). Setting $\tilde{u}_i = X(x_k) U_i(t)$, equation (4) becomes:

$$L_{ijkl} X_{,lj} U_k = \rho X \ddot{U}_i. \tag{5}$$

The general solution of equation (5) in terms of U_i is $U_i = U_i(t) = g_k e^{st}$ leading to:

$$\left(L_{ijkl} X_{,lj} - \rho X s^2 \delta_{ik}\right) g_k = 0, \tag{6}$$

where δ_{ij} is the Kronecker delta.

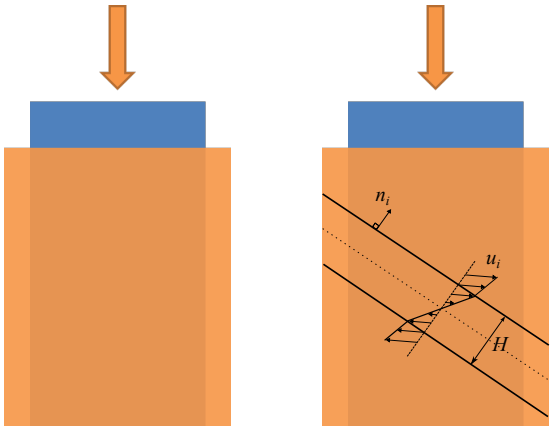


Figure 4: Schematic representation of homogeneous (left) and localized deformation (right) in the form of a band. Blue indicates the initial undeformed state and orange the final form. The loading is applied slowly in a quasistatic manner such as to always satisfy equilibrium $\sigma^*_{ij,j} = 0$. A bifurcation from the equilibrium state with homogeneous deformation to another equilibrium state of non-homogeneous deformation can occur under the conditions of described by equation (8).

2.3 Deformation bands

Figure 4 shows the formation of a deformation band. As explained below, shear, compaction and dilation bands are all deformation bands. The kinematics for the formation of a deformation band determine the form of the perturbation $\tilde{u}_i$ and consequently X, which has to be a planar wave propagating in direction n_i, i.e. $X(x_k) = e^{\mathrm{i}kn_ix_i}$. $k = \frac{2\pi}{\lambda}$ is the wave number of the perturbation, λ its wavelength and $\mathrm{i} = \sqrt{-1}$. Therefore, the perturbation $\tilde{u}_i$ takes the form:

$$\tilde{u}_i = g_i e^{st+\mathrm{i}kn_jx_j} \tag{7}$$

and equation (6) becomes:

$$\left(\Gamma_{ik} - \rho c^2 \delta_{ik}\right) g_k = 0, \tag{8}$$

where $\Gamma_{ik} = n_j L_{ijkl} n_l$ is the so-called *acoustic tensor*, $c = \mathrm{i}\frac{\lambda s}{2\pi}$ is the propagation velocity of the sinusoidal plane wave described by $\tilde{u}_i = U_i e^{\mathrm{i}kn_ix_i+st}$. The above condition for strain localization coincides with the bifurcation conditions determined in [RR75] (see also Appendix A) and takes the form of a classical eigenvalue problem. The above eigenvalue problem has three eigenvalues $q^{(i)}$ corresponding to three eigenvectors $\{g_k^{(m)}\}$ $(m = 1, 2, 3)$. Given the eigenvalues and solving for the propagation velocity we get $c^{(m)} = \sqrt{\frac{q^{(m)}}{\rho}}$. The Lyapunov exponent is then $s^{(m)} = -\mathrm{i}\frac{2\pi}{\lambda}\sqrt{\frac{q^{(m)}}{\rho}}$. If an $s^{(m)}$ with positive real part exists, i.e. $Re(s^{(m)}) > 0$, then the homogeneous

solution u_i^* is unstable and the system bifurcates to a non-uniform solution, a deformation band, with direction n_i. Strain localization takes place. The type of the defor-

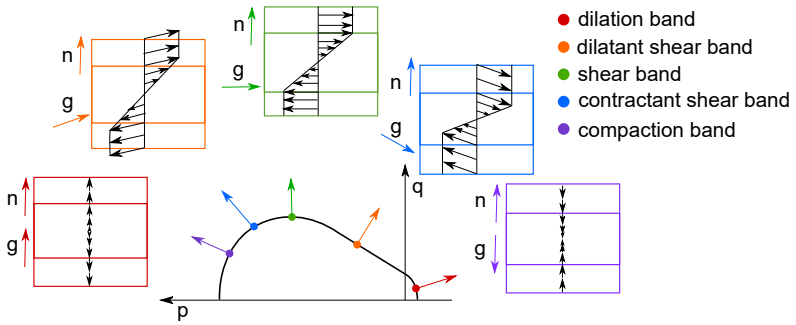


Figure 5: n_i and g_i for different types of deformation bands.

mation band (compaction, shear, dilation band) is determined by the product $n_i g_i$. If $n_i g_i = 0$, the deformation is a shear band, if $n_i g_i = -1$ a pure compaction band and if $n_i g_i = +1$ a pure dilation (extension) band. The intermediate states, $-1 < n_i g_i < 0$ and $0 < n_i g_i < +1$ correspond to contractant and dilatant shear bands, respectively. This is schematically shown in figure 5. More precisely, for an elastoplastic material whose plastic behavior is a function of the first and second invariants of the stress tensor (figure 6), it can be shown that under axisymmetric compression conditions of loading, strain localization occurs when the hardening modulus becomes lower than a critical value h_{cr} for given values of friction coefficient μ and dilatancy β (see figure 7 and [IR00]).

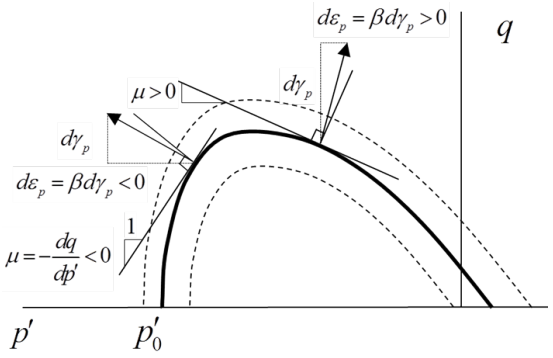


Figure 6: Elastoplastic yield envelope with hardening/softening (dotted lines). Compression is considered negative.

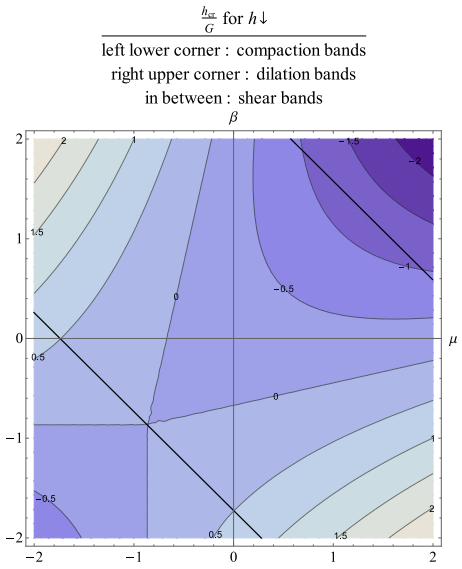


Figure 7: Critical hardening values in function of friction coefficient μ and dilatancy β for strain localization [IR00]. Notice that for non-associate plastic flow rule, localization can occur even with hardening ($h_{cr} > 0$).

2.4 Mesh dependency

With the exception of some special cases of constitutive laws that are out of the scope of the present chapter, the acoustic tensor Γ_{ik} does not depend on the (perturbation) wavelength, λ. Therefore, its eigenvalues, $q^{(m)}$ will not depend on λ either. Consequently, the Lyapunov exponent s becomes maximum for decreasing λ ($s^{(m)} = -\mathrm{i}\frac{2\pi}{\lambda}\sqrt{\frac{q^{(m)}}{\rho}}$). In particular $s \to \infty$ for $\lambda \to 0$. This means that the dominant perturbation in time is the one with the smallest wavelength ($\tilde{u}_i = U_i e^{\mathrm{i}kn_i x_i + st}$). In other words the minor imperfection in size in te medium will grow faster and dominate over the other imperfections of larger wavelength. This is why in the classical Cauchy continuum, which has no internal lengths (the acoustic tensor Γ_{ik} does not depend on λ), the deformation band thickness is zero. This means that strain localization takes place on a mathematical plane. The fact that the smallest perturbation propagates faster justifies also the mesh dependency in Finite Element calculations, if one associates the mesh size with the characteristic wavelength of the perturbation. For instance, in the frame of classical simulations in elastoplasticity of Cauchy rate-independent continua with softening behavior (or even in perfect plasticity), the numerically predicted shear band thickness depends on the finite element discretization and on the element size (figure 8).

It is worth emphasizing that the above condition for strain localization‘is independent

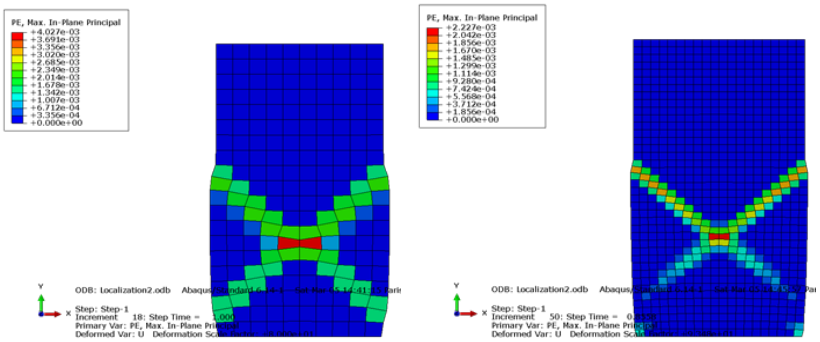


Figure 8: Shear band formation and mesh dependency for a rate-independent elasto-plastic, von Mises, Cauchy medium with strain softening. The shear band thickness is always 1-2 elements thick and therefore mesh dependent. The plastic strains and the global energy dissipation are also mesh dependent.

of the specific constitutive law, provided that the material is rate-independent and that equation (3) can be written. Rate-dependent materials are treated in the next section where a similar approach is followed for studying strain localization. The above methodology is quite general and can be applied in many problems, including problems with multiphysical couplings, such as thermo-poro-chemo-mechanical couplings (e.g. [Ste14, Sul15]). Moreover, even though a Cauchy (Boltzmann) continuum was considered here, the same approach can be applied in Cosserat or even higher order continua (e.g. [Müh88, Sul11]) as shown in the next Sections.

2.5 1D example

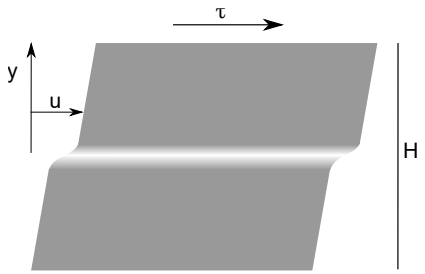


Figure 9: Simple shear of an infinite layer.

In this paragraph we present a simple one dimensional example in order to illustrate the above theoretical notions. We consider a layer that is sheared as shown in figure 9. For the material of the layer we consider an elastoplastic constitutive behavior with

mechanical softening. The yield surface is defined as:

$$F = \sigma_{12} - \tau_0 \leq 0, \tag{9}$$

The strain increments are split in elastic and plastic parts as follows (small deformations):

$$\dot{\varepsilon}_{ij} = \dot{\varepsilon}^{el}_{ij} + \dot{\varepsilon}^{pl}_{ij}. \tag{10}$$

In a linear, elastic, isotropic Cauchy medium, the stresses are related to the elastic deformations as follows:

$$\sigma_{ij} = K\varepsilon^{el}_{kk}\delta_{ij} + 2G\left(\varepsilon^{el}_{ij} - \frac{1}{3}\varepsilon^{el}_{kk}\delta_{ij}\right), \tag{11}$$

where K is the bulk modulus and G is the shear modulus. In this 1D example the system is invariant in x_1 and x_3 directions and, therefore, the momentum balance equations become:

$$\frac{\partial\sigma_{12}}{\partial x_2} = \rho\ddot{u}_1; \quad \frac{\partial\sigma_{22}}{\partial x_2} = \rho\ddot{u}_2. \tag{12}$$

We assume that at steady state (equilibrium) we have homogeneous shear. In particular, $\sigma_{12} = \sigma^*_{12} = \tau_0$, $\sigma_{22} = \sigma^*_{22} = \sigma_0$. This state will be stable as long as any perturbation does not grow in time. By perturbing the displacement fields ($u_i = u^*_i + \tilde{u}_i$) Equations (12) become:

$$\frac{\partial\tilde{\sigma}_{12}}{\partial x_2} = \rho\ddot{\tilde{u}}_1; \quad \frac{\partial\tilde{\sigma}_{22}}{\partial x_2} = \rho\ddot{\tilde{u}}_2. \tag{13}$$

For elastoplasticity with mechanical softening (equation (9)):

$$\begin{aligned} \tilde{\sigma}_{12} &= 2G\frac{h}{1+h}\tilde{\varepsilon}_{12} \\ \tilde{\sigma}_{22} &= M\tilde{\varepsilon}_{22}, \end{aligned} \tag{14}$$

where $M = K + \frac{4G}{3}$ is the p-wave elastic modulus and $h = \frac{1}{G}\frac{d\tau_0}{dq} > -1$ is the hardening modulus, with $\dot{q} = \dot{\gamma}^{pl}_{(12)}$. $h < 0$ denotes softening.

The perturbations $\tilde{u}_i$ have to fulfill the boundary conditions: $\tilde{\sigma}_{12}\left(x_2 = \pm\frac{H}{2}\right) = \tilde{\sigma}_{22}\left(x_2 = \pm\frac{H}{2}\right) = 0$. H is the height of the sheared layer. Equations (13) and (14) together with the above boundary conditions form a linear system, which admits solutions of the form of Eqs.(7) with $\{n_i\} = \{0, 1, 0\}$. Replacing into equations (13) and solving for s as described in the previous sections, we obtain:

$$s = \mathrm{i}kv_p \quad \text{or} \tag{15}$$

$$s = \pm\mathrm{i}kv_s\sqrt{\frac{h}{h+1}}, \tag{16}$$

where $v_p = \sqrt{\frac{M}{\rho}}$ is the p-wave velocity. The system is unstable when $Re[s] > 0$ or, equivalently when $h < 0$ (softening). As expected, the growth coefficient s becomes infinite for $\lambda \to 0$, which leads to mesh dependency.

3 Viscous regularization - characteristic time

Materials whose mechanical response depends on the rate of deformation, as well as on the deformation itself, are referred to as viscous or rate-dependent. The general expression without considering path dependence reads $\sigma_{ij} = \sigma_{ij}\left(\varepsilon_{ij}, \dot{\varepsilon}_{ij}\right)$. The linearized form of the constitutive law around a reference state of homogeneous deformation (see above) reads:

$$\tilde{\sigma}_{ij} = L_{ijkl}\tilde{\varepsilon}_{kl} + M_{ijkl}\dot{\tilde{\varepsilon}}_{kl}. \tag{17}$$

Injecting in the balance equation we obtain:

$$L_{ijkl}\tilde{u}_{k,lj} + M_{ijkl}\dot{\tilde{u}}_{k,lj} = \rho\ddot{\tilde{u}}_i. \tag{18}$$

Using separation of variables (or Fourier transform) we can solve the above linear equation. Limiting our analysis to deformation bands, the perturbation field is given by equation (7). The balance equation now becomes:

$$-n_j L_{ijkl} n_l k^2 g_k XU - n_j M_{ijkl} n_l s k^2 g_k XU - \rho s^2 g_k XU = 0 \tag{19}$$

and finally:

$$\left[n_j L_{ijkl} n_l + n_j M_{ijkl} n_l s + \rho\left(\frac{s}{k}\right)^2 \delta_{ik}\right] g_k = 0 \tag{20}$$

Drawing a parallel to the acoustic tensor, a corresponding second order tensor can be defined for the quantity $n_j M_{ijkl} n_l$. It should however be noted that there is a difference in units. The parallel for the viscous response can then be introduced as $\Delta_{ik} = n_j M_{ijkl} n_l$ leading to:

$$\left(\Gamma_{ik} + \Delta_{ik} s - \rho c^2 \delta_{ik}\right) g_k = 0. \tag{21}$$

3.1 Scaling: Characteristic time and length

To determine the characteristic times, the following quantities are introduced

$$\tau = \frac{t}{T}, \quad \chi_k = \frac{x_k}{L}, \tag{22}$$

where T is a characteristic time for the problem and L a characteristic length (e.g. the height H of the sheared layer of figure 9. Introducing these quantities into equation (21) and dividing by the shear modulus G yields:

$$\left[\frac{\Gamma_{ik}}{G} + \frac{\Delta_{ik}}{GT}\hat{s} + \left(\frac{L}{v_s \hat{k} T}\right)^2 \hat{s}^2 \delta_{ik}\right] g_k = 0, \tag{23}$$

where v_s is the shear-wave velocity, $v_s = \sqrt{\frac{G}{\rho}}$, $\hat{s} = sT$ and $\hat{k} = kL$.

Case 1: Characteristic time due to viscosity

Let us assume that viscosity is dominant. In other words $\frac{\Gamma_{ik}}{G}$ and $\frac{\Delta_{ik}}{GT}$ are terms of $O(1)$ and $\frac{L^2}{v_s^2\hat{k}^2T^2}$ of $O(\varepsilon)$, $\varepsilon \ll 1$. We will determine the characteristic time T_{visc} for this case. $\frac{\Delta_{ik}}{GT_{visc}} = c_{ik} \approx O(1)$ leads to:

$$T_{visc} = c_{ik}\frac{\Delta_{ik}}{G}. \tag{24}$$

Moreover, by hypothesis $\frac{L^2}{v_s^2\hat{k}^2T^2} \ll 1$ and therefore:

$$T_{visc} \gg \frac{L}{v_s\hat{k}} \Rightarrow c_{ik}\frac{\Delta_{ik}}{G} \gg \frac{L\hat{\lambda}}{v_s 2\pi} \Rightarrow$$

$$\hat{\lambda} \ll 2\pi v_s \frac{c_{ik}\Delta_{ik}}{GL} \equiv \hat{\lambda}^*. \tag{25}$$

In other words, when $\lambda \ll \lambda^*$ the inertia terms in equation (23) can be dropped, resulting in:

$$\left(\frac{\Gamma_{ik}}{G} + \frac{\Delta_{ik}}{G}\frac{\hat{s}}{T_{visc}} + \varepsilon\hat{s}^2\delta_{ik}\right) g_k = 0 \Rightarrow \tag{26}$$

$$\left(\frac{\Gamma_{ik}}{G} + c_{ik}s\right) g_k = 0. \tag{27}$$

Assuming strain localization in an isotropic rock with $G \approx 30\text{GPa}$, $c_{ij}\Delta_{ij} = \eta \approx 20\text{MPas}$ and $v_s \approx 2000\text{m/s}$, $\lambda^* \simeq 8m$, which is much bigger than the localization thickness in a deformation band that is of the order of some millimeters or even smaller. Therefore, for typical applications, one would expect viscosity effects to be dominant over inertial ones. This is also shown in the numerical examples in the next paragraph.

Case 2: Characteristic time due to inertia

Suppose that inertia terms are dominant over viscosity. In this case $\frac{\Gamma_{ik}}{G}$ and $\frac{L^2}{v_s^2\hat{k}^2T^2}$ are terms of $O(1)$ and $\frac{\Delta_{ik}}{GT}$ of $O(\varepsilon)$.

From $\frac{L^2}{v_s^2\hat{k}^2T^2} \approx O(1)$ it results that:

$$T_{iner} = \frac{L}{v_s\hat{k}} = \frac{\hat{\lambda}L}{2\pi v_s}. \tag{28}$$

$\frac{\Delta_{ik}}{GT} \ll 1$ yields:

$$c_{ik}\frac{\Delta_{ik}}{G} \ll T_{iner} = \frac{\hat{\lambda}L}{2\pi v_s} \Rightarrow$$

$$\hat{\lambda} \gg \hat{\lambda}^* = 2\pi v_s \frac{c_{ik}\Delta_{ik}}{GL}. \tag{29}$$

Equation (29) means that for very large wave lengths, inertia effects are dominant to viscosity and the second term of equation (23) can be dropped:

$$\left(\frac{\Gamma_{ik}}{G} + s^2\delta_{ik}\right) g_k = 0. \tag{30}$$

Based on the above scalings one -practically- characteristic length, λ^*, was identified and two time scales T_{visc} and T_{iner}. However, there is one more length scale that could be identified.

Case 3: Time scale of negligible rate-independent terms

Of interest is the third and final combination, where both viscosity and inertia are dominant over rate-independent behavior, i.e. $\frac{L^2}{v_s^2\hat{k}^2T^2}$ and $\frac{\Delta_{ik}}{GT}$ are terms of $O(1)$ and $\frac{\Gamma_{ik}}{G}$ of $O(\varepsilon)$, $\varepsilon \ll 1$.

Suppose first that $\hat{\lambda} \ll \lambda^*$ (or $T_{visc} \gg T_{iner}$). We assume a new time-scale, such that $\tau_{v\&i} = \varepsilon^{-1}\tau_{visc}$, with $\varepsilon = \frac{L^2}{v_s^2\hat{k}^2T^2} = T^2_{inertia}$. The characteristic time scale related to viscosity is expected to be slow. Therefore, this new time scale is expected to be fast. After change of variables, equation (26) becomes:

$$\left(\frac{\varepsilon\Gamma_{ik}}{G} + c_{ik}\hat{s} + \hat{s}^2\delta_{ik}\right) g_k = 0 \Rightarrow \tag{31}$$

$$\left(c_{ik}\hat{s} + \hat{s}^2\delta_{ik}\right) g_k = 0 \tag{32}$$

and

$$T_{v\&i} = \varepsilon T_{visc} = \frac{T^2_{iner}}{T_{visc}} \tag{33}$$

For $\lambda \gg \lambda^*$ (or $T_{visc} \ll T_{iner}$), $\varepsilon = \frac{c_{ij}\Delta_{ik}}{G} = T_{visc}$, equation (32) holds again, but in this case:

$$T_{v\&i} = \frac{T^2_{visc}}{T_{iner}}. \tag{34}$$

Which time scale is more important than the others, depends on the application at hand (e.g. loading conditions, material parameters, localization thickness, imperfections etc.). For a given application, the above scaling laws can considerably simplify the calculations and give physical insight to the results.

3.2 Effects of viscosity and inertia

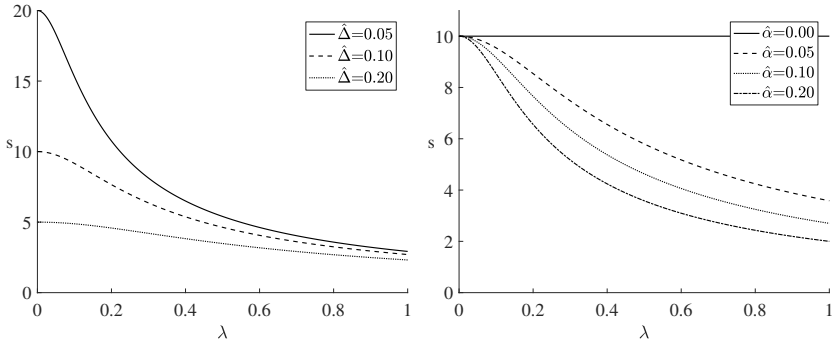
In one dimension equation (21) reduces to:

$$\Gamma + \Delta s + \frac{\rho}{k^2}s^2 = 0. \tag{35}$$

The question to be answered is for which value of k does s receive its maximum value. The solution to the above equation for $\Gamma = 0$ is $s = 0$ for any value of k and

$$s = -\frac{\Delta k^2}{\rho} \tag{36}$$

is negative for any value of k other than 0, which corresponds to an infinite wavelength.



(a) $\hat{\alpha}$=0.1. (b) $\hat{\Delta}$=0.1.

Figure 10: s versus λ for different values of $\hat{\alpha}$ and $\hat{\Delta}$.

If Γ is not equal to zero, there is no loss in generality in dividing by Γ. Substituting in $\lambda = 2\pi/k$ then yields:

$$-1 + \frac{\Delta}{\Gamma}s + \frac{\rho\lambda^2}{4\pi^2\Gamma}s^2 = 0 \Rightarrow \tag{37}$$

$$-1 + \hat{\Delta}s + \hat{\alpha}\lambda^2 s^2 = 0. \tag{38}$$

The solutions to the above quadratic equation are:

$$s_{1,2} = \frac{-\hat{\Delta} \pm \sqrt{\hat{\Delta}^2 + 4\hat{\alpha}\lambda^2}}{2\hat{\alpha}\lambda^2}. \tag{39}$$

Of the two solutions only the one, corresponding to the plus sign, has a positive real part. This is plotted in figure 10 for different values of the quantities $\hat{\Delta}$ and $\hat{\alpha}$. It can be observed that the maximum value of s corresponds in all cases to $\lambda = 0$. In contrast to what was discussed in section 2.4 for the case without any regularization, the maximum value of s is now finite and in fact equal to $1/\hat{\Delta}$. This is illustrated in figure 10a.

It can be further observed that larger values of inertia, represented by the term $\hat{\alpha}$, correspond to a more rapid decrease in the value of s with increasing values of α. This is particularly easy to note in figure 10b.

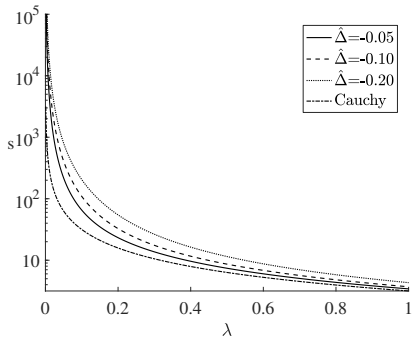


Figure 11: s versus λ for different negative values of $\hat{\Delta}$ and for $\hat{\alpha}$=0.1.

Up to this point it has been tacitly assumed that the contribution of the viscosity has been positive, in the sense that the material exhibits strain rate hardening, rather than strain rate softening. This is the case when the eigenvalues of Δ_{ij} in two- or three-dimensions or the value of $\hat{\Delta}$ in one dimension are positive. It is however also possible that a material may exhibit strain rate softening. For the one dimensional case its effect is illustrated in figure 11. As may be observed, the maximum value of the growth coefficient s goes to infinity as $\lambda \to 0$ and in fact does so faster than in the case of the Cauchy continuum.

3.3 1D Example

The example considered in section 2.5 is revisited here, incorporating a viscous material response. The sheared layer of figure 9 is considered. The material response is assumed to be elasto-viscoplastic with the yield function given in equation (9) and the elastic response given in equation (11). The strain increments are split in elastic and viscoplastic parts as follows (small deformations):

$$\dot{\varepsilon}_{ij} = \dot{\varepsilon}^{el}_{ij} + \dot{\varepsilon}^{vpl}_{ij}. \tag{40}$$

with the viscoplastic strain increments described by a Perzyna type model [Per66]:

$$\dot{\varepsilon}^{vp}_{ij} = \dot{\lambda}\frac{\partial F}{\partial \sigma_{ij}} = \frac{F}{\eta f_0}\frac{\partial F}{\partial \sigma_{ij}}, \tag{41}$$

where η is a constant with units of time, indirectly expressing viscosity, and f_0 is a constant with units of stress, commonly the initial value of the material parameter τ_0. The balance equations are the same as in equations (12) and (13) for the unperturbed and the perturbed states respectively.

From the definition of the plastic multiplier $\dot{\lambda}$ in equation (40) it results that:

$$\dot{F} = \eta f_0 \ddot{\lambda} \Rightarrow \tag{42}$$

$$\dot{\sigma}_{12} = 2G\frac{h}{1+h}\dot{\varepsilon}_{12} + \frac{\eta f_0}{1+h}\ddot{\lambda}. \tag{43}$$

From the definition of the viscoplastic strain and the form of the yield function it results that $\ddot{\lambda} = \ddot{\varepsilon}_{12}^{vp}$. Moreover, the first time derivative can be substituted with a perturbation by considering time integration and successive perturbation. The result reads:

$$\tilde{\sigma}_{12} = 2G\frac{h}{1+h}\tilde{\varepsilon}_{12} + 2\frac{\eta f_0}{1+h}\dot{\tilde{\varepsilon}}_{12}^{vp} \Rightarrow \tag{44}$$

$$\tilde{\sigma}_{12} = 2G\frac{h}{1+h}\tilde{\varepsilon}_{12} + 2\frac{\eta f_0}{1+h}\dot{\tilde{\varepsilon}}_{12} - \frac{\eta f_0}{G(1+h)}\dot{\tilde{\sigma}}_{12}. \tag{45}$$

Making use of equation (43) in equation (45) in the form of successive substitutions yields

$$\tilde{\sigma}_{12} = 2G\frac{h}{1+h}\tilde{\varepsilon}_{12} + 2\frac{\eta f_0}{(1+h)^2}\dot{\tilde{\varepsilon}}_{12} - 2\frac{(\eta f_0)^2}{G(1+h)^3}\ddot{\tilde{\varepsilon}}_{12} + 2\frac{(\eta f_0)^3}{G^2(1+h)^4}\dddot{\tilde{\varepsilon}}_{12} - \dots \tag{46}$$

It can be observed that the above is an infinite series with alternating sign of coefficients. The coefficients follow a geometric progress where the multiplier is equal to $\frac{\eta f_0}{G(1+h)}$. To maintain an analogy to equation (17) only the first two terms on the right hand side of the equation are retained. On the whole the constitutive law now reads:

$$\begin{aligned} \tilde{\sigma}_{12} &= 2G\frac{h}{1+h}\tilde{\varepsilon}_{12} + 2\frac{\eta f_0}{(1+h)^2}\dot{\tilde{\varepsilon}}_{12} \\ \tilde{\sigma}_{22} &= M\tilde{\varepsilon}_{22}, \end{aligned} \tag{47}$$

where, again, $M = K + \frac{4G}{3}$ is the p-wave elastic modulus and $h = \frac{1}{G}\frac{d\tau_0}{dq} > -1$ is the hardening modulus.

Following the analysis presented in section 2.5 from the first balance equation we obtain:

$$G\frac{h}{1+h}k^2 + \frac{\eta f_0}{(1+h)^2}k^2 s + \rho s^2 = 0 \Rightarrow \tag{48}$$

$$v_s^2\frac{h}{1+h}k^2 + v_s^2\frac{\eta f_0}{G(1+h)^2}k^2 s + s^2 = 0 \tag{49}$$

Solving for s yields:

$$s = -\frac{1}{2}v_s^2\frac{\eta f_0}{G(1+h)^2}k^2 \pm \frac{1}{2}\sqrt{\left(v_s^2\frac{\eta f_0}{G(1+h)^2}k^2\right)^2 - 4v_s^2\frac{h}{1+h}k^2}. \tag{50}$$

On the whole, from both balance equations we obtain:

$$s = \mathrm{i}kv_p \quad \text{or} \tag{51}$$

$$s = -v_s^2\frac{\eta f_0}{2G(1+h)^2}k^2 \pm \sqrt{\left(v_s^2\frac{\eta f_0}{2G(1+h)^2}k^2\right)^2 - v_s^2\frac{h}{1+h}k^2}, \tag{52}$$

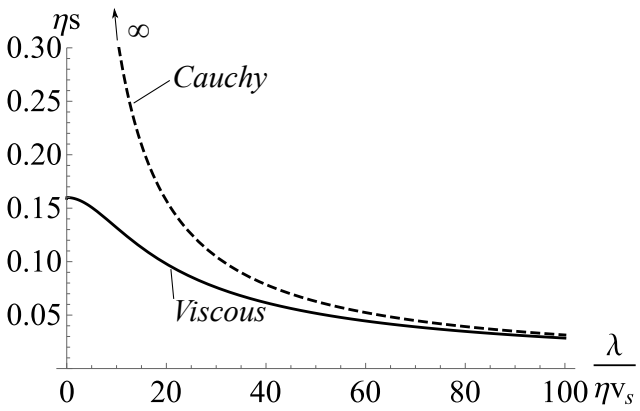


Figure 12: Growth coefficient in function of the perturbation wavelength. Contrary to Cauchy continuum, the maximum value of the growth coefficient is finite. For $\hat{\lambda} \gg \hat{\lambda}^* \approx 20$ inertia is dominant.

where $v_p = \sqrt{\frac{M}{\rho}}$ is the p-wave velocity and $v_s = \sqrt{\frac{G}{\rho}}$ is the s-wave velocity. The system is unstable when $Re[s] > 0$ and it reduces to the solution given in section 2.5 when $\eta = 0$.

A comparison to the Cauchy continuum is presented in figure 12. For $\eta \neq 0$, we observe that the growth coefficient s becomes maximum for $\lambda = 0$, but non infinite as in the case of rate-independent materials. In other words the presence of a characteristic time due to viscosity limits the growth coefficient and consequently perturbations propagate in finite time. Moreover, in the absence of inertia (or if it is very small, see scaling, paragraph 3.1) the growth coefficient is finite and independent of the wave length of the perturbation (see figure 10 for $\hat{\alpha}$). The consequence of this latter observation is that in numerical analyses the results are mesh-independent [Nee88, WSD96]. However, they depend on existing perturbations related to the material parameters (e.g. imperfections) or the loading conditions which might favor one wave-length or another.

4 Micromorphic continua regularization - characteristic lengths

The theory of Micromorphic continua is a general continuum theory that can represent various heterogeneous systems with microstructure of non-negligible size and take into account various length and time scales (internal lengths) that the classical Cauchy continuum fails to represent. The various features of the Micromorphic continuum theory were studied by many researchers in the past, showing several advan-

tages compared to the classical continuum approach. Intrinsic wave dispersion, non-singular fields in fracture mechanics, interesting properties related to the design of metamaterials, are some of the applications that emerge from the deep study of these continua. Regularization in strain localization problems is another feature of these continua due to the characteristic lengths they embody.

According to Germain [Ger73b] the Cauchy continuum is a continuous distribution of particles, each of them being represented geometrically by a point and characterized kinematically by a velocity V_i. In a theory that takes microstructure into account each particle has kinematic properties that are defined in a more detailed way.

At the microscopic level of observation, a particle appears itself as a continuum $P(M)$ of small extent. Let M be the center of mass of the particle $P(M)$, M' a point of $P(M)$, u_i the displacement of M (V_i its velocity), x'_i the coordinates of M' in a Cartesian frame parallel to the given, global frame and M its origin, u'_i the displacement of M' with respect to the given frame (V'_i its displacement) and x_i the coordinates of M in the given frame (see figure 13). D denotes the control volume. As $P(M)$ is of small extent, it is natural to look at the asymptotic expansion of V'_i with respect to x'_i:

$$u'_i = u_i + \chi_{ij}x'_j + \chi_{ijk}x'_jx'_k + \chi_{ijkl}x'_jx'_kx'_l + \dots, \tag{53}$$

where χ_{ij} is a micro-deformation tensor, which expresses the gradient of the relative displacements u'_i and $\chi_{ij\dots m}$ are higher order micro-deformation rate tensors. In three-dimensions: $i, j, \dots, m = 1, 2, 3$. The tensors $\chi_{ij\dots m}$ are assumed to be fully symmetric with respect to the indices $j, \dots, m$.

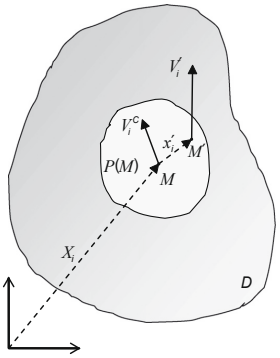


Figure 13: Continuum with microstructure.

Applying the principle of virtual power and using the divergence theorem (see [Ger73b]

and [Ste18] for more details on derivation), we obtain:

$$\begin{aligned} \tau_{ij,j} + f_i = 0, \qquad & t_i = \tau_{ij} n_j \\ \nu_{ijk,k} + s_{ij} + \psi_{ij} = 0, \qquad & \mu_{ij} = \nu_{ijk} n_k \\ \nu_{ijkl,l} + s_{ijk} + \psi_{ijk} = 0, \qquad & \mu_{ijk} = \nu_{ijkl} n_l \\ \dots, & \end{aligned} \tag{54}$$

where, again, n_i is the outward pointing unit normal vector field of the boundary of the solid. The above system of equations represents the equilibrium equations of a Micromorphic continuum of order n (strong form).

The additional degrees of freedom of Micromorphic continua introduce microinertia terms, whose presence leads to interesting wave dispersion properties, especially at short wavelengths (optic branch) [SSV10] and finite Lyapunov exponents in localization problems [SSV11].

In figure 14 we outline the various higher order (Micromorphic) continuum theories and their special cases. Besides the classical continuum and the Cosserat continuum (called also micropolar continuum, see [Var09]), a special case of Micromorphic continuum is also the second gradient and the indeterminate couple stress theory (called also restrained Cosserat medium).

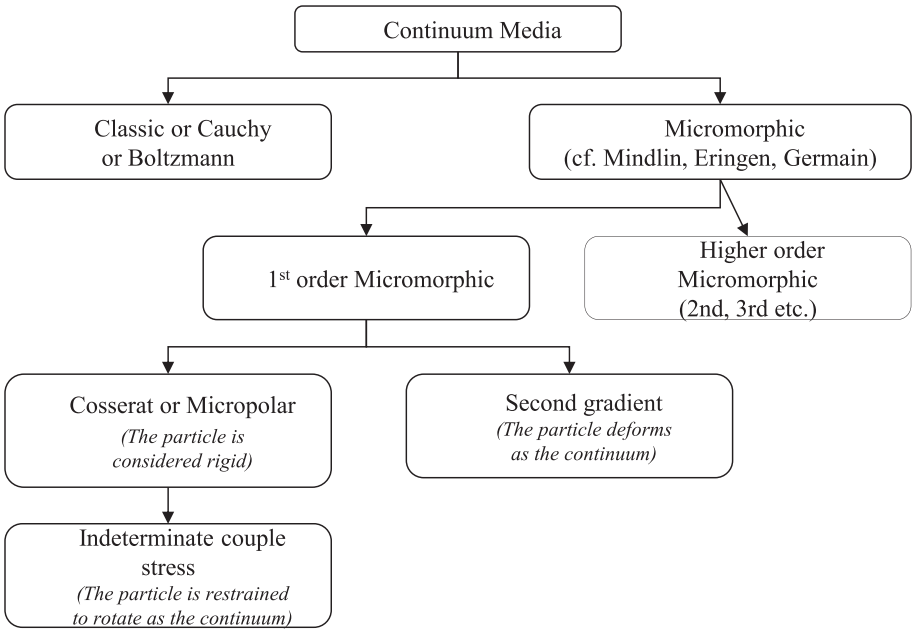


Figure 14: Higher order continuum theories according to Germain's terminology [Ger73b]; see also [Min64, Eri99].

Retrieving the classical, Boltzmann continua, is straightforward by setting χ_{ij} and the higher order microdeformation rate tensors null. In this case, $s_{ij} = 0$ and $\tau_{ij} = \sigma_{ij}$, i.e. equal to the Cauchy stress tensor, which is symmetric.

