

Diss. ETH No. 15860

MODELLING APPROACHES TO GEODYNAMIC PROCESSES

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY, ZÜRICH
for the degree of Doctor of Natural Sciences

Presented by
BORIS JOZEF PAUL KAUS
Dipl. Natw. ETH
born 4. December, 1976
from the Netherlands

Accepted on the recommendation of

Prof. Dr. J.-P. Burg	ETH Zürich	examiner
Prof. Dr. Y.Y. Podladchikov	PGP, University of Oslo	co-examiner
Prof. Dr. P.J. Tackley	University of California, Los Angeles	co-examiner

2004

Contents

Abstract	9
Zusammenfassung	11
1 Introduction	13
1.1 Methods	14
1.2 Thesis structure	15
1.3 Final remarks	18
Bibliography	19
2 Basin subsidence	21
2.1 Introduction	22
2.2 Phase diagrams	24
2.3 Subsidence model	26
2.4 Results	33
2.4.1 Basin subsidence with and without phase transitions	33
2.4.2 Effect of stretching factor on basin subsidence	35
2.4.3 Parameterized density models	37
2.4.4 Deviations from the TDD pure-shear stretching model	38
2.5 Discussion and conclusions	40
2.6 Acknowledgements	44
2.7 Simplified phase diagrams	44
Bibliography	47

3	Folding versus Diapirism	53
3.1	Introduction	54
3.2	Crustal-scale folds	56
3.2.1	Geological information	56
3.2.2	Mechanical background	57
3.3	Magmatic/diapiric domes	60
3.3.1	Geological Information	60
3.3.2	Mechanical background	61
3.4	Analytical perturbation method	64
3.4.1	Method	64
3.4.2	Phase diagram: folding versus diapirism	65
3.4.3	Discussion	69
3.5	Numerical simulations	72
3.5.1	Numerical technique	72
3.5.2	Comparison between numerics and analytics	73
3.5.3	Geometries developing at finite amplitude stages	73
3.5.4	Dynamics of dome-formation	77
3.6	Strain in and around domes	79
3.7	Discussion	80
3.8	Conclusions	85
3.9	Acknowledgements	85
3.10	Perturbation analysis	85
	Bibliography	89
4	Shear Localization	97
4.1	Introduction	98
4.2	Mathematical model	102
4.2.1	Governing equations	102
4.2.2	Non-dimensionalization	103
4.3	Numerical methods	106
4.3.1	0-D model	106
4.3.2	1-D model	107
4.3.3	2-D model	109

4.4	Initiation stage	111
4.4.1	0-D model	111
4.4.2	1-D model	116
4.4.3	Link to classics	123
4.4.4	2-D model	125
4.5	Shear localization on planets	131
4.6	Conclusions	137
4.7	Appendix A	138
4.7.1	Powerlaw creep	138
4.7.2	Diffusion creep	139
4.7.3	Peierls low temperature plasticity	140
4.7.4	Conversion to simplified model	141
4.8	Appendix B	142
4.9	Acknowledgements	142
	Bibliography	147
5	3D Rayleigh Taylor instability	153
5.1	Introduction	154
5.2	Mathematical model and numerical method	154
5.3	Growthrate calculations	155
5.4	Forward modeling results	157
5.5	Reverse modeling results	160
5.6	Conclusions	161
	Bibliography	163
6	Finite amplitude diapirism	165
6.1	Introduction	165
6.2	Model and methods	166
6.3	Initial stages	167
6.4	Finite amplitude effects	169
6.5	2D vs 3D	171
6.6	Discussion and conclusions	172
	Bibliography	175

7	Diapiric patterns	177
7.1	Introduction	178
7.2	Model and numerical methods	180
7.3	2D simulations	182
7.3.1	Linear stability analysis	182
7.3.2	Initial random perturbations	184
7.3.3	Initial step-like perturbation	186
7.4	3D simulations	187
7.4.1	Linear and nonlinear stability	187
7.4.2	Initial random noise	189
7.4.3	Initial step-like perturbations	194
7.5	Discussion	194
7.6	Conclusions	198
7.7	Acknowledgements	199
7.8	Appendix	199
	Bibliography	203
8	Numerical Method	209
8.1	Introduction	210
8.2	Mathematical model	210
8.3	Numerical method	212
8.3.1	Discretization	212
8.3.2	Iterative approach	217
8.3.3	Plasticity formulation	219
8.3.4	Further numerical issues	221
8.4	Verification of the numerical code	222
8.4.1	0D rheology tests	222
8.4.2	Linear stability analysis of folding and RT-instability	225
8.4.3	Viscoelastic buckling	226
8.4.4	Stress around a weak inclusion	227
8.4.5	Diapirism	227
8.4.6	Shearband inclination	230
8.4.7	Thermal advection and diffusion	231