In the case that the particle $P(M)$ is deformable and its microdeformation coincides with the deformation of the (macro-)continuum, i.e. $\chi_{ij} = V_{i,j}$, we obtain the so-called second gradient continuum theory. As in this case the microdeformation rate tensor is no more an independent generalized virtual velocity, one has to start from the very beginning and apply the principle of virtual power for deriving the strong form of the equilibrium equations and the appropriate boundary conditions. For more details we refer to [Ger73b] and for some interesting applications of the theory to [DSMP93, CCE98, ZPV01, CCC06, SDC07, KABC08, PZ16, DAD$^+$17], among others. Alternatively, second gradient continua can be derived by assuming that the internal energy depends explicitly on the second gradient of the displacement field [Ger73a, Min65]. As fat it concerns strain localization, second gradient continuum leads to deformation bands of finite thickness and remedies mesh dependency in finite element analyses. For a more detailed study of second gradient theories related to strain localization we refer to [CCM01, CCC06].

4.1 Cosserat continuum

The derivation of the Cosserat continuum is more direct than the second gradient. The basic assumption is that the particle $P(M)$ behaves as a rigid body and so it can not only translate, but also rotate. In this case the microdeformation rate tensor has to be anti-symmetric and the rest higher-order microdeformation tensors zero.

Adding inertia effects and neglecting volumic forces, equilibrium equations (equations (54)) become:

$$\begin{aligned} \tau_{ij,j} + f_i &= \rho \ddot{u}_i, \qquad & t_i &= \tau_{ij} n_j \\ m_{ij,j} - \epsilon_{ijk}\tau_{jk} + \psi_i &= I\ddot{\omega}^c_i, \qquad & \mu_i &= m_{ij} n_j. \end{aligned} \tag{55}$$

This is the strong form of the Cosserat continuum equations. τ_{ij} is the Cosserat stress tensor, which is not symmetric and m_{ij} is the Cosserat moment (couple stress) tensor. ϵ_{ijk} is the Levi-Civita symbol. f_i and ψ_i are respectively volumic (body) forces and moments and u_i and ω^c_i are respectively the Cosserat displacements and rotations. t_i and μ_i denote boundary tractions and I is the microinertia.

A constitutive law connects the generalized stresses τ_{ij} and m_{ij} with the generalized deformations γ_{ij} and κ_{ij}:

$$\begin{aligned} \gamma_{ij} &= u_{i,j} + \epsilon_{ijk}\omega^c_k \\ \kappa_{ij} &= \omega^c_{i,j}, \end{aligned} \tag{56}$$

i.e. $\tau_{ij} = \tau_{ij}(\gamma_{ij}, \kappa_{ij})$ and $m_{ij} = m_{ij}(\gamma_{ij}, \kappa_{ij})$. We assume an equilibrium state of homogeneous deformation and search for the conditions where this state becomes

unstable leading to the formation of a deformation band. To this extend, we perturb the kinematic fields u_i and ω_i as follows:

$$\begin{aligned}\tilde{u}_i =& u_i - u_i^* = U_i e^{st+k_j n_j} \\ \tilde{\omega}_i^c =& \omega_i^c - \omega_i^{c*} = \Omega_i e^{st+k_j n_j}.\end{aligned} \tag{57}$$

Linearization of the constitutive law yields (see [RSS18] for an application in elasto-plasticity):

$$\begin{aligned}\tilde{\tau}_{ij} =& C_{ijkl}^{TT}\tilde{\gamma}_{kl} + C_{ijkl}^{TM}\tilde{\kappa}_{kl} \\ \tilde{m}_{ij} =& C_{ijkl}^{MT}\tilde{\gamma}_{kl} + C_{ijkl}^{MM}\tilde{\kappa}_{kl}.\end{aligned} \tag{58}$$

Notice that κ_{ij} has units of deformation over length. Consequently, any ratio of the various tensors C^{XX} produces a characteristic length for the problem at hand. Inserting equations (57) and equations (58) into equations (55) we obtain the following system of algebraic equations:

$$\begin{bmatrix} \Gamma_{ik} + \rho\left(\frac{s}{k}\right)^2 \delta_{ik} & \Delta_{ik} \\ \Xi_{ik} & \Pi_{ik} + I\left(\frac{s}{k}\right)^2 \delta_{ik} \end{bmatrix} \begin{bmatrix} U_k \\ \Omega_k \end{bmatrix} = 0, \tag{59}$$

where

$$\Gamma_{ik} = n_j C_{ijkl}^{TT} n_l \tag{60}$$

$$\Delta_{ik} = -\mathrm{i}\frac{1}{k} n_j e_{qlk} C_{ijql}^{TT} + n_j C_{ijkl}^{TM} n_l \tag{61}$$

$$\Xi_{ik} = n_j C_{ijkl}^{MT} n_l + \mathrm{i}\frac{1}{k} e_{ijr} C_{jrkq}^{TT} n_q \tag{62}$$

$$\Pi_{ik} = n_j C_{ijkl}^{MM} n_l - \mathrm{i}\frac{1}{k} e_{rnk} C_{ilrn}^{MT} n_l + \frac{1}{k^2} e_{ilr} C_{lrnq}^{TT} e_{nqk} + \mathrm{i}\frac{1}{k} e_{ilr} C_{lrkq}^{TM} n_q. \tag{63}$$

The strain localization condition for deformation bands in the framework of Cosserat continuum is:

$$\mathrm{Det}\left(\begin{bmatrix} \Gamma_{ik} - \rho c^2 \delta_{ik} & \Delta_{ik} \\ \Xi_{ik} & \Pi_{ik} - I c^2 \delta_{ik} \end{bmatrix}\right) = 0. \tag{64}$$

The singularity of the above tensor is similar to the condition found in [IW98, SW91] for the onset of localization ($s = 0$). In these papers, the authors derive the localization condition from the kinematic and static compatibility conditions across the shear band as done classically for strain localization analysis [MV87, VS95]. Note that if no Cosserat effects are considered the classical condition of localization for rate-independent materials with a Cauchy continuum is retrieved, i.e. $\mathrm{Det}(n_j C_{ijkl}^{TT} n_l) = 0$. For more details we refer to [RSS18].

4.2 1D example of regularization with Cosserat continuum

An elastoplastic constitutive behavior with mechanical softening is considered in this example. More advanced Cosserat constitutive models such as the Mühlhaus-Vardoulakis

Cosserat plasticity model [MV87, VS95, RSS18] might be used, but the advantage of this simple and unrealistic for granular material models is that analytical derivations can be performed easily. The yield surface is defined as:

$$F = \tau_{(12)} - \tau_0 \leq 0, \tag{65}$$

where $\tau_{(ij)}$ denotes the symmetric part of the stress tensor τ_{ij}. The strains and curvatures of the Cosserat medium are split in elastic and plastic parts as follows:

$$\begin{aligned} \dot{\gamma}_{ij} &= \dot{\gamma}^{el}_{ij} + \dot{\gamma}^{pl}_{ij} \\ \dot{\kappa}_{ij} &= \dot{\kappa}^{el}_{ij} + \dot{\kappa}^{pl}_{ij}. \end{aligned} \tag{66}$$

In a centrosymmetric, linear elastic isotropic Cosserat medium, the stresses are related to the generalized elastic deformation measures according to the following constitutive relations [Var09]:

$$\begin{aligned} \tau_{ij} &= K\gamma^{el}_{kk}\delta_{ij} + 2G\left(\gamma^{el}_{(ij)} - \frac{1}{3}\gamma^{el}_{kk}\delta_{ij}\right) + 2\eta_1 G\gamma^{el}_{[ij]} \\ m_{ij} &= 4GR^2\left(\kappa^{el}_{(ij)} + \eta_2\kappa^{el}_{kk}\delta_{ij}\right) + 4\eta_3 GR^2\kappa^{el}_{[ij]}, \end{aligned} \tag{67}$$

where η_1, η_2, η_3 are positive material constants and R is an internal length parameter, which for a granular material can be identified with the mean radius of the grains of the Representative Volume Element (RVE). For more details on homogenization approaches tailored to Cosserat continuum and upscaling, both in elasticity and plasticity, the reader is referred to [BV01, GSSS16, RC16]. $\gamma_{(ij)}$ and $\gamma_{[ij]}$ denote respectively the symmetric and anti-symmetric parts of γ_{ij}. The Cosserat shear modulus, which expresses the stiffness related to the relative rotation of the particle (e.g. of a grain) with respect to the macro-rotation of the continuum (e.g. of the assemblage of grains) is defined as $G_c = \eta_1 G$. In this 1D example the system is invariant in x_1 and x_3 directions and, therefore, the momentum balance equations become:

$$\begin{aligned} &\frac{\partial \tau_{12}}{\partial x_2} = \rho\ddot{u}_1; \quad \frac{\partial \tau_{22}}{\partial x_2} = \rho\ddot{u}_2 \\ &\frac{\partial m_{32}}{\partial x_2} + \tau_{21} - \tau_{12} = I\ddot{\omega}^c_3. \end{aligned} \tag{68}$$

At steady state we have a Cauchy continuum under homogeneous shear. In particular, $\tau_{(12)} = \tau^*_{(12)} = \tau_0$, $\tau_{22} = \tau^*_{22} = \sigma_0$, $\tau_{[12]} = \tau^*_{[12]} = 0$ and $m_{32} = m^*_{32} = 0$. This state will be stable as long as any perturbation does not grow in time. By perturbing the displacement and the rotation fields at steady state ($u_i = u^*_i + \tilde{u}_i$, $\omega_3 = \omega^{c*}_3 + \tilde{\omega}^c_3$) equations (68) yield:

$$\begin{aligned} &\frac{\partial \tilde{\tau}_{12}}{\partial x_2} = \rho\ddot{\tilde{u}}_1; \quad \frac{\partial \tilde{\tau}_{22}}{\partial x_2} = \rho\ddot{\tilde{u}}_2 \\ &\frac{\partial \tilde{m}_{32}}{\partial x_2} + \tilde{\tau}_{21} - \tilde{\tau}_{12} = I\ddot{\tilde{\omega}}^c_3. \end{aligned} \tag{69}$$

For elastoplasticity with mechanical softening (equation (65)):

$$\begin{aligned}\tilde{\tau}_{(12)} &= 2G\frac{h}{1+h}\tilde{\gamma}_{(12)}\\ \tilde{\tau}_{[12]} &= 2G_c\tilde{\gamma}_{[12]}\\ \tilde{\tau}_{22} &= M\tilde{\gamma}_{22}\\ \tilde{m}_{32} &= 4GR^2\tilde{\kappa}_{32}.\end{aligned} \tag{70}$$

The perturbations $\tilde{u}_i$ and $\tilde{\omega}_3^c$ have to fulfill the boundary conditions: $\tilde{\sigma}_{12}\left(x_2=\pm\frac{H}{2}\right)=\tilde{\sigma}_{22}\left(x_2=\pm\frac{H}{2}\right)=\tilde{m}_{32}\left(x_2=\pm\frac{H}{2}\right)=0$. H is the height of the sheared layer. Equations (69) and (70) together with the above boundary conditions form a linear system which admits solutions of the form of equations (57) with $\{n_i\}=\{0,1,0\}$. Replacing into equations (69) and solving for s as described in the previous sections, we obtain:

$$s = \mathrm{i}kv_p \quad \text{or} \tag{71}$$

$$s = \pm\mathrm{i}kv_s\sqrt{\frac{h}{h+1}}\sqrt{\frac{\eta_1\left(1+\frac{1}{k^2R^2}\right)+\frac{h+1}{h}}{\frac{\eta_1}{k^2R^2}+1}}, \tag{72}$$

where I was taken equal to zero for simplicity. The system is unstable when $Re[s]>0$ or, equivalently when $h<0$ (softening) and $\eta_1\left(1+\frac{1}{k^2R^2}\right)+\frac{h+1}{h}>0$. The latter condition leads to a critical wavelength λ_{cr}:

$$\lambda > \lambda_{cr} = 2\pi R\sqrt{-\frac{1+h}{h}-\frac{1}{\eta_1}}. \tag{73}$$

The wavelength of the perturbation has to be larger than this critical value for localization to occur. Notice that λ_{cr} is proportional to the Cosserat internal length, R. If $R\to 0$ we retrieve the same condition for strain localization with the 1D example presented in paragraph 2.5 for a Cauchy continnum (see figure 15).

5 Regularization and multiphysics couplings - characteristic length/time

It is often the case that the mechanical response of a given material depends on other physical or chemical processes taking place. Such processes can in turn be influenced by the mechanical response of the material to the changing conditions and to the load, leading for example to changes in porosity or internal structure and producing heat through internal friction. If that is the case one speaks of multi-physical couplings. The processes most commonly taken into account are hydraulic, thermal or chemical.

Materials whose mechanical response depends on a number of additional physical quantities θ_i, $i=1,\ldots,N$, as well as on the deformation, obey the following general expression, when the path dependence is not considered: $\sigma_{ij}=\sigma_{ij}\left(\epsilon_{ij},\theta_m\right)$. The

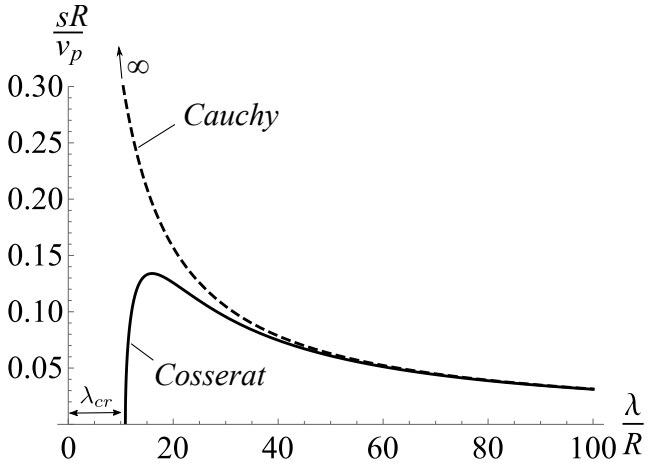


Figure 15: Growth coefficient in function of the perturbation wavelength. Contrary to Cauchy continuum, strain localization is not possible for $\lambda < \lambda_{cr}$ in the case of Cosserat continuum.

linearized form of the constitutive law around the a reference state of homogeneous deformation reads:

$$\tilde{\sigma}_{ij} = L_{ijkl}\tilde{\varepsilon}_{kl} + A_{mij}\tilde{\theta}_m. \tag{74}$$

Injecting in the balance equation we obtain:

$$L_{ijkl}\tilde{u}_{k,lj} + A^m_{ij}\tilde{\theta}_{m,j} = \rho\ddot{\tilde{u}}_i. \tag{75}$$

Each of the quantities θ_i obey in turn their own balance equations, which, with little loss in generality we can assume to be of the advection-diffusion type:

$$\left(D^{(i)}\left(u_{m,n},\theta_k\right)\theta_{i,j}\right)_{,j} - \left(v_j\theta_i\right),j + R^{(i)}\left(u_{m,n},\theta_k\right) = \dot{\theta}_i, \tag{76}$$

where $D^{(i)}$ is the diffusion coefficient of the quantity θ_i, $R^{(i)}$ is the source or sink term of the same quantity and v_j is the velocity field controlling the advection. When the diffusion coefficient is constant and the advective flow is incompressible, the advection-diffusion equation simplifies to:

$$D^{(i)}\theta_{i,jj} - v_j\theta_{i,j} + R^{(i)}\left(u_{m,n},\theta_k\right) = \dot{\theta}_i. \tag{77}$$

Expansion of the balance equation (76) leads to the following formulation:

$$\begin{aligned} D^{(i)}\theta_{i,jj} + \frac{\partial D^{(i)}}{\partial u_{m,n}}\theta_{i,j}u_{m,nj} + \frac{\partial D^{(i)}}{\partial\theta_k}\theta_{i,j}\theta_{k,j} \quad &+ \\ -v_{jj}\theta_i - v_j\theta_{i,j} + R^{(i)}\left(u_{m,n},\theta_k\right) \quad &= \quad \dot{\theta}_i, \end{aligned} \tag{78}$$

where the dependencies of the function $D^{(i)}\left(u_{m,n},\theta_k\right)$ have been omitted for the sake of brevity. Linearization in turn yields:

$$\begin{aligned} & D^{(i)}\tilde{\theta}_{i,jj} \quad + \\ +\frac{\partial D^{(i)}}{\partial u_{m,n}}\theta_{i,j}\tilde{u}_{m,nj} + \frac{\partial D^{(i)}}{\partial u_{m,n}}u_{m,nj}\tilde{\theta}_{i,j} \quad & + \\ +\frac{\partial D^{(i)}}{\partial \theta_k}\theta_{i,j}\tilde{\theta}_{k,j} + \frac{\partial D^{(i)}}{\partial \theta_k}\theta_{k,j}\tilde{\theta}_{i,j} \quad & + \\ -v_{jj}\tilde{\theta}_i - v_j\tilde{\theta}_{i,j} \quad & + \\ +\frac{\partial R^{(i)}}{\partial u_{m,n}}\tilde{u}_{m,n} + \frac{\partial R^{(i)}}{\partial \theta_k}\tilde{\theta}_k \quad & = \quad \dot{\tilde{\theta}}_i, \end{aligned} \tag{79}$$

where the dependencies of the function $R^{(i)}\left(u_{m,n},\theta_k\right)$ have also been omitted. Collecting the terms with the same perturbation components results in:

$$\begin{aligned} D^{(i)}\tilde{\theta}_{i,jj} + P_j^{(i)}\tilde{\theta}_{i,j} - v_{jj}\tilde{\theta}_i + Q_{kj}^{(i)}\tilde{\theta}_{k,j} + \frac{\partial R^{(i)}}{\partial \theta_k}\tilde{\theta}_k \quad & + \\ +S_{mnj}^{(i)}\tilde{u}_{m,nj} + \frac{\partial R^{(i)}}{\partial u_{m,n}}\tilde{u}_{m,n} \quad & = \quad \dot{\tilde{\theta}}_i. \end{aligned} \tag{80}$$

The stability of the system of equations consisting of the above system of equations and equation (75) obviously depends on the values of the prefactors, which however are not known at this point.

The corresponding linearized form for the balance equation (77) reads:

$$D^{(i)}\tilde{\theta}_{i,jj} - v_j\tilde{\theta}_{i,j} + \frac{\partial R^{(i)}}{\partial u_{m,n}}\tilde{u}_{m,n} + \frac{\partial R^{(i)}}{\partial \theta_k}\tilde{\theta}_k = \dot{\tilde{\theta}}_i. \tag{81}$$

The perturbation $\tilde{u}_i$ is assumed to be of the form given in equation (7), while $\tilde{\theta}_i$ is assumed to be given by:

$$\tilde{\theta}_i = h_i e^{st+\mathrm{i}kn_jx_j}. \tag{82}$$

Equations (75), (80) and (81) become:

$$-\left(k^2n_jL_{ijkl}n_l + s^2\rho\delta_{ik}\right)g_k + \mathrm{i}kA_{kij}n_jh_k = 0 \tag{83}$$

$$\begin{aligned} \left[-k^2D^{(i)}n_i\delta_{ik} + \mathrm{i}k\left(P_j^{(i)}n_j\delta_{ik} + Q_{kj}^{(i)}n_j\right) + \frac{\partial R^{(i)}}{\partial \theta_k} - \left(v_{jj}+s\right)\delta_{ik}\right]h_k + & \\ + \left(-k^2S_{kjl}^{(i)}n_jn_l + \mathrm{i}k\frac{\partial R^{(i)}}{\partial u_{k,j}}n_j\right)g_k = 0 & \end{aligned} \tag{84}$$

and

$$\left[\left(-k^2D^{(i)} - \mathrm{i}kv_jn_j - s\right)\delta_{ik} + \frac{\partial R^{(i)}}{\partial \theta_k}\right]h_k + \mathrm{i}k\frac{\partial R^{(i)}}{\partial u_{k,j}}n_jg_k = 0. \tag{85}$$

respectively. Irrespectively of whether the simplified or the full form of the advection-diffusion equation is used for the multiphysical processes, the system of equations now takes the form:

$$\begin{bmatrix} C_{ik}^{gg} & C_{il}^{gh} \\ C_{ik}^{hg} & C_{il}^{hh} \end{bmatrix} \begin{bmatrix} g_k \\ h_l \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{86}$$

with

$$C_{ik}^{gg} = -\left(k^2 \Gamma_{ik} + s^2 \rho \delta_{ik}\right) \tag{87}$$

$$C_{ik}^{gh} = \mathrm{i} k A_{kij} n_j \tag{88}$$

$$C_{ik}^{hg} = -k^2 S_{kjl}^{(i)} n_j n_l + \mathrm{i} k \frac{\partial R^{(i)}}{\partial u_{k,j}} n_j \tag{89}$$

$$C_{ik}^{hh} = -k^2 D^{(i)} n_i \delta_{ik} + \mathrm{i} k \left(P_j^{(i)} n_j \delta_{ik} + Q_{kj}^{(i)} n_j\right) + \frac{\partial R^{(i)}}{\partial \theta_k} - \left(v_{jj} + s\right) \delta_{ik}, \tag{90}$$

where the last two equations correspond to the full form of the advection-diffusion equation.

Requiring the determinant of the system of equations to be equal to zero results in a cubic equation in terms of s. The three roots clearly depend on the value of the various multipliers, but some general remarks may be made without introducing numerical values.

No coupling

When the coupling terms C_{ik}^{gh} and C_{ik}^{hg} are ignored, the possible roots are the two roots resulting from the LSA of the stress balance without any regularization and an additional one from the advection-diffusion equation. When considering the advection-diffusion equation with simplifications, it is clear that the real part of this root will be negative as long as the diffusivity coefficient is positive. This is usually the case with some exceptions, such as the consolidation equation when considering a collapsible solid matrix.

One way coupling

Assuming only the physical process to have an effect on the stress balance equation and itself not to be affected, one speaks of one way coupling. Then it is enough to consider the solution of:

$$-k^2 \Gamma - s^2 \rho + \mathrm{i} k A = 0 \tag{91}$$

with respect to s. The solution reads:

$$s = \pm \frac{\sqrt{-k^2 \Gamma + \mathrm{i} k A}}{\sqrt{\rho}} = \pm \frac{\sqrt{-4\pi^2 \Gamma + \mathrm{i} 2\pi A \lambda}}{\sqrt{\rho \lambda^2}} = \pm \frac{\sqrt{-\hat{\Gamma} + \mathrm{i} \hat{A} \lambda}}{\lambda}, \tag{92}$$

where solutions with positive real part are principally of interest. It is clear from the above equation that for $\lambda \to 0$ the real part of s goes to infinity, as for the case without regularization.

Two way coupling

When the coupling terms are taken into account a cubic equation results with the coupling terms contributing to the constant term with respect to s. To illustrate the effect of the coupling, the problem is considered in one dimension for a single coupled process for the simplified form of the advection-diffusion equation. For the sake of simplicity the advective term and the influence of θ on the source term are ignored:

$$-\left(k^2\Gamma + s^2\rho\right) g + \mathrm{i}kAh = 0, \tag{93}$$

$$-\left(k^2 D^{(i)} + s\right) h + \mathrm{i}kdRg = 0. \tag{94}$$

The determinant then reads:

$$\left(k^2\Gamma + s^2\rho\right)\left(k^2 D^{(i)} + s\right) + k^2 AdR = 0. \tag{95}$$

To investigate the effect of the coupling, we examine the influence of the various terms. The equation to consider is a cubic polynomial:

$$\rho s^3 + \rho k^2 D^{(i)} s^2 + k^2\Gamma s + k^4 D^{(i)}\Gamma + k^2 AdR = 0. \tag{96}$$

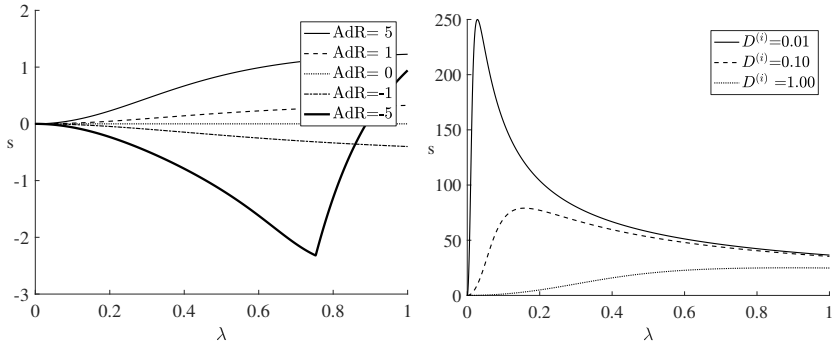
where the values of the prefactors are not known, but their signs can be deducted with the exception of the last term.

Table 1: Values used for the effect of multiphysical coupling on stability.

ρ [kg/m^3]	$D^{(i)}$ [m^2/s]	Γ [Pa]	AdR [Pa/(ms)]
1.0	0.1	-10^{-4}	10^4

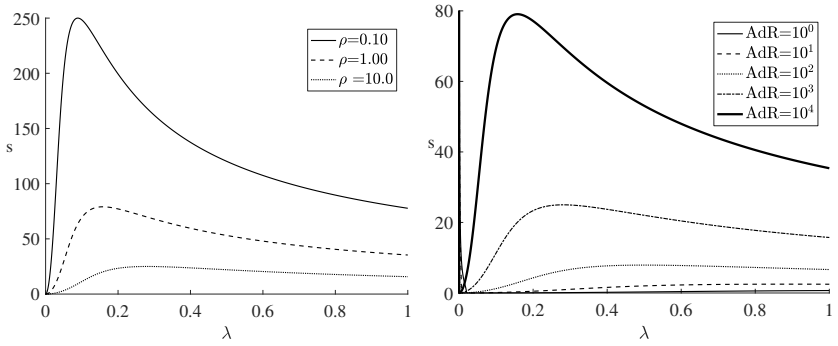
An attempt is made herein to visualize the effect of the different terms. For the graphs that follow the values given in table 1 are used unless otherwise stated. In figure 16a the effect of the coupling term on the mechanical problem is illustrated, when the latter is stable. It is clear that positive coupling terms and negative coupling terms with a high absolute value can lead to loss of stability, though the value of s when positive is an increasing function of λ. In fact s is in all cases equal to zero when λ is equal to zero.

For the values given in table 1 the effect of the diffusivity is illustrated in figure 16b. It is clear that the maximum value of s corresponds to a nonzero value of λ, which increases with increasing diffusivity, leading one to expect a wider localized zone for higher diffusivity values. On the other hand, changes in the inertia, in the form of changes in the material density, have a very similar effect, changing both the magnitude and location of the maximum value of s, as shown in figure 17a. The effect on the location of the maximum seems to be less pronounced than that of the diffusivity, but this may well be linked to the values selected here.



(a) Γ=1.0. (b) Effect of $D^{(i)}$.

Figure 16: s versus λ for different values of parameters.



(a) Γ=1.0. (b) $\hat{\Delta}$=0.1.

Figure 17: s versus λ for different values of $\hat{\alpha}$= and $\hat{\Delta}$.

In figure 17b the effect of the coupling term is illustrated. The most obvious result is an increase in the maximum for increasing values of the coupling term. A detail of this figure is shown in figure 18, illustrating the existence of a first positive branch of s, which goes to infinity for zero λ. This area becomes smaller with increasing values of the coupling term. This results from the mechanical instability and can be alleviated by employing a mechanical response accounting for the materials intrinsic characteristic time or length.

On the whole it may be concluded that (two-way) multiphysical couplings can have both a stabilizing and a destabilizing effect on problem, but in any case introduce a finite width for the localization zone.

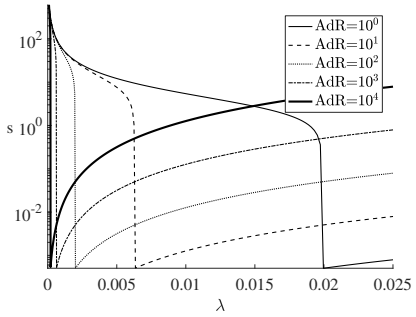


Figure 18: Detail of figure 17b.

5.1 1D Example

The example considered in section 2.5 is revisited here, incorporating thermal coupling. It is assumed that the material exhibits either thermal softening or hardening and that the shearing process generates heat. The constitutive law in its linearized form will read:

$$\begin{aligned} \tilde{\sigma}_{12} &= 2G\frac{h}{1+h}\tilde{\varepsilon}_{12} + AT \\ \tilde{\sigma}_{22} &= M\tilde{\varepsilon}_{22} + BT, \end{aligned} \tag{97}$$

where T stands for the temperature, $M = K + \frac{4G}{3}$ is the p-wave elastic modulus and $h = \frac{1}{G}\frac{d\tau_0}{dq} > -1$ is the hardening modulus.

The temperature on the other hand will obey the equation:

$$\frac{\partial T}{\partial t} = \kappa\frac{\partial^2 T}{\partial x^2} + 2\beta\tilde{\varepsilon}_{12}, \tag{98}$$

where κ and β are assumed to be constants.

From the first balance equation and using the usual forms for the perturbations we obtain

$$-(k^2 G\frac{h}{1+h} + s^2\rho)g + \mathrm{i}kA\theta = 0 \Rightarrow \tag{99}$$

$$-(k^2 v_s^2\frac{h}{1+h} + s^2)g + \mathrm{i}k\frac{A}{\rho}\theta = 0. \tag{100}$$

Correspondingly from the heat balance we obtain

$$\mathrm{i}k\beta g - \left(k^2\kappa + s\right)\theta = 0. \tag{101}$$

For more than one non-trivial solutions to exist, the determinant of the system consisting of equations (100) and (101) must be equal to zero:

$$s^3 + k^2\kappa s^2 + k^2 v_s^2\frac{h}{1+h}s + k^4 v_s^2\frac{h}{1+h}\kappa + k^2\frac{A\beta}{\rho} = 0 \Rightarrow \tag{102}$$

$$\hat{s}^3 + \hat{k}^2\hat{s}^2 + \hat{k}^2\frac{h}{1+h}\hat{s} + \hat{k}^4\frac{h}{1+h} + \hat{k}^2\frac{A\beta\kappa}{\rho v_s^4} = 0, \tag{103}$$

with

$$\hat{k} = \frac{\kappa}{v_s}k,\ \hat{s} = \frac{v_s^2}{\kappa}s. \tag{104}$$

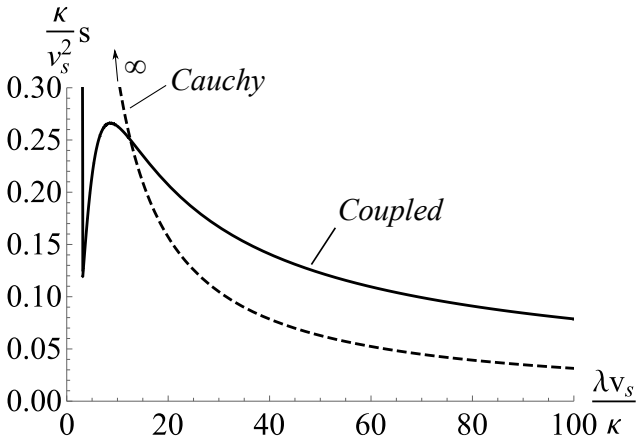


Figure 19: Growth coefficient in function of the perturbation wavelength. While an infinite value is observed for zero wavelength, a second maximum is present.

The growth coefficient as a function of the perturbation length is shown in figure 19. A branch tending to infinity for $\lambda \to 0$ can be observed in a way similar to figure 18. As already mentioned, this is a result of the lack of an internal length or a characteristic time for the material.

6 Conclusions

This chapter focuses on providing the basic tools to graduate students for studying strain localization in solids. The fundamental notions of loss of uniqueness, bifurcation, stability, ill-posedness and mesh dependency are explained through simple examples. Without overlooking classical approaches in bifurcation analysis, we study strain localization by using the systematic mathematical framework that provides Lyapunov stability. More specifically, we use the first Lyapunov method for exploring the conditions for which equilibria of homogeneous deformation become unstable leading to strain localization. In this method we determine the growth in time of perturbations of arbitrary wavelength from the equilibrium (steady-)state. One-dimensional examples are systematically given to help understanding keeping calculus to the minimum.

First we study strain localization in a classical Cauchy, Boltzmann continuum. We limit our analysis in deformation bands, i.e. compaction, shear, dilation bands and

their combinations. This type of strain localization is often observed in several scales, starting from laboratory experiments, such as shear band formation in a granular material or compaction bands in porous rocks, to geological settings, such as faults and landslides. We derive the conditions for strain localization for a general rate-independent constitutive law. Under these assumptions we retrieve the acoustic tensor and we study the dependence of the Lyapunov exponent (growth coefficient) in terms of the perturbation wave length in order to determine the thickness of the localization zone. We show that the perturbations that evolve faster in time (and dominate over the others are characterized by asymptotically zero perturbation wave length and have infinite Lyapunov exponent (singularity in time). This means that deformation bands have zero thickness, which is in contrast with observations. This mathematical artifact explains also the observed mesh-dependency in finite element analysis of strain localization when rate-independent Cauchy continua are used.

The aforementioned pathology is partially remedied when rate-dependent Cauchy continua are used. This viscous regularization introduces a characteristic time into the system. Due to the presence of the aforementioned characteristic time, the Lyapunov exponent remains finite (regularization in time). Moreover, when inertia effects are negligible, it renders the system independent of the perturbation wave length. As a result, the behavior of the system during strain localization (e.g. stress-strain profile and deformation band thickness) is determined only by existing imperfections. Such imperfections can be parasitic stresses or material defects. When inertia is not negligible, the dominant perturbation is again the one characterized by the smallest wavelength. Scale analysis shows that inertia terms are important when the perturbation wave lengths are larger than a characteristic wavelength, which depends on the material parameters. Three characteristic times are also identified showing when inertia, viscosity or rate-independent behavior can be neglected. An one-dimensional example using Perzyna viscoplasticity illustrates in a simple way most of the above mathematical findings.

An alternative regularization technique, is the use of Micromorphic continua, such as strain gradient and Cosserat continua, which enrich the continuum description with characteristic lengths. The presence of these lengths remove mesh-dependency and determine the thickness of the localization zone, which is proportional to the aforementioned internal lengths. Moreover, in the presence of inertia terms the Lyapunov exponent if finite. Therefore, Micromorphic continua remedy both spatial and time singularities, in the expense however of a more complex theory and complicated mathematical derivations. As an example we study strain localization in a rate-independent Cosserat continuum. We use first a general constitutive law in three-dimensions and then we give an one-dimensional example of a sheared infinite layer in order to clarify the mathematical derivations and to illustrate how Cosserat continuum regularizes the problem. Notice, that Cosserat continuum was successfully used by Mühlhauss and Vardoulakis for predicting the shear band thickness of granular materials [MV87].

The chapter closes with a section dedicated to multi-physics couplings and their effects of strain localization. Thermo-Hydro-Chemo-Mechanical effects (THMC, among oth-

ers) introduce several length and time scales to the system and consequently regularize in a physical manner the underlying mathematical problem. Our analysis is again general, in three-dimensions and considers n-couplings. We show that two-way coupling is necessary for regularization. An one-dimensional example of an infinite layer with thermo-mechanical couplings is then presented for making clear the effects of the various physical mechanism. A rate-independent Cauchy material was used in this last example.

Following this theoretical results and examples the following question is raised:
Which is the best way and theory for best describing strain localization in solids and in particular in geomaterials?
The answer is always found by the modeler and depends on the physical/engineering problem at hand. For instance, for modeling the stress-strain response and the thickness of the principal slip zone of seismic faults (i.e. a narrow shear band formed during seismic slip) a THMC Cosserat continuum was recently used providing realistic predictions [RSS18, RSS18, VSS13]. For studying the damage zone during gallery excavation in the context of radioactive waste disposal a double-scale, poro-mechanical, strain-gradient model was employed [EBC$^+$16].

Acknowledgements

The author, I.S., would like to acknowledge the support of the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program (Grant agreement no. 757848 CoQuake).

Appendix

A Classical bifurcation analysis and acoustic tensor

Consider a homogeneous, homogeneously deformed solid subjected to quasi-static increments of deformation. Let's assume that after an increment, a deformation band of thickness H is formed, which breaks the aforementioned homogeneity of the deformation field (and consequently of the stress field) as shown in figure 20. The displacement field remains continuous across the boundaries of the band, but its gradient does not (different strains inside the band):

$$[\![\Delta u_i]\!] = 0 \quad \text{and} \quad [\![\Delta u_{i,j}]\!] = g_i n_j \tag{105}$$

where $[\![.]\!]$ denotes discontinuity across the deformation band boundary (e.g. $[\![\alpha]\!] = \alpha^+ - \alpha^-$), n_i is the orientation vector of the deformation band with $i = 1, 2, 3$ in the three-dimensional space, u_i the displacement field and Δ denotes the increment of a field. $(.)_{,i}$ denotes derivation in terms of x_i.

The jump of the shear stresses at the boundary of the shear band is not zero due to

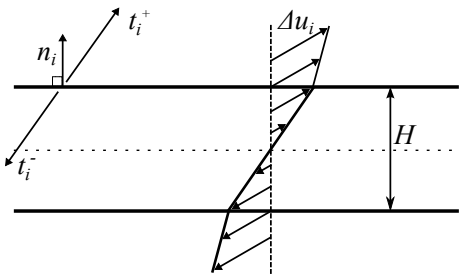


Figure 20: Schematic representation of a deformation band and of the discontinuity of the strain's field.

acceleration (not in equilibrium). From the linear momentum balance we obtain:

$$[\![\Delta t_i]\!] = [\![\Delta \sigma_{ij}]\!] n_j = -\rho c [\![\gamma_i]\!] \tag{106}$$

where c is the velocity of a propagating discontinuity in direction n_i such that $[\![\gamma_i]\!] = [\![\Delta v_i]\!] = -cg_i$, with v_i the velocity field (see Hadamard conditions on propagating discontinuities [Had03, LCBD09]). Consider the class of materials that for a small increment Δ, the constitutive law can be written (linearized) as follows:

$$[\![\Delta \sigma_{ij}]\!] = L_{ijkl} \Delta u_{k,l} \tag{107}$$

The tensor L_{ijkl} can be continuous across the boundary of the band ($[\![L_{ijkl}]\!] = 0$) or discontinuous in the sense that elastic unloading can occur outside the band, while continued inelastic loading continues within the band. In the first case we say that we have *continuous bifurcation*, while in the second *discontinuous bifurcation*. It is shown that continuous bifurcation precedes discontinuous bifurcation [RR80]. Inserting Eq.(108) into (106) and using (105) we get:

$$\left(n_j L_{ijkl} n_l - \rho c^2 \delta_{ik}\right) g_k = 0 \tag{108}$$

where $\Gamma_{ij} = n_j L_{ijkl} n_l$ is the acoustic tensor. This equation coincides with equation (8).

References

[AHV+12] Edward Andò, Stephen A. Hall, Gioacchino Viggiani, Jacques Desrues, and Pierre Bésuelle. Grain-scale experimental investigation of localised deformation in sand: A discrete particle tracking approach. *Acta Geotechnica*, 7(1):1–13, 2012.

[Ben05] Ahmed Benallal. On localization modes in coupled thermo-hydro-mechanical problems. *Comptes Rendus Mécanique*, 333(7):557–564, jul 2005.

[BH91a] Davide Bigoni and Tomasz Hueckel. Uniqueness and localization—I. Associative and non-associative elastoplasticity. *International Journal of Solids and Structures*, 28(2):197–213, 1991.

[BH91b] Davide Bigoni and Tomasz Hueckel. Uniqueness and localization—II. Coupled elastoplasticity. *International Journal of Solids and Structures*, 28(2):215–224, jan 1991.

[BN69] F. Brauer and J.A. Nohel. *The Qualitative Theory of Ordinary Differential Equations: An Introduction*. Dover Publications, 1969.

[Bow09] Allan F. Bower. *Applied Mechanics of Solids*. CRC Press, 2009.

[BSS17] Nicolas Brantut, Ioannis Stefanou, and Jean Sulem. Dehydration-induced instabilities at intermediate depths in subduction zones. *Journal of Geophysical Research: Solid Earth*, 122(8):6087–6107, aug 2017.

[BV01] J. P. Bardet and Ioannis Vardoulakis. The asymmetry of stress in granular media. *International Journal of Solids and Structures*, 38(2):353–367, jan 2001.

[CCC06] F. Collin, R. Chambon, and R. Charlier. A finite element method for poro mechanical modelling of geotechnical problems using local second gradient models. *International Journal for Numerical Methods in Engineering*, 65(11):1749–1772, mar 2006.

[CCE98] R. Chambon, D. Caillerie, and N. El Hassan. One-dimensional localisation studied with a second grade model. *European Journal of Mechanics, A/Solids*, 17(4):637–656, 1998.

[CCM01] René Chambon, Denis Caillerie, and Takashi Matsuchima. Plastic continuum with microstructure, local second gradient theories for geomaterials: Localization studies. *International Journal of Solids and Structures*, 38(46-47):8503–8527, 2001.

[CCV04] René Chambon, Denis Caillerie, and Gioacchino Viggiani. Loss of uniqueness and bifurcation vs instability: some remarks. *Revue Française de Génie Civil*, 8(5-6):517–535, 2004.

[DAD+17] Jacques Desrues, A Argilaga, S. Dal Pont, G Combe, Denis Caillerie, and T. kein Nguyen. Restoring Mesh Independency in FEM-DEM Multi-scale Modelling of Strain Localization Using Second Gradient Regularization. In E. Papamichos, P. Papanastasiou, E. Pasternak, and A. Dyskin, editors, *Bifurcation and Degradation of Geomaterials with Engineering Applications. IWBDG 2017. Springer Series in Geomechanics and Geoengineering*, pages 453–457. Springer, Cham, 2017.

[DSMP93] R. De Borst, Lambertus J Sluys, H.B. Mühlhaus, and J. Pamin. Fundamental issues in finite element analyses of localization of deformation. *Engineering Computations*, 10(2):99–121, feb 1993.

[EBC+16] A P Van Den Eijnden, P Bésuelle, F. Collin, René Chambon, and Jacques Desrues. Modeling the strain localization around an underground gallery with a hydro-mechanical double scale model ; effect of anisotropy. *Computers and Geotechnics*, 2016.

[Eri99] A. Cemal Eringen. *Microcontinuum Field Thoeries*. Springer Verlag, 1999.

[Ger73a] P. Germain. La méthode des puissances virtuelles en mécanique des milieux continus - 1. Théorie du second gradient. *Journal de Mecanique Theorique et Appliquee*, 12(2):235–274, 1973.

[Ger73b] P. Germain. The Method of Virtual Power in Continuum Mechanics. Part 2: Microstructure. *SIAM Journal on Applied Mathematics*, 25(3):556–575, nov 1973.

[GSSS16] Michele Godio, Ioannis Stefanou, Karam Sab, and Jean Sulem. Multisurface plasticity for Cosserat materials: Plate element implementation and validation. *International Journal for Numerical Methods in Engineering*, 108(5):456–484, nov 2016.

[Had02] Jacques Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 49-52, 1902.

[Had03] Jacques Hadamard. *Leçcons sur la propagation des ondes et les équations de l'hydrodynamique*. Librairie Scientifique A.Hermann, Paris, 1903.

[IR00] Kathleen A. Issen and John W. Rudnicki. Conditions for compaction bands in porous rock. *Journal of Geophysical Research*, 105(B9):21529, 2000.

[IW98] Maria-Magdalena Iordache and Kaspar J Willam. Localized failure analysis in elastoplastic Cosserat continua. *Computer Methods in Applied Mechanics and Engineering*, 151(3-4):559–586, 1998.

[KABC08] P. Kotronis, S. Al Holo, P. Bésuelle, and R. Chambon. Shear softening and localization: Modelling the evolution of the width of the shear zone. *Acta Geotechnica*, 3(2):85–97, mar 2008.

[LCBD09] Jean Lemaitre, Jean-Louis Chaboche, Ahmed Benallal, and Rodrigue Desmorat. *Mécanique des matériaux solides*. Dunod, Paris, 2009.

[Lya92] Aleksandr Mikhailovich Lyapunov. The general problem of the stability of motion. *International Journal of Control*, 55:531–773, 1992.

[Min64] R.D. Mindlin. Micro-structure in linear elasticity. *Archive for Rational Mechanics and Analysis*, 16(1), 1964.

[Min65] R.D. Mindlin. Second gradient of strain and surface-tension in linear elasticity. *International Journal of Solids and Structures*, 1(4):417–438, nov 1965.

[MV87] Hans-Bernd Mühlhaus and Ioannis Vardoulakis. The thickness of shear bands in granular materials. *Géotechnique*, 37(3):271–283, sep 1987.

[ND11] François Nicot and Félix Darve. Diffuse and localized failure modes: Two competing mechanisms. ... *for Numerical and Analytical Methods in* ..., 35(5):586–601, apr 2011.

[Nee88] A. Needleman. Material rate dependence and mesh sensitivity in localization problems. *Computer Methods in Applied Mechanics and Engineering*, 67(1):69–85, 1988.

[Per66] P. Perzyna. Fundamental problems in viscoplasticity. *Recent Advances in Applied Mechanics*, 9:243–377, 1966.

[PZ16] Panos C. Papanastasiou and Antonios Zervos. Numerical modelling of strain localization. In *Modelling of instabilities and bifurcation in Geomechanics, ALERT geomaterials doctoral school*, 2016.

[RC16] Roozbeh Rezakhani and Gianluca Cusatis. Asymptotic expansion homogenization of discrete fine-scale models with rotational degrees of freedom for the simulation of quasi-brittle materials. *Journal of the Mechanics and Physics of Solids*, 88:320–345, 2016.

[RR75] John W. Rudnicki and James R. Rice. Conditions for the localization of deformation in pressure-sensitive dilatant materials. *Journal of the Mechanics and Physics of Solids*, 23(6):371–394, dec 1975.

[RR80] James R. Rice and John W. Rudnicki. A note on some features of the theory of localization of deformation. *International Journal of Solids and Structures*, 16:597–605, 1980.

[RSS18] Hadrien Rattez, Ioannis Stefanou, and Jean Sulem. The importance of Thermo-Hydro-Mechanical couplings and microstructure to strain localization in 3D continua with application to seismic faults. Part I: Theory and linear stability analysis. *Journal of the Mechanics and Physics of Solids*, 115:54–76, jun 2018.

[SA16] Ioannis Stefanou and S. Alevizos. Fundamentals of bifurcation theory and stability analysis. In Jean Sulem, Ioannis Stefanou, Euripides Papamichos, and Emmanuil Veveakis, editors, *Modelling of instabilities and bifurcation in Geomechanics, ALERT geomaterials Doctoral School 2016*. Aussois, France, 2016.

[SDC07] Giulio Sciarra, Francesco Dell'Isola, and Olivier Coussy. Second gradient poromechanics. *International Journal of Solids and Structures*, 44(20):6607–6629, oct 2007.

[SSV10] Ioannis Stefanou, Jean Sulem, and Ioannis Vardoulakis. Homogenization of interlocking masonry structures using a generalized differential expan-

sion technique. *International Journal of Solids and Structures*, 47(11-12):1522–1536, jun 2010.

[SSV11] Jean Sulem, Ioannis Stefanou, and Emmanuil Veveakis. Stability analysis of undrained adiabatic shearing of a rock layer with Cosserat microstructure. *Granular Matter*, 13(3):261–268, feb 2011.

[Ste18] Ioannis Stefanou. Definition and Uses of the Principle of Virtual Power The Principle of Virtual Power. In *ALERT geomaterials Doctoral School 2018*. 2018.

[SW91] Paul Steinmann and Kaspar J Willam. Localization within the Framework of Micropolar Elasto-Plasticity. *Advances in Continuum Mechanics*, pages 296–313, 1991.

[Var96a] Ioannis Vardoulakis. Deformation of water-saturated sand: I. uniform undrained deformation and shear banding. *Géotechnique*, 46(3):441–456, jan 1996.

[Var96b] Ioannis Vardoulakis. Deformation of water-saturated sand: II. effect of pore water flow and shear banding. *Géotechnique*, 46(3):457–472, jan 1996.

[Var09] Ioannis Vardoulakis. *Lecture notes on Cosserat continuum mechanics with application to the mechanics of granular media*. http://geolab.mechan.ntua.gr/teaching/lectnotes/CCM2009.pdf, 2009.

[VS95] Ioannis Vardoulakis and Jean Sulem. *Bifurcation Analysis in Geomechanics*. Blackie, Glascow, 1995.

[VSS13] Emmanuil Veveakis, Ioannis Stefanou, and Jean Sulem. Failure in shear bands for granular materials: thermo-hydro-chemo-mechanical effects. *Géotechnique Letters*, 3(April-June):31–36, may 2013.

[WSD96] W.M. Wang, Lambertus J Sluys, and René De Borst. Interaction between material length scale and imperfection size for localisation phenomena in viscoplastic media. *European Journal of Mechanics - A/Solids*, 15(3):447–464, 1996.

[ZPV01] Antonios Zervos, Panos C. Papanastasiou, and Ioannis Vardoulakis. A finite element displacement formulation for gradient elastoplasticity. *International Journal for Numerical Methods in Engineering*, 50(6):1369–1388, 2001.

Petroleum Geomechanics

Euripides Papamichos[1], Panos Papanastasiou[2]

[1]*SINTEF Industry, N-7465 Trondheim, Norway*
Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece
epapamic@civil.auth.gr

[2] *University of Cyprus, 1678 Nicosia, Cyprus*
panospap@ucy.ac.cy

This paper presents basic knowledge, analytical, physical, mathematical and computational modelling for the solution of research and advanced practical problems in petroleum engineering related to geomechanics. Petroleum geomechanics introduces the in-situ stresses and rock-fabric in coupled flow-stress analysis in order to give more accurate predictions and to account for problems such as reservoir compaction and surface subsidence, casing collapse, fault activation and other. In all these problems the knowledge of the in-situ stresses and rock strength is required. For the applications of petroleum geomechanics we present modelling approaches to the problems of wellbore stability, sand prediction analysis and hydraulic fracturing.

1 Introduction

In the last three decades the oil & gas industry has witnessed what can be called the 'petroleum geomechanics revolution'. Geomechanics has become a regular consideration in all the phases of a field development from exploration to development and production. For the operator's perspective regarding the commercial value of geomechanics, evidence of the significance role of geomechanics is drawn from many field cases in the areas of drilling, production and reservoir management. Within the oil & gas service industry, geomechanics became the fastest growing commercial area for technical investment and payed revenue. Among the main drivers for the promotion and advances of petroleum geomechanics are the high rig daily rates in in excess of $500,000/day in the deep water of the Gulf of Mexico, in offshore West Africa and recently in the East Mediterranean, the drilling in harsh environments

such as tectonic fields, salt-domes, high-pressure high-temperature fields, and the drilling of more horizontal and extended reach wells spanned with multilateral junctions.

The last decade geomechanics plays a critical role in the shale oil and gas revolution successfully used in optimising shale hydrocarbon exploitation. Geomechanics knowledge in shale prospects enables an engineer or geoscientist to make better field development and operational decisions in spacing and drilling of stable long horizontal sections and in multistage fracturing of stiff naturally fractured rock masses. Heterogeneity, mechanical anisotropy, and natural fractures in shale formations and their influence on drilling and stimulation are critical elements of a successful shale stimulation programme.

Environmental concern and restrictions for the disposal of contaminated cuttings and produced water promoted new applications for reinjection and hydraulic fracturing. The demand for higher hydrocarbon recovery, leading to extreme reservoir depletions, causes other problems such as surface subsidence and wellbore collapse, and opened a new area called 'reservoir geomechanics'. Like in any other field, advances in the information technology such as software, visualization tools, the internet, and recently the handling of 'big data' with cloud computing and storage have their share in the promotion of petroleum geomechanics.

In this paper we will focus on the three major problems related to petroleum geomechanics presenting basic knowledge, analytical, physical, mathematical and computational modelling in the solution of practical problems in petroleum engineering. In section 1 we will present the problem of wellbore stability and on the methods used to for calculating the optimum mud density for drilling wellbores. In section 2 we will extend the analysis for determining the optimum wellbore pressure for sand-free production of hydrocarbons. In section 3 we will present a modeling approach to hydraulic fracturing based on FEM which is employed for examining the influence of different coupled non-linear processes, such as plasticity, cohesive cracking and pore pressure which are involved in some applications of hydraulic fracturing.

We recognize the work done and published in numerous papers on petroleum geomechanics by many researchers and practitioners over the last half century. Some of the references, but not limited to those, will be given next within the sections. For a general reference on petroleum geomechanics the readers can consider the books [Fja01], [Zob01].

2 Borehole Stability

Borehole instabilities during drilling cause substantial problems and result according to the industry about 5–10% of the drilling costs in exploration and production, incorporating loss of time and sometimes also of equipment [Fja01]. These numbers mean that the worldwide is in the hundreds of million dollars per year. A borehole

stability problem is an example of what drillers refer to as a tight hole or stuck pipe but also when they refer to mud losses. There are many possible causes of wellbore stability but usually the fundamental reason is mechanical collapse of the borehole. Table 1 lists after Pasic et al. [Pas01], possible causes of wellbore stability. Most instabilities occur in the overburden shale formations but sometimes they may occur within the reservoir. Sometimes the culprit is a lack of hole cleaning ability. Table 2 lists indicators of borehole instabilities.

Table 1. Causes of borehole stability (after Pasic et al. [Pas01]).

Uncontrollable (natural) factors	Controllable factors
Naturally fractured or faulted formations	Bottom hole pressure (mud density)
Tectonically stresses formations	Well inclination and azimuth
High in situ stresses	Transient pore pressures
Mobile formations	Physico/chemical rock-fluid interaction
Unconsolidated formations	Drill string vibrations
Naturally over-pressured shale collapse	Erosion
Induced over-pressured shale collapse	Temperature

Table 2. Indicators of borehole instabilities.

Direct indicators	Indirect indicators
Oversize hole	High torque and drag (friction)
Undergauge hole	Hanging up of drillstring, casing or coiled tubing
Excessive volume of cuttings	Increased circulating pressures
Excessive volume of cavings	Stuck pipe
Cavings at surface	Excessive drillstring vibrations
Hole fill after tripping	Drillstring failure
Excess cement volume required	Deviation control problems
	Inability to run logs
	Poor logging response
	Annular gas leakage
	Keyhole seating
	Excessive doglegs

Traditionally, the oil industry has looked at borehole instability as being caused by clay swelling, which can be treated by chemical additives (e.g. salt) to the drilling mud. The selection of mud weight has been governed by the pore pressure and the fracture gradient profiles: In order to prevent influx of fluids (in particular gas) it has been considered necessary to keep the mud weight above the pore pressure gradient. In order to prevent loss of mud into fractures it has been found necessary to keep the mud weight below the fracture gradient. This results in a maximum and minimum

mud weight that is called mud window. The mud window can also be estimated with rock mechanics analysis.

In the analysis the stresses at the well are calculated form elastic or elastoplastic analysis. These stresses are counteracted by the drilling mud in the well which prevent flow of pore fluid in the wellbore, prevents hole failure, transports drill cuttings to surface and cools the drill-bit. The mud pressure p_w at a certain depth D of the well is controlled by the mud density ρ_w with $p_w = \rho_w g D$. The oil well drilling language refers to mud weight in density units and gradients of stress or pressure equivalently in density gradient units. When the mud is circulating the Equivalent Circulating Density (ECD) is used because the effective mud pressure is 5-10% higher than the static mud pressure due to friction in the annulus during flow.

The borehole stability rock mechanics analysis can be easier demonstrated in a vertical well. Horizontal and inclined boreholes can then be treated accordingly but applying the same methodology. Figure 1 shows the stresses in the formation as a function of normalized radial distance r/R at a vertical impermeable borehole based on linear elastic rock and isotropic horizontal stresses. The well radius is R, and σ_r, σ_θ, σ_z are the radial, tangential and axial stresses in the cylindrical coordinate system of the well. The far field horizontal stress is σ_h and the vertical stress is σ_v. The wall is impermeable wall when we have a perfect mud cake or we drill in shale. The formation pressure is denoted p_{fm}.

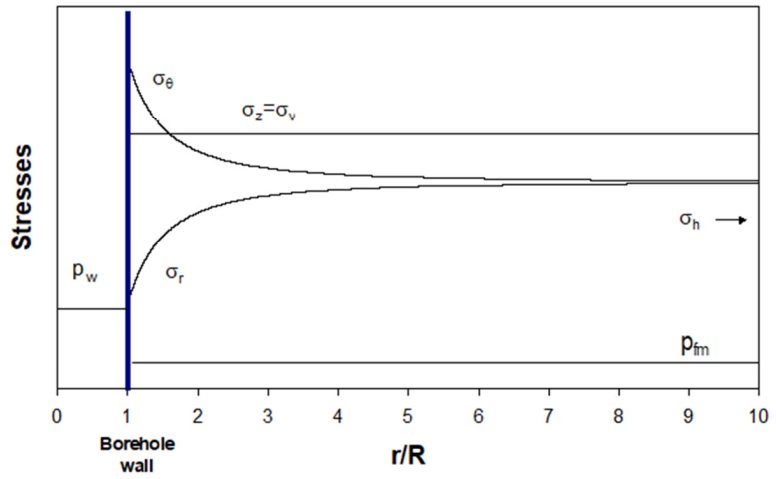
Depending on the stress magnitude we may have different cases:
(a) The stresses are $\sigma_\theta > \sigma_z > \sigma_r$. In this case $\sigma_r = p_w, \sigma_\theta = 2\sigma_h - p_w, \sigma_z = \sigma_v$.
(b) The stresses are $\sigma_z > \sigma_\theta > \sigma_r$ and $\sigma_r = p_w, \sigma_\theta = 2\sigma_h - p_w, \sigma_z = \sigma_v$.
The borehole stability analysis uses the borehole stresses and the Mohr-Coulomb (MC) failure criterion to determine the minimum permitted well pressure to prevent shear failure at borehole wall (hole collapse). The MC failure criterion is written in terms of the maximum and minimum effective stress, the Unconfined Compressive Strength (UCS), the cohesion c and the friction angle φ as

$$\sigma_1' = UCS + m\sigma_3', \qquad UCS = \frac{2c\cos\varphi}{1-\sin\varphi}, \qquad m = \tan^2\beta = \frac{1+\sin\varphi}{1-\sin\varphi} \tag{1}$$

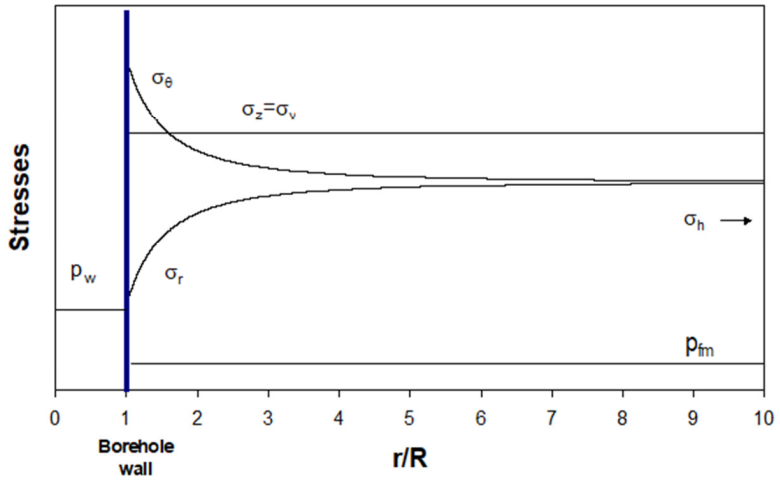
Substituting the stresses results in

$$p_{w,\min}^{(a)} = \frac{2\sigma_h - UCS + (m-1)\,p_{fm}}{m+1}, \quad p_{w,\min}^{(b)} = \frac{\sigma_v - UCS + (m-1)\,p_{fm}}{m} \tag{2}$$

for cases (a) and (b), respectively. The mud weight is then calculated as $\rho_{w,\min} = \frac{p_{w,\min}}{gD}$.



(a)



(b)

Figure 1. *Stresses at vertical impermeable borehole wall based on linearly elastic rock and isotropic horizontal stresses. Cases (a) and (b) depending on the magnitude of the stresses in the borehole wall.*

In general, for the various stress combinations Table 3 gives the minimum mud pressure to prevent shear failure instabilities. Figure 2 shows the shape of the failure modes after Maury [Mau01]. Based on the shape of the cavings in the surface we may determine the failure mechancism.

Table 3. Minimum mud pressure for different stress magnitudes

Case	$\sigma_1 \geq \sigma_2 \geq \sigma_3$	Borehole failure occurs if
a	$\sigma_\theta \geq \sigma_z \geq \sigma_r$	$p_w \leq p_f + \frac{2(\sigma_h - p_f) - C_0}{1 + \tan^2 \beta}$
b	$\sigma_z \geq \sigma_\theta \geq \sigma_r$	$p_w \leq p_f + \frac{\sigma_v - p_f - C_0}{\tan^2 \beta}$
c	$\sigma_z \geq \sigma_r \geq \sigma_\theta$	$p_w \geq p_f + 2(\sigma_h - p_f) - \frac{\sigma_v - p_f - C_0}{\tan^2 \beta}$
d	$\sigma_r \geq \sigma_z \geq \sigma_\theta$	$p_w \geq p_f + \frac{2(\sigma_h - p_f)\tan^2 \beta + C_0}{1 + \tan^2 \beta}$
e	$\sigma_r \geq \sigma_\theta \geq \sigma_z$	$p_w \geq p_f + (\sigma_v - p_f)\tan^2 \beta + C_0$
f	$\sigma_\theta \geq \sigma_r \geq \sigma_z$	$p_w \leq p_f + 2(\sigma_h - p_f) - (\sigma_v - p_f)\tan^2 \beta - C_0$

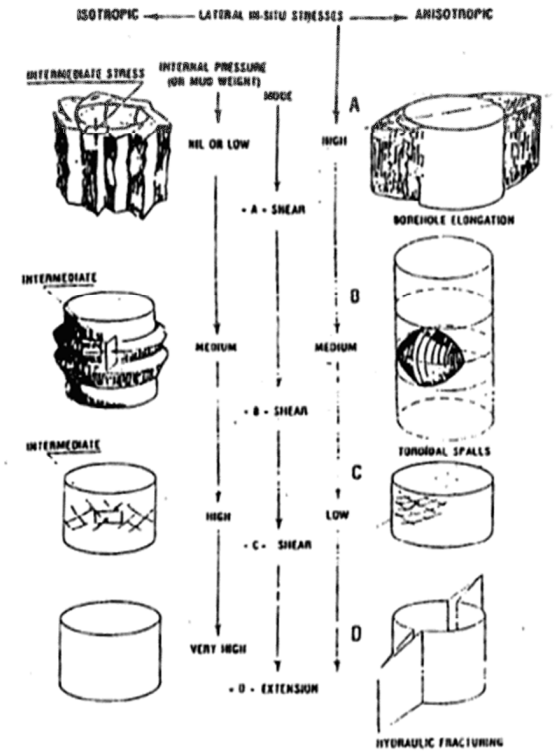


Figure 2. *Borehole failure modes depending on the magnitude of the borehole stresses (after Maury [Mau01]).*

Al low well pressure we may run into tensile failure of the well due to tensile radial stresses. This may occur at low well pressure when we for example have underbalanced drilling with $p_w < p_{fm}$. The condition for this is that

$$\sigma'_r = -T_s \quad \Rightarrow \quad p_{w,\min}^{rad,tens} = p_{fm} - T_s \tag{3}$$

where T_s is the tensile strength which of shale is rather low. This failure mode gives sharp, blade-shaped cavings and can lead to tight hole. The stress situation for tensile radials tress failure is shown in Figure 3.

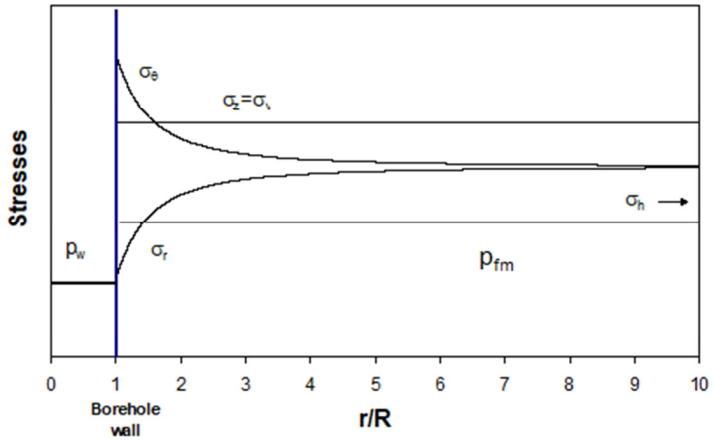


Figure 3. *Stresses at vertical impermeable borehole wall for tensile radial stress failure.*

Tensile failure may also occur at high well pressure and lead to hydraulic fracturing of the formation, as shown in the schematic in Figure 4.

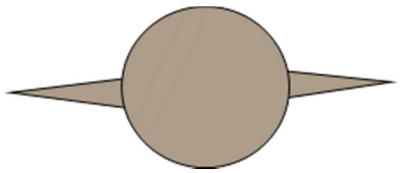


Figure 4. *Hydraulic fracturing due to high well pressure.*

The condition in that case is

$$\sigma'_\theta = -T_s \quad \Rightarrow \quad p_{w,\max}^{frac} = 2\sigma_h - p_{fm} + T_s \tag{4}$$

Depending on the magnitude of the other borehole stresses three cases are possible as illustrated in Figure 5.

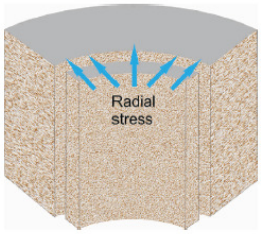
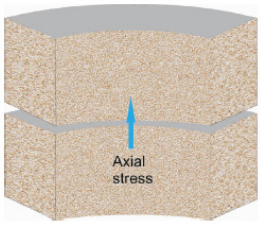
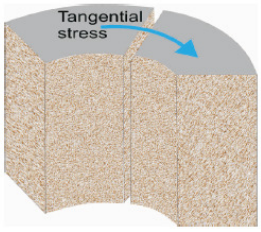
Failure type	Geometry and Orientation	Figure
Tensile Failure Cylindrical $\sigma_r \leq -T_o$	This failure is concentric with the borehole. A low mud weight would favor the failure due to the magnitude of σ_r being lower.	
Tensile Failure Horizontal $\sigma_a \leq -T_o$	This failure creates horizontal fractures.	
Tensile Failure Vertical $\sigma_t \leq -T_o$	This failure creates a vertical fracture parallel with the maximum horizontal stress direction. This is because, this orientation is the tangential stress has to overcome the smallest formation tensile strength.	

Figure 5. *Hydraulic fracturing modes due to high well pressure. The second case is not likely in vertical wells.*

This analysis provides the mud weight window for the drilling program. It can summarized as follows:

(a) The minimum mud weight is determined from the maximum values as determined by (i) Hole collapse in shale (shear failure case (a) or (b)), (ii) Radial tensile failure in shale, (iii) Pore pressure in the case that underbalanced drilling is prohibited.

(b) The Maximum mud weight is determined form the minimum value as determined by (i) σ_h the minimum horizontal stress in case of pre-existing natural fractures, (b) fracturing of borehole wall.

Figure 6 provides after Hawkes and McLellan [Haw01] a schematic diagram of the effect of increasing mud weight in the failure mode of the borehole.

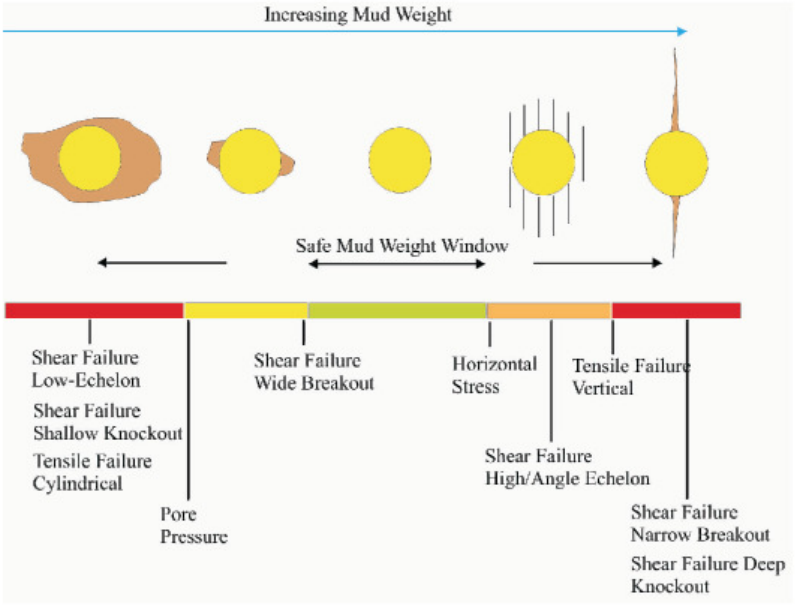


Figure 6. *Hydraulic fracturing modes due to high well pressure. The second case is not likely in vertical wells.*

Besides the basic rock mechanics approach that was described here other issues enter also the borehole stability analysis such as (i) anisotropic stresses, (ii) deviated holes, (iii) rock anisotropy where shales have natural bedding which acts as weak plane, (iv) Plasticity effects, (v) Time-delayed borehole failure due to e.g. creep, consolidation, cooling or due to chemical interaction with the mud fluid.

3 Sand production

Sand production often occurs in the petroleum industry as the result of erosion of reservoir sandstones during hydrocarbon production. In such cases, the rock around a wellbore or perforation is plastified, decohesioned and weakened due to the stress concentration of the in-situ stresses around the cavity. When weakened and decohesioned, the sandstone may erode by the produced fluid. The prediction of the sand production initiation and the amount of produced sand and how these are affected by the applied stresses and flow rates over time are important for safe and economical hydrocarbon production. It has been calculated that about 70% of the oil reserves are in reservoir sandstones that are prone to produce sand at some time during their production life.

In sandstone reservoirs, the wellbore itself may be used to produce hydrocarbons or perforations may be shot from the well vertically towards the formation. In the first case, sand production results from the failure of the wellbore while in the second case from the failure of the perforation itself. A perforated well is illustrated in Figure 7a where the well is first drilled and then a steel liner is put in place. The annulus between the liner and the formation is cemented to support the formation and perforations are shot from the wellbore to the formation to allow the production of the hydrocarbons when a positive drawdown is applied. The perforations are small tunnels with typical length ca. 0.5 m and typical diameter 1.5 cm. The drawdown is the difference between the reservoir pore pressure p_{res} and the well fluid pressure p_w. The production flow rate increases when the drawdown increases, i.e. when the well pressure p_w reduces. Figure 7b shows an example of massive sand production in an Indonesian field to demonstrate the problem.

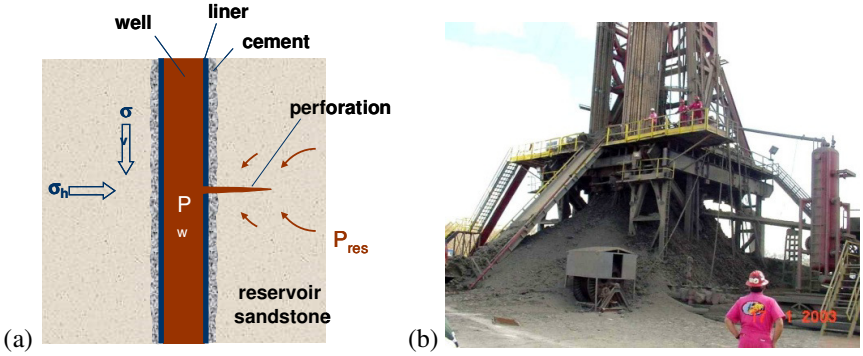


Figure 7. *(a) Schematic of a vertical perforated well in a sandstone reservoir under in situ total vertical stress σ_v, horizontal stress σ_h, reservoir pore pressure p_{res} and well fluid pressure p_w, and (b) Massive sand production in an Indonesian field.*

Sand production is a coupled hydro-mechanical process that involves two mechanisms: (i) the applied stresses fail the rock around the cavity, and (ii) the fluid flow removes or erodes the failed rock. The hydrodynamic forces themselves are too weak to erode the intact rock. In fact, the total hydrodynamic forces per unit volume of the rock are equal to the pore pressure gradient. The sand production process is shown in Figure 8 in a sequence of X-ray CT scans of a specimen under compression and radial fluid flow towards the cavity. The specimen initially fails close to the cavity. With increasing stress more material fails, and it is eroded or produced by the flowing fluid resulting in a larger cavity.

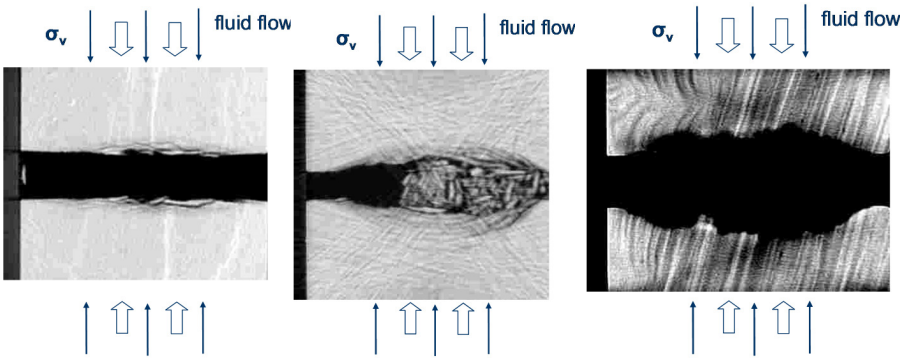


Figure 8. *Sequence (from left to right) of perforation failure and sand erosion under external compression and fluid flow towards the perforation cavity. X-ray CT scan sections along the hole axis of a hollow cylinder specimen tested in the laboratory [Pap05a].*

Sand production involves three regimes, as illustrated in Figure 9, which plots the sand mass rate vs. the applied stress (or drawdown): (i) the no sand regime, (ii) the manageable sand regime, and (iii) the catastrophic sand regime. These three regimes are separated from the sand initiation stress and the catastrophic sand stress. Thus in sand production studies it is important to (i) predict the sand initiation stress that delineates the no sand regime and corresponds to the stress for initial failure of the cavity, (ii) quantify the sand mass versus time, stress and flow rate (or drawdown and depletion) in the manageable sand regime where sand is produced at a controlled rate, and (iii) predict the catastrophic sand stress which marks the start of catastrophic sand production, i.e. a sand rate either unacceptable for the life and safety of the operations or so high that will cause plugging (sanding) of the well. In the last few years, despite the obvious disadvantages of sand production, the industry in many cases chooses to operate in the manageable sand regime because of the increased production rates. Often this is a question of designing the appropriate facilities to handle and dispose the produced sand.

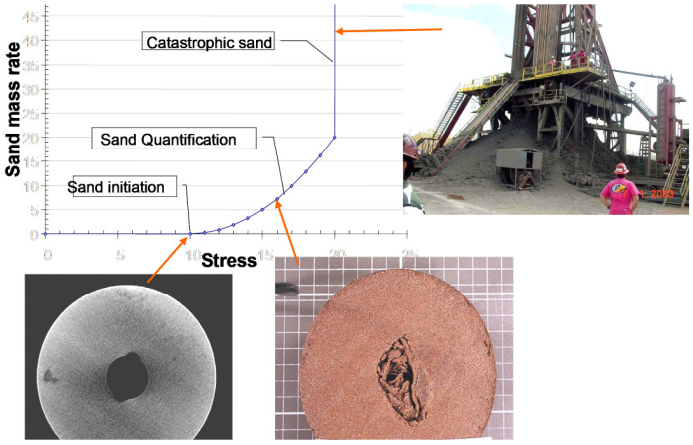


Figure 9. *Sand production regimes: (a) sand initiation, (b) manageable sand production, and (c) catastrophic sand production.*

3.1 Sand initiation

Analytical sand initiation or sand onset prediction models use the near wellbore or perforation stresses to predict formation failure. For shear stress failure, which is most often the case in sand production, the simplest failure criterion compares the tangential stress at the hole with a suitable formation strength, e.g. [Kes01]. More sophisticated criteria include all stress components to improve the predictions in field conditions under various anisotropic stress fields [Pal01, Pap01a, Pap02a, Pap03a]. Experimental results on Hollow Cylinder (HC) and Hollow Prism (HP) sandstone specimens have shown that such models are necessary to replicate experimental results [Pap02a, Pap03a]. Failure is assumed to be in shear due to the stress concentration at the hole. Failed rock is subsequently transported by the weak hydrodynamic forces of the flowing fluid leading to sand production. Sand onset is assumed to coincide with hole failure and therefore the two terms, hole failure strength and sand onset stress will be used in the following interchangeably.

Hole failure criteria based on classical rock mechanics failure models, i.e. the Mohr-Coulomb (MC) and the Drucker-Prager (DP), have been developed for better predictions under the anisotropic stresses in the field. The models were developed by applying the MC and DP failure criteria to the stresses at the hole which are calculated assuming linear poroelasticity. The MC criterion contains one material strength parameter which for better predictions is calibrated on the hole failure/sand onset stress σ_S. The sand onset stress corresponds to the isotropic loading stress for hole failure in a HC test, and can be obtained experimentally. The DP model includes an additional frictional material parameter k_1 that can be calibrated on anisotropic loading HP tests. The parameter k_1 influences the effect of axial stress, i.e. of the stress parallel to the hole axis. The effect increases with increasing value of k_1. It can be

shown that $k_1 \leq 0.5$ such that hole failure is predicted under all loading conditions. In addition, the criterion that uses only the tangential stress is considered in the following. This simplified criterion can be viewed as a simplified MC (sMC) criterion.

Analytical sand onset criteria for field applications can be formulated for the sMC, the MC and the DP hole failure criteria. The formulation provides the critical formation strength, i.e. either sand onset stress σ_S or uniaxial compression strength (*UCS*), for given in situ conditions and given drawdown and depletion. Alternatively, it can provide the critical drawdown for given formation strength and depletion or the critical depletion for given formation strength and drawdown. Using these results, the typical sand onset triangle plots can be constructed. They plot the critical bottom hole pressure for sand onset as a function of reservoir pressure.

Sand production onset models are derived from the HC hole failure models based on the assumption that wellbore failure corresponds to onset of sand production. The expressions for the open hole wellbore can be applied to perforated completions by considering the perforations as open holes of small diameter. In such a case, it is assumed that the wellbore does not influence the stress field around the perforations which is supported by numerical investigations [Pap03a]. In the following compressive stresses strains are taken positive as usual in rock mechanics. Compressive pore fluid pressures are also positive.