8.5	Conclusions	233
8.6	Appendix A: stress rotation	234
8.7	Appendix B: finite element formulation	235
	Bibliography	237
A	CO₂ driven aftershocks	241
A.1	Introduction	242
A.2	Aftershock patterns	244
A.3	Modeling results	246
A.4	Aftershock triggering	248
A.5	Conclusions	249
A.6	Methods	251
	Bibliography	253
	Acknowledgements	257
	Curriculum vitae	259

Abstract

This thesis addresses problems from a geodynamic perspective. Numerical methods and analytical techniques are used to obtain insight in the mechanics of geodynamic processes.

The interaction between petrology and tectonics, for example, is studied in a model of sedimentary basin subsidence. It is demonstrated that metamorphic phase transitions, which affect rock density, have a first order effect on the basin subsidence. In particular mantle phase transitions reduce the amount of synrift-subsidence, in agreement with observations.

The mechanics of a model that allows for erosion, compressional and density-driven deformation are studied in detail. The concept of mechanical phase diagrams is employed to study the key parameters controlling the dynamics of the system. In the model, up to eight deformation modes exist, as a function of two non-dimensional parameters. Good agreement is demonstrated between analytical and numerical results.

Mechanical phase diagrams are also applied to shear-localization in visco-elasto-plastic materials. Localization is initiated through shear-heating-induced increase of temperature, whereas it may be inhibited by thermal diffusion. Results are derived for 0D, 1D and 2D settings with constant stress and constant velocity boundary conditions. Up to six deformation modes exist as a function of four nondimensional parameters. Scaling laws, derived for the various deformation modes demonstrate that the model may be relevant to natural conditions.

The Rayleigh-Taylor (RT) instability is formed if a fluid of low density underlies a fluid of higher density. The deformation of the fluid interface is investigated analytical and numerically in both two and three dimensions. It is shown that the purely three-dimensional normal modes amplify faster than 2D modes. The superposition of different normal modes may result in relatively complex patterns. The 2D-3D transition is studied for various boundary conditions. The limits of existing linear stability theory are examined and a new finite-amplitude theory is proposed, which yields a better description of finite-amplitude structures growing from thin source layers. Reversibility issues of the RT-instability are discussed for different initial structures.

The tools used throughout this thesis include analytical techniques (thick- and thin-

plate methods), and numerical methods (finite-difference, finite-element and hybrid finite-difference/spectral methods) that have been specifically tailored to the investigated problems. Progress has been made in the incorporation of solid-like (elasto-plastic) and fluid-like (viscous) behavior in a single computational framework.

Zusammenfassung

Diese Dissertation befasst sich mit Problemen aus dem Gebiet der Geodynamik. Numerische und analytische Methoden werden angewandt, um Einsicht in die Mechanik verschiedener geodynamischer Prozesse zu bekommen.

Die Interaktion zwischen Petrologie und Tektonik, zum Beispiel mit Hilfe eines Modelles eines Sedimentbeckens studiert. Es wird gezeigt, dass metamorphe Phasenübergänge, die die Dichte eines Gesteines beeinflussen, einen grossen Effekt auf die Subsidenz des Beckens haben. Insbesondere die Phasenübergänge im Erdmantel führen zu einer starken Verringerung der Synriftsubsidenz, was mit Beobachtungen übereinstimmt.

Die Mechanik eines Modelles, welches die Effekte von Erosion, kompressiver Deformation und dichtegetriebener Deformation berücksichtigt, wird detailliert studiert. Das Konzept mechanischer Phasendiagramme wird angewandt, um die wichtigsten Parameter des dynamischen Systems zu charakterisieren. Im Modell existieren acht Deformationsarten, die eine Funktion von zwei dimensionslosen Parametern sind. Eine gute Übereinstimmung zwischen analytischen und numerischen Resultaten wurde festgestellt.

Mechanische Phasendiagramme werden auch für die Scherlokalisierung in viskoelasto-plastischen Materialien berechnet. In diesem Modell wird die Lokalisierung durch Scherung initiiert, die einen Anstieg der Temperatur zur Folge hat. Gleichzeitig kann die Lokalisierung durch den Effekt von thermischer Diffusion verhindert werden. Resultate sind gültig für den 0D, 1D und 2D Fall, für sowohl eine konstante Geschwindigkeit als auch eine konstante Spannung als Randbedingung. Bis zu sechs Deformationsarten existieren als Funktion von vier nichtdimensionalen Parametern. Skalierungsgesetzmässigkeiten, hergeleitet für die verschiedenen Deformationsarten, zeigen, dass das Modell für natürliche Bedingungen relevant ist.