The hole failure criteria are expressed through an equivalent cavity stress σ_C that is compared with the sand onset stress or hole failure strength σ_S of the formation, i.e.

$$\sigma_C - \sigma_S \begin{cases} < 0 & \text{No failure, No sand} \\ = 0 & \text{Hole failure, Sand onset} \\ > 0 & \text{Sand production} \end{cases} \tag{5}$$

The equivalent cavity stress σ_C is given for the various criteria as

$$\sigma_C = \frac{\sigma'_{\theta i}}{2} \qquad \text{Simplified Mohr-Coulomb}$$

$$\sigma_C = \frac{1}{4}\left[\sigma'_{\theta i} + \sigma'_{zi} + \sqrt{\left(\sigma'_{\theta i} - \sigma'_{zi}\right)^2 + 4\sigma^2_{\theta zi}} \right] \qquad \text{Mohr-Coulomb} \tag{6}$$

$$\sigma_C = \frac{-\left(\sigma'_{zi} + \sigma'_{\theta i}\right) k_1 + \sqrt{\sigma'^2_{\theta i} + \sigma'_{zi}\left(\sigma'_{zi} - \sigma'_{\theta i}\right) + 3\sigma^2_{\theta zi}}}{-3k_1 + \sqrt{3}} \qquad \text{Drucker-Prager}$$

In the case of a deviated wellbore with fluid pressure p_w, the non-zero effective stresses σ'_{mni} at the hole wall ($r = r_i$) are obtained as

$$
\begin{aligned}
\sigma'_{\theta i} &= \sigma_{Xo} + \sigma_{Yo} - 2\left(\sigma_{Xo} - \sigma_{Yo}\right)\cos 2\theta - 4\sigma_{XYo}\sin 2\theta - 2p_{reso} + \\
&+ \left[2 - \chi_X - \chi_Y + 2\left(\chi_X - \chi_Y\right)\cos 2\theta + 4\chi_{XY}\sin 2\theta\right]\Delta p_{dep} + \\
&+ \left(2 - 2\eta_B - \frac{\eta_B}{\ln r_i / r_e}\right)\Delta p_{dd} \\
\sigma'_{zi} &= \sigma_{Zo} - 2\nu\left(\sigma_{Xo} - \sigma_{Yo}\right)\cos 2\theta - 4\nu\sigma_{XYo}\sin 2\theta - p_{reso} + \\
&+ \left[1 - \chi_Z + 2\nu\left(\chi_X - \chi_Y\right)\cos 2\theta + 4\nu\chi_{XY}\sin 2\theta\right]\Delta p_{dep} + \\
&+ \left(1 - 2\eta_B - \frac{\nu\eta_B}{\ln r_i / r_e}\right)\Delta p_{dd} \\
\sigma_{\theta zi} &= 2\left(\sigma_{YZo}\cos\theta - \sigma_{ZXo}\sin\theta\right) - 2\left(\chi_{YZ}\cos\theta - \chi_{ZX}\sin\theta\right)\Delta p_{dep}
\end{aligned}
\tag{7}
$$

where σ_{IJo} are the original formation stresses in the (x, y, z) coordinate system of the wellbore and p_{res} the reservoir pore pressure. For failure, an effective stress coefficient equal to 1 has been used in Eq.(3) for the calculation of effective stresses. In the above the drawdown Δp_{dd} and the depletion Δp_{dep} have been introduced as

$$
\begin{aligned}
\Delta p_{dep} &= p_{reso} - p_{res} \\
\Delta p_{dd} &= p_{res} - p_w
\end{aligned}
\tag{8}
$$

where p_{reso} is the original reservoir pore pressure. The stress ratios χ_{IJ} describe how the in situ total stresses change with depletion Δp_{dep}. The equivalent cavity stress σ_C for each model is obtained by substituting the effective stresses into expression Eq.(2) for σ_C for the sMC, MC and DP hole failure criteria. The equivalent cavity stress depends on the angle θ, and thus the most critical orientation for hole failure and sand onset will be at an angle θ that maximizes σ_C. Thus, the sand production criterion Eq.(1) becomes

$$
Max\left(\sigma_C\right)_\theta - \sigma_S = 0 \tag{9}
$$

For given in situ conditions and sand onset stress σ_S, criterion Eq.(9) can be applied to calculate whether the stress and pore pressure conditions are such that sand production will occur. Alternatively, for given in situ conditions, Eq.(9) can be solved for the critical σ_S^{cr} for sand onset. A scaling law for σ_S must be employed to account for the strengthening of a hole with decreasing hole diameter. A scaling law that has been proposed based on experiments on sandstones is written as [Pap02a]

$$
\frac{\sigma_S}{\sigma_{Sref}} = \frac{1}{3} + \frac{2}{3}\left(\frac{D_{ref}}{D}\right)^{2/5}, \qquad D_{ref} = 2 \text{ cm} \tag{10}
$$

where σ_S is the sand onset stress of a hole with diameter D and σ_{Sref} is the sand onset stress of a reference hole with diameter D_{ref} = 2 cm. The isotropic loading sand onset stress σ_{SRef} can be related to the UCS, i.e. $\sigma_{Sref} = f\left(UCS\right)$ ([Pap01a], [Cer01]) in which case, a critical UCS^{cr} for sand onset can also be calculated.

For the sMC criterion, it is possible to obtain the maximum equivalent cavity stress and the critical drawdown without resorting to a search for the critical angle. This is possible because the model is simple and involves only the tangential effective stress. Substitution of the first of Eq.(5) into the first of Eq.(2) and analytically obtaining the maximum gives

$$\begin{aligned} Max\left(\sigma_c\right)_\theta = \frac{\sigma_{Xo}+\sigma_{Yo}}{2} + \sqrt{\left[\sigma_{Xo}-\sigma_{Yo}-\left(\chi_X-\chi_Y\right)\Delta p_{dep}\right]^2 + 4\left(\sigma_{XYo}-\chi_{XY}\Delta p_{dep}\right)^2} - \\ - p_{reso} + \left(1-\frac{\chi_X+\chi_Y}{2}\right)\Delta p_{dep} + \left(1-\eta_B-\frac{\eta_B}{2\ln r_i/r_e}\right)\Delta p_{dd} \end{aligned} \tag{11}$$

Substitution of Eq.(11) in criterion Eq. (9) for sand onset can yield the critical UCS^{cr} for sand onset using the relation between UCS and σ_S. Alternatively, the critical drawdown for sand onset for given initial in situ stresses and reservoir pressure, formation strength and depletion can be obtained by solving with respect to the drawdown Δp_{dd} to obtain

$$\begin{aligned} \Delta p_{dd}^{cr} = \frac{1}{1-\eta_B-\frac{\eta_B}{2\ln r_i/r_e}}\Bigg[\eta^p\sigma_S - \frac{\sigma_{Xo}+\sigma_{Yo}}{2} + p_{reso} - \left(1-\frac{\chi_X+\chi_Y}{2}\right)\Delta p_{dep} - \\ - \sqrt{\left[\sigma_{Xo}-\sigma_{Yo}-\left(\chi_X-\chi_Y\right)\Delta p_{dep}\right]^2 + 4\left(\sigma_{XYo}-\chi_{XY}\Delta p_{dep}\right)^2}\Bigg] \end{aligned} \tag{12}$$

The sMC is often used in sand onset analyses due to its simplicity. However, it performs poorly under anisotropic stresses because it does not consider the effect of axial stress (stress parallel to the hole axis) and shear stress on sand onset. Shear stresses develop when the hole is inclined with respect to vertical. The sMC gives straight lines in the sand onset triangle plot for the critical bottom hole pressure vs. reservoir pressure. It is the least conservative of all models giving the higher sand onset stress, except for the DP model which may give, depending on k_1, a higher sand onset stress for some values of axial stress.

The MC model considers the effect of the axial and the shear stresses. The axial stress effect, however, materializes when the axial stress becomes the major principal stress. For lower values, there is no effect of the axial stress. A change in the major principal stress gives a bilinear line in the sand onset triangle plot. The MC model coincides with the sMC model, for low axial stress values and vertical or

horizontal holes. It is the second least conservative after the sMC model, except again for the DP which may give, depending on k_1, a sand onset stress for values of $K_z > 1$.

The DP model considers the effect of the axial and the shear stresses. This gives a curved line in the sand onset triangle plot. The DP model shows the largest axial stress effect on sand onset stress. This effect is amplified with increasing the k_1 value. It is the most conservative for axial stress less than the lateral stress (the stress normal to the hole axis), but it can become the least conservative for some values of axial stress greater than the lateral stress depending also on the k_1. Note that the DP model predicts the same sand onset stress under isotropic stresses independently of the friction parameter k_1, since the model is calibrated at these conditions. The k_1 parameter influences the effect of axial stress anisotropy on hole failure. This means that vertical holes are usually stronger with the DP model as compared to the sMC or MC while horizontal holes are weaker.

3.2. Sand quantification

In the manageable sand regime, the interest lies in predicting the amount of produced sand or sand rate. These are functions of the applied stresses, the fluid flow rate and time in the sense that the sand rate is not constant over time. The quantification of produced sand has been the focus of investigations since the mid-90s when the operators began to realize the benefits of sand production in terms of increased hydrocarbon production. The increase in production results mainly from the possibility of applying higher drawdowns but also from improved near-well flow characteristics, such as improved near-well permeability, removal of damaged or plugged zones, etc. In fact, in some extreme cases, such as in the Canadian heavy oil reservoirs, oil cannot be produced economically without sand production. The benefits of sand production are of course been weighted against its negatives such as possibility of well failure, erosion and maintenance of pipelines and facilities, and sand separation and environmental disposal.

Sand quantification is a transient, coupled mechanical and fluid flow problem in the post-failure regime of material behavior. Experiments were thus needed to elucidate the mechanisms involved and the important parameters at play. Models and analytical and numerical tools were then developed and a methodology for sand quantification was established including material parameter identification and calibration [Pap05a], [Pap06a], [Pap07a], [Pap08a].

An important aspect in sand production is related to multiphase flow due to water breakthrough into a producing well. Water breakthrough leads to increased sand production mainly in already sand producing wellbores. Water below the hydrocarbons may break into a well as a result of two reservoir-drive mechanisms: (i) Water drive where oil is driven through the reservoir by an active aquifer and whereas the reservoir depletes, the water moves in from the aquifer below and displaces the oil,

and (ii) Water injection where water is injected into the reservoir to increase pressure and sweep the oil from the reservoir and thereby stimulate production. Water injection increases the reservoir recovery factor and maintains the production rate for a longer period. Three water-driven mechanisms of increased sand production have been identified, namely:

- *Water-sensitive strength*. In general, a decrease in rock strength (due to e.g. smectite cement, capillary strength, chemical effects etc.) is observed with increasing water saturation with the most significant part of strength reduction occurring within the first 3-4% of water saturation and usually up to 20%.

- *Capillary cohesion in the near-cavity failed rock*. In the failed and disaggregated zone, sand grains are held together by capillary cohesion forces due to connate water present in the reservoirs (e.g. sandcastle made with damp sand). When water saturation increases with water breakthrough, the capillary cohesion is eliminated, and sand is produced (e.g. sandcastle collapses as tide comes in).

- *Multiphase flow effects with a high pore pressure gradient front* moving together with the water front. The pressure gradient front may destabilize the sand in the failed and disaggregated zone close to the cavity.

4 Hydraulic Fracturing

The hydraulic fracturing (HF) technique involves the pumping of a viscous-fluid from a well into the rock formation under high fluid pressure to fracture the reservoir. The pumping of fluid is maintained at a rate high enough for the fluid pressure to overcome the flow friction losses, the minimum in-situ stress or closure stress, the resistance to splitting the rock and hence to propagate the fracture (Figure 4.1). The initiated fracture as propagating in a complex stress field near the wellbore will reorient itself to propagate further in the direction of lease resistance which is perpendicular to the minimum insitu compressive stress. During the pumping process, material like sand, called proppant, is gradually mixed with the fracturing fluid to ensure that the fracture will remain propped open after the pumping stops. A permeable channel of high conductivity will hence be formed for oil or gas to flow from the reservoir in the well. For the fundamentals and details on HF technique, design and execution the reader is refer to [Eco01]. This contribution focusses on the more advanced modelling of HF in weak formations.

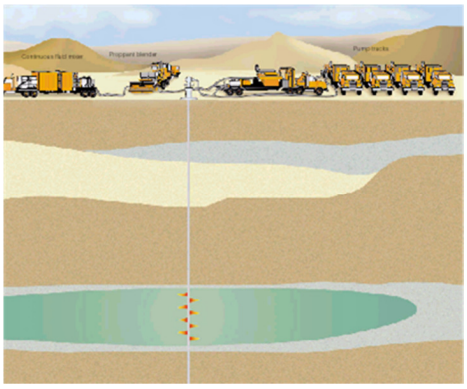


Figure 4.1. Illustration of the hydraulic fracturing technique

Hydraulic fracturing modeling involves the coupling of various complex physical processes including (i) the viscous flow of the fluid in the fracture and the leak off in the formation (ii) the rock deformation of the surrounding medium induced by the fluid pressure on the fracture surfaces in the presence of high confining stresses and (iii) the rock splitting and fracture propagation (Figure 4.2). In most models, including analytical solutions and commercial design codes, the solid deformation is modeled with the elasticity theory, represented by an integral equation that determines the non-local relationship between the fracture width and the fluid pressure. The fluid flow is modeled by lubrication theory, expressed by a non-linear partial differential equation that relates the local fluid flow velocity, the fracture width and the gradient of the pressure. The fracture propagation is assumed to take place when the stress intensity factor at the tip reaches a critical value equal with the rock fracture toughness which in many cases is ignored assuming to be close to zero.

In field operations, attention is focused on the prediction of the wellbore pressure which is normally measured during the treatment and is the only parameter available to evaluate or to redesign in real time the operation. Classical hydraulic fracturing simulators often underestimate the measured down-hole pressures. Research work involving surveying on net-pressures (difference between the fracturing fluid pressure and the far-field confining stress) indicated that the net pressures encountered in the field are on average 50-70% higher than the predicted by conventional models [Dam01]. These observations have triggered a series of new ideas and dedicated studies which looked into the importance of the rock plastic deformation in hydraulic fracturing [Pap01, Pap02, Pap03, Pap04, Pap05, Pap06, Pap07, Dam01]. Sarris and Papanastasiou [Sar01, Sar02, Sar03] extended these studies to account for the pressure diffusion and porous behavior of the rock deformation. A common characteristic of these studies is the use of the cohesive zone law to propagate fractures in order to investigate the inelastic behavior of rocks in hydraulic fracturing. The influence of the parameters of the process zone on hydraulic fracturing results was studied in [Sar01, Car01]. Some early studies utilizing the cohesive zone model in hydraulic fracturing include the important work in [Boo01, Boo02].

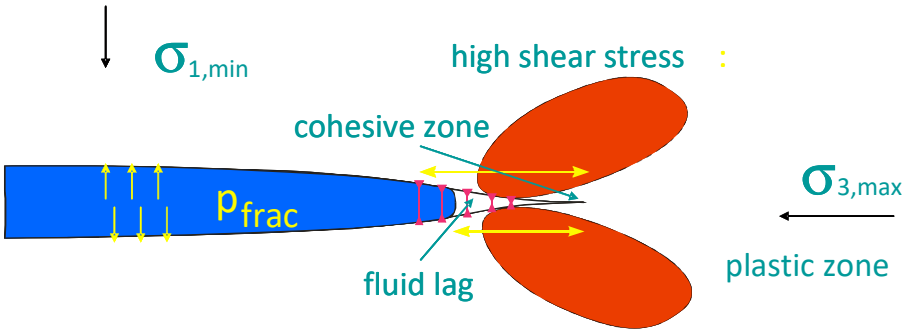


Figure 4.2. Geometry of a plane strain hydraulic fracture with the near tip processes

The models presented were developed for plain strain geometry taking into consideration the symmetry conditions. This geometry is appropriate for modelling short fractures with fracture height relatively greater than the fracture length. Furthermore, this geometry is also appropriate for examining tip effects since the deformation of any arbitrary fracture shape is approximately planar near the tip. The fracture propagates perpendicular to the minimum in situ stress and remains planar. This predefined path for the propagation is also convenient with the cohesive zone numerical approach. For the stress-deformation we assume that rock obeys the equation of plasticity. For the sake of completeness we describe next the involved physical processes: the fluid flow, the rock deformation, the fracture propagation and the methodology that was adopted in the numerical model.

4.1 Fluid-flow

The physical process of the fluid driven fracture involves the pumping of a viscous fluid that pressurizes the fracture surfaces which deform. Increasing the pressurization, critical loading conditions will be reached ahead of the tip splitting the rock and driving hydraulically the fracture. Thus, this process reveals that there is a strong coupling between the moving fluid, rock deformation and fracture propagation.

The fluids that are used in hydraulic fracturing are normally power-law with shear-thinning behavior which means that the viscosity decreases with increasing shear rate. In order to avoid this complex fluid behavior, a simple appropriate model for fluid flow in a fracture is assumed to follow the lubrication approximation. It assumes laminar flow of an incompressible uniformly viscous Newtonian fluid and accounts for the time dependent rate of crack opening. The continuity equation which imposes the conservation of mass in one dimensional flow is

$$\frac{dq}{dx} - \frac{dw}{dt} + q_l = 0 \tag{4.1}$$

where q is the local flow rate along the fracture in direction x, q_l is the local fluid loss in rock formation and w is the local crack opening. Eq. (4.1), which accounts for the fluid leak-off from the fracture surface into the rock formation, can be used to determine the local flow rate q.

The second equation is derived from the conservation of momentum balance. For a flow between parallel plates the lubrication equation, which relates the pressure gradient to the fracture width for a Newtonian fluid of viscosity μ, yields

$$q = u.w = -\frac{w^3}{12\mu}\frac{dp}{dx} \tag{4.2}$$

where p denotes the fluid pressure and u the average velocity of the fluid over a cross-section of the fracture. Eq. (4.2) determines the pressure profile along the fracture from the local width and the local flow rate. According to Eq. (4.2), the pressure gradient and hence the solution, is very sensitive to fracture width. Therefore, the largest part of the pressure drop takes place within a small area near the tip where the width decreases significantly before it vanishes at the tip.

4.2 Rock deformation

In weak rocks, large inelastic deformation is expected to take place in the area near the crack tip due to high shear stress concentration (Figure 4.2). In such a case one should use plasticity theory which properly describes the irreversible deformation. When the fracture propagates the plastic zones unload elastically behind the advancing crack and the new area near the current tip deforms plastically. In summary, the rock mass remote from the fracture may deform elastically, whereas the area near the body of the fracture is initially elastic but then deforms plastically and finally unloads elastically after the fracture tip has advanced. Under such conditions, the plasticity model must be capable of dealing with non-proportional loading. Such capabilities are provided, of course, by an incremental flow theory of plasticity and finite element analysis.

Among the different yield criteria, the Mohr-Coulomb model adequately describes the pressure-sensitive behaviour of rocks which exhibit dilatancy when sheared. Unlike most cases of classical fracture mechanics, the remote stress field in the hydraulic fracturing problem is compressive, due to the presence of the in situ stresses. In such a case the use of Mohr-Coulomb yield criterion, which is usually employed in cases of compressive stresses, is fully justified. The tensile failure along the propagation line is modeled as described in the next section by a cohesive-type model. The Mohr-Coulomb yield criterion can be expressed in terms of maximum and minimum principal stresses, σ_1 and σ_3, respectively, (compression is negative)

$$\sigma_e = \sigma_1 \frac{1+\sin\phi}{1-\sin\phi} - \sigma_3 \tag{4.3}$$

where φ is the friction angle and σ_e is the effective stress which is related to the cohesion c via

$$\sigma_e = 2c \frac{\cos\phi}{1-\sin\phi} \tag{4.4}$$

Post yield strengthening with deformation can be modeled using a cohesion hardening model. According to this model the equivalent stress, σ_e, increases with the equivalent plastic strain, ε^p,

$$\sigma_e = \sigma_e^0 + he^p \tag{4.5}$$

where σ_e^0 is the value of effective stress at initial yield. The linear hardening plasticity modulus h has been derived in [Pap02] from the loading and unloading moduli as measured in a uniaxial compression test.

$$h = \frac{E_{loading}}{1 - \left(E_{loading} \Big/ E_{unloading} \right)} \tag{4.6}$$

In the flow theory of plasticity the strain increment $d\varepsilon_{ij}$ is decomposed into an elastic $d\varepsilon_{ij}^e$ and a plastic part $d\varepsilon_{ij}^p$

$$d\varepsilon_{ij} = d\varepsilon_{ij}^e + d\varepsilon_{ij}^p \tag{4.7}$$

The elastic strain increment $d\varepsilon_{ij}^e$ can be obtained from Hooke's law. The plastic strain increments, $d\varepsilon_{ij}^p$, are generated when the stress state reaches the yield surface and can be expressed by a non-associated flow rule in the form

$$d\varepsilon_{ij}^p = dk \frac{\partial Q}{\partial \sigma_{ij}} \tag{4.8}$$

where Q is the plastic potential and the scalar function $d\kappa$ is the plastic multiplier. The plastic potential Q can have a similar form to yield surface σ_e, if in equation (4.3) the dilation angle ψ replaces the friction angle, φ.

As mentioned before, the yield surface and plastic potential are generally functions of stresses and the hardening parameter ε_p. The hardening parameter ε_p is calculated from the principle of plastic power equivalence,

$$\sigma_e d\varepsilon^p = \sigma_{ij} d_{ij}^p \tag{4.9}$$

In general, weak rocks obey a non-linear yield criterion and exhibit non-associative behaviour. Experimental results from triaxial compression tests show (a) that the dilation angle increases slightly with increasing plastic strain when low confining pressures are used but remains approximately constant in samples under high confining pressure and (b) the value of dilation angle is a strong function of the confining pressure. In tests with low confining pressure the measured dilation angle is greater than the friction angle but decreases rapidly with increasing confining pressure, eventually becoming negative. This indicates compaction at very high confining pressure. In hydraulic fracturing, compaction is excluded because the initial in-situ mean pressure near the crack tip decrease during propagation. Furthermore, earlier computations in [Pap01] showed that the non-associative solution for zero dilation was bounded by the associative solution and the elastic solution.

It is out of the scope of this paper to present details on the discretization of the fluid-flow in the fracture and its coupling with the rock deformation. For more details on the discretization, coupling, iterative and continuation algorithms for the problem of hydraulic fracturing the reader is referred to [Pap03].

4.3 The cohesive zone model as fracture propagation criterion

The cohesive zone model implies that normal stress continues to be transferred across a discontinuity which may or may not be visible as shown in Figure 4.3. This stress is determined from the softening stress-strain relation that various rocks exhibit in calibrations tests. This transferred normal stress is a function of the separation and falls to zero at a critical opening and then the fracture propagates. The evolution of the crack is governed by energy balance between the work of the external loads and the sum of the bulk energy of the uncracked part and the energy dissipated in the fracture process. The main mathematical difficulty is given by the fact that the fracture energy depends on the opening of the distributed micro-cracks. To simplify the mathematical difficulties, it is assumed that the cohesive zone localizes, due to its softening behavior, into a narrow band ahead of the visible crack [Hil01].

The constitutive behavior of the cohesive zone is defined by the traction-separation relation derived from laboratory tests. The traction-separation constitutive relation for the surface is such that, with increasing separation, the traction across this cohesive surface reaches a peak value and then decreases and eventually vanishes, permitting for a complete separation. Simple cohesive zone models can be described by two independent parameters which are usually, for mode-I plane strain, the normal work of separation or the fracture energy G_{IC} and either the tensile strength σ_t or the complete separation length δ_{IC} [Pap03,Pap04]. In those models, even though the implemented cohesive constitutive equation followed a simple rigid-linear softening response, the calculated normal to the propagation direction stress distribution ahead of the open crack exhibited a smooth non-linear response.

Nevertheless, an additional parameter in these models is the slope of the initial loading which may define a range from rigid-softening to elastic-softening response

under tensile stress-state. In order to investigate the main characteristics of the full curve in hydraulic fracturing, Sarris and Papanastasiou [Sar01] carried out computations for different initial slopes to simulate a rigid-softening to elastic softening behavior. The transition from the elastic softening to the rigid-softening was carried out by increasing the initial slope of the constitutive cohesive law by five times in each model. The case of the most rigid behavior corresponds to twenty times the stiffness of the most elastic case. In all cases the area under the curve which is related to the work of separation is maintained the same (Figure 4.3b)

The area under the traction-separation curve equals with fracture energy G_{IC} which is the work needed to create a unit area of fully developed crack. For elastic solids this energy is related to the rock fracture toughness K_{IC} through [Ric01, Kan01]

$$K_{IC}^2 = \frac{G_{IC} E}{1-\nu^2} \tag{4.10}$$

where E is the young modulus and ν is the Poisson ratio. The rock fracture toughness can be calculated from laboratory tests. For the case of the rigid-softening behavior the traction-separation relation is uniquely determined by

$$\sigma = \sigma_t (1 - \delta / \delta_{IC}) \tag{4.11}$$

where σ_t is the uniaxial tensile strength of the rock and δ_{IC} is the critical opening displacement at which σ falls to zero. The value of δ_{IC} is given in [Kan01]

$$\delta_{IC} = \frac{2K_{IC}^2 (1-\nu^2)}{\mathrm{E}\sigma_t} \tag{4.12}$$

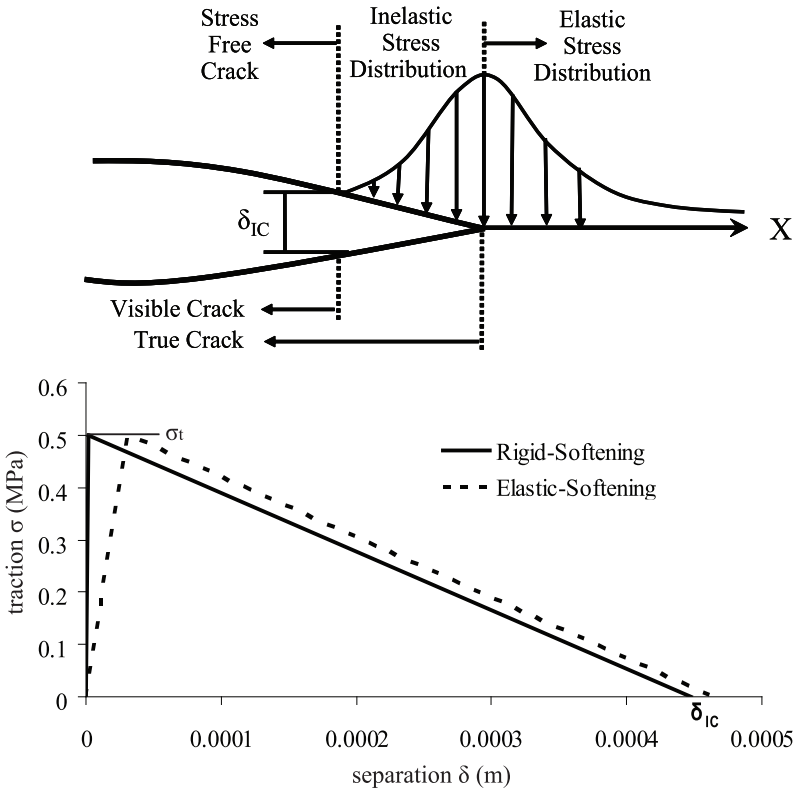


Figure 4.3 Representation of the fracture process zone and the constitutive cohesive zone law.

For the case of the elastic loading the cohesive constitutive relations were augmented and modified to take into account the initial part of the curve as follows

$$\sigma = \sigma_t \left(\frac{\delta}{\delta_{el}} \right) \tag{4.13}$$

with the limit of elastic deformation given by

$$\delta_{el} = \frac{\sigma_t}{k_n} \tag{4.14}$$

where k_n is the stiffness of the traction-separation relation in the loading regime with units of [MPa/m]. In the post-peak softening regime the cohesive constitutive relation is given by

$$\sigma = \sigma_t \left[1 - \frac{\left(\delta - \delta_{el}\right)}{\left(\delta_{IC} - \delta_{el}\right)} \right] \tag{4.15}$$

In order to investigate further the influence of the cohesive zone law in hydraulic fracturing, Sarris and Papanastasiou [Sar01] have also studied different forms of softening behavior.

For the numerical implementation of the cohesive zone model interface elements are employed along the propagation direction. The interface elements are two-dimensional, isoparametric, 6-node or 4-node interface element. A consistent isoparametric formulation permits modelling of curved crack surfaces and provides an element that is compatible with the 8-node quadratic displacement or 4-node linear finite elements that are used to discretize the internal domain. The numerical algorithm requires initial conditions to be specified at t =0 for the initial fracture length $\ell(0)$ the width profile $w(x)=0$ and fluid pressure $p(x)=0$. In setting these initial conditions we recommend the analytic solution derived in [Des01] for an elastic material with zero fracture toughness. A meshing/remeshing scheme is employed in order to carry out longer propagations with fine mesh near the fracture tip. The meshing/remeshing scheme was based on the following steps: The sensitive area near the tip where high gradients exist was discretized using a fine mesh and the region away from the crack-tip was discretized with a coarser mesh.

4.4 Computational results

In this section we present sample results for hydraulically driven fractures in an elastic and an elastoplastic media to show the efficiency of the proposed algorithm. The parameters upon which the numerical computations were based are given in Table 1. With the chosen material parameters and in-situ stresses the rock is initially elastic but very close to a yielding state.

Figure 4.4 shows the profiles of propagating elastic (solid-lines) and elasto-plastic (dashed-lines) fractures for the same fracture lengths. The cusping of the crack tips with zero slope is a result of the cohesive model which was incorporated as the propagation criterion. As mentioned earlier, the model of the elastoplastic fracture requires an initial fracture length which was set to 0.5 m; the influence of this can be seen in the width profiles in Figure 4.4 but as mentioned earlier the results are unaffected in the area where the fracture was propagated. If we compare the fracture openings in the region where the fractures were propagated (i.e., distance from wellbore between 0.5 and 2.2 m) we see that the width profile of the elastoplastic fracture is much wider than the width profile of the elastic fracture.

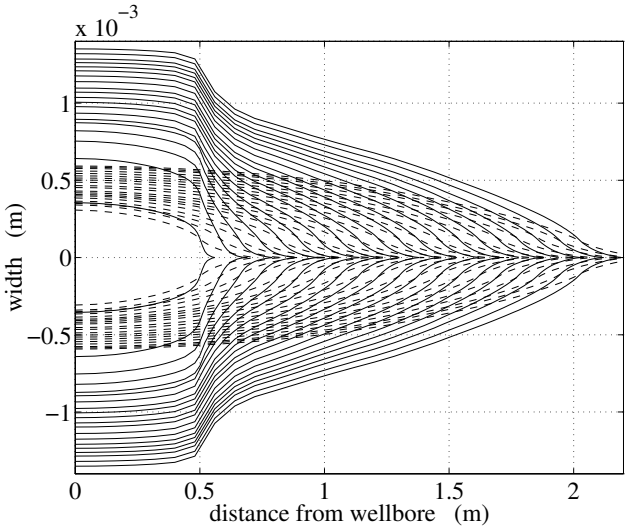


Figure 4.4. Fracture profiles for elasto-plastic (solid lines) and elastic (dashed-lines)

Figure 4.5 shows the comparison of the net-pressure profiles in the fracture for the same fracture length. The very narrow opening of the elastic fracture results in greater pressure drop near the fracture tip and significant fluid-lag. Higher energy is required for propagating the elastoplastic fracture. Most of this energy is coming from the smaller fluid lag region. Plastic yielding near the tip of a propagating fracture provides an effective shielding, resulting in an increase in the effective rock fracture toughness [Pap02, Pap04, Pap07]. This is shown in Figure 4.6 where we plotted the increase in the effective fracture toughness as a function of the fracture growth. The effective fracture toughness was determined using the calculated, path independent, J-integral [Ric01]. The value of the effective fracture toughness is directly related to the size of the plastic zones.

The elastoplastic fracture was propagated further to reach a length of 8 m before examining the closure pattern (Figure 4.7). Figure 4.8 shows the width profile of a receding elastoplastic fracture. It was assumed that there was no fluid-flow in the fracture during closure, therefore the pressure acting along the fracture was constant. The elastoplastic fracture makes contact initially near the tip and then the closure moves towards the wellbore [Pap05, Pap06]. This closure pattern agrees with the experimental results reported by [Dam 01]. Figure 4.9 shows the corresponding net-pressure at the wellbore vs fracture length during propagation and closure after the fracture has reached a) 5 m and b) 8 m length. These results show that the net-pressure drops to zero while the fracture is still wide-open along a large proportion of the original length. The decrease of the fracture surface during closure will result in the underestimation of the insitu leak-off coefficient at the late stage of the pressure decline analysis. More accurate leak- off predictions can be obtained in the

early stage of pressure decline analysis, immediately after shut-off of the pumping stage. The results of Figure 4.9 suggest that the fracture will close completely at negative net-pressures (fluid pressure less than the far-field stress). Application of classical analysis, which assumes that the fracture closes completely when the fluid-pressure drops to the value of the far-field stress, would lead to the underestimation of the minimum insitu stress.

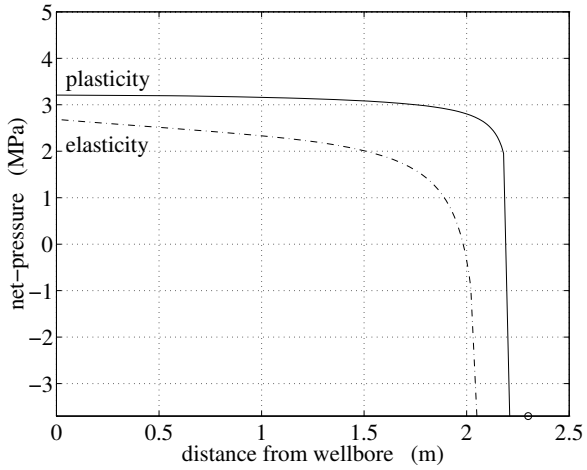


Figure 4.5. Net-pressure profiles on fractures.

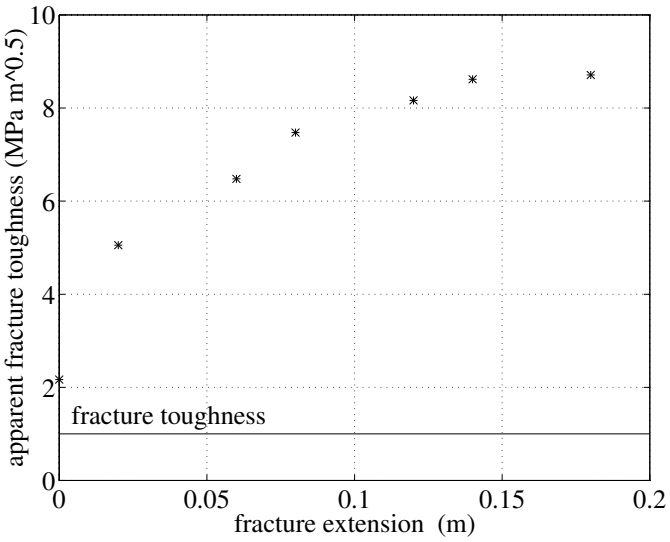


Figure 4.6. Apparent fracture toughness vs fracture extension.

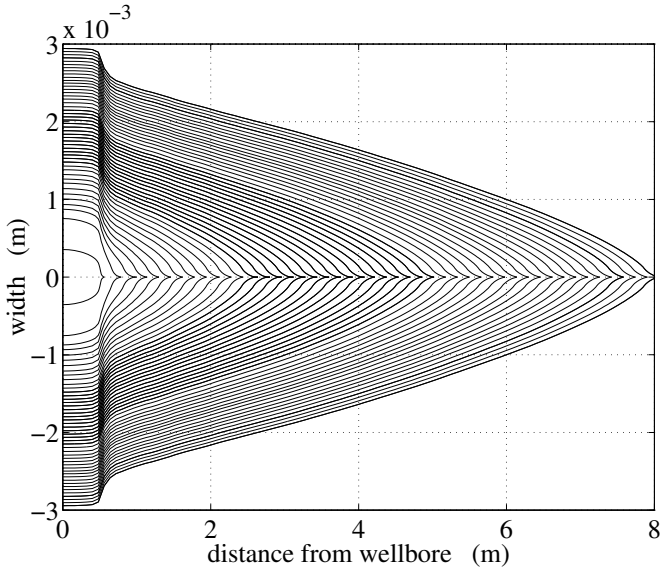


Figure 4.7. Propagation of elasto-plastic fracture

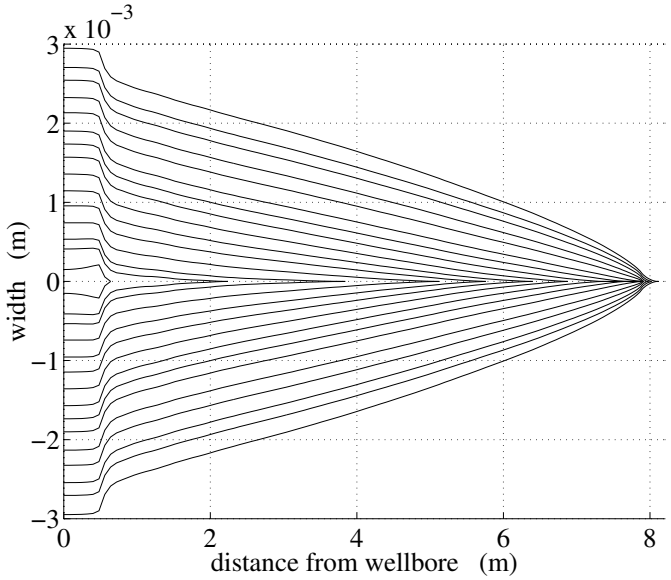


Figure 4.8. Closure of elasto-plastic fracture

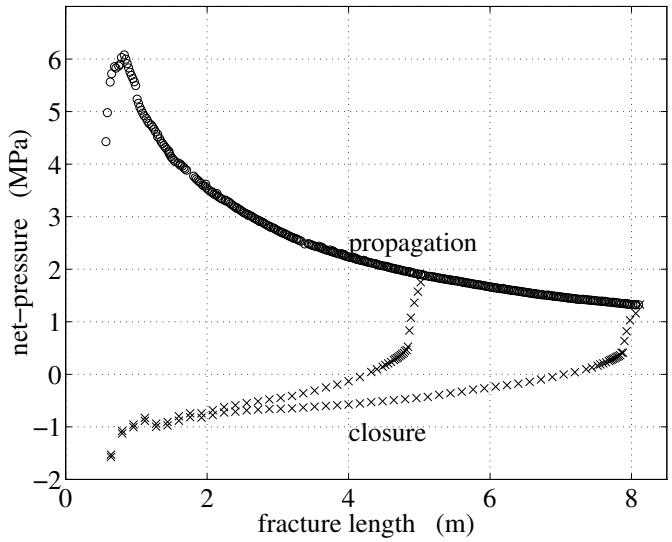


Figure 4.9. Net-pressure at wellbore vs fracture length

4.5 Summary of results and conclusions

In summary, we found that plastic yielding provides a shielding mechanism near the tip resulting in an increase of the effective fracture toughness [Pap04, Pap07]. Higher pressure is needed to propagate an elastoplastic fracture and the created fracture is shorter and wider than an elastic fracture [Pap01, Pap02]. We demonstrated that the standard hydraulic fracturing simulators, which are all based on elasticity, will yield better results if the unloading modulus is used as the Young's modulus. We have also modelled hydraulic fracture closure [Pap05, Pap06]. It was found that the assumption made that the fracture closes completely once the fluid-pressure in the fracture drops to the value of the far-field stress, is not valid. In addition, we showed that the formation will be more stable after it is fractured due to the redistribution of the stresses [Pap05]. In [Pap07] it was shown that an induced HF due to CO2 injection is likely to propagate horizontally than vertically to the higher effective toughness in the vertical direction. The reported results were found to be in good agreement with the findings of hydraulic fracturing experiments carried out on large blocks in Delft University [Dam01].

5 Conclusion

Meeting unpredicted problems in petroleum geomechanics such as wellbore instability, sanding during production and failures in hydraulic fracturing design and execution, is almost inevitable, attributed mainly to the uncertainties on the values of dominant parameters mainly the in situ stresses, rock strength and heterogeneities. This realization motivated the development of real time techniques for facing these problems. The main idea behind the new approach is to use real time measurements and monitoring in drilling, to evaluate the drilling job and to update the design in order to resolve the encountered problems. In the heart of the design and updated phases an accurate modelling for estimating the optimum mud-pressure is always needed. Similar approach is followed in hydraulic fracturing where downhole measurements of the pressure and acoustic emissions, tracers and tilt-meters are deployed to estimate the fracture geometry and update the hydraulic fracturing simulators to the real time data. Sanding in unconsolidated formations, the transport of sanding from the reservoir to surface facilities and the estimation of the volumes of the produced sand are still outstanding issues in petroleum engineering.

Finally, a new area in petroleum engineering, called reservoir geomechanics has developed in the recent years. The reservoir simulators are coupled with stress analysis in order to give more accurate predictions and to account for problems such as reservoir compaction and surface subsidence, casing collapse, fault activation and other. In all these problems the knowledge of the in situ stresses is required. The finite element analysis can be combined with local measurements for obtaining the in situ stresses near complex geological structures.

6. References

[Boo01] Boone T.J, Wawrzynek P.A, Ingraffea A.R. Simulation of fracture process in rock with application to hydrofracturing, *International Journal of Rock Mechanics and Mining Science and Geomechanics Abstracts, 23, 255-265*, 1986.

[Boo02] Boone J.T, Ingraffea A.R. A numerical procedure for simulation of hydraulically driven fracture propagation in poroelastic media, *International Journal for Numerical and Analytical Methods in Geomechanics, 14, 27-47,* 1990.

[Car01] Carrier, B., and Sylvie, G. Numerical modeling of hydraulic fracture problem in per-meable medium using cohesive zone model. Engineering Fracture Mechanics, 79, 312-328, 2012.

[Cer01] Cerasi PR, Berntsen AN, Stenebråten JF, Papamichos E (2010). Use of volumetric sand production tool on field data to help plan oil production strategy. 44th US Rock Mechanics Symp and 5th U.S.-Canada Rock Mechanics Symp, Salt Lake City, UT.

[Des01] Desroches J., Detournay E., Lenoach B., Papanastasiou P., Pearson JRA, Thiercelin M., Cheng A. The crack tip region in hydraulic fracturing, *Proc. Royal Society. London A, vol. 447, 39-48,* 1994.

[Eco01] Economides M., Nolte K. *Reservoir Stimulation*, Chichester, England: Wiley, 2000.

[Fja01] Fjaer E., Holt R.M., Raaen A.M. *Petroleum Related Rock Mechanics,* 2nd Edition, Elsevier Science, 2008.

[Hil01] Hillerborg A., Modeer M., Petersson P.E.Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements, *Cement and Concrete Research*, vol. 6, 773-782,1976;

[Kan01] Kanninen M.F, Popelar C.H. *Advanced fracture mechanics*. Oxford University Press, 1985.

[Kes01] Kessler N, Wang Y, Santarelli F (1993). A simplified pseudo 3D model to evaluate sand production risk in deviated cased holes, SPE 26541, 68th Annual Technical Conf. and Exhibition of the SPE, Houston.

[Mor01] Morita N (2004) Well orientation effect on borehole stability. SPE 89896, 2004 SPE Annual Technical Conference and Exhibition, http://dx.doi.org/10.2118/89896-MS.

[Pal01] Palmer I, Higgs N, Moschovidis Z, Cameron J (2005). How 3D Stresses can change the standard Thick-Walled Cylinder approach to sand prediction. ARMA/USRMS 05-792, 40th US Symposium on Rock Mechanics: Rock Mechanics for Energy, Mineral and Infrastructure Development in the Northern Regions, Anchorage, AL, USA.

[Pap01a] Papamichos E (2002). Sand mass prediction in a North Sea reservoir. SPE/ISRM 78166, OilRock 2002, Irving, Texas.

[Pap02a] Papamichos E, Furui K (2013). Sand production initiation criteria and their calibration. 47th US Rock Mechanics/ Geomechanics Symp, San Francisco, CA.

[Pap03a] Papamichos E, Tronvoll J, Skjærstein A, Unander TE (2010). Hole stability of Red Wildmoor sandstone under anisotropic stresses and sand production criterion. Journal of Petroleum Sciences and Engineering 72, 78-92, http://dx.doi.org/10.1016/j.petrol.2010.03.006.

[Pap04a] Papamichos E, Furui K (2019). Analytical models for sand onset under field conditions. Journal of Petroleum Sciences and Engineering 172, 171-189, doi:10.1016/j.petrol.2018.09.009.

[Pap05a] Papamichos E (2010). Erosion and multiphase flow in porous media. Application to sand production. Special issue of European Journal of Environmental and Civil Engineering 14, 8-9, 1129-1154, doi:10.3166/ejece.14.1129-1154.

[Pap06a] Papamichos E (2010). Analysis of borehole failure modes and pore pressure effects. Computers and Geotechnics 37, 141-152, doi:10.1016/j.compgeo.2009.08.006.

[Pap07a] Papamichos E, Vardoulakis I (2005). Sand erosion with a porosity diffusion law. Computers and Geotechnics 32, 47-58, doi:10.1016/j.compgeo.2004.11.005.

[Pap08a] Papamichos E, Vardoulakis I, Tronvoll J, Skjærstein A (2001). Volumetric sand production model and experiment. Intl Journal for Numerical and Analytical Methods in Geomechanics 25, 8, 789-808, doi:10.1002/nag.154.

[Pap01] Papanastasiou P, Thiercelin M. Influence of inelastic rock behavior in hydraulic fracturing, *International Journal of Rock Mechanics and Mining Science and Geomechanics Abstracts, 30, 1241-1247*, 1993.

[Pap02] Papanastasiou P. The influence of plasticity in hydraulic fracturing, *International Journal of Fracture, 84, 61-79*, 1997.

[Pap03] Papanastasiou P. An efficient algorithm for propagating fluid driven fractures, *Computational Mechanics, 24, 258-267*, 1999.

[Pap04] Papanastasiou P. The effective fracture toughness approach in hydraulic fracturing, *International Journal of Fracture, 96, 127-147,*1999.

[Pap05] Papanastasiou P. Formation stability after hydraulic fracturing', *International Journal for Numerical and Analytical Methods in Geomechanics*, 23, 1927-1944, 1999.

[Pap06] Papanastasiou P., Hydraulic Fracture closure in pressure-sensitive geomaterials, *Int. J. Fracture, 103, 149-161*, 2000.

[Pap07] Papanastasiou P, Papamichos E. and Atkinson C. On the risk of hydraulic fracturing in CO2 geological storage, *International Journal for Numerical and Analytical Methods in Geomechanics, 40, 1472-1484,* 2016.

[Pas01] Pasic B, Gaurina-Međimurec N, Matanović D (2007) Wellbore instability: causes and consequences. Rudarsko-geološko-naftni zbornik 19, 87-98.

[Rah01] Rahmati H, Jafarpour M, Azadbakht S, Nouri A, Vaziri H, Chan D, Xiao Y (2013). Review of sand production prediction models. J of Petroleum Engineering, Article 864981, http://dx.doi.org/10.1155/2013/864981.

[Ric01] Rice JR.A path-independent integral and the approximate analysis of strain concentration by notches and cracks, *Journal of Applied Mechanics, Trans. ASME, vol.35, 379-386*, 1968

[Sar01] Sarris E., Papanastasiou P. The influence of the cohesive process zone in hydraulic fracturing, *International Journal of Fracture, 167,33-45*, 2011.

[Sar02] Sarris E., Papanastasiou P. Modelling of hydraulic fracturing in a poroelastic cohesive formation*, ASCE, International Journal of Geomechancs, 12, 160-167*, 2012.

[Sar03] Sarris E., Papanastasiou P. The influence of pumping parameters in hydraulic fracturing. *International Journal for Numerical and Analytical Methods in Geomechanics,* 39, 635-654, 2015.

[Dam01] van Dam D.B., *The influence of inelastic rock behaviour on hydraulic fracture geometry*, Ph.D thesis, Delft University, 1999.

[Hoe01] van den Hoek PJ, Hertogh GMM, Kooijman AP, Bree Ph de, Kenter CJ, Papamichos E (2000). A new concept of sand production prediction: Theory and laboratory experiments. SPE 65756, SPE Drilling & Completion 15, 4, 261-273.

[Zob01] Zoback M.D. *Reservoir Geomechanics, Cambridge University Press, 201*

Grainsize dynamics: mixing, segregation, crushing and their heterarchy

Itai Einav, Benjy Marks and Pierre Rognon

School of Civil Engineering, The University of Sydney 2006, NSW, Australia

This Chapter introduces the field of 'grainsize dynamics' – the mechanics dealing with the evolution of particle size distributions in space and time, and their governing forces. Typical forces in grainsize dynamics include mixing, segregation, crushing, attrition, agglomeration and thermal expansion. Here, we focus on the first three forces, starting from stochastic particle scale physics and how/whether these physics could be upscaled to frame enriched continuum models. A discussion will follow on why 'open-system' dynamics by mixing and segregation do not lend themselves for hierarchical approach that artificially identifies scales and treat them separately. Although the physics of grain crushing can be understood using a 'closed-system' idealisation, when coupled with mixing and segregation, the modelling of grain crushing also requires an open-system description within a heterarchical approach that does not separate scales (yet benefits from non-hierarchical organisational rules). Although this Chapter focuses on particle size, much of the presented philosophy may be adapted for other shape descriptors such as elongation and sphericity.

1 Introduction

Our journey into grainsize dynamics began with Prof Ioannis Vardoulakis [Var], whom the first author visited in Athens on June 2007. During that visit, and triggered by the breakage mechanics theory [Ein07] of grain crushing, Ioannis (who masterfully used to identify good problems) proposed a cooperation into what he thought would yield the first mathematical solution to the oldest industrial problem in human history – the problem of grinding flour using stonemills (see Figure 1), which involves simultaneous grain crushing, segregation and mixing. Similarly, motivated by Ioannis [Var], such a solution should also help to address other large flow problems involving brittle granular media, let them be pyroclastic flows or dry snow avalanches.

Figure 1: The Vardoulakis's challenge [Var]. Resolve mathematically the oldest industrial problem in human history – the simultaneous crushing, mixing and segregation of flour in stonemills.

To date, the grainsize dynamics in neither the stonemills nor the pyroclastic flow problems have been resolved! But we got much closer. This Chapter describes how.

In order to solve those problems framed by Prof Vardoulakis, his initial proposition was to combine the breakage mechanics theory with Gray & Thornton's [GT05] (GT's) theory for bi-mixture flows down chutes involving both segregation and mixing. During that visit, however, it became clear that these two theories adopt assumptions that are not amenable for a theoretical unification.

In particular, while GT's original mixing-segregation theory can deal with open-system flows of two sizes of grains [GT05], in brittle granular systems grain crushing always introduce some degree of polydispersity. Discussion on this point along with a proposed solution for modelling flows in polydispersed media will be offered in the sections focusing on mixing and segregation. Similarly, while the breakage mechanics theory can capture the evolving polydispersity in brittle granular media, its intrinsic assumption of an ultimate grainsize distribution is only appropriate for closed-systems, and not for open-system problems such as stonemills and pyroclastic flows where particles can swap places through preferential advection. We shall come back to this point in the section dealing with grain crushing, along with proposed heterarchical solution that is free from the ultimate grainsize distribution assumption. We will emphasise: (1) why hierarchical models – that are built on scale separation – cannot be used for open system problems of stonemills and pyroclastic flows; and (2) the advantages of heterarchical models for coupling crushing, segregation and mixing. In conclusions, we will emphasise the remaining obstacles for establishing a smooth, reliable mathematical solution of grainsize dynamics in stonemills and pyroclastic flows based on a continuum heterarchical model that could finally address the original challenge by Vardoulakis [Var].

2 Grainsize

By now you would have noticed the term 'grainsize' throughout article. Where the traditional 'particle size' is used to denote the representative size of a single particle, the term 'grainsize' is used to express a continuous mathematical coordinate. With this in mind, we can define the probability density function $\phi(s)$ as the grainsize distribution, where s is the coordinate representing particle size. In general media, the grainsize distribution varies over time and across the space. Similarly, one can define the mean velocity $\mathbf{u}(s)$ of all the grains of a given grainsize s at a certain point in space $\mathbf{x}$ and time t. Most generally, we can think of a five-dimensional (5D) continuua with:

$$\phi \equiv \phi(\mathbf{x}, t, s), \quad \mathbf{u} \equiv \mathbf{u}(\mathbf{x}, t, s), \tag{1}$$

where a representative volume element is defined by the three spatial coordinates in $\mathbf{x}$, the time t and grainsize s. In addition, the normalisation condition for probability density functions requires that

$$\int_0^\infty \phi(s)\, ds = 1. \tag{2}$$

To recover a certain value of the volume fraction of a range of particle sizes Φ, say between s_a and s_b, we can take the integral of the grainsize distribution to give

$$\Phi(s_a < s \leq s_b) = \int_{s_a}^{s_b} \phi(s)\, ds. \tag{3}$$

3 Mixing

Mixing is the equitable process by which matter spreads evenly in space. In a perfectly mixed granular system, the nature of the grains including their particle size distribution and their mechanical properties are the same anywhere we look. Mixing granular systems can be achieved by two processes, namely *shear-induced diffusion* and *chaotic advection*. None of these come for free. They both require some flow or agitation to develop, meaning the input of some mechanical work.

This is a fundamental difference between mixing processes in granular materials and molecular fluids. To highlight this difference, consider a first Thought Experiment from the kitchen. Poor some flower into a bowl, add some raw sugar too, then wait. As we know from experience, the flower and the sugar particles will not mix, at least until we would start stirring them, for example with a spatula, or in other words, at least until we would add a certain amount of mechanical energy into the bowl.

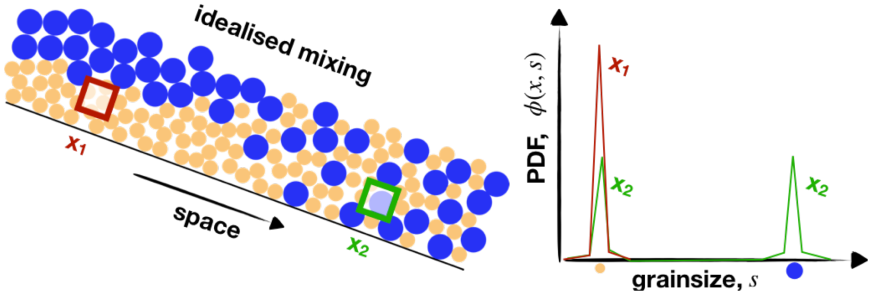


Figure 2: Schematic representation of the mixing process. Initially, particles are ordered by size, but in this hypothetical scenario they mix as they flow down a slope. Note that this is rather unphysical behaviour in this particular instance.

Next consider a second Thought Experiment from the kitchen. Open a container of vanilla bean, then go and sit at the end of your lounge, a few metres away. After some time, the fragrance will reach you without needing to mechanically stir the air. This means that the vanillin molecules our noses detect have been somewhat spontaneously mixed through the air molecules in order to get to our location.

So, if not like molecular fluids, how do granular materials mix? This section introduces some elements of answers to this question, which still very much remain an active field of research. The focus will be on both the known particle-scale mechanisms at the origin of granular mixing, and on the available continuum models that may capture mixing.

3.1 Physics at the particle scale

Mixing in granular media requires the grains to move relative to their neighbours. On the other hand, in classical fluids mixing is associated with the relative motion of molecules. Therefore, in both cases the origin of mixing is inherently microscopic, with the particle size defining the dimension of the microscale in granular media, and the molecule size defining the microscale in classsical fluids. Most generally, mixing can be understood in terms of the elementary trajectories, of either grains in granular media or molecules in fluids. The basic question that is underlying the physics of mixing is then: will two initially close particles be separated, or will they stay close to each other?

In the following we discuss how chaotic advection and diffusion could be related to particle separation.

3.1.1 Chaotic advection

An effective way for separating particles is to subject the material that carries them a highly heterogeneous advective field. The velocity difference between two particles distanced by $\Delta\mathbf{x}_p$ is given as $\Delta\mathbf{u}_p = \mathbf{u}_p(\mathbf{x}_p) - \mathbf{u}_p(\mathbf{x}_p + \Delta\mathbf{x}_p)$. (The subscript p was added to distinguish particle kinematics from continuum kinematics). This means that the distance between those particles $\Delta\mathbf{x}_p(t) = \int_{t'=0}^{t} \Delta\mathbf{u}_p dt'$ could quickly grow over time t. Stirring grains with a spatula does precisely this, since some grains are pushed along and around the spatula while others stay still. If one applies a non-steady stirring pattern, the advective field will not only be heterogeneous but will also evolve over time. The evolution of the relative distance between particles driven by many stirring patterns may be described using an exponential law $\Delta\mathbf{x}_p(t) \propto \Delta\mathbf{x}_p(t=0)e^{\lambda t}$, with λ being known as the Lyapounov exponent. The advective field for scenarios with a positive $\lambda > 0$ is called chaotic. In those fields pairs of particles (exponentially) quickly separate and the evolution of their distances should practically be impossible to predict – hence the adjective 'chaotic'. Chaotic advective field are good for mixing.

Back to the kitchen, a famous example for chaotic advective fields is the blinking vortex field, which bakers can produce using egg bitters with two counter rotative blades. Another way to produce chaotic advection is the baker's map, which could be followed by cutting bread dough in half, then stacking and compressing those halves onto one another, then repeating this operation over and over. Finally, rotating drums that are used as industrial granular mixers, produce intermittent recirculating flows, which may also form some chaotic advective field.

3.1.2 Diffusion as a stochastic process

Diffusion of particles is typically associated with the random walk of their trajectories. In molecular gas, this happens even without advective field. This is because molecules or atoms in gas are thermally agitated. Kinetic theory applied for an ideal gas predicts that the root-mean-squared of their velocities is given by: $\delta u_p = \frac{3k_bT}{m}$, where T is the temperature expressed in Kelvins, m is the particle mass and k_b the Boltzmann constant.

Gas particle trajectory schematically involves a ballistic free-flight over a distance ϵ at speed δu_p, followed by a binary collision with another particle and a bounce in a different direction. Series of such steps forms a random walk with an elementary step distance ϵ and frequency of direction change $\delta u_p/\epsilon$. Under these conditions the particle trajectories would be characterised by a mean square displacement growing linearly in time, as:

$$\lim_{\Delta t \to +\infty} \left[\mathbf{x}_p(t+\Delta t) - \mathbf{x}_p(t)\right]^2 = \beta D t, \quad \text{where } D = \epsilon \delta u_p, \tag{4}$$

where D is the self-diffusivity of the particles which has a dimension of length squared

divided by time. In this formulation β, the constant controlling the growth, receives the value $\beta = 2$, 4 or 6 depending on the system's dimension (1D, 2D or 3D, respectively). This means that the distance between the particles increases with time as: $\Delta \mathbf{x}_p(t) = \sqrt{\beta D t}$. This is typically slower than the exponential separation in chaotic advection.

The effect of mixing on the trajectory of one particle in a one-dimensional system could therefore be represented as:

$$x_p(t + \Delta t) = x_p(t) \pm \sqrt{2D\Delta t}. \tag{5}$$

Next, consider three adjacent representative volume elements (RVEs) in space $\{x + \Delta x,\ x,\ x - \Delta x\}$, which are separated by a distance $\Delta x = \Delta x_p = \sqrt{2D\Delta t}$ that is set over the incremental time difference Δt. We imagine many particles in each of these RVEs being captured by the concentration function $c(x)$ as the number of particles per unit volume Δx. The average exchange of particles through one RVE in space at point x is governed by the surcharges in the concentrations that the adjacent RVEs have relative to that x-RVE, which are given as $\Delta c_{\text{up}} = c(x + \Delta x) - c(x)$ and $\Delta c_{\text{down}} = c(x - \Delta x) - c(x)$. From a stochastic point of view, the change over an incremental time Δt in the x-RVE's concentration (that is $c(x)\Delta t$) can then be given by averaging the fluxes from the neighbouring RVE's, as given by $\dot{c}(x)\Delta t = \frac{1}{2}\left(\Delta c_{\text{up}} + \Delta c_{\text{down}}\right)$, and therefore:

$$\partial_t c(x) = D\frac{c(x + \Delta x) - 2c(x) + c(x - \Delta x)}{\Delta x^2} \approx D\nabla^2 c(x), \tag{6}$$

where the approximation comes by identifying the second central derivative of the concentration $\nabla^2 c(x)$; ∇ is the Del operator (in 1D: $\nabla = \frac{\partial}{\partial x}$); and $\partial_t = \frac{\partial}{\partial t}$ is the time derivative.

3.1.3 Diffusivity in granular media

In molecular gas, particle diffusion is thus driven by temperature-induced velocity fluctuations. In granular materials, particles are large (typically larger than one micron) and therefore 'athermal' in that grains motion is not affected by their temperature, irrespective to how fast the molecules within them move. However, in sheared granular materials, grains do exhibit strong velocity fluctuations. In diluted systems, these velocity fluctuations scale like:

$$\delta u = \dot{\gamma} f(n) d, \tag{7}$$

where $\dot{\gamma}$ is the shear strain rate (with dimension of inverse of time), d the particle size and $f(n)$ a function of the porosity n. Schematically, grains follow a trajectory comprised of free flight step of size $\ell \propto f(n)d$ at a speed δu_p before colliding with another

grain. This is similar to the trajectory of particles in molecular gas. Accordingly, the coefficient of self diffusion in diluted granular system is given by [HLCY08, UB04]:

$$D \propto \dot{\gamma} d^2. \tag{8}$$

The fact that the diffusivity D is proportional to the shear rate $\dot{\gamma}$ implies that granular diffusion only happens if there is some shear flow ($\dot{\gamma} > 0$). This is why it is referred to as *shear-induced* diffusion.

In dense granular flows, the gas analogy breaks down. Grains are so densely packed that there are hardly any spaces between them for free flights. Instead, grains endure sustained and multiple contacts at all times. Nonetheless, they exhibit strong velocity fluctuations scaling like [DCEP$^+$05, KR17]:

$$\delta u = \dot{\gamma} d \frac{1}{\sqrt{I}}, \quad \text{where } I = \dot{\gamma} d \sqrt{\rho_g / \sigma} \tag{9}$$

is the inertial number that expresses the ratio between the shear time $t_s = \dot{\gamma}^{-1}$ and the inertial time $t_i = d\sqrt{\rho_g/\sigma}$. As the dense flow regime corresponds to small values of the inertial number ($I \ll 1$), the velocity fluctuations are much larger than the elementary velocity scale $\dot{\gamma} d$.

To complete a random walk vision of grain trajectory in dense flows, one needs to consider the typical length ϵ on which grains can move in a free-flight manner. While the notion of free-flight is probably wrong given that grains are touching their neighbours at all times, the natural length scale for a pseudo-free flight is a particle size d. This leads to the following scaling for the diffusivity in dense granular flows [KR17, KR18]:

$$D \propto \dot{\gamma} d^2 \frac{1}{\sqrt{I}}. \tag{10}$$

It is possible to associate this scaling with a random walk trajectory whereby grains move a distance at the order of their size d at a frequency $\dot{\gamma}/\sqrt{I}$ much higher than the shear frequency $\dot{\gamma}$ (for the small I's associated with dense flows). This implies that within one shear deformation, grains exhibit a succession of fast displacements of the order of d.

Such trajectories are made possible by the development of transient clusters of jammed grains in dense granular flows [GRME13, MR19]. These clusters have a size scaling like $\ell_c \propto d/\sqrt{I}$ and a typical life time $t_c = \sqrt{t_i/\dot{\gamma}}$. The diffusivity in Eq. (10) can be expressed in terms of these time scales: $D = \ell_c^2/t_c$.

The connection between granular clusters and shear-induce diffusion is important, as it shows that the diffusivity in Eq. (10) results from the existence of large structures. Near solid boundaries (at a distance lower than l_c), the size of these structures is

truncated and the diffusivity is lower than that in similar flow conditions without walls [MRME13, RMME15].

3.2 Continuum mechanics

Continuum models for mixing in granular flow lump the micro-processes of particle trajectories into mass fluxes, and build the local budget of those fluxes coming in and out of an infinitesimal representative volume at any point in the flow, and at any time. At its simplest, this budget takes the following ' diffusion advection equation' form:

$$\partial_t c = \nabla \cdot (D \nabla c - \mathbf{u} c) \,, \tag{11}$$

where c refers to the concentration of a single species (for instance, particles with a certain grainsize s), and $\partial_t c$ is its temporal evolution. The first term in the right hand side is the divergence of the diffusive flux, while the second denotes the advective flux, D the diffusivity of the given species in the corresponding location, and $\mathbf{u}$ the local velocity.

The second term in the right hand side, which represents advective fluxes, can be expanded using the chain rule of differentiation to yield:

$$\mathring{c} = \nabla \cdot (D \nabla c) + c \dot{\varepsilon}_v, \quad \mathring{c} = \partial_t c + \mathbf{u} \cdot \nabla c, \quad \dot{\varepsilon}_v = -\nabla \cdot \mathbf{u}, \tag{12}$$

where $\mathring{c}$ is the material time derivative of the concentration, and $\dot{\varepsilon}_v$ is the volumetric strain rate that is here defined to be positive in compression, along with the conventional sign convention in geomechanics.

During volumetric compression we have $\dot{\varepsilon}_v > 0$, and thus the second term in the first relation above increases the concentration c. Conversely, volumetric dilation reduces the concentration. For conciseness, we focus in the following on the diffusive part, and thus consider zero volumetric strain rates with $\dot{\varepsilon}_v = 0$. In this way, the evolution equation for the concentration reduces to the classical diffusion equation:

$$\mathring{c} = \nabla \cdot (D \nabla c) \,. \tag{13}$$

If the diffusivity D was spatially invariant, one could take it out of the divergence operator and this budget becomes: $\mathring{c} = D \nabla^2 c$, which is in agreement with the result obtained from the stochastic averaging in Eq. (6).

3.2.1 Grainsize mixing

There are numerous open questions regarding the mixing by shear-induced diffusion in granular materials. Perhaps the most relevant for the current discussion is how

the grains diffuse in a polydisperse mixture. So far, our knowledge of shear-induce diffusion is mostly valid for quasi-monodisperse systems. A naïve extension to polydispersed media can follow by replacing c in Eq. 13 with the distribution $\phi(s)$ of a grainsize coordinate s using

$$\mathring{\phi} = \nabla \cdot [D(\dot{\gamma}, \phi) \nabla \phi] \,, \tag{14}$$

but at the moment we do not know the precise dependence of the diffusivity D on the shear rate $\dot{\gamma}$ and grainsize distribution ϕ, when the media is polydispersed. For example, do all grainsize species diffuse similarly in a mixture? Probably not. Does the diffusivity of a given species depend only on their own size s or also on the whole shape of the grainsize distribution $\phi(s)$? If yes, can the mean grainsize $\bar{s} = \int \phi(s) ds$ be taken as the most relevant parameter to capture those characteristics, or should we consider higher order weighed averages?

At any rate, inspired by Eqs. 8 and 10 we would expect D to depend on the shear rate $\dot{\gamma}$, itself dependent on the position. However, research is still required to quantify how strongly the diffusivity D depends on the full characteristics of the grainsize distribution.

4 Segregation

As described above, mixing will cause a system that is initially heterogeneous to become homogeneous. To do the opposite — to create order from disorder — is the process of segregation. In society, segregation is the outcome of the personal choices of individuals. Those who wish to live near good schools, or places of worship, or expensive cafes, tend to cluster geographically [Sch71]. When examining a map of the demography of a region, it is rare to find homogeneity. This basic observation implies that housing preferences are not random, and that segregation (which occurs when occupants move homes) is driven externally towards heterogeneity. For instance, when one of the authors purchased a home recently, it was partly to be close to work, partly to be close to shopping, and largely driven by cost. These various external driving factors are similar to those of his neighbours, who are typically in similar age, salary and socio-economic brackets.

4.1 Physics at the particle scale

The same segregation processes are observed in both humans and grains. Grains, while not necessarily having the individual capacity to relocate themselves, are often forced to do so by external factors such as construction work and earthquakes. When jostled, these grains transition from being in a stable environment with persistent contact with grains around them, to a more fluid-like state, where they are constantly

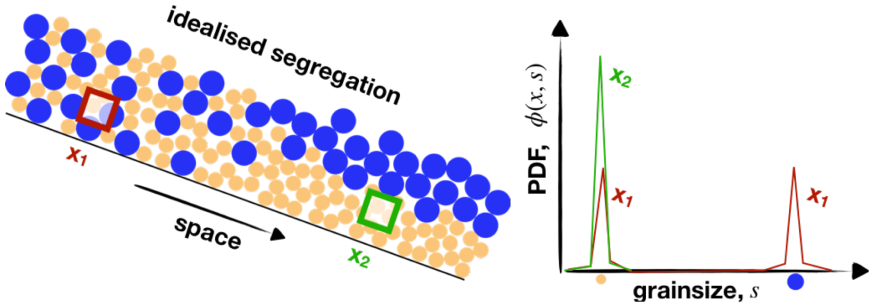


Figure 3: Schematic representation of the segregation process [ME11]. During flow down an inclined plane, when particles are initially homogeneously mixed, it is common to observe large particles rising to the free surface while small particles settle towards the base of the flow.

battering into those grains around them. As described in Section 3, these repeated collisions in general lead to mixing. If, however, collisions between grains are asymmetric (e.g. some grains consistently gain more energy than their neighbours), then such collisions may not in fact lead to mixing, but may cause particles to arrange themselves geographically.

In reality, particles have been observed to segregate because of differences in at least size, density, shape, roughness and elasticity [TBT14]. If any of these properties differs between grains, there is the possibility that the jostling (variously described by granular temperature, fluiditiy or kinetic stress) will not be shared equally by the particles and segregation may occur. Physicists call this a breakdown of the equipartition of energy in granular systems.

4.1.1 Segregation as a stochastic process

In Section 3.1.2 above, we described how diffusion could be represented as a stochastic, random walk process. Similarly, we can describe segregation as a stochastic process, but now there is a direction to the walking that depends on the grainsize s, which can be represented by a segregation velocity $\hat{u}(x,s,t)$. The effect of this velocity on the trajectory of one particle in a one-dimensional system could therefore be represented as:

$$x(s,t+\Delta t) = x(s,t) + \hat{u}(x,s,t)\Delta t. \tag{15}$$

Therefore, unlike the stochastic Eq. 5 for mixing, here the trajectory of the particles depend on their grainsize.

In a 'bidisperse' 1D system with only two grainsizes, large s_l and small s_s, over one time step Δt, the flux of small particles which moves out of a point (x,t) is the product of the segregation velocity at that height $\hat{u}$, the amount of that small size present at that height $\Phi_s(x,t)$ and the available space to move into which is controlled by the available space in the neighbouring space $\Phi_l(x+\Delta x,t) = 1-\Phi_s(x+\Delta x,t)$. Given the segregation velocity $\hat{u}$ at point x, we see small particles trickling down to $x-\Delta x$ and large ones surging up to $x+\Delta x$. Expressing the net effect of these motions on the conservation of the mass of small particles at the RVE at point x over the time step Δt, we get [ME11]:

$$\Big[\Phi_s(x,t+\Delta t)-\Phi_s(x,t)\Big]\Delta x = \Big[\hat{u}_s(x+\tfrac{1}{2}\Delta x)\Phi_s(x+\Delta x,t)\,(1-\Phi_s(x,t)) \\ -\hat{u}_s(x-\tfrac{1}{2}\Delta x)\Phi_s(x,t)\,(1-\Phi_s(x-\Delta x,t))\Big]\Delta t,$$

which in words reads as "the difference in the mass of small particles during Δt equals the difference in the flux of small particles from above minus those leaving down over the small RVE distance Δx". Therefore, given the numerical approximations of first derivatives, this stochastic segregation process takes the following continuum form:

$$\partial_t \Phi_s = \frac{\partial(\hat{u}_s \Phi_s(1-\Phi_s))}{\partial x}, \tag{16}$$

which could be compared [ME11] to the original segregation theory by Gray & Thornton [GT05]. This result is indeed limited to systems with only two particle sizes, though could be extended to polydispersed media. Furthermore, it does not disclose the structure of the segregation velocity $\hat{u}$. This is where the momentum equation of continuum mechanics becomes very handy.

4.2 Continuum mechanics

Over the past several decades there have been many approaches to describing segregation systems using continuum mechanics. One such approach has been to employ our models for the mixing of systems, but to use a negative diffusivity (which will cause de-mixing, rather than mixing). This may contradict thermodynamic laws (see previous ALERT School), and thus we do not explore this idea further. Others have used statistical approaches to try to relate the frequency of collisions with material properties [SL88]. The most common methods, however, follow the continuum approach with advective forces between the species, and describe the segregation of two (or more) types of material with a well defined segregation velocity [DU95, GT05].

Typically, models are based around systems of two sizes of particles (large s_l and small s_s), which are called 'bidisperse'. For example, such a system was described

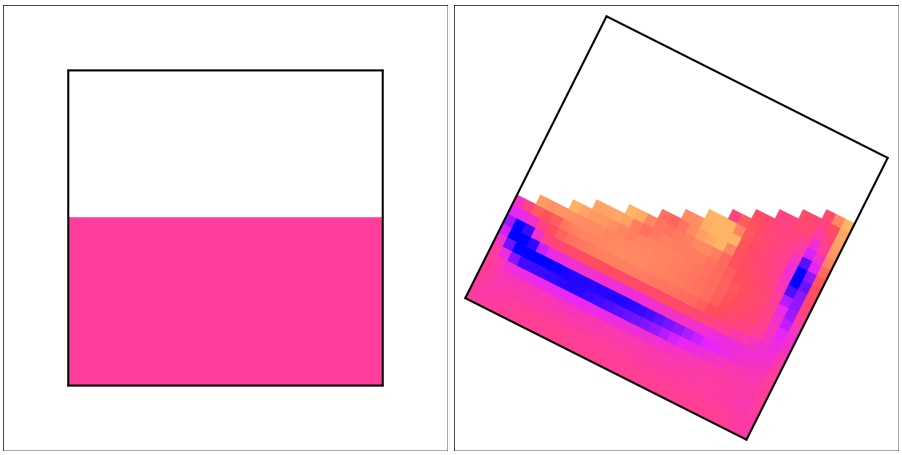


Figure 4: Segregation in a rotating square drum filled with a bimixture. Colour represents the average particle size, with blue being pure large particles, yellow pure small particles and pink a mixture of the two. *Left*: Initial sample that is homogeneously mixed. *Right*: Partially segregated sample after 50s of rotation.

stochastically using Eq. 16. We are then interested in tracking changes in the local volume fraction, Φ, of each component $\Phi_l = \delta(s-s_l)\phi(s)$ and $\Phi_s = \delta(s-s_s)\phi(s)$ (with $\delta(s)$ being the Dirac delta function). We then have the requirement that $\Phi_s + \Phi_l = 1$. While this may seem to be a reasonable starting point to understand the behaviour of more complex systems, it has recently been shown that this two-component approach (and its generalisations to many components) have significant drawbacks [MRE12]. We then turn to a different approach, which requires a complete description of the grainsize distribution $\phi(s)$.

Continuum models for segregation typically begin with some statement of conservation of mass and momentum for a mixture, largely based on mixture theory, which can be stated as

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{17}$$

$$\partial_t (\rho \mathbf{u}) + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) = \mathbf{F}_\rho, \tag{18}$$

where $\rho(s) = \phi(s)(1-n)\rho_g$ is the bulk density of the s species, $1-n$ is the solid fraction given the porosity n, and ρ_g the intrinsic density of the grains (which is here assumed similar for all grainsizes). The term $\mathbf{F}_\rho(s)$ contains all of the physics related to mixing and segregation, and the choice of this term is largely still open to intense discussion. Various forms have been proposed, but largely the focus has been to provide an analytical expression for the segregation velocity. Similarly, dividing by

$\phi(s)(1-n)$, the differential equations for the grainsize distribution $\phi(s)$ and grainsize velocities $\mathbf{u}(s)$ become

$$\partial_t \phi + \nabla \cdot (\phi \mathbf{u}) = 0, \tag{19}$$

$$\partial_t (\phi \mathbf{u}) + \nabla \cdot (\phi \mathbf{u} \otimes \mathbf{u}) = \mathbf{F}_\phi, \tag{20}$$

with $\mathbf{F}_\rho = \mathbf{F}_\phi (1-n)\rho_g$.

A recent work [MED+17] presents the current state of the art in modelling of segregation in polydisperse systems. In this context, the trajectory of a volume-grainsize element over time is described by

$$\mathbf{u} = \bar{\mathbf{u}} + \hat{\mathbf{u}} + \mathbf{u}', \tag{21}$$

where $\mathbf{u}(s)$ is the true velocity of species s, $\bar{\mathbf{u}}$ is the barycentric velocity, $\hat{\mathbf{u}}(s)$ the segregation velocity and $\mathbf{u}'(s)$ the fluctuating velocity due to diffusion. The segregation velocity, $\hat{\mathbf{u}}(s)$, can be related to the particle size s, the bulk density $\bar{\rho}$ and the kinetic stress $\bar{\boldsymbol{\sigma}}_k$ as

$$\hat{\mathbf{u}} = \frac{1 - \frac{\bar{s}}{s}}{\bar{\rho}\eta} \nabla \cdot \bar{\boldsymbol{\sigma}}_k, \tag{22}$$

where $\bar{s}$ is the hyperbolic mean grainsize and η is a single fitting parameter which sets the rate of segregation. One crucial outstanding issue in this context is a lack of predictability of the kinetic stress. This term is an additional stress that occurs due to the jiggling of particles. While there has been significant recent effort for predictions of this jiggling, a well established general model is yet to be widely accepted.

Recently, such models have begun to be included in numerical solvers for arbitrary problems, typically using the material point method [ME17b, FSHZ17]. One such code could be found in [Mar]. With these new capabilities, it is possible to predict the time evolution of the grainsize distribution in arbitrarily complex geometries. For example, the evolution of the mean grainsize in a rotating square drum is pictured in Figure 4.