Eine Rayleigh-Taylor (RT) Instabilität tritt auf, wenn eine Flüssigkeit niedriger Dichte unter einer Flüssigkeit höherer Dichte liegt. Die Deformation der Grenzlinie zwischen den beiden Flüssigkeiten wird sowohl analytisch als auch numerisch für zwei und drei Dimensionen untersucht. Es wird gezeigt, dass reine dreidimensionale Perturbationen ("normal modes") schneller wachsen als zweidimensionale Perturbationen. Eine Überlagerung verschiedener Perturbationen resultiert in komplexen dreidimen-

sionalen Strukturen. Der Übergang von 2D zu 3D Strukturen wird für verschiedenste Randbedingungen studiert. Die Grenzen existierender linearer Theorien werden aufgezeigt und eine neue finite-amplituden Theorie wird vorgeschlagen, welche das Verhalten von Strukturen, die aus dünnen Grenzschichten wachsen, besser beschreibt. Die Reversibilität der RT-Instabilität wird für verschiedene Anfangsbedingungen aufgezeigt.

Die Methoden, die in dieser Dissertation benutzt werden, umfassen analytische Methoden (Theorie der dünnen und dicken Platten), und numerische Methoden (Finite-Differenzen-, Finite-Elemente und Spektral-Methoden) die speziell für diese Arbeit entwickelt und angepasst wurden. Fortschritte wurden auch bei der Implementierung von festem (elasto-plastisch) und flüssigem (viskoses) Materialverhalten in ein einziges numerisches Programm erzielt.

Chapter 1

Introduction

This thesis addresses geodynamical problems. What does geodynamics mean? *Geo* (Earth) stands for geology, for example sedimentological, petrological and structural observations. Geology, initially a descriptive science, becomes an increasingly quantitative science. *Dynamics* stands for time evolution, mechanics, physics, rheology. Dynamics tries to constrain the observations and interpretations that geologists have made. Such constraints cannot be stated without a thorough mechanical understanding of geological processes. Laboratory experiments on the rheology of rocks have demonstrated that diverse behaviors exist, ranging from brittle failure at low temperatures to (nonlinear) creep-like behavior at higher temperatures. Knowing the rheology of rocks is essential in understanding deformation, but is only one part of geodynamics. The other part involves understanding of mechanical processes that are somewhat different than problems solved in the engineering community (e.g. turbulence in gas, or car-crash tests), since the timescales and the amount of strain involved in many geodynamic processes (e.g. mantle convection, mountain building) are much larger. Geological processes evolve slowly and have negligible inertial terms, which results in specific governing equations. Thus, results obtained by the engineering community are not directly applicable to geodynamic settings, and advances in this direction have to be made by the geodynamic community itself. Different approaches have been employed, including (1) scaled laboratory experiments, (2) analytical techniques and (3) numerical experiments. Laboratory experiments employ analogue materials like honey, silicon putty or sand, to simulate flow of the mantle and crust, thereby taking care that key nondimensional parameters of the experiments are identical to natural values.

Whereas much insight has been obtained by these experiments, it is difficult to incorporate processes like phase transitions or temperature-dependent rheologies. Analytical techniques can be used to derive solutions to the governing mechanical equations, but are typically limited to small deformations and/or relatively simple rheologies. Numerical experiments do not suffer from these drawbacks but have the disadvantage of being computational expensive, especially in three dimensions. Moreover, it is sometimes not easy to understand results of numerical experiments in terms of the controlling parameters, since typically many different (nonlinear) processes are incorporated and active simultaneously in these models. It is clear, for example, that erosion is active during the formation of a mountain belt. At the same time, an orogeny would not exist without horizontal compression, and gravity may also influence its formation. But what are the respective roles of gravity and compression? What is the effect of erosion? In this thesis such questions are addressed by a combination of numerical (FEM, finite difference, spectral) and analytical (thin-plate, thick-plate) modelling techniques.

1.1 Methods

Continuum mechanics is employed throughout this work. The governing equations are solved analytically, if possible, and numerically. The analytical solutions, together with an appropriate non-dimensionalization, allow the identification of controlling parameters. In many cases, mechanical 'phase diagrams' are derived as a function of few non-dimensional parameters that characterize the dynamics of a given setup. In addition, the effects of realistic densities, known from petrological studies, on large-scale geodynamic processes are studied.

Several numerical codes have been developed specifically for this work. One of the novel aspects of these codes is that they allow to couple solid-like (elasto-plastic) behavior with creep-like (viscous) behavior in a single computational domain. The codes also allow relatively large deformations, and are tuned for geodynamic (slowly-moving) processes. The first code is a mixed Eulerian spectral/finite difference technique, whereas the second code is a finite element method (FEM). The rheology employed in both cases is temperature-dependent visco-elasto-plastic with non-associated Mohr-Coulomb plasticity and Maxwell visco-elasticity. The reason to develop several codes is that each is tuned for different problems. The spectral method, for example, is very fast in

cases where the initial setup is periodic and/or has limited lateral variation in effective viscosity (e.g. folding, diapirism). Some disadvantages are that lateral variations in viscosity require special treatment (see Chapter 8) and that non-standard boundary conditions or model domains are more difficult to implement. The FEM method has the advantage of being very general and easy to modify, but the disadvantage of being relatively slow, especially in 3D.

An advantage of combining the different numerical techniques with analytical methods, is that this gives additional crosschecks for each of the methods.

1.2 Thesis structure

Chapter 2 - The effect of mineral phase transitions on sedimentary basin subsidence and uplift.

Chapter 2 focusses on vertical movements in sedimentary basins. The pure-shear stretching model [McKenzie, 1978] predicts that the amount of subsidence rifting is equal than, or larger than the amount of post-rift subsidence. This is at odd with observations that indicate that the thickness of both syn-rift and post-rift sediments are approximately equal. Moreover, syn-rift uplift is frequently observed, which cannot be explained with the pure-shear stretching model (for reasonable parameter values). In this chapter, it is shown that this mismatch may be caused by the fact that the pure-shear model assumes density to be a function of temperature only. Metamorphic petrology, however, showed that density is a function of the mineral assemblage of the rocks. We have computed realistic density models for a range of crustal and mantle mineralogies from thermodynamic data by free energy minimization. Incorporation of these density structures into the pure-shear stretching model results both in syn-rift uplift and a reduction of total syn-rift subsidence compared to post-rift subsidence.

Chapter 3 - The interplay between folding, diapirism and erosion in compressional settings

The interaction between erosion, gravity and compression is studied for a relatively simple, two-layer setup. This model allows for two main modes of vertically amplified structures: folding (compression-controlled) and diapirism (gravity-controlled). Previous workers have typically focussed on one of the two modes, letting aside the other

one (for example ignoring the effect of gravity). By using an analytical perturbation technique, it has been possible to derive solutions for several end-member cases. By matching these solutions, it could be demonstrated that two non-dimensional parameters exist, which control the dynamics of the system. A mechanical 'phase-diagram' has been derived, which exhibits eight folding-diapirism modes, five of which are geologically relevant. A drawback of these analytical methods, however, is that they are limited to small strain. Therefore, numerical simulations have been performed to study how analytical results are applicable to later stages. They demonstrated that the various deformation modes detected with the analytical method result in distinct late-stage patterns.

Chapter 4 - Initiation of localized shear in visco-elasto-plastic rocks

This chapter systematically studies the initiation of localized shear in a visco-elasto-plastic rheology. Shear localization is induced by shear-heating. Mechanical phase diagrams are derived in 0D, 1D and 2D settings under both constant velocity and constant stress boundary conditions. Up to six deformation modes exist as a function of four non-dimensional parameters. Localization may be inhibited by thermal diffusion. The derived scaling-laws give a first-order explanation of previously observed shear-localization effects in numerical models which employed more complex rheologies. Examples of shear-localization at the lithospheric scale are presented. The simple rheological model presented here is thus an attractive candidate for incorporation in next-generation global convection models that study plate-like behavior.

Chapter 5 - Forward and reverse modelling of the three-dimensional viscous Rayleigh-Taylor instability

The Rayleigh-Taylor (RT) instability arises when a low density fluid underlies a higher density fluid. Here, the RT instability is studied in three-dimensions using a newly developed spectral/finite difference code. Detailed comparisons are made between 3D and previous 2D results. Numerically computed growth rate spectra show a wavelength selection towards 3D perturbations at finite amplitudes. Systematic computations are performed to study the survival of initially imposed 2D structures during later stages. Their survival depends on the initial 2D amplitude and wavelength and the amplitude of background noise. It is also shown that reverse (backward) modelling is capable of

restoring the initial geometry as long as overhangs are not developed.

Chapter 6 - The transition from exponential to buoyancy-controlled diapirism.