5 Crushing

Where grainsize dynamics by mixing (Fig. 2) and segregation (Fig. 3) intrinsically relate to mass transfer through space, it is possible to imagine 'comminution' (a term used specifically for grainsize dynamics by crushing) without notable mass transfer. This does not mean that one can capture all the problems related to comminution without considering mass transfer, but that a closed system approach may be a good

starting point prior to addressing open system problems that involve significant mass transfers. This is particularly appealing for many geotechnical problems, where the soil mass only slightly deforms due to the deep-situ, confined ground conditions. In these cases one may imagine comminution as occurring within idealised insulated boxes that prevent neighbouring particles to pass through them, as shown in Fig. 5(a).

However, the closed system idealisation seizes to work for large deformation problems where particles often move preferentially by size (see Figure 5(b)). This is the case related to the challenge set by Vardoulakis in Fig. 1, as it involves simultaneous grain crushing, segregation and mixing. In the following we distinguish the physics of comminution in closed and open systems.

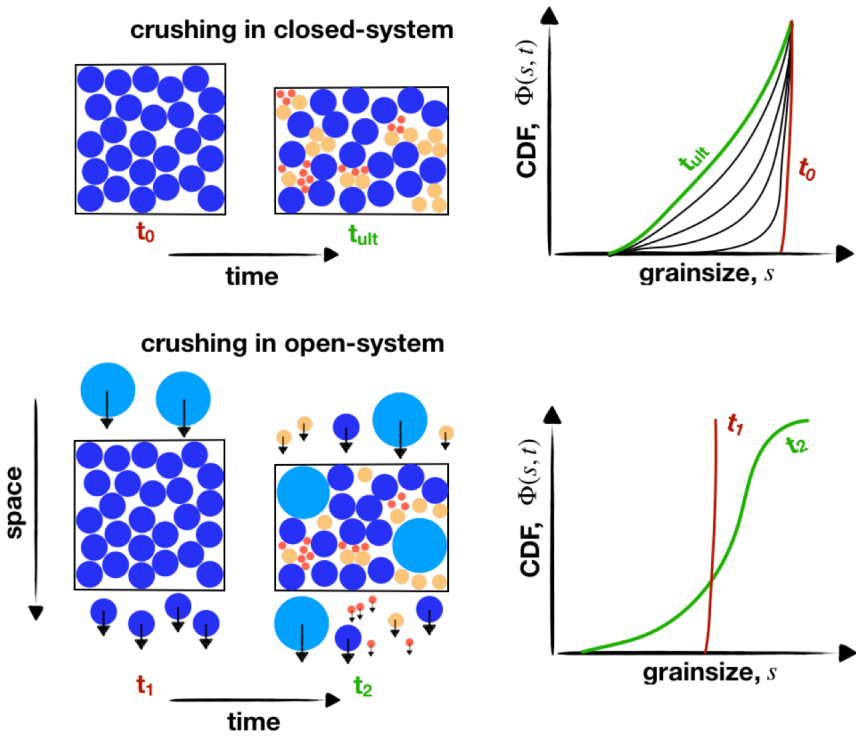


Figure 5: Schematic representation of continuously deforming crushable grains. *Up*: In closed-systems the grainsize distribution typically tends towards an ultimate distribution. *Down*: In open-systems the grainsize distributions do not necessarily tend to an ultimate state due to the exchange of particles with the surrounding.

5.1 Physics at the particle scale

Comminution has been studied over many decades, starting with the earlier work of von Rittinger in mineral processing [vR67] who defined a simple scaling law of comminution, which neglects the actual state of the particles as a function of loading and boundary conditions. The effect of these conditions on comminution was considered more recently in geomechanics, through the development of constitutive models in continuum mechanics [Har85, PVO93, Ein07]. The development of constitutive comminution models is mostly based on observations from insulated experimental devices or field conditions that typically prevent large mass flow through them, and are thus relevant for closed systems.

In closed systems comminution tend to reveal unique deterministic physical properties that are used in these developments. In particular, in earthquake fault physics research has motivated the idea that fragments self-organise into self-similar topological configurations that support the notion of a power law grainsize distribution at an ultimate state [Tur86, SOA$^+$86, SKB87] (or so-called 'fractal' configurations). This idea of an ultimate (typically power law) grainsize distribution after extremely large deformations in confined conditions has been supported in experimental geomechanics, after both extreme confined compressive [NHH$^+$01] and shearing conditions [CSBFG04]. Considering the existence of a minimum grainsize s_{min} due to the limit of fracture mechanics [Ken78] and long-range attraction forces (that bring smaller pieces to coalesce), and since there is always a maximum grainsize s_{max} within an RVE, the ultimate power law grainsize distribution takes the form:

$$\phi_u(s) = \frac{(3-\alpha)s^{2-\alpha}}{s_{\text{max}}^{3-\alpha} - s_{\text{min}}^{3-\alpha}}, \tag{23}$$

where α is the fractal dimension, which was motivated to range between 2.5 and 3 using on different models [SKB87, ME15].

Using idealised closed-system geometrically tessellated models of stochastic subvolumes in 2D and 3D [Tur86, SOA$^+$86], the reasons behind the ultimate power law grainsize distribution were linked to a competition between the crushing strength of individual particles that is typically larger for smaller particles, and the cushioning of the larger particles due to a larger number of surrounding contacting grains. Specifically, Mogi [Mog62] proposed to adopt the Weibull's weakest link distribution theory [W$^+$51] to describe the likelihood of grains to crush as a function of their size due to growing defects in larger particles. Similar conclusion was recently derived entirely based on fracture mechanics and non-linear elasticity originating from the non-linearity of contacts [ZBE15].

Let us return to the notion of a maximum grainsize s_{max} as shown in Eq. 23. Place many brittle grains within an insulated box with a maximum grainsize s_{max}. Next, shear and compress that box as much as you want. If we let grain crush, but not grow the maximum grainsize will never go above s_{max}. In fact, if the box contained many

grains, the competition mentioned above between Weibull's smaller-is-stronger and cushioning bigger-is-stronger will mean that, statistically speaking, particles with a maximum grainsize $s_{\max}$ will always remain in that box.

Next, open the box, get some particles in, put some other particles from another box nearby. Will the maximum grainsize remain unchanged? Of course not, since we did not say anything about the neighbouring box. This point is illustrated in Fig. 5, which explains why the fractal grainsize distribution model in Eq. 23 does not actually work for comminuting open systems. Indeed, from geological surveys of pyroclastic flows and dry snow avalanches with brittle grains, we do not normally find power law distributions, but rather, quite often, log-normal distributions [Sch67, BM09].

5.1.1 Crushing as a stochastic process

How can we capture this difference in the grainsize distributions between closed- and open systems involving grain crushing? Current treatments of comminution in open systems rely on stochastic information that goes beyond the deterministic information gained from the idealised experimental closed system conditions. Similar to mixing and segregation, comminution could also be described as a stochastic process. In this case there is no direct effect from advection, only indirect one. For this purpose the process could be understood without making reference to the continuum location. A crushing event occurs in a grain of size s at a microstructural position m if its the tensile stress σ_m exceeds its crushing strength σ_{cr} determined by its own size s_m and its nearest neighbours' sizes s_{m-1} and s_{m+1}. This is specified as [ME15, ME17a]:

$$s_m(t+\Delta t) = \zeta(s_m) s_m(t) \quad \text{if} \quad \sigma_i \geq \sigma_{\text{cr}}(s_{m-1}, s_m, s_{m+1}), \tag{24}$$

where $\zeta \in [0,1]$ is an independent random variable pulled from a fragment size distribution ranging from 0 to 1. In reality, when a single grain crushes it turns into many different sized fragments, but here this is taken stochastically using the variable ζ. A key point in this model is the use of the microstructural coordinate m that allows to preserve the information of grainsize of neighbours, which is said before controls the crushing process.

In isolation, this model [ME15, ME17a] was shown to well explain the emergence of power law grainsize distributions [ME15, ME17a]. By connecting the grainsize distribution to available porosities, it could be used to explain the constitutive behaviour of brittle granular media [GEMC18]. Finally, by combining this model with the two other stochastic models mentioned before for mixing in Sec. 3.1.2 and segregation in Sec. 4.1.1, it can also explain why in pyroclastic flows the grainsize distribution approaches a log-normal function [ME15, ME17a].

In order to describe the grainsize dynamics in this stochastic model one can establish the population balance. Following [Ram00] and formulating the conservation of mass

at the $\{s,t\}$-RVE (for a fixed point in x) the rate of the grainsize distribution can be calculated as [ME17a]:

$$\dot{\phi} = h^+(s,t) - h^-(s,t), \tag{25}$$

$$h^+ = \int_s^\infty b(s')P(s|s')\phi(s',t)ds', \tag{26}$$

$$h^- = b(s)\phi(s,t) \tag{27}$$

where h^+ is the incoming mass flux from all the s sized fragments originating from the bigger grains, while h^- is the outgoing mass flux representing the loss of mass from the corresponding RVE by crushing of s particles that become smaller. Here, $P(s'|s)$ is the conditional probability density function which dictates the probability of creating grainsize s from crushing a particle of grainsize s, which could be related directly to the fraction ζ in the stochastic model [GEMC18]. Finally, $b(s)$ is the breakage rate which governs the frequency at which particles of grainsize s break into smaller fragments.

5.2 Continuum mechanics

5.2.1 Closed systems

As described earlier, in closed systems during continuous loading the grainsize distribution approaches an ultimate grainsize distribution, which is typically given by the fractal power law of Eq. 23. Moreover, since in such systems grains can only reduce in size, the grainsize distribution tend to develop in an almost predictable monotonic path from initial states towards the ultimate state (see Fig. 5). This idea was adopted in the breakage mechanics theory [Ein07], with which one can express the evolution of the grainsize distribution ϕ using a breakage B internal variable:

$$\phi(s,B) = \delta(s - s_{\max})(1-B) + \phi_u(s)B, \tag{28}$$

where $\delta(x)$ is the Dirac's delta function. In the breakage mechanics theory the rate of breakage $\dot{B}$ is then derived within a thermodynamically consistent description, such that the grainsize distribution could be updated with $\dot{\phi} = (\phi_u(s) - \delta(s - s_{\max}))\dot{B}$.

5.2.2 Open systems

There is currently no mathematical solution to transform the stochastic model expressed in Eqs. 25 into a smooth exact analytic differential equation for the time evolution of the grainsize distribution. For example, we do not currently have an exact

analytic form for the birth function $b(s)$ in the population balance representation of the the stochastic model in Eq. 24, see [ME17a]. Obtaining such a smooth differential description of stochastic comminution is left for future research.

6 From hierarchicy to heterarchy in multiscale models

As shown in Fig. 6, typical multiscale models are *hierarchical*, coupling distinctly different models at different length scales. Take for example the definition of coupled FEM–DEM models [AT09, NCDD11, GZ14], which replace a constitutive model in a finite element simulation with a discrete element simulation at each gauss point. Unfortunately, this hierarchical approach does not lend itself for open-system processes such as mixing and segregation, since those erase the resolution required for tracking grainsize dynamics across the boundaries of neighbouring RVEs due to the scale separation. For example, considering the hierarchical FEM–DEM approach, the passage of grainsizes from one Gauss-DEM-point to its neighbouring Gauss-DEM-points and vice versa cannot be done without information on the exact locations of the particles on the boundaries of the other Gauss-DEM-points.

Recently, Marks & Einav [ME17a] proposed to resolve this issue using a new paradigm of multiscale models – a *heterarchical* model. This stochastic model, in contrast to a hierarchical one, does not separate between the two scales, as it allows information to be passed from the microscale within one RVE to another RVE. Put simply, this model simply integrates the closed-system stochastic crushing model with the open-system stochastic mixing and segregation models. In so doing, for general comminuting problems one cannot impose an ultimate grainsize distribution, since there could alway be a way for neighbouring particles of unknown sizes to flow into the RVE. However, in closed systems, where that flow is prevented, this model does recover an ultimate fractal grainsize distribution achieved after much loading.

The disadvantage of this heterarchical multi-scale model is that it is not analytically described using smooth differential equations due to its stochastic ingredient of grain crushing, albeit the possible smooth descriptions of the other stochastic ingredients of mixing and segregation.

7 Conclusions

This Chapter reviewed the particle-scale stochastic physical processes that control the evolution of grainsize distributions in time and space. These processes include open-system mixing and segregation linked with mass transfer and closed-system grain crushing where mass transfer is not required for description. The modelling of each of these three dynamics is discussed in the context of both stochastic and continuum mechanics. In the cases of mixing and segregation we demonstrated that the stochastic

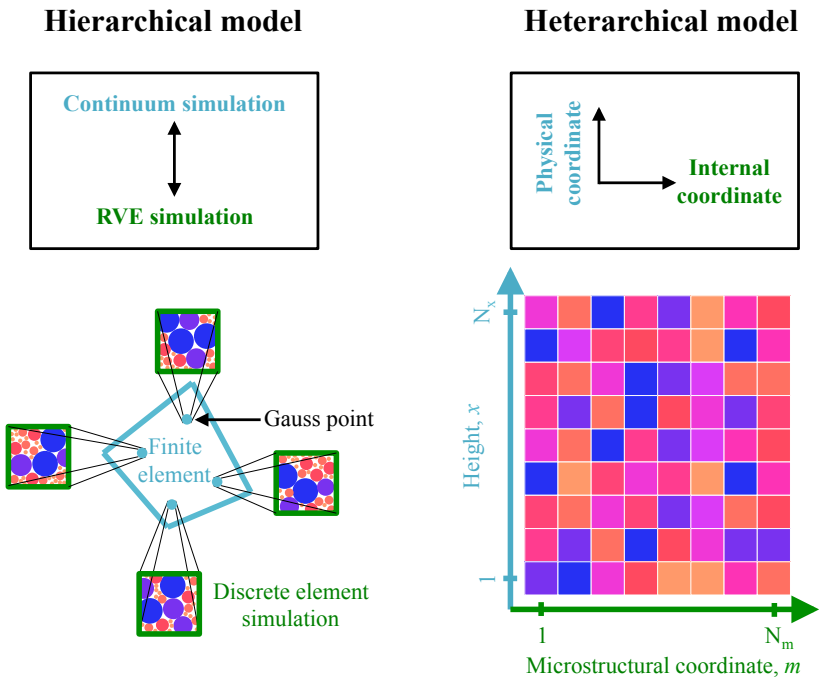


Figure 6: Alternative structures for multiscale models (see [ME17a]). *Top left* a hierarchical model, where continuum information is passed between two distinct simulations, each at different length scales. *Top right* a heterarchical model, where the two length scales are coupled directly in a simple model that allows discrete information to pass across scales. *Bottom left* a typical example of a hierarchical model, where a constitutive model in a FEM simulation is given by a separate DEM simulation. *Bottom right* a heterarchical model where the two scales coexist in the same framework. In all cases, *light blue* represents the continuum scale, and *green* represents the representative volume element scale.

physics at the particle scale could be upscaled to derive smooth differential equations for continuum model realisations. So far we do not have a good equivalent model for grain crushing that can homogenise from stochastic to continuum levels.

What is clear, though, is that the complete modelling of grainsize dynamics cannot be achieved by adopting hierarchical models due to grainsize dependent mass transfers. The alternative way forward is to adopt an open-system heterarchical approach. With a homogenised stochastic solution for grain crushing to be derived in the future, we envisage that the whole spectrum of mixing-segregation-crushing dynamics could be tracked within a continuum mechanics framework, with one additional microstructural coordinate to represent transfers through the grainsize scales. In this way, we would hopefully be able to resolve mathematically the oldest industrial problem in the world,

and in so doing address the challenge set by Vardoulakis [Var] (Figure 1).

References

[AT09] José E Andrade and Xuxin Tu. Multiscale framework for behavior prediction in granular media. *Mechanics of Materials*, 41(6):652–669, 2009.

[BM09] Perry Bartelt and Brian W McArdell. Granulometric investigations of snow avalanches. *Journal of Glaciology*, 55(193):829–833, 2009.

[CSBFG04] MR Coop, KK Sorensen, T Bodas Freitas, and G Georgoutsos. Particle breakage during shearing of a carbonate sand. *Géotechnique*, 54(3):157–163, 2004.

[DCEP+05] Frédéric Da Cruz, Sacha Emam, Michaël Prochnow, Jean-Noël Roux, and François Chevoir. Rheophysics of dense granular materials: Discrete simulation of plane shear flows. *Physical Review E*, 72(2):021309, 2005.

[DU95] VN Dolgunin and AA Ukolov. Segregation modeling of particle rapid gravity flow. *Powder Technology*, 83(2):95–103, 1995.

[Ein07] Itai Einav. Breakage mechanics – part i: theory. *Journal of the Mechanics and Physics of Solids*, 55(6):1274–1297, 2007.

[FSHZ17] Minglong Fei, Qicheng Sun, Kimberly Hill, and Gordon GD Zhou. Simulation of size segregation in granular flow with material point method. In *EPJ Web of Conferences*, volume 140, page 11010. EDP Sciences, 2017.

[GEMC18] Giulia Guida, Itai Einav, Benjy Marks, and Francesca Casini. Linking micro grainsize polydispersity to macro porosity. *International Journal of Solids and Structures*, 2018.

[GRME13] D Griffani, Pierre Rognon, B Metzger, and I Einav. How rotational vortices enhance transfers. *Physics of Fluids*, 25(9):093301, 2013.

[GT05] JMNT Gray and AR Thornton. A theory for particle size segregation in shallow granular free-surface flows. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461(2057):1447–1473, 2005.

[GZ14] Ning Guo and Jidong Zhao. A coupled fem/dem approach for hierarchical multiscale modelling of granular media. *International Journal for Numerical Methods in Engineering*, 99(11):789–818, 2014.

[Har85] Bobby O Hardin. Crushing of soil particles. *Journal of geotechnical engineering*, 111(10):1177–1192, 1985.

[HLCY08] Shu-San Hsiau, Li-Shin Lu, Cheng-Yu Chou, and Wen-Lung Yang. Mixing of cohesive particles in a shear cell. *International Journal of Multiphase Flow*, 34(4):352–362, 2008.

[Ken78] K Kendall. The impossibility of comminuting small particles by compression. *Nature*, 272(5655):710, 1978.

[KR17] Prashidha Kharel and Pierre Rognon. Vortices enhance diffusion in dense granular flows. *Physical review letters*, 119(17):178001, 2017.

[KR18] Prashidha Kharel and Pierre Rognon. Shear-induced diffusion in nonlocal granular flows. *EPL (Europhysics Letters)*, 124(2):24002, 2018.

[Mar] Benjy Marks. Material point method for grainsize segregation. `https://github.com/benjym/poly-mpm`. Accessed: 2019-06-10.

[ME11] Benjy Marks and Itai Einav. A cellular automaton for segregation during granular avalanches. *Granular Matter*, 13(3):211–214, 2011.

[ME15] Benjy Marks and Itai Einav. A mixture of crushing and segregation: the complexity of grainsize in natural granular flows. *Geophysical Research Letters*, 42(2):274–281, 2015.

[ME17a] Benjy Marks and Itai Einav. A heterarchical multiscale model for granular materials with evolving grainsize distribution. *Granular Matter*, 19(3):61, 2017.

[ME17b] Benjy Marks and Itai Einav. Material point method modelling of landslides with coupled segregation. In *International Workshop on Bifurcation and Degradation in Geomaterials*, pages 213–218. Springer, 2017.

[MED+17] Benjy Marks, Jon Alm Eriksen, Guillaume Dumazer, Bjørnar Sandnes, and Knut Jørgen Måløy. Size segregation of intruders in perpetual granular avalanches. *Journal of Fluid Mechanics*, 825:502?514, 2017.

[Mog62] Kiyoo Mogi. The influence of dimensions of specimens of the fracture strength of rocks-comparison between the strength of rock specimens and that of the earth's crust. *Bull. earthq. Res. Inst., Univ. Tokyo*, 40:175–185, 1962.

[MR19] Matthew Macaulay and Pierre Rognon. Shear-induced diffusion in cohesive granular flows: effect of enduring clusters. *Journal of Fluid Mechanics*, 858, 2019.

[MRE12] Benjy Marks, Pierre Rognon, and Itai Einav. Grainsize dynamics of polydisperse granular segregation down inclined planes. *Journal of Fluid Mechanics*, 690:499–511, 2012.

[MRME13] Thomas Miller, P Rognon, B Metzger, and I Einav. Eddy viscosity in dense granular flows. *Physical review letters*, 111(5):058002, 2013.

[NCDD11] Michał Nitka, Gaël Combe, Cristian Dascalu, and Jacques Desrues. Two-scale modeling of granular materials: a dem-fem approach. *Granular Matter*, 13(3):277–281, 2011.

[NHH+01] Yukio Nakata, Masayuki Hyodo, Adrian FL Hyde, Yoshinori Kato, and Hidekazu Murata. Microscopic particle crushing of sand subjected to high pressure one-dimensional compression. *Soils and foundations*, 41(1):69–82, 2001.

[PVO93] E Papamichos, I Vardoulakis, and H Ouadfel. Permeability reduction due to grain crushing around a perforation. *International Journal of Rock Mechanics and Mining Sciences*, 30(7):1223–1229, 1993.

[Ram00] Doraiswami Ramkrishna. *Population balances: Theory and applications to particulate systems in engineering*. Elsevier, 2000.

[RMME15] Pierre G Rognon, Thomas Miller, Bloen Metzger, and Itai Einav. Long-range wall perturbations in dense granular flows. *Journal of Fluid Mechanics*, 764:171–192, 2015.

[Sch67] Hans-Ulrich Schmincke. Graded lahars in the type sections of the ellensburg formation, south-central washington. *Journal of Sedimentary Research*, 37(2):438–448, 1967.

[Sch71] Thomas C Schelling. Dynamic models of segregation. *Journal of mathematical sociology*, 1(2):143–186, 1971.

[SKB87] Charles Sammis, Geoffrey King, and Ronald Biegel. The kinematics of gouge deformation. *Pure and Applied Geophysics*, 125(5):777–812, 1987.

[SL88] SB Savage and CKK Lun. Particle size segregation in inclined chute flow of dry cohesionless granular solids. *Journal of Fluid Mechanics*, 189:311–335, 1988.

[SOA+86] Charles G Sammis, Robert H Osborne, J Lawford Anderson, Mavonwe Banerdt, and Patricia White. Self-similar cataclasis in the formation of fault gouge. *Pure and Applied Geophysics*, 124(1-2):53–78, 1986.

[TBT14] Deepak Raju Tunuguntla, Onno Bokhove, and Anthony Richard Thornton. A mixture theory for size and density segregation in shallow granular free-surface flows. *Journal of fluid mechanics*, 749:99–112, 2014.

[Tur86] DL Turcotte. Fractals and fragmentation. *Journal of Geophysical Research: Solid Earth*, 91(B2):1921–1926, 1986.

[UB04] Brian Utter and Robert P Behringer. Self-diffusion in dense granular shear flows. *Physical Review E*, 69(3):031308, 2004.

[Var] Ioannis Vardoulakis. Personal communication on June 2007.

[vR67] RP von Rittinger. Textbook of mineral dressing. *Ernst and Korn, Berlin*, 1867.

[W+51] Waloddi Weibull et al. A statistical distribution function of wide applicability. *Journal of applied mechanics*, 18(3):293–297, 1951.

[ZBE15] YD Zhang, Giuseppe Buscarnera, and I Einav. Grain size dependence of yielding in granular soils interpreted using fracture mechanics, breakage mechanics and weibull statistics. *Géotechnique*, 66(2):149–160, 2015.

Hydro-mechanics of porous and granular material – Poro-elasticity and beyond

Holger Steeb

University of Stuttgart, Institute of Applied Mechanics (CE) & SimTech, Pfaffenwaldring 7, D-70569 Stuttgart, Germany

We review classical linear poro-elasticity and discuss in detail kinematical assumptions, constitutive relations, and the derivation of field equations. It will be pointed out that for an isotropic homogeneous porous material, Biot's poro-elastic constitutive equations include three elastic bulk moduli and one shear modulus. The elastic constitutive equations are complemented by Darcy's equation resulting from a linear assumption for the (viscous) momentum exchange. It is shown how the resulting boundary value problem could be numerically solved, e.g. by means of Finite Element Methods for various application problems in geomechanics and geophysics. Further, the quasi-static poro-elastic equations are subsequently extended towards Biot's set of equations describing acoustic waves in porous media. We discuss how effective material properties of heterogeneous porous materials distinguish from properties of homogeneous media and present various typical examples of heterogeneities in geomechanical and geophysical applications.

1 Introduction

1.1 The matter of scales

The physical behaviour of porous and granular media can be (and is) observed, modelled and numerically simulated on various length (and time) scales. On geological length scales (km), hydro-mechanical coupling in (fractured) hydro-carbon or geothermal reservoirs could lead to phenomena like the well-known Noordbergum effect also denoted as an inverse pumping effect [Rod89]. Inherent structural properties of the reservoir rock on the investigated length scale, like single fractures or fracture networks, have to be resolved in a discrete way in order to study the above mentioned phenomenon. Significantly smaller details, like heterogeneities on the meter scale

are geometrically not resolved in such investigations. Small scale properties are included through an inherently made homogenization step and enter into the simulations as effective material properties. On that scale hydro-mechanical coupling effects could be well desribed by coupled continuum models like (linear) quasi-static poro-elasticity [Bio41]. Samples with characteristic lengths in the range of 10 cm - 1 m, often found in laboratory investigations of soils/rocks or engineered geostructures, involve more detailed morphological features. Again, continuum approaches, like poro-elasticity [Wan00, Che16, Ver10] and models based on continuum mixture theory [Tru57, Cou10, Ehl02a] are numerically implemented and used in simulations. On the basis of advanced image-based characterization tools, like X-Ray Computed Tomography (XRCT), Magnetic Resonance Imaging (MRI), Neutron Tomography (NT) etc, pore-scale resolved investigations and simulations are getting more and more popular. Digital Rock Physics (DRP), with resolution on the μm scale [KLM$^+$09] is a rapidely evolving discipline where effective physical properties like elastic moduli, acoustic wave velocities, intrinsic permeability, electric or thermal conductivity etc. are numerically calculated on the basis of voxel-based μXRCT scans of the pore space [SEKS11]. Thus, the morphology of the pore space of a typical reservoir rock or of a granular packing of poly-disperse particles is inherently taken into account in simulations. Obviously, this leads to additional demands on simulations tools which will be briefly mentioned for the calculation of two physical properties:

Effective intrinsic permeability: The effective hydraulical properties of granular or porous media can be numerically obtained by Direct Numerical Simulations (DNS) of fluid flow through the pore space of the porous medium. In DRP, the (complex) pore geometry is described by the post-processed (segmented) voxel data set of the XRCT scan. On the basis of a binarized data set which consists of a) pore space and b) skeleton, computationally demanding flow simulations on the basis of a voxel (cartesian) grid are performed. Often, this leads to 3-dim simulations for grids with more then 1000^3 voxels [YMP$^+$15], cf. Stokes simulations in Fig. 1. If we are interested in the (scalar) intrinsic permeability of a sample, i.e. if the velocities in the pore space are small and "creeping flow" conditions can be assumed which is related to small Reynolds numbers[1] $Re \ll 1$, the flow behaviour can be predicted by Lattice Boltzman simulations or by solving the stationary Stokes equations with Finite Differences, Finite Element or Finite Volume Methods [YMP$^+$15, OURS15]. The effective intrinsic permeability $k^{\mathfrak{s}}$ of the numerically investigated domain is then calculated from the mean volume fluxes and the effective pressure loss. Obviously, this technique is very similar to an experimental permeability test in the laboratory. Compared to (allways necessary) experimental investigations, one benefit of DNS simulations based on XRCT data is the possibility to perform a series of tests with modified boundary conditions or to avaluate all fluid flow components. This information could be used for instance to investigate the amount of anisotropy and, if anisotropy is relaevant, to calculate the diagonal and off-diagonal components of a permeability tensor and its

[1] Here we use a "multiscale" definition of the Reynolds number. $Re :=$ viscous forces / inertia forces $= (\rho_0^{\mathfrak{f}R} \, \|\mathbf{w}\| \, r)/\eta^{\mathfrak{f}R}$. $\rho_0^{\mathfrak{f}R}$ is the effective density of the fluid in the initial state, $\|\mathbf{w}\|$ is the mean bulk velocity, r is a microscopic length scale (grain or particle size), and $\eta^{\mathfrak{f}R}$ is the effective dynamic viscosity of the fluid.

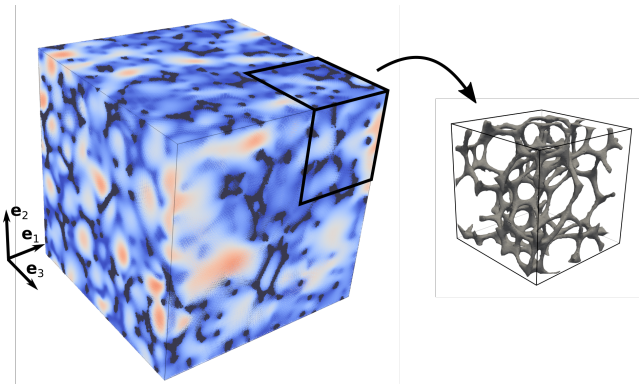


Figure 1: Pore morphology of a highly porous volcanic rock (reticulite) and velocity magnitudes numerically obtained by high-resolution Stokes simulations from μXRCT data.

principal axis.
Elastic properties based on coarse-grained granular media: Discrete Elements Methods (DEM) are one of the most prominent numerical methods taking into account the "discrete" phenomena of particle-particle interactions [CS79] in a granular medium. Thus, load- or deformation-dependent local properties of granular packings like "force-chains" are implicitely taken into account in numerical DEM simulations. Additionally, the non-linear evolution of the granular pore morphology, caused e.g. by local shear banding, is taken into account by DEM simulations. Unfortunately, due to the numerical expense of discrete methods, coarse graining of "realistic" (polydisperse) particle morphologies is often necessary in a pre-processing step. Pore scale resolved DNS simulations of fluid flow or mechanical DEM investigations are only two examples of numerical simulations which are taking the discrete or pore structure of porous media into account. They could give us an insight into the governing physical processes on the small (pore) scale and could provide effective material properties which could be used in subsequent continuum scale investigations.
Porous materials which are fully saturated with a viscous pore fluid can be effectively described on that continuum scale with the theory of poro-elasticity. Historically, two-phase poro-elastic theories are tracing back to the seminal work of Karl von Terzaghi [Ter43] establishing the (linear) theory of consolidation and the generalization launched by Maurice Anthony Biot for the quasi-static case [Bio41] and the case of acoustic waves in porous media [Bio56a, Bio56b]. In the textbook of Reint de Boer [dB00] one can find a detailed historical discussion about poro-elastic modelling including the tragic dispute of Paul Fillunger and Karl von Terzaghi. It was Clifford Truesdell [Tru57] who generalized the concepts of so-called super-imposed continua (and therefore also poro-elastic media) introducing the concept of continuum mixture theory which was in the last decades further developed and especially adopted for specific (e.g. non-linear) applications in porous media by de Boer [dB00], Ehlers

[Ehl02a] and Coussy [Cou95, Cou04].
In the following sections we discuss linear (quasi-static) poro-elasticity in the extended form of Biot [Bio41]. In these fields, various review articles and books have been published in the last decades. Only a part of them should be mentioned here: [RC76, ZSK86, K̈91, DC93, Wan00, Ver10, Leh11, RS15, Che16, Mer16, SR19] and [Zim91, Chapter 7.].

The paper is organized as follows: In the next section a brief summary of the basic concepts of poro-elasticity will be given. We mainly focus on basic assumptions and their consequences related to kinematics, balance relations, and constitutive assumptions. After introducing the set of governing partial differential equations of quasi-static linear poro-elasticity we extend the model towards the description of acoustic waves. The paper is closed with a discussion of heterogeneities in porous media, their physical consequences, and some applications.

2 Linear quasi-static poro-elasticity

From a modelling point of view, the theory of poro-elasticity is a based on the concepts of superimposed continua. Thus, at the material point $\mathcal{P}(\mathbf{x},\, t)$ (beeing the smallest mathematical object or pint in a continuum formulation) the fluid phase and the solid constituent co-exist. Like other well-established continuum theories (cf. the Lamé-Navier equations for Hookean elastic solids or the Navier-Stokes equations for Newtonian fluids) poro-elasticity describes processes on a macroscopical scale. In such classical continuum model approaches, all microscopical details are only taken into account in in a coarse grained or homogenized (smeared-out) approach. One immanent consequence of that modelling concept is sketched in Fig. 2. We observe that details of the pore morphology of a porous medium which are for instance the morphological details of the pores or the pore throats are disregared. In poro-elasticity only porosity, i.e. the (averaged) volume occupied by the fluid phase in a Representative Volume Element (RVE) is taken into account. Here, a RVE is a well-defined unit volume which is assumed to be large enough to capture quantities like e.g. porosity or the volume occuped by solid particles (Fig. 2).

2.1 Concepts, assumptions, definitions and some notations

On the poro-elastic modelling scale (i.e. the continuum scale) the REV is the smallest modelling unit, called the "material point" $\mathcal{P}(\mathbf{x}, t)$ where all the local properties like local volume $\mathrm{d}v$, volume occupied by the solid phase $\mathrm{d}v^{\mathfrak{s}}$ and the fluid phase $\mathrm{d}v^{\mathfrak{f}}$, respectively as well as local mass $\mathrm{d}m^{\mathfrak{s}}$ of the solid and the fluid phase $\mathrm{d}m^{\mathfrak{f}}$ are attached. Thus, the local mass and volume elements are uniquely split into their phase contribution

$$\mathrm{d}m = \mathrm{d}m^{\mathfrak{s}} + \mathrm{d}m^{\mathfrak{f}} \quad \text{and} \quad \mathrm{d}v = \mathrm{d}v^{\mathfrak{s}} + \mathrm{d}v^{\mathfrak{f}}. \tag{1}$$

Obviously, the material point and therefore the volume elements can also be evaluated at initial time $t = t_0$. In this reference state, we denote the volume elements as $\mathrm{d}v(\mathbf{x}, t_0) =: \mathrm{d}v_0$, $\mathrm{d}v^{\mathfrak{s}}(\mathbf{x}, t_0) =: \mathrm{d}v_0^{\mathfrak{s}}$ and $\mathrm{d}v^{\mathfrak{f}}(\mathbf{x}, t_0) =: \mathrm{d}v_0^{\mathfrak{f}}$. From this geometrical consideration, we could define field variables on the continuum scale (note that field variables are locally defined and therefore functions of the current position vector $\mathbf{x}$ and time t). Porosity is such a field variable which is denoted by Coussy [Cou04, p. 5] as Eulerian porosity[2] $n^{\mathfrak{f}}(\mathbf{x}, t)$

$$n^{\mathfrak{f}} := \frac{\mathrm{d}v^{\mathfrak{f}}}{\mathrm{d}v} = \frac{\mathrm{d}v - \mathrm{d}v^{\mathfrak{s}}}{\mathrm{d}v}. \tag{2}$$

It's a (non-linear) Eulerian quantity as the current volume of the fluid phase $\mathrm{d}v^{\mathfrak{f}}$ is related to the current volume element $\mathrm{d}v$. Similar, a Lagrangian porosity $\phi(\mathbf{x}, t)$ could be introduced. Note that (in contrast to the Eulerian porosity), ϕ is a linear measure where only the numerator is evolving in time

$$\phi := \frac{\mathrm{d}v^{\mathfrak{f}}}{\mathrm{d}v_0} = \frac{\mathrm{d}v - \mathrm{d}v^{\mathfrak{s}}}{\mathrm{d}v_0} \quad \text{with} \quad \phi_0 = \frac{\mathrm{d}v_0^{\mathfrak{f}}}{\mathrm{d}v_0}. \tag{3}$$

Here, the current volume of the fluid phase $\mathrm{d}v^{\mathfrak{f}}$ is related to the volume element $\mathrm{d}v_0$ in it's reference configuration. Have in mind that in linear poro-elasticity the Lagrangian and the Eulerian porosities are identical if the reference state is strain-free which can be shown by a formal mathematical linearization step of the Eulerian porosity

$$\mathrm{lin}(n^{\mathfrak{f}}) = \phi. \tag{4}$$

Consequently, we define further field variables like the effective densities $\rho^{\mathfrak{s}R}$, $\rho^{\mathfrak{f}R}$, the partial densities $\rho^{\mathfrak{s}}$, $\rho^{\mathfrak{f}}$ and the density of the biphasic mixture ρ

$$\rho^{\mathfrak{s}R} := \frac{\mathrm{d}m^{\mathfrak{s}}}{\mathrm{d}v^{\mathfrak{s}}} \quad \text{and} \quad \rho^{\mathfrak{f}R} := \frac{\mathrm{d}m^{\mathfrak{f}}}{\mathrm{d}v^{\mathfrak{f}}}, \tag{5}$$

$$\rho^{\mathfrak{s}} := \frac{\mathrm{d}m^{\mathfrak{s}}}{\mathrm{d}v} \quad \text{and} \quad \rho^{\mathfrak{f}} := \frac{\mathrm{d}m^{\mathfrak{f}}}{\mathrm{d}v}, \tag{6}$$

$$\rho := \frac{\mathrm{d}m}{\mathrm{d}v}. \tag{7}$$

These quantities can also be evaluated in the reference configuration at time t_0. Following the notation introduced earlier, these initial quantities will be denoted with a subscript (e.g. for the effective density of the fluid $\rho^{\mathfrak{f}R}(\mathbf{x},\, t_0) =: \rho_0^{\mathfrak{f}R}$). If the pore space is saturated with more then one pore fluid, this set of variables and their notations could be extended in a straightforward sense.

[2] In continuum mixture theory, this concept could be generalized by introducing the so-called volume fractions n^{α} for each constituent φ^{α} with $n^{\alpha} := \mathrm{d}v^{\alpha}/\mathrm{d}v$. We easily observe that a saturation condition $\sum_{\alpha} n^{\alpha} \equiv 1$, is fulfilled cf. [Cou04, Ehl02b]. Here, we restrict ourselves to biphasic media, thus we could imit ourselves to one volume fraction $n^{\mathfrak{f}} \equiv \phi$.

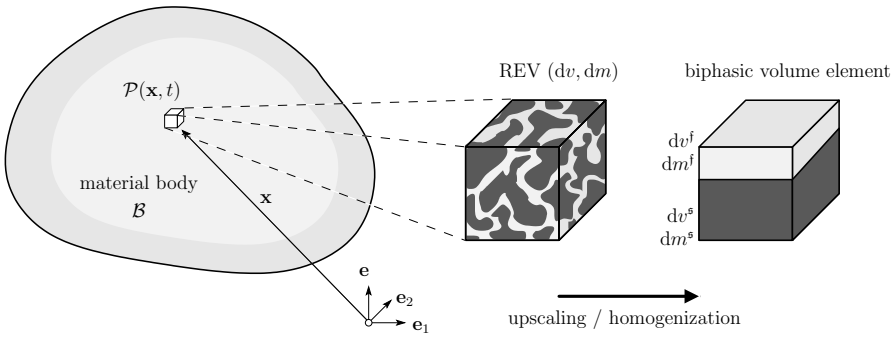


Figure 2: The concept of superimposed continua and the properties of constituents on the macroscopical scale.

2.2 Kinematics

The displacement of the porous body (skeleton) is described by the displacement vector $\mathbf{u}_{\mathfrak{s}}$.[3]. Further, we introduce velocities and accelerations denoted by $\mathbf{v}_{\mathfrak{s}}$, $\mathbf{v}_{\mathfrak{f}}$ and $\mathbf{a}_{\mathfrak{s}}$, $\mathbf{a}_{\mathfrak{f}}$. Note that we restrict ourselves here to a linear theory, thus all time derivatives ("dot" derivatives) are understood as partial derivates [4]:

$$\ddot{\mathbf{u}}_{\mathfrak{s}} = \dot{\mathbf{v}}_{\mathfrak{s}} = \mathbf{a}_{\mathfrak{s}} \quad \text{and} \quad \ddot{\mathbf{u}}_{\mathfrak{f}} = \dot{\mathbf{v}}_{\mathfrak{f}} = \mathbf{a}_{\mathfrak{f}}. \tag{8}$$

A relative or seepage velocity could be introduce as $\mathbf{w}_{\mathfrak{f}} = \mathbf{v}_{\mathfrak{f}} - \mathbf{v}_{\mathfrak{s}}$. From the solid displacement we are able to introduce the solid strains as the symmetric part of the displacement gradients

$$\boldsymbol{\varepsilon}_{\mathfrak{s}} = \frac{1}{2}\left(\operatorname{grad}\mathbf{u}_{\mathfrak{s}} + \operatorname{grad}^T\mathbf{u}_{\mathfrak{s}}\right). \tag{9}$$

Further, we know that each (symmetric) second order tensor could be split into a deviatoric and a volumetric part expressing shear and volumetrical deformations, respectively

$$\boldsymbol{\varepsilon}_{\mathfrak{s}} = \operatorname{dev}(\boldsymbol{\varepsilon}_{\mathfrak{s}}) + \operatorname{vol}(\boldsymbol{\varepsilon}_{\mathfrak{s}}) =: \boldsymbol{\gamma}_{\mathfrak{s}} + e_{\mathfrak{s}}\,\mathbf{I}, \tag{10}$$

where $\mathbf{I} = \delta_{ij}\,\mathbf{e}_i \otimes \mathbf{e}_j$ is the second order unity tensor (identity map with $\mathbf{a} = \mathbf{I}\cdot\mathbf{a}$) and δ_{ij} is the so-called Kronecker delta with properties $\delta_{ij} \equiv 1$ for $i = j$ and $\delta_{ij} \equiv 0$ for $i \neq j$. It should be noted (details in [SR19, Cou04]) that the linear volumetrical strain

[3] Here, we use the following notation for subscripts and superscripts: Kinematical quantities have subscripts $(-)_{\mathfrak{s},\mathfrak{f}}$ while all other non-kinematical quantities will be denoted with superscripts $(-)^{\mathfrak{s},\mathfrak{f}}$, cf. details in [Ehl02b]

[4] In continua, the so-called material time derivative of a vectorial field variable $\boldsymbol{\Psi}^{\alpha}$ in a mixture is given by $(\boldsymbol{\Psi}^{\alpha})'_{\alpha} = \partial_t\boldsymbol{\Psi} + \operatorname{grad}\boldsymbol{\Psi}\cdot\mathbf{v}_{\alpha}$. The first term is denoted as (linear) local or partial time derivative while the 2nd term is a convective (non-linear) term which vanishes in linear models.

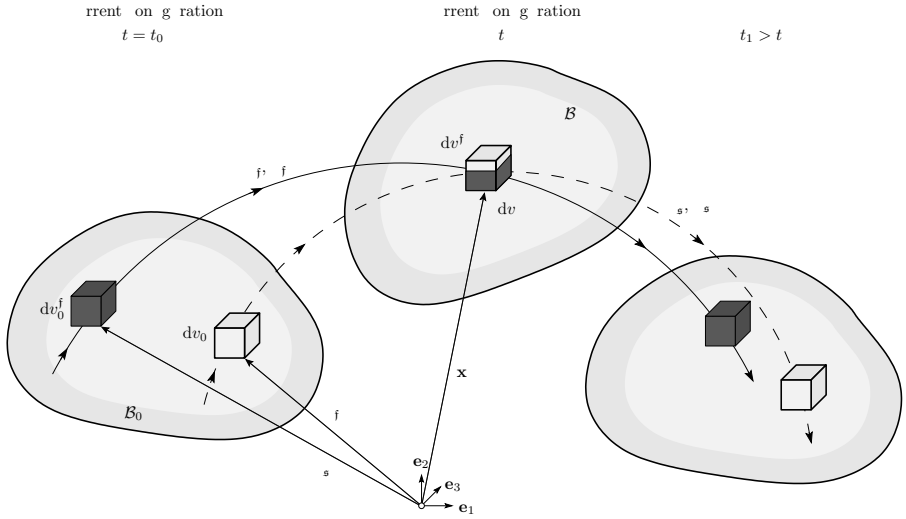


Figure 3: Partial motion function of the fluid and the solid constituent in a poro-elastic medium.

of the solid skeleton $e_{\mathfrak{s}}$ and the one of the fluid phase $e_{\mathfrak{g}}$ could also be interpreted with the previously introduced volume elements[5]

$$e_{\mathfrak{s}} = \frac{\mathrm{d}v - \mathrm{d}v_0}{\mathrm{d}v_0} \quad \text{and} \quad e_{\mathfrak{f}} = \frac{\mathrm{d}v^{\mathfrak{f}} - \mathrm{d}v_0^{\mathfrak{f}}}{\mathrm{d}v_0^{\mathfrak{f}}}. \tag{11}$$

A further (linear) "relative" kinematical measure ζ, denoted as the increment of fluid content and introduced by Biot and Willis [BW57], can be defined as

$$\zeta = \phi_0 \, (e_{\mathfrak{s}} - e_{\mathfrak{f}}). \tag{12}$$

Please note that, besides the definition of the increment of fluid content given in (12), we are able to find alternative definitions in the literature, e.g. [Wan00, RC76], which is leading sometimes to notational confusion.

A simple example: In order to highlight the concept of "deformations" in poro-elasticity, we discuss the following example sketched in Fig. 5. We assume that the

[5]These linear strain measures can be obtained from the linearized map of volume elements classically introduced in continuum mechanics (for mixtures). Here we should have in mind, that the skeleton volume elements are mapped from the initial configuration to the current configuration with the Jacobian of the solid phase, i.e. $\mathrm{d}v = J_{\mathfrak{s}} \, \mathrm{d}v_0$ with $\det \mathbf{F}_{\mathfrak{s}} =: J_{\mathfrak{s}}$ while the bulk volumes of the fluid phase are mapped with the Jacobian of the fluid, i.e. $\mathrm{d}v^{\mathfrak{f}} = J_{\mathfrak{f}} \, \mathrm{d}v_0^{\mathfrak{f}}$ and $\det \mathbf{F}_{\mathfrak{f}} =: J_{\mathfrak{f}}$. The Jacobian $J_{\mathfrak{s}}$ links the *Eulerian* with the *Lagrangian* porosity through $\phi = J_{\mathfrak{s}} \, n^{\mathfrak{f}}$. Further, the linearized Jacobian [SR19] is given by $\mathrm{lin}(J_{\mathfrak{s}}) = e_{\mathfrak{s}} + 1 = \mathrm{div} \, \mathbf{u}_{\mathfrak{s}} + 1$.

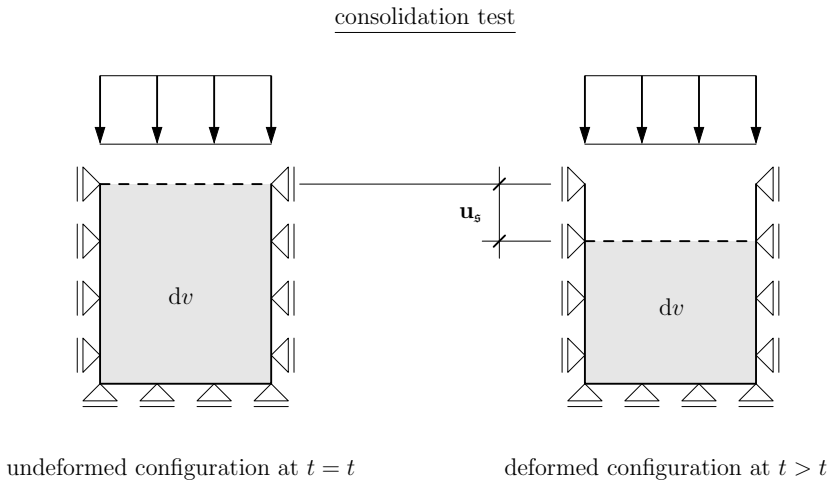


Figure 4: Consolidation test and skeleton deformations

sample is fully undrained and undergoes (elastic) hydrostatic loading, i.e. shear deformations are zero and only volumetrical deformations can be observed. Thus, we could apply the volumetric strain measures introduced in Eqs. 11. Furthermore, we assume that the observed state of deformation (at time t) is homogeneous. Thus, the macroscopically "observable" strains of the skeleton and the volumetrical deformation of the fluid phase (at all fluid volumes in the pores) should be identical. It is shown, that the introduced strain definitions lead exactly to this result:

$$e_{\mathfrak{s}} = \frac{\mathrm{d}v - \mathrm{d}v_0}{\mathrm{d}v_0} = \frac{1/2\,\mathrm{d}v_0 - \mathrm{d}v_0}{\mathrm{d}v_0} = -\frac{1}{2},$$

and

$$e_{\mathfrak{f}} = \frac{\mathrm{d}v^{\mathfrak{f}} - \mathrm{d}v_0^{\mathfrak{f}}}{\mathrm{d}v_0^{\mathfrak{f}}} = \frac{1/3\,\mathrm{d}v - 1/3\,\mathrm{d}v_0}{1/3\,\mathrm{d}v_0} = \frac{\mathrm{d}v - \mathrm{d}v_0}{\mathrm{d}v_0} = -\frac{1}{2}$$

Note that in this experiment we do not allow for any local fluid flow (assumption of homogeneity). Thus the fluid increment is also zero, $\zeta \equiv 0$. □

2.3 Balance of mass

For the solid and the fluid constituents composing the poro-elastic body $\mathcal{B}$, we could formulate the global conservation of mass for the constituents composing the body as

$$\mathcal{M}^{\mathfrak{s}} = \int_{\mathcal{B}} \rho^{\mathfrak{s}}\,\mathrm{d}v = \mathcal{M}_0^{\mathfrak{s}} = \text{const.} \quad \text{and} \quad \mathcal{M}^{\mathfrak{f}} = \int_{\mathcal{B}} \rho^{\mathfrak{f}}\,\mathrm{d}v = \mathcal{M}_0^{\mathfrak{f}} = \text{const.}. \tag{13}$$

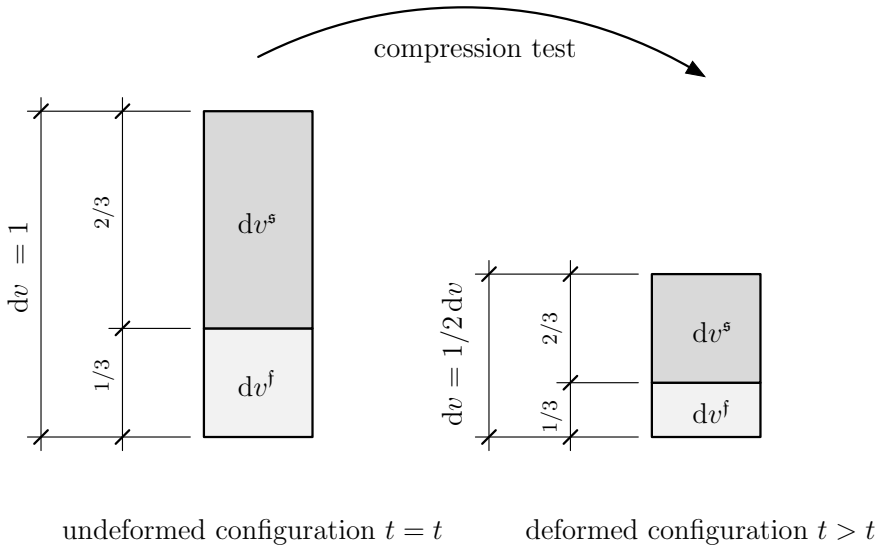


Figure 5: Homogeneous deformation of an undrained (with infinitesimal small volume) poro-elastic sample.

Applying "standard" arguments of continuum mechanics, details can be found in any textbook of continuum mechanics like [Hau00], we finally obtain the local form of the balance of mass of she solid constituent which can be expressed as

$$\partial_t(n^{\mathfrak{s}}\,\rho^{\mathfrak{s}R}) + \operatorname{div}(n^{\mathfrak{s}}\,\rho^{\mathfrak{s}R}\,\mathbf{v}_{\mathfrak{s}}) = 0, \tag{14}$$

or

$$n^{\mathfrak{s}}\,\partial_t(\rho^{\mathfrak{s}R}) + \rho^{\mathfrak{s}R}\,\partial_t(n^{\mathfrak{s}}) + n_{\mathfrak{s}}\,\rho^{\mathfrak{s}R}\,\operatorname{div}\mathbf{v}_{\mathfrak{s}} + \mathbf{v}_{\mathfrak{s}}\cdot\operatorname{grad}(n^{\mathfrak{s}}\,\rho^{\mathfrak{s}R}) = 0. \tag{15}$$

We should have in mind that Eqs. (14) and (15) include non-linear terms (volume fractions, effective densities and velocities are field functions evolving in time). Thus, the partial balance of mass (15) has to be re-formulated (linearized) in order to apply it in linear poro-elasticity. After the formal linarization step of equation (15) around the initial state ($t = t_0$), details can be found in [SR19], we obtain the linearized form of the mass balance (Note again: convective terms are vanishing; the "dot" derivative is identical to the partial time derivative "∂_t")

$$n_0^{\mathfrak{s}}\,\dot{\rho}^{\mathfrak{s}R} + \rho_0^{\mathfrak{s}R}\,\dot{n}^{\mathfrak{s}} + n_0^{\mathfrak{s}}\,\rho_0^{\mathfrak{s}R}\,\dot{e}_{\mathfrak{s}} = 0. \tag{16}$$

The mass balance of the solid (16) could also be re-arranged and expressed with respect to porosity change

$$\dot{\phi} = (1-\phi_0)\,\frac{\dot{\rho}^{\mathfrak{s}R}}{\rho^{\mathfrak{s}R}} + (1-\phi_0)\,\dot{e}_{\mathfrak{s}}. \tag{17}$$

Without showing further details, a similar results could be obtained for the mass balance of the fluid

$$\dot{\phi} = -\phi_0 \, \frac{\dot{\rho}^{\mathfrak{f}R}}{\rho^{\mathfrak{f}R}} - \phi_0 \, \dot{e}_{\mathfrak{f}}. \tag{18}$$

Eqs. (17) and (18) could be time-integrated. These resulting equations are interesting as they are stating that the field variable porosity $\phi(\mathbf{x},\, t)$ can be replaced by two other field variables. In the following sections, after introducing constitutive equations, we will come back to these equations and derive an expression for porosity as a function of volumetric deformation of the solid and the fluid constituent [SR19]

$$\phi = \phi(\rho^{\mathfrak{s}R},\, e_{\mathfrak{s}}) \quad = \quad 2\,\phi_0 - 1 + (1-\phi_0)\,\left(\frac{\rho^{\mathfrak{s}R}}{\rho_0^{\mathfrak{s}R}} + e_{\mathfrak{s}}\right), \tag{19}$$

$$\phi = \phi(\rho^{\mathfrak{f}R},\, e_{\mathfrak{f}}) \quad = \quad 2\,\phi_0 - \phi_0\,\left(\frac{\rho^{\mathfrak{f}R}}{\rho_0^{\mathfrak{f}R}} + e_{\mathfrak{f}}\right). \tag{20}$$

2.4 Balance of momentum

The second set of balance equations which have to be discussed in linear poro-elasticity is the set of balances of (linear) momentum. We restrict our discussion to an absolut minimum and refer to excellent contributions [dB05, Ehl02b, SH09, Cou10] for detailed derivations. One important result from the partial balances of moment of momentum is the symmetry of Cauchy's stress tensor of the solid phase $\boldsymbol{\sigma}^{\mathfrak{s}} \equiv \boldsymbol{\sigma}^{\mathfrak{s},T}$ (or in index notation for the tensor components: $\sigma^{\mathfrak{s}}_{ij} = \sigma^{\mathfrak{s}}_{ji}$). The symmetry condition is also valid for the total stress tensor of the mixture, i.e. $\boldsymbol{\sigma} \equiv \boldsymbol{\sigma}^T$. For the further discussion, it is convenient to split the (partial) stress tensor of the solid phase additively into a volumetric and a deviatoric part

$$\boldsymbol{\sigma}^{\mathfrak{s}} = \mathrm{vol}(\boldsymbol{\sigma}^{\mathfrak{s}}) + \mathrm{dev}(\boldsymbol{\sigma}^{\mathfrak{s}}) = \frac{1}{3}\,\mathrm{tr}(\boldsymbol{\sigma}^{\mathfrak{s}})\,\mathbf{I} + \mathrm{dev}(\boldsymbol{\sigma}^{\mathfrak{s}}) := s^{\mathfrak{s}}\,\mathbf{I} + \boldsymbol{\tau}^{\mathfrak{s}}. \tag{21}$$

Note that the shear stress term of the fluid constituent is assumed to be zero $\boldsymbol{\tau}^{\mathfrak{f}} = \mathbf{0}$ (in classical linear poro-elasticity viscous shear stresses of the fluid phases are assumed to vanish [HG87]). Thus, the partial stress tensor of the fluid phase is only composed by the pore pressure p, i.e. $\boldsymbol{\sigma}^{\mathfrak{f}} = s^{\mathfrak{f}}\,\mathbf{I} = -p\,\mathbf{I}$. The total stress tensor $\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\mathfrak{s}} + \boldsymbol{\sigma}^{\mathfrak{f}}$ of the mixture is split into the total mean stress $\sigma^M = s^{\mathfrak{s}} + s^{\mathfrak{f}}$ and the shear stresses of the solid phase

$$\boldsymbol{\sigma} = (s^{\mathfrak{s}} + s^{\mathfrak{f}})\,\mathbf{I} + \boldsymbol{\tau}^{\mathfrak{s}} =: \sigma^M\,\mathbf{I} + \boldsymbol{\tau}. \tag{22}$$

The global balance of momentum for the solid constituent for linear poro-elasticity states that the momentum $\boldsymbol{\mathcal{J}}^{\mathfrak{s}}$ is changed by the sum of the body forces $\boldsymbol{\mathcal{F}}^{\mathfrak{s}}_{\mathcal{B}}$, contact forces $\boldsymbol{\mathcal{F}}^{\mathfrak{s}}_{\partial\mathcal{B}}$, and interaction forces $\hat{\boldsymbol{\mathcal{P}}}^{\mathfrak{s}} = -\hat{\boldsymbol{\mathcal{P}}}^{\mathfrak{f}}$ as

$$\frac{\partial}{\partial t}(\boldsymbol{\mathcal{J}}^{\mathfrak{s}}) = \boldsymbol{\mathcal{F}}^{\mathfrak{s}}_{\mathcal{B}} + \boldsymbol{\mathcal{F}}^{\mathfrak{s}}_{\partial\mathcal{B}} + \hat{\boldsymbol{\mathcal{P}}}^{\mathfrak{s}}, \tag{23}$$

or in local form

$$\rho^{\mathfrak{s}}\,\mathbf{a}_{\mathfrak{s}} - \operatorname{div}\boldsymbol{\sigma}^{\mathfrak{s}} = \rho^{\mathfrak{s}}\,\mathbf{b} - \hat{\mathbf{p}}^{\mathfrak{f}}. \tag{24}$$

Here, $\rho^{\mathfrak{s}}\,\mathbf{b}$ are the body force densities and $\hat{\mathbf{p}}^{\mathfrak{f}}$ is the local momentum interaction term, i.e. the hydro-mechanical momentum interaction action between the fluid and the solid phase. Later we observe, that we need constitutive equations for the stress tensor $\boldsymbol{\sigma}^{\mathfrak{s}}$ and the momentum interaction $\hat{\mathbf{p}}^{\mathfrak{f}}$ in order to close the set of equations. Additionally, we obtain the local form of the balance of momentum for the fluid as

$$\rho^{\mathfrak{f}}\,\mathbf{a}_{\mathfrak{f}} + \operatorname{div}(\phi\,p\,\mathbf{I}) = \rho^{\mathfrak{f}}\,\mathbf{b} + \hat{\mathbf{p}}^{\mathfrak{f}} \tag{25}$$

2.5 Constitutive equations

The boundary value problem of linear poro-elasticity can be formulated when the model is well-posed. Thus we need as many equations as we have unknowns. This is achieved by formulating constitutive equations for the equilibrium and the non-equilibrium case. Equilibrium in this sense is defined as the case when the fluid is at rest, i.e. when the relative velocities $\mathbf{w}_{\mathfrak{f}}$ (and pore pressure gradients) are vanishing. In the non-equilibrium case, a non-zero viscous momentum interaction term exists $\hat{\mathbf{p}}^{\mathfrak{f}}_{neq} \neq \mathbf{0}$. This state is characterized by the fact that the porous medium is may be consolidated or pressure diffusion effects occur. Energy is dissipated by the viscous momentum interaction.

In this section, we do not discuss a thermodynamically rigorous derivation of the equilibrium and non-equilibrium contributions of the constitutive equations. To do so, the structure of constitutive equations have to be obtained by an evaluation procedure of the entropy inequality following the formalism proposed by Coleman-Noll or Liu-Müller, cf. details and examples in [Ehl02b, SH09]. Here, the results of a formal evaluation procedure is shortly summarized. Specific details related to poro-elasticity can be found in [RS15, SR19].

2.5.1 Equilibrium: fluid phase

In linear poro-elasticity the pore fluid is assumed to be a linear barotropic fluid. Thus, for every material point $\mathcal{P}$, the pore pressure $p(\mathbf{x},\,t)$ is a linear function of the effective fluid density $\rho^{\mathfrak{f}R}(\mathbf{x},\,t)$

$$p \propto \rho^{\mathfrak{f}R} \qquad \text{or} \qquad p = K^{\mathfrak{f}}\left(\frac{\rho^{\mathfrak{f}R}}{\rho_0^{\mathfrak{f}R}} - 1\right). \tag{26}$$

Note that $K^{\mathfrak{f}}$ is the bulk modulus of the fluid.

2.5.2 Equilibrium: solid phase

In the equilibrium case, a linear poro-elastic material could be understood as a simple elastic composite material. The solid skeleton has a shear and a volumetric stiffness, respectively while the fluid phase has only a volumetric stiffness. Like for standard elastic materials, the shear stiffness of the skeleton is taken into account by the shear modulus G. In Eq. 26 we have allready introduced the bulk modulus of the fluid which takes into account the volumetric stiffness of the fluid. What remains is a discussion of the volumetric stiffness of the solid constituent. Let us assume for the beginning that the material composing the solid skeleton is incompressible $\rho^{\mathfrak{s}R} = \rho_0^{\mathfrak{s}R}$ and, additionally, the pore space is "empty" (or under vacuum). If we would conduct a volumtric deformation experiment for that case, we could measure the skeleton stiffness. The related bulk modulus will be introduced as $K^{\mathfrak{s}}$. Besides the assumption of an incompressible pore fluid expressed by $\rho^{\mathfrak{f}R} = \rho_0^{\mathfrak{f}R}$ exactly this incompressibility constraint (or assumption) was made in the consolidation theory of Terzaghi [Ter43]. Biot [Bio41] neither assumes an incompressible pore fluid nor an incompressible material which is composing the porous skeleton. Thus, he introduced one additional bulk modulus $K^{\mathfrak{s}}$ for the solid material composing the skeleton (it could be understood as the bulk modulus of single particles in case of porous granular media). These assumptions are the basic differences between linear (quasi-static) poro-elastic theories of Biot and Terzaghi.

In the following we sketch the basic modelling ideas of Biot's linear poro-elasticity and show the special case of Terzaghi. From the thermodynamical evaluation (not shown) of the entropy inequality, one knows that the total deviatoric and volumetric stresses and the pore pressure could be obtained from thermodynamical potentials. Here, we use the strain energy function $W = W(\boldsymbol{\gamma}_{\mathfrak{s}},\, e_{\mathfrak{s}},\, \zeta)$ which is a function of the (chosen set of) governing kinematical process variables as the thermodynamical potential [Sme92, Wan00, Che16, SR19]. Thus we get

$$\boldsymbol{\tau} = \frac{\partial W}{\partial \boldsymbol{\gamma}_{\mathfrak{s}}}, \qquad \sigma^M = \frac{\partial W}{\partial e_{\mathfrak{s}}}, \qquad p = \frac{\partial W}{\partial \zeta}. \tag{27}$$

It should be noted that the chosen set of process variables here is only one possible choice out of a many other possible ones. Other combinations of process variables lead to similar constitutive relations. However, the inherent material parameters will become different to the one introduced in Eq. (28). Constitutively, these dynamic variables are well described if the strain energy function is specified. The most general formulation of a strain energy function leading to a (material) linear model is a quadratic function in its arguments. This yields

$$W = \mathrm{dev}(W) + \mathrm{vol}(W) = G\,\boldsymbol{\gamma}_{\mathfrak{s}} : \boldsymbol{\gamma}_{\mathfrak{s}} + \frac{1}{2} K_u\, e_{\mathfrak{s}}^2 - \alpha\, M\, e_{\mathfrak{s}}\, \zeta + \frac{1}{2} M\, \zeta^2. \tag{28}$$

We observe that linear poro-elasticity is governed by four elastic parameters here introduced as $\mathcal{M} = (G,\, K_u,\, \alpha,\, M)$. By means of "Gedankenexperimente", these parameters can be expressed as functions of other elastic parameters, e.g. the experimentally observable and well defined parameters $\tilde{\mathcal{M}} = (G,\, K,\, K^{\mathfrak{s}},\, K^{\mathfrak{f}})$. Such relations

can be often found in the literature e.g. [Wan00, MMD09, Che16, SR19] as a starting point. For the (scalar) volumetric quantities, the "stress-strain" relations can be summarized in a symmetric matrix-like notation

$$\begin{bmatrix} \sigma^M \\ p \end{bmatrix} = \begin{bmatrix} K_u & -\alpha\, M \\ -\alpha\, M & M \end{bmatrix} \begin{bmatrix} e_{\mathfrak{s}} \\ \zeta \end{bmatrix}, \tag{29}$$

or

$$\begin{bmatrix} \sigma^M \\ \zeta \end{bmatrix} = \begin{bmatrix} K_u - \alpha\, M & -\alpha \\ \alpha & 1/M \end{bmatrix} \begin{bmatrix} e_{\mathfrak{s}} \\ p \end{bmatrix}. \tag{30}$$

It should be noted that an effective stress principle can be observed in Eq. (30)

$$\sigma^M = (K_u - \alpha\, M)\, e_{\mathfrak{s}} - \alpha\, p =: \sigma_E^{M,\mathfrak{s}} - \alpha\, p. \tag{31}$$

This is an important concept of linear poro-elasticity stating that a weighted balance of mean stress and fluid pressure is loading the solid skeleton and causes volumtric deformation. The weighting factor is the Biot-Willis parameter with $0 \leq \alpha \leq 1$. Note that Eq. (31) includes Terzaghi's effective stress principle [Ter43] as a special case for $\alpha \equiv 1$. Thus, in Terzaghi's consolidation poro-elastic theory the weighting factor can not be "adjusted" to the volumetric stiffness of the solid constituent. Additionally, we have introduced here the undrained bulk modulus K_u which is also denoted as the Gassmann modulus [Gas51]. The Gassmann modulus could be obtained by a volumetric experiment under undrained conditions

$$K_u = \left. \frac{\partial \sigma^M}{\partial e_{\mathfrak{s}}} \right|_{\zeta=0}. \tag{32}$$

Further, α is the so-called Biot-Willis parameter and obtained

$$\alpha = \left. \frac{\partial \zeta}{\partial e_{\mathfrak{s}}} \right|_{p=const.}. \tag{33}$$

The storage modulus M is the inverse of the specific storage capacity $s_{e_{\mathfrak{s}}}$ under constant volumetric deformation of the solid skeleton and can be obtained via

$$\frac{1}{M} = s_{e_{\mathfrak{s}}} = \left. \frac{\partial \zeta}{\partial p} \right|_{e_{\mathfrak{s}}=const.}. \tag{34}$$

The mentioned relations to the bulk moduli K, $K^{\mathfrak{s}}$, $K^{\mathfrak{f}}$ are given by

$$\alpha = 1 - \frac{K}{K^{\mathfrak{s}}}, \quad K_u = K + \alpha^2\, M, \quad \frac{1}{M} = \frac{\phi_0}{K^{\mathfrak{f}}} + \frac{\alpha - \phi_0}{K^{\mathfrak{s}}}. \tag{35}$$

If we regard Terzaghi's consolidation theory as a special case of Biot's model, we could observe that in Terzaghi incompressible limit given by $K^{\mathfrak{f}} \to \infty$ and $K^{\mathfrak{f}} \to \infty$ these moduli are reducing to

$$\tilde{\alpha} = 1, \quad \tilde{K}_u = \infty, \quad \frac{1}{\tilde{M}} = 0. \tag{36}$$

In Eq. (36) we observe that one of these parameters (here the undrained bulk modulus $\tilde{K}_u$) is not defined while a second one $(1/\tilde{M})$ is zero. Thus, the number of elastic parameters is reduced to only two, which are are the shear modulus G and e.g. the (dry) bulk modulus of the solid skeleton K.

2.5.3 Non-equilibrium

Introducing the partial balance of momentum of the solid and the fluid phase, we have mentioned that we need constitutive relations for the viscous momentum interaction term introduced as $\hat{\mathbf{p}}^{\mathfrak{f}} = -\hat{\mathbf{p}}^{\mathfrak{s}}$. This dynamic variable is split into an equilibrium and a non-equilibrium term

$$\hat{\mathbf{p}}^{\mathfrak{f}} = \hat{\mathbf{p}}^{\mathfrak{f}}_{eq} + \hat{\mathbf{p}}^{\mathfrak{f}}_{neq}, \tag{37}$$

with

$$\hat{\mathbf{p}}^{\mathfrak{f}}_{eq} = p\ \mathrm{grad}\,\phi \qquad \text{and} \qquad \hat{\mathbf{p}}^{\mathfrak{f}}_{neq} = -\frac{\phi_0^2\,\gamma_0^{\mathfrak{f}R}}{k^{\mathfrak{f}}}\,\mathbf{w}_{\mathfrak{f}} = -\frac{\phi_0^2\,\eta^{\mathfrak{f}R}}{k^{\mathfrak{s}}}\,\mathbf{w}_{\mathfrak{f}}. \tag{38}$$

In Eq. (38) we have introduced the effective weight of the fluid $\gamma_0^{\mathfrak{f}R}$ with unit [N/m^3], the effective dynamic viscosity $\eta^{\mathfrak{f}R}$ with unit [Pa s], the hydraulic conductivity or Darcy permeability $k^{\mathfrak{f}}$ with unit [m/s], and the intrinsic permeability $k^{\mathfrak{f}}$ with units [m^2] or Darcy [D]. After formulating the constitutive relation for the momentum interaction, we could insert this equation (38) into the local balance of momentum of the fluid phase (25) and derive the well-known Darcy relation for the "quasi-static" case (inertia forces $\rho^{\mathfrak{f}}\,\mathbf{a}_{\mathfrak{f}} \equiv \mathbf{0}$)

$$\mathrm{grad}\,p = -\frac{\phi_0\,\gamma_0^{\mathfrak{f}R}}{k^{\mathfrak{f}}}\,\mathbf{w}_{\mathfrak{f}} + \rho^{\mathfrak{f}R}\,\mathbf{b} = -\frac{\phi_0\,\eta^{\mathfrak{f}R}}{k^{\mathfrak{s}}}\,\mathbf{w}_{\mathfrak{f}} + \rho^{\mathfrak{f}R}\,\mathbf{b}. \tag{39}$$

Obviously, Darcy's relation expresses the proportionality between the pore pressure gradient and the seepage velocity. Is has to be remarked, that any extended, e.g. non-linear, relation between pore pressure gradient and seepage velocity has to be formally derived through a re-formulation of the constitutive equation (38) and not of (39). One prominent example is the non-linear Forchheimer relation which could be obtained by adding a quadratic term $\beta\,|\mathbf{w}_{\mathfrak{f}}|\,\mathbf{w}_{\mathfrak{f}}$ to the linear seepage term in the non-equilibrium momentum exchange (38). Obviously, the additional constitutive "material" parameter β needs to be determined by higher Re-number experiments.

2.6 The poro-elastic boundary value problem

The set of linear PDEs in Table 1 are formulated in the primary variables $\{\mathbf{u}_{\mathfrak{s}},\ p\}$ (note that the effective stress $\boldsymbol{\sigma}^{\mathfrak{s}}_E = \boldsymbol{\sigma} + \alpha\,p\,\mathbf{I}$ is a function of the solid strains $\boldsymbol{\varepsilon}_{\mathfrak{s}}$ which are itself calculated from the gradients of $\mathbf{u}_{\mathfrak{s}}$). It consists of the quasi-static form of the balance of momentum of the mixture and the quasi-static balance of momentum of the fluid phase Eq. (25). For a detailed derivation we refer to [RS15, SR19]. The $\{\mathbf{u}_{\mathfrak{s}},\ p\}$-formulation can be efficiently solved by numerical techniques like the Finite Element

Biot's model	**Terzaghi's model**
equations in the domain, i.e., $\forall\, \mathbf{x} \in \mathcal{B}$	
$-\operatorname{div}(\boldsymbol{\sigma}^{\mathfrak{s}}_E - \alpha\, p\, \mathbf{I}) = \rho\, \mathbf{b}$	$-\operatorname{div}(\boldsymbol{\sigma}^{\mathfrak{s}}_E - p\, \mathbf{I}) = \rho\, \mathbf{b}$
$\frac{\dot{p}}{M} - \frac{k^{\mathfrak{f}}}{\gamma^{\mathfrak{f}R}} \operatorname{div} \operatorname{grad} p + \alpha \operatorname{div} \mathbf{v}_{\mathfrak{s}} = 0$	$-\frac{k^{\mathfrak{f}}}{\gamma^{\mathfrak{f}R}} \operatorname{div} \operatorname{grad} p + \operatorname{div} \mathbf{v}_{\mathfrak{s}} = 0$
boundary conditions, i.e. $\forall\, \mathbf{x} \in \partial\mathcal{B}$	
$\mathbf{u}_{\mathfrak{s}} = \bar{\mathbf{u}}_{\mathfrak{s}}$ on $\Gamma^{\mathfrak{s}}_D$	
$p = \bar{p}$ on $\Gamma^{\mathfrak{f}}_D$	
$\boldsymbol{\sigma} \cdot \mathbf{n} = \bar{\mathbf{t}}$ on $\Gamma^{\mathfrak{s}}_N$	
$\mathbf{w}_{\mathfrak{f}} \cdot \mathbf{n} = \bar{w}_{\mathfrak{f}}$ on $\Gamma^{\mathfrak{f}}_N$	

Table 1: The set of governing Partial Differential Equations (PDEs) of consolidation for Biot's [Bio41] and Terzaghi's [Ter43] formulation of poro-elasticity, cf. [RS15]. We have not inserted the stress tensor $\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\mathfrak{s}}_E - \alpha\, p\, \mathbf{I} = 2\, G\, \boldsymbol{\gamma}_{\mathfrak{s}} + (K_u - \alpha\, M)\, e_{\mathfrak{s}}\, \mathbf{I} - \alpha\, p\, \mathbf{I}$.

Method (FEM). Therefore, the strong form of the equations in Table 1 are multiplied with a test function and integrated in the domain $\mathcal{B}$ (and integration by parts) in the sense of standard Bubnov-Galerkin approaches. In order to fulfill necessary mathematical conditions for the solution of the mixed set of equations (the LBB condition), special discrete function spaces have to be chosen for the shape/test functions. The so-called Taylor-Hood element [Bra97] with quadratic shape functions for the displacements $\mathbf{u}_{\mathfrak{s},h}$ and linear shape functions for the pressure p_h is one possible choice out of many others more sophisticated ones. Further details about Finite Elements for poro-elastic problems could be found in [ZCP+99].

3 Acoustic waves in linear poro-elasticity

In two seminal contributions by Maurice Biot [Bio56a, Bio56b], linear poro-elasticity was extended towards acoustic wave propagation. In contrast to quasi-static linear poro-elasticity the additional contribution takes into account the effect of inertia forces $\rho^{\mathfrak{s}}\, \mathbf{a}_{\mathfrak{s}}$ and $\rho^{\mathfrak{f}}\, \mathbf{a}_{\mathfrak{f}}$ in the partial balances of linear momentum, cf. Eqs. (24) and (25). One of the major physical effects which are proposed by Biot's set of coupled PDEs is the existence of two dispersive, i.e. frequency-dependent compressional waves (also called longitudinal- or P-waves) and one frequency-dependent shear wave (also denoted as transversal- or S-wave). Within an artificial porous sample made out of sin-

tered glass beads, Plona discovered the existence of the 2nd P-wave (the slow wave or Biot's wave) at ultrasound frequencies and, therefore, experimentally validated Biot's theoretical proposal around 25 years later [Plo80]. Further, Biot [Bio56a, Bio56b] extended the contribution of inertia forces in order to include the effect of tortuosity (or "added mass") with densities ρ_{ij}, cf. Eqs. (43). Additionally, Biot was distinguishing physical effects occuring in the "lower" and the "higher" frequency domain (he even split his 1956 papers in a low and a high frequency part). In the lower frequency domain, the velocity profile of the fluid constituent within the pores is assumed to be parabolic, i.e. it is dominated by viscous momentum interaction (a "viscous boundary layer"). Above a transition frequency (in the higher frequency domain), the velocity profile of the fluid is more and more dominated by inertia forces and deviates from a parabolic shape. Biot took that effect into account making the momentum interaction frequency dependent. According to the introduced (non-equilibrium part of the) momentum interaction in Eq. (38), this could be achieved by making the coefficient (e.g. the Darcy permeability $k^{\mathfrak{f}}$ or the intrinsic permeability $k^{\mathfrak{s}}$) frequency dependent.

3.1 Governing set of equations

In order to derive the governing equations for acoustic waves in poro-elastic media, we first take into account the inertia forces (acceleration terms) in the partial balances of momentum (24) and (25). As a result, which is in contrast to the quasi-static field equations of poro-elasticity, the governing set of PDEs is now given by the partial balances of momentum. Inerting the derived constitutive equations in Eqs. (24) and (25) we observe, that both equations are now becoming wave equations (with 2nd order derivatives in time plus 2nd order derivatives in space) plus a diffusive/dispersive part which is taking into account the viscous momentum intercation between the fluid and the solid phase. For the mathematical formulation of the coupled set of PDEs, we prefer a compact notation which was introduced in [Kel97, Sme05, Ste10]. The drawback of that compact notation is that we have to introduce another set of material parameters which could of course be related to the material parameters introduced in the previous section. An extensive discussion about waves in poro-elastic media, including effects and detailed derivations, is given in excellent textbooks of [Sto89, BCZ87, AA09, Car01, MMD09] and others. Here, we present mainly the resulting equations here and skip (important) technical details:

$$\begin{aligned} \rho_{11}\,\ddot{\mathbf{u}}_{\mathfrak{s}} + \rho_{12}\,\ddot{\mathbf{u}}_{\mathfrak{f}} + b_0\,F(\dot{\mathbf{u}}_{\mathfrak{s}} - \dot{\mathbf{u}}_{\mathfrak{f}}) &= N\ \mathrm{div}\ \mathrm{grad}\,\mathbf{u}_{\mathfrak{s}} + (A+N)\,\mathrm{grad}\ \mathrm{div}\,\mathbf{u}_{\mathfrak{s}} \\ &\quad + Q\ \mathrm{grad}\ \mathrm{div}\,\mathbf{u}_{\mathfrak{f}}, \\ \rho_{12}\,\ddot{\mathbf{u}}_{\mathfrak{s}} + \rho_{22}\,\ddot{\mathbf{u}}_{\mathfrak{f}} - b_0\,F(\dot{\mathbf{u}}_{\mathfrak{s}} - \dot{\mathbf{u}}_{\mathfrak{f}}) &= Q\ \mathrm{grad}\ \mathrm{div}\,\mathbf{u}_{\mathfrak{s}} + R\ \mathrm{grad}\ \mathrm{div}\,\mathbf{u}_{\mathfrak{f}}. \end{aligned} \tag{40}$$

Various "new" coefficients have been introduced here and need to be discussed in further details. We will be able to observe that all the coefficients could be linked to material parameters introduced in the previous section or phenomena which are newly introduced. First, a viscous damping factor $b_0 = \eta^{\mathfrak{f}R}\,\phi_0^2/k^{\mathfrak{s}}$ have been introduced. An

additional frequency-dependent pre-factor $F(\omega)$ appears in the term of the viscous momentum interaction

$$F = \sqrt{1 + \frac{1}{2}\, i\, M\, \omega/\omega_r}. \tag{41}$$

As allready mentioned, that term takes into account the macroscopical effect of the frequency dependency of the shape of the velocity profile of the fluid, i.e. the frequency dependent fluid-solid coupling. Here, we use the original notation introduced in [JKD87] with the shape factor M, usually taken as $M \approx 1$, and the critical (or roll-over / transition) frequency ω_{crit}.

$$\omega_{crit} = \frac{\eta^{\mathfrak{f}R}\, \phi_0}{\alpha_\infty\, \rho^{\mathfrak{f}R}\, k^{\mathfrak{s}}}. \tag{42}$$

We observe that the critical frequency ω_{crit} (physical unit $[1/s]$) is a material property of the porous medium and separates the low- from the high-frequency domain. In the lower frequency domain, the coupling between the fluid and the solid phase is purely viscous and dominated by the viscous skin depth. This results for ideal (e.g. cylindrical) pores in a parabolic velocity profile. In contrast, inertia becomes more and more dominant for higher frequencies above the critical one. The shape of the velocity profile become "more-and-more" non-parabolic [KS12]. As a consequence the introduction of effective material properties, like a dynamic or frequency-dependent intrinsic permeability was proposed by [JKD87]. Here, the tortuosity α_∞ is expressed by its extreme high frequency limit ($\omega \mapsto \infty$), cf. Kelder [Kel97]. An additional physical consequence of the tortuous pore structure are added mass effects which are captured by the introduced densities

$$\rho_{11} = (1-\phi_0)\, \rho^{\mathfrak{s}R} - \rho_{12}, \tag{43}$$

$$\rho_{12} = (1-\alpha_\infty)\, \phi_0\, \rho^{\mathfrak{f}R}, \tag{44}$$

$$\rho_{22} = \alpha_\infty\, \phi_0\, \rho^{\mathfrak{f}R}. \tag{45}$$

According to Berryman [Ber80], the morphological and material-dependent tortuosity can be geometrically estimated as

$$\alpha_\infty = 1 - r(1 - 1/\phi_0), \qquad \text{with} \qquad r = 1/2. \tag{46}$$

The parameters A, Q, R, and P introduced in the field equations (40) are related to the parameter of "quasi-static" poro-elasticity introduced in the previous section [BW57, Gas51]:

$$N = G, \tag{47}$$

$$A = K - 2\,N/3 + K^{\mathfrak{f}}(1 - \phi_0 - K/K^{\mathfrak{s}})^2/\phi_0^R, \tag{48}$$

$$Q = \phi_0\, K^{\mathfrak{f}}(1 - \phi_0 - K/K^{\mathfrak{s}})/\phi_0^R, \tag{49}$$

$$R = \phi_0^2\, K^{\mathfrak{f}}/\phi_0^R, \tag{50}$$

$$P = A + 2\,N. \tag{51}$$

In Eqs. (48)-(50) we have additionally introduced the so-called effective porosity ϕ_0^R with $\phi_0^R = \phi_0 + K^{\mathfrak{f}}/K^{\mathfrak{s}}(1 - \phi_0 - K/K^{\mathfrak{s}})$. Having discussed modelling aspects and

the notation of the field equations (40), we could discuss in the next step the frequency dependent properties of these equations.

3.2 Dispersion relations

3.2.1 Shear waves

Using a classical harmonic ansatz for the displacement fields (technical details can be found in [Ste10]), we end up in the dispersion relation for transversal wave modes, i.e. shear waves or S-waves. Here, we present the dispersion relation for the wave number squared ($\xi = k^2$), where $k(\omega)$ is the complex wave number

$$\xi = \frac{\tilde{\rho}_{11}\,\tilde{\rho}_{22} - \tilde{\rho}_{12}^2}{N\,\tilde{\rho}_{22}}. \tag{52}$$

For notational purposes, it is convinient to introduce the frequency-dependent densities

$$\tilde{\rho}_{12} = \rho_{12} + i\,b_0\,F/\omega, \tag{53}$$

$$\tilde{\rho}_{11} = \rho_{11} - i\,b_0\,F/\omega, \tag{54}$$

$$\tilde{\rho}_{22} = \rho_{22} - i\,b_0\,F/\omega. \tag{55}$$

3.2.2 Compressional waves

The dispersion relation for compressional waves, i.e. P-waves, have two physical solutions. Using again the wave number squared introduced as ξ, we get

$$\xi_{1,2} = \frac{\Delta \pm \sqrt{\Delta^2 - 4\,(P\,R - Q^2)(\tilde{\rho}_{11}\,\tilde{\rho}_{22} - \tilde{\rho}_{12}\,\tilde{\rho}_{12})}}{2\,(P\,R - Q^2)}, \tag{56}$$

where we have introduced $\Delta = P\,\tilde{\rho}_{22} + R\,\tilde{\rho}_{11} - Q\,\tilde{\rho}_{12}$.

For a typical reservoir sandstone (Berea, [Wan00, Table C.1]), we plot now the frequency-dependent phase velocities as well as the intrinsic attenuation captured by the so-called quality factor. The inverse quality factor is defined as $1/Q = 2\,|\mathrm{Im}(k)/\mathrm{Re}(k)|$. The phase velocity is calculated from the real part of the complex wave number $c = 1/\mathrm{Re}(k)$. For the calculation of the disprsive wave properties we use the material data listed in Table 2 supplemented by the effective density of the solid skeleton $\rho_0^{fR} = 2650$ kg/m^3 and the effective density of the fluid phase $\rho_0^{fR} = 1000$ kg/m^3. We did not take into account tortuosity effects ($\alpha_\infty = 1$) and did not apply any frequency correction ($F = 1$). In Fig. 6 we show the typical behaviour of shear waves in a porous medium. The phase velocities are only "slightly" dispersive. Thus, deviations from the low-frequency limit of the shear wave velocities given by $v_s = \sqrt{N/\rho}$ are small. Also the predicted intrinsic attenuation $1/Q$ of shear wave is small. The

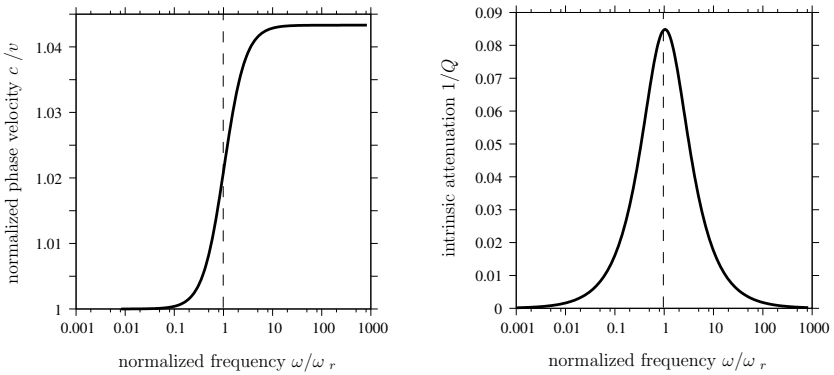


Figure 6: Left: Phase velocity c_s of shear wave as function of frequency. Right: Intrinsic attenuation of shear wave $1/Q$ as function of frequency. Phase velocity is normalized by the low-frequency limit of $v_s = \sqrt{N/\rho}$ with $\rho = \rho^{\mathfrak{s}} - \rho^{\mathfrak{f}}$. The frequency domain is normalized with the critical (roll-over) frequency ω_{crit}, cf. Eq. (42).

maximum attenuation is observed at the transition from the low to the high frequency domain predicted by ω_{crit}. Besides the low frequency limit of shear waves we also observe that a high frequency limit is predicted by Biot's equations.
We do not show phase velocitiy and attenuation curves for the 1st and the 2nd P-wave. Especially the 2nd P-wave is strongly frequency-dependent. In the low frequency regime $\omega < \omega_{crit}$, it is a diffusive 'wave" mode while for $\omega > \omega_{crit}$ travelling waves could be observed. Not that even in the low frequency range, the 2nd P-wave is highly attenuated. For $\omega \to \infty$, the phase velocity is reaching again a high frequency limit.

3.3 Discussion

Discussing the intrinsic attenuation depicted with the inverse quality factor $1/Q$ for S- and P-waves, we observe that attenuation in porous media predicted by a (homogeneous) poro-elastic model is small for seismic frequencies of around $\omega \approx 100$ 1/s or lower, cf. Fig. 7. This is in contrast to field observations where a significant higher amount of attenuation is observed [MMD09]. What is obviously neglected in the present discussion of attenuation in poro-elastic media is the allways inherent heterogeneity of geological structures like faults, sedimentary layers etc. If a wave is hitting the interface of two layers with different material properties, part of the wave is reflected and part of the wave is transmitted. For highly dispersive wave modes, such as the 2nd P-wave, this leads to a significant internal loss of energy.
To be more precise, the inherent attenuation mechanism in poro-elasticity is most pronounced around the critical frequency which is in the MHz-regime, cf. the comparison in Fig. 7 or the models discussed in the Rock Physics Handbook [MMD09]. Thus, in order to model attenuation in porous media in a more realistic way, we have to take

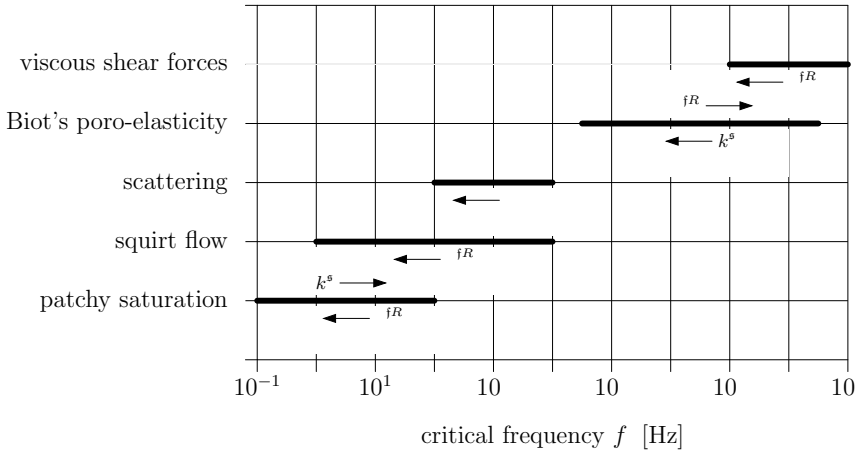


Figure 7: Attenuation/dissipation in fluid-saturated granular rocks, modified after [MMD09]. The critical frequency is evaulated based on typical material properties (e.g. $\{\eta^{\mathfrak{f}R}, k^{\mathfrak{s}}\}$) of reservoir rocks. Have in mind that these critical parameters in general depend also on characteristic length scales (e.g. L).

heterogeneities in our models or simulations implicitely or explicitely into account. In the last years, various scientific approaches have been published for the investigation of that topic, especially in a geopyhsical context. In the following section, we are aiming to give a short insight into some approaches concentrating on effects on the seismic frequency range where the 2nd wave mode is diffusive and inertia effects play a minor role regarding effective properties of the porous media.

4 The role of heterogeneities - observations and consequences

4.1 Consolidation again

Let us first discuss a configuration which could be understood as a prototype problem of a heterogeneous poro-elastic medium. We could call it a setup of internal or local "consolidation". In the geopyhsical context such problems are also denoted as patchy-saturation, interlayer flow, or meso-scopic loss depending on the choice of inherent material properties and configurations. In contrast to Terzaghi's [Ter43] original formulation of consolidation, we apply the quasi-static poro-elastic equations of

Biot [Bio41] introduced in the previous sections. Inertia terms are neglected. Note that also for Biots's [Bio41] formulation an analytical solution exists for the homogenous consolidation test which could be used to validate numerical (FEM) implementations, e.g. [Ver10]. The Boundary Value Problem (BVP) is depicted in Fig. 8. On the left and in the middle, Terzaghi's classical consolidation problem is sketched. In that classical 1-dim consolidation formulation, the poro-elastic domain is assumed to be homogeneous. Thus, at all material points $\mathcal{P}(\mathbf{x},\, t)$, the introduced material parameters used in linear poro-elasticity are identical. Through the choice of the boundary conditions (the sample is loaded with a total surface force ($\mathbf{t} = \bar{\mathbf{t}}$) acting on the solid skeleton and the fluid phase) including a drained boundary ($\bar{p} = 0$) condition, the pore fluid is "squeezed out" of the sample. Viscous dissipation ("intrinsic attenuation" in terms of acoustics) occurs. It is caused by the viscous momentum interaction between the pore fluid and the solid skeleton. In the right sketch of Fig. 8, we slightly modified the consolidation problem. Compared to the orgiginal BVP (Fig. 8, middle), we have chosen undrained boundary conditions ($\bar{w}_{\mathfrak{f}} = 0$) at the top and the bottom of the sample. Further, we assumed that the upper half of the sample has material properties (A) while the lower part has properties (B). Obviously, this makes the problem heterogeneous. If this sample is loaded, the deformation field, and thus also the pressure field according to the effective stress principle, is heterogeneous for $t < t_\infty$. The pressure gradients lead to fluid flow and, therefore again, dissipation ("intrinsic attenuation") occurs. Not that this would not happen if the sample would be homogeneous (compare the discussed homogeneous "Gedankenexperiment" in the previous section). In Fig. 12 we show a set of numerical solutions for the heterogeneous boundary value problem of "internal consoldation". Basically, we depict the numerical solution for the solid displacement components in vertical direction $u \equiv u_{\mathfrak{s}}$, the pore pressure p, the fluxes $q = \phi_0\, w_{\mathfrak{f}}$ (with vertical component of the seepage velocity $w_{\mathfrak{f}}$), the total and the effective stress components in vertical direction (σ_y and σ_y^E) and an effective (locally determined) pseudo Skempton coefficient $B^\star$. It could be observed that shortly after applying the vertical load (which is applied in one single step as a Heaviside function) at time t_0, the pore pressure gradients have a maximum value at the material interface. Thus, viscous momentum interaction (and flow) is also maximized. If time is elapsing (e.g. for t_1 and t_2), the pore pressure gradients are spatially smeared out and are decreasing; the fluxes are getting smaller and attenuation is also getting lower. At time t_3 we are close to a new equilibrium state $\hat{\mathbf{p}}_{neq} \approx \mathbf{0}$; the pressure gradients and the fluid fluxes are vanishing. Note that we are able to calculate local material properties which are, due to the evolving response of the material (e.g. $p(\mathbf{x},\, t)$), effective material properties which are also time dependent quantities. Often, these effective material properties are calculated for a heterogeneous unit cell which is representative for the considered medium. Such unit cell are denoted as Representative Volume Elements or RVEs. Still, the obtainable material properties are effective, i.e. time-dependent quantities which are sometimes transformed to the frequency domain by means of a FFT. A detailed discussion, about geopyhsical applications, configurations and various numerical solutions schmemes could be found e.g. in [QSFS11, QSF$^+$12, QCHS19].

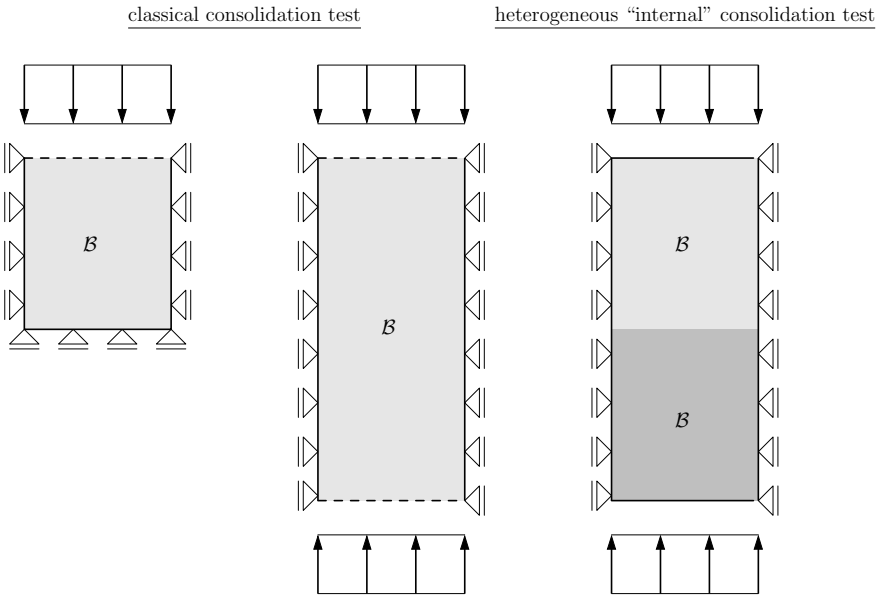


Figure 8: The homogeneous and the heterogeneous (internal) consolidation test.

	K	$K^{\mathfrak{s}}$	$K^{\mathfrak{f}}$	G	ϕ	$\gamma^{\mathfrak{f}R}$	$\eta^{\mathfrak{f}R}$	$k^{\mathfrak{s}}$
$\mathcal{B}^A$	8.0 GPa	36.0 GPa	2.2 GPa	6.0 GPa	0.19	9.81 kN/m^3	1 mPas	160 mD
$\mathcal{B}^B$	1.6 GPa	7.2 GPa	2.2 GPa	1.2 GPa	0.19	9.81 kN/m^3	1 mPas	32 mD

Table 2: Material parameters used in the heterogeneous consolidation problem, cf. Figs. 8 and 12. A is the top/bottom part of the sample; B is the middle part.

4.2 More heterogeneities

It should be noted here, that various types of heterogeneities are found in porous materials. In the previous paragraph, we basically discuss the effect of layered media. Often the pore morphology is inherent heterogeneous two. Micro- or meso-fractures and/or pores exists leading potentially to local pore pressure gradients if the material is mechanically loaded (or if a wave is passing through). Especially a physical-based understanding of the role of fractures on all kind of scales on effective material properties (like phase velocities/attenuation) is of great interest for geomaterials like porous reservoir rocks. We do not discuss any distinct physical problem in detail here but aim to show some fundamental effects caused by fractures.

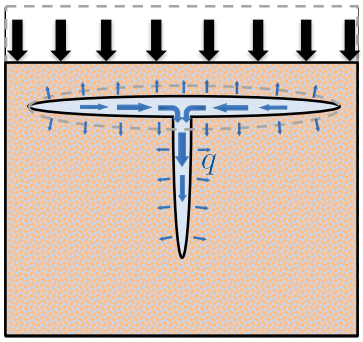


Figure 9: Prototype problem of two intersecting fractures in a porous matrix. Here, the boundary value problem consists of vertically and horizontally undrained boundaries conditions. The problem is vertically loaded with a total force $\bar{\mathbf{t}}$.

4.3 Hydro-mechanics of fractures

In Fig. 9 we show a simple configuration of two intersecting fractures in a porous medium. Intersected fractures often occur in geothermal or petroleum applications, e.g. in crystalline or sedimentary rocks after a hydraulic stimulation process. If the stress state of such a configuration is disturbed, e.g. by a passing P-wave or a mechanical load as depicted in the sketch, pore pressure gradients are invoked. Interestingely, two superimposed physical effects could be observed which are caused by the hydrodynamical interaction of the fluid-filled fractures and the fluid-saturated porous matrix. Due to the deformation of the fracture, pressure gradients in direction of the fractures are evolving. Thus, pressure diffusion in direction of the fractures could be observed, cf. bold blue arrows in Fig. 9. This diffusion process is fast and the characteristic diffusion time of the process is characteried by the length scale and effective hydro-mechanical properties of the system. A typical characteristic frequency could be numerically determined by solving an appropriate quasi-static poro-elastic boundary value problem and transfering the result to frequency domain. In Fig. 10, the blue curve is the characteristic attenuation response of pressure diffusion within the fracture. On top of pressure diffusion within the fracture, we could also observe pressure gradients perpendicular to the fracture-matrix interfaces. These pressure gradient are the driving forces for a slower (smaller frequencies) leak-off process which is sketched with the red curves in Fig. 10. Obviously, in such fractured systems both physical processes occur simultaniously leading to attenuation with two characteristic frequency peaks (dashed cruve), cf. details in [VRS14]. More complex hydro-mechanical coupled scenarios like porous media with inherent fracture networks could be studied in order to investigate the effect of heterogeneous fracture distributions on effective properties in further detail, cf. the 2-dim fracture network shown in Fig. 11. Efficient numerical solutions schemes for such complex fracture networks are still challenging. If

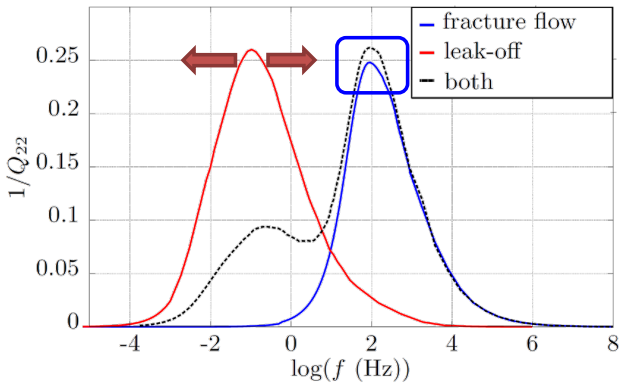


Figure 10: Attenuation in intersecting fractures embedded in a porous matrix, after [VRS14].

Finite Element Methods are applied, one reason for the complexity is meshing of high aspect ratio fractures. Therefore more sophisticated numerical schemes have been developed for such problems. One promising solution strategy is the use of an hybrid dimensional discretization approach. Therein, the 2-dim matrix is discretizized by 2-dim elements while the fractures are embedded 1-dim objects. Even weak-coupling schemes between the 1- and the 2-dim elements can then be applied allowing for very efficient non-conformal FE meshes [SS19]

Acknowledgement

Holger Steeb would like to thank the German Research Foundation (DFG) for supporting the work within projects B05 and C05 of SFB 1313, and within the Cluster of Excellence EXC 2075 “Data-integrated Simulation Science (SimTech)” at the University of Stuttgart.

References

[AA09] J. F. Allard and N. Attala. *Propagation of Sound in Porous Media*. John Wiley & Sons, Ltd, 2009.

[BCZ87] T. Bourbié, O. Coussy, and B. Zinszner. *Acoustics of Porous Media*. Editions Technip, Paris, 1987.

[Ber80] J. G. Berryman. Confirmation of biot’s theory. *Appl. Phys. Lett.*, 37(4):382–384, 1980.

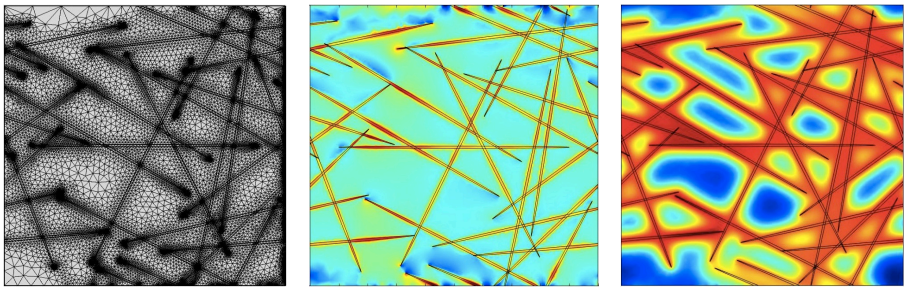


Figure 11: FEM discretization of a complex 2-dim fracture network in a porous matrix. The hydro-mechanical interaction phenomena are shown in the pore pressure distributions for two characteristic times. Middle: Pressure distribution close to the initial state ($t \rightarrow t_0$). Pressure gradients (and thus pressure diffusion) in fracture direction are observed. Right: For $t > t_0$ pressure gradients perpendicular to the fractures (and thus leak-off) are dominating.

[Bio41] M. A. Biot. General theory of three-dimensional consolidation. *J. Appl. Phys.*, 12:155–164, 1941.

[Bio56a] M. A. Biot. Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range. *J. Acoust. Soc. Am.*, 29:168–178, 1956.

[Bio56b] M. A. Biot. Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range. *J. Acoust. Soc. Am.*, 29:179–191, 1956.

[Bra97] D. Braess. *Finite Element Methods*. Cambridge University Press, Cambridge, 1997.

[BW57] M.A. Biot and D. G. Willis. The elastic coefficients of the theory of consolidation. *J. Appl. Mech.*, 24:594–601, 1957.

[Car01] J. M. Carcione. *Wave Fields in Real Media: Wave Propagation in Anisotropic, Anelastic and Porous Media*. Pergamon Press, 2001.

[Che16] A. H.-D. Cheng. *Poroelasticity*. Springer-Verlag, Berlin, 2016.

[Cou95] O. Coussy. *Mechanics of Porous Continua*. John Wiley & Sons, Chichester, 1995.

[Cou04] O. Coussy. *Poromechanics*. John Wiley & Sons, New York, 2004.

[Cou10] O. Coussy. *Mechanics and Physics of Porous Solids*. John Wiley & Sons, West Sussex, UK, 2010.

[CS79] P. A. Cundall and O. D. L. Strack. A discrete numerical model for granular assemblies. *Géotechnique*, 29:47–65, 1979.

[dB00] R. de Boer. *Theory of porous media*. Springer-Verlag, Berlin, 2000.

[dB05] R. de Boer. *Trends in Continuum Mechanics of Porous Media*. Springer-Verlag, Berlin, 2005.

[DC93] E. Detournay and A. H.-D. Cheng. Fundamentals of poroelasticity. In C. Fairhurst, editor, *Comprehensive Rock Engineering: Principles, Practice and Projects, Analysis and Design Method*, volume 2, chapter 5, pages 113–171. Pergamon Press, 1993.

[Ehl02a] W. Ehlers. Foundations of multiphasic and porous materials. In W. Ehlers and J. Bluhm, editors, *Porous Media: Theory, Experiments and Numerical Applications*, pages 3–86. Springer-Verlag, Berlin, 2002.

[Ehl02b] W. Ehlers. Foundations of multiphasic and porous materials. In W. Ehlers and J. Bluhm, editors, *Theory, Experiments and Numerical Applications*. Springer, 2002.

[Gas51] F. Gassmann. Über die Elastizität poröser Medien. *Vierteljahresschrift d. Naturf. Ges. Zürich*, 96:1–23, 1951.

[Hau00] P. Haupt. *Continuum Mechanics and Theory of Materials*. Springer-Verlag, Berlin, 2000.

[HG87] S. M. Hassanizadeh and W. G. Gray. High velocity flow in porous media. *Transport Porous Med.*, 2:521–531, 1987.

[JKD87] D.L. Johnson, J. Koplik, and R. Dashen. Theory of dynamic permeability and tortuosity in fluid-saturated porous-media. *J. Fluid Mech.*, 176:379–402, 1987.

[K$\ddot{9}$1] H.-J. Kümpel. Poroelasticity: parameters reviewed. *Geophys. J. Int.*, 105:783–799, 1991.

[Kel97] O. Kelder. *Frequency-dependent wave propagation in water-saturated porous media*. PhD-thesis, Technical University Delft, NL, 1997.

[KLM$^+$09] M. A. Knackstedt, S. Latham, M. Madadi, A. Sheppard, T. Varslot, and C. Arns. Digital rock physics: 3D imaging of core material and correlations to acoustic and flow properties. *The Leading Edge*, 28:28–33, 2009.

[KS12] P. S. Kurzeja and H. Steeb. About the transition frequency in Biot's theory. *J. Acoust. Soc. Am.*, 131:EL454–EL460, 2012.

[Leh11] F. K. Lehner. A review of the linear theory of anisotropic poroelastic solids. In Y. M. Leroy and F. K. Lehner, editors, *Mechanics of Crustal Rocks*, number 533 in CISM Courses and Lectures, chapter 1, pages 1–41. Springer, Wien, NewYork, 2011.

[Mer16] A. Merxhani. An introduction to linear poroelasticity. preprint on arXiv:1607.04274, 2016.

[MMD09] G. M. Mavko, T. Mukerji, and J. Dvorkin. *The Rock Physics Handbook*. Cambridge University Press, Cambridge, UK, 2009.

[OURS15] M. Osorno, D. Uribe, O.E. Ruiz, and H. Steeb. Finite difference calculations of permeability in large domains in a wide porosity range. *Arch. Appl. Mech.*, 85:1043–1054, 2015.

[Plo80] T. Plona. Observation of a second bulk compressional wave in a porous medium at ultrasonic frequencies. *Appl. Phys. Lett.*, 36:259–261, 1980.

[QCHS19] B. Quintal, E. Caspari, K. Holliger, and H. Steeb. Numerically quantifying energy loss caused by squirt flow. *Geopyhs. Prospect.*, 2019. doi:10.1111/1365-2478.12832.

[QSF+12] B. Quintal, H. Steeb, M. Frehner, S. M. Schmalholz, and E. H. Saenger. Pore fluid effects on S-wave attenuation caused by wave-induced fluid flow. *Geophysics*, 77:L13–L23, 2012.

[QSFS11] B. Quintal, H. Steeb, M. Frehner, and S. M. Schmalholz. Quasi-static finite element modeling of seismic attenuation and dispersion due to wave-induced fluid flow in poroelastic media. *J. Geophys. Res*, 116:B01201, 2011.

[RC76] J. R. Rice and M. P. Cleary. Some basic stress diffusion solutions for fluid saturated elastic porous media with compressible constituents. *Rev. Geophys. Space Phys.*, 14:227–241, 1976.

[Rod89] J. D. Rodrigues. The Noordbergum effect and characterization of aquitards at the Rio Maior mining project. *Ground Water*, 21:200–2007, 1989.

[RS15] J. Renner and H. Steeb. Modeling of fluid transport in geothermal research. In W. Freeden, M.Z. Nashed, and T. Sonar, editors, *Handbook of Geomathematics*, pages 1443–1500. Springer-Verlag Berlin, Heidelberg, 2015.

[SEKS11] E.H. Saenger, F. Enzmann, Y. Keehm, and H. Steeb. Digital rock physics: Effect of fluid viscosity on effective elastic properties. *J. Appl. Geophys.*, 74:236–241, 2011.

[SH09] L. Schneider and K. Hutter. *Solid-Fluid Mixtures of Frictional Materials in Geophysical and Geotechnical Context*. Springer-Verlag, Berlin, 2009.

[Sme92] D. M. J. Smeulders. *On wave propagation in saturated and partially saturated porous media*. 1992. PhD thesis, TU Eindhoven.

[Sme05] D. M. J. Smeulders. Experimental evidence for slow compressional waves. *J. Eng. Meth-ASCE.*, 31:908–917, 2005.

[SR19] H. Steeb and J. Renner. Mechanics of poro-elastic media: A review with emphasis on foundational state variables. *Transport Porous Med.*, 2019. under review.

[SS19] P. Schmidt and H. Steeb. Numerical aspects of hydro-mechanical coupling of fluid-filled fractures using hybrid-dimensional element formulations and non-conformal meshes. *GEM - International Journal on Geomathematics*, 10, 2019. doi:10.1007/s13137-019-0127-5.

[Ste10] H. Steeb. Ultrasound propagation in cancellous bone. *Arch. Rat. Mech. Anal.*, 80:489–502, 2010.

[Sto89] R. D. Stoll. *Sediment Acoustics*. Lecture Notes in Earth Sciences. Springer-Verlag, Berlin, 1989.

[Ter43] K. Terzaghi. *Theoretical Soil Mechanics*. Wiley, New York, 1943.

[Tru57] C. A. Truesdell. Sulle basi della thermomeccanica. *Rend. Lincei*, 22:33–38, 158–166, 1957. Translated into english in: C. Truesdell (Editor), Continuum Mechanics II-IV, Gordon & Breach, New York, 1966.

[Ver10] A. Verruijt. *An Introduction to Soil Dynamics*. Springer-Verlag, Berlin, 2010.

[VRS14] C. Vinci, J. Renner, and H. Steeb. On attenuation of seismic waves associated with flow in fractures. *Geophys. Res. Lett*, 41:7515–7523, 2014.

[Wan00] H. F. Wang. *Theory of Linear Poroelasticity*. Princeton University Press, Princeton & Oxford, 2000.

[YMP+15] X. Yang, Y. Mehmani, W. A. Perkins, A. Pasquali, M. Schönherr a K. Kim, M. Perego, M. L. Parks, N. Trask, M. T. Balhoff, M. C. Richmond, M. Geier, M. Krafczyk, L.-S. Luo, A. M. Tartakovsky, and T. D. Scheibe. Intercomparison of 3D pore-scale flow and solute transport simulation methods. *Adv. Water Resources*, 95:176–189, 2015.

[ZCP+99] O. C. Zienkiewicz, A. H. C. Chan, M. Pastor, B. A. Schrefler, and T. Shiomi. *Computational Geomechanics with Special Reference to Earthquake Engineering*. John Wiley & Sons, Chichester, 1999.

[Zim91] R. Zimmerman. *Compressibility of Sandstones*. Elsevier, Academic Press, 1991.

[ZSK86] R. W. Zimmerman, W. H. Somerton, and M. S. King. Compressibility of porous rocks. *J. Geophys. Res*, 91:765–777, 1986.

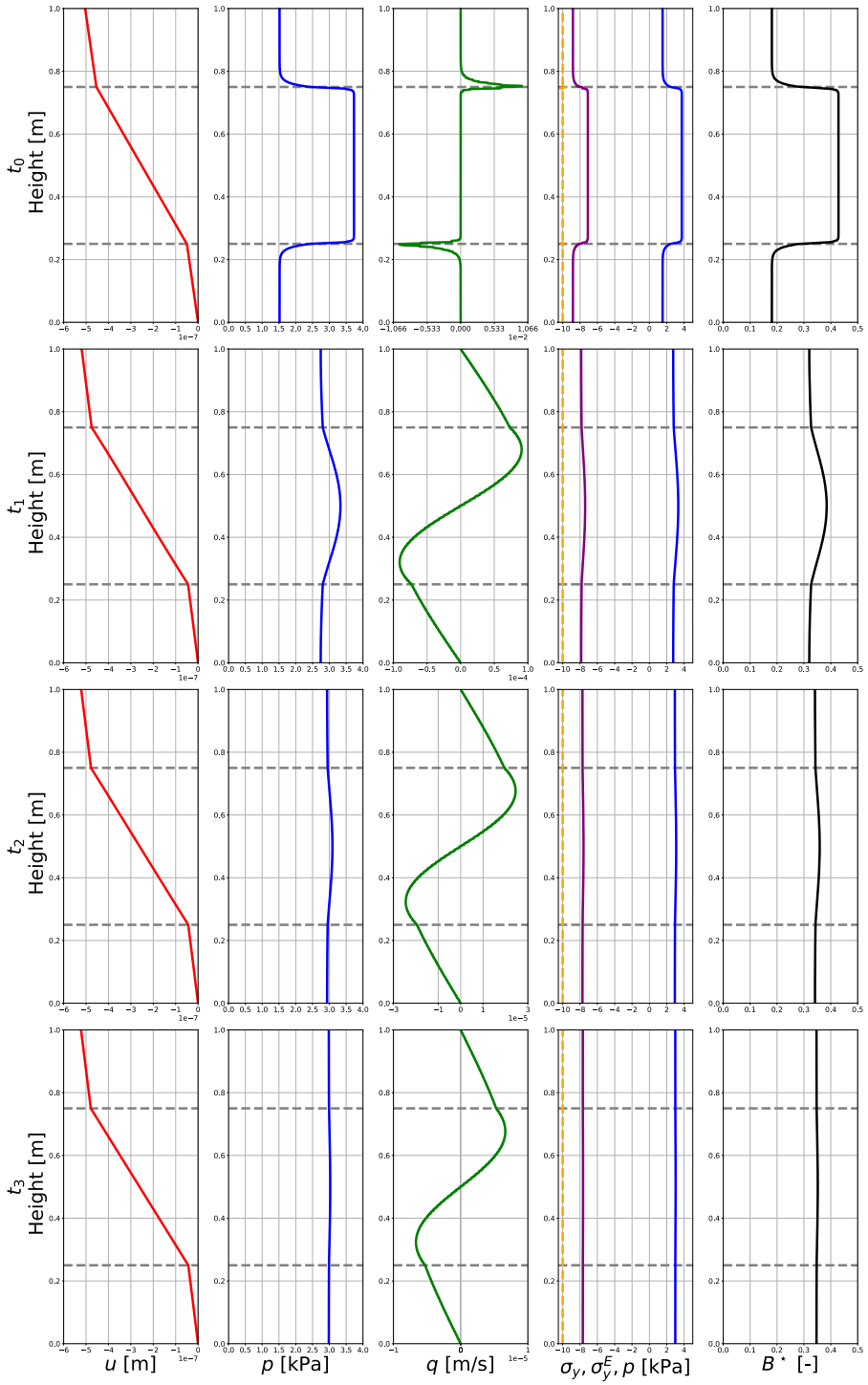


Figure 12: Numerical result of the internal or heterogeneous (internal) consolidation test depicted in Fig. 8. Here we use the material parameters for Berea sandstone [Wan00].

Laboratoire 3SR / Bâtiment Galilée
CS 40700
38 058 Grenoble cedex 9
France

ISBN 978-2-9561359-4-4

Fon: +33 (0) 456 528 621
Fax: +33 (0) 476 827 043
president@alertgeomaterials.eu
http://alertgeomaterials.eu

ALERT Doctoral School 2019

The legacy of Ioannis Vardoulakis to Geomechanics

Editors: J. Sulem, C. Viggiani

ISBN 978-2-9561359-4-4