During the initiation stages of the RT instability, the growth of its amplitude versus time is approximately exponential. During later stages, however, the growth becomes linear, and the exponential theory loses its validity. A modification of the exponential theory is proposed to link the initial with the late stages. The improvement is based on the fact that the late-stage vertical amplification is determined by the Stokes velocity of a rising sphere. The effective radius of this sphere is computable from the initial wavelength of the perturbation. The transition between exponential and linear growth only occurs for initially thin layers. The new theory is valid for aspect ratios of 3D dominant modes of up to 10. Purely 2D modes rise slower than 3D modes, but aspect ratios in excess of 100 are required for a 3D perturbation to behave effectively like a 2D structure.

Chapter 7 - 3D diapiric pattern formation.

A detailed study is made to understand the influence of boundary and initial conditions on the geometry and on the pattern of 2D and 3D RT instabilities. Whereas no-slip, free-slip and no-stress boundary conditions all result in similarly looking, mushroom-shaped geometries, a fast-erosion boundary condition results in chimney-like structures, both in 2D and 3D cases. The spacing of evolved structures is predicted to a large extent by linear stability theory, which gives a characteristic spacing (the dominant wavelength) in 2D cases. In the general 3D case, however, an infinite number of dominant modes exist. Linear superposition of these dominant modes results in complex patterns ranging from circular to curved and elongated in map view. Wall-like structures are initiated in a 3D setup if an initial step-like perturbation is present at the interface of the two fluids. Natural salt-dome patterns range from approximately regularly spaced finger-like diapirs to wall-like diapirs. The results obtained in this chapter show that the spacing of finger-like salt pillows and domes may be consistent with a RT-model. At the same time, wall-like diapirs cannot initiate spontaneously in this model. An initial 2D perturbation of the salt-overburden interface is required, which could, for example represent brittle faulting of the overburden.

Chapter 8 - A finite difference/spectral method for large deformation modelling of visco-elasto-plastic geomaterials

This chapter describes a spectral/finite-difference numerical method that can deal with a Maxwell viscoelastic rheology combined with Mohr-Coulomb plasticity in an Eulerian framework. The method is essentially a further-development of the method presented in Schmalholz et al. [2001] and allows thermomechanical simulations of geodynamic processes. Benchmark studies are presented, testing various aspects of the code.

Appendix A - Aftershocks driven by a high pressure CO₂ source at depth

An unusual strong series of aftershocks followed a magnitude 5.7 earthquake in 1997 in the Assisi-region (Italy). Models of elastic stress transfer fail to explain this normal-faulting sequence. We show that this sequence may have been driven by a fluid pressure pulse generated from the co-seismic release of a known deep source of trapped CO₂. The flow of CO₂ is governed by a Darcy-type model, in which the permeability has a non-linear dependency on the effective stress (i.e. the high-pressure fluid generates its own path). There is a correlation between the high pressure front and the aftershock hypocenters over a two week period. The 10-20 MPa triggering amplitude of the pressure pulse overwhelms the typical 0.1-0.2 MPa range from stress changes in the usual stress triggering arguments.

1.3 Final remarks

The models presented here give insight into various geodynamic processes. Any model, whether numerical, analytical, or analogue has the drawback of being a simplification of nature. Ultimately, whether the simplifications are justified can only be tested through close collaboration between the different fields of geosciences. From my perspective, the duty of modelers is to provide a thorough understanding of processes that happen in their highly idealized world. If possible, scaling laws with predictive power should be derived. The examples presented in this thesis realize this philosophy through combined application of analytical and numerical methods.

Bibliography

- D. McKenzie. Some remarks on the development of sedimentary basins. *Earth and Planetary Science Letters*, 40:25–32, 1978.
- S. M. Schmalholz, Y. Y. Podladchikov, and D. W. Schmid. A spectral/finite difference method for simulating large deformations of heterogeneous, viscoelastic materials. *Geophysical Journal International*, 145(1):199–208, 2001.

Chapter 2

The effect of mineral phase transitions on sedimentary basin subsidence and uplift.¹

Abstract Metamorphic phase transitions influence rock density, which is a major parameter affecting lithosphere dynamics and basin subsidence. To assess the importance of these effects we have computed realistic density models for a range of crustal and mantle mineralogies from thermodynamic data by free energy minimization. These density distributions are incorporated into one- and two-dimensional kinematic models of basin subsidence. The results demonstrate that, compared to models in which density is solely temperature dependent, phase transitions have the effect of increasing post-rift subsidence while decreasing syn-rift subsidence. Discrepancies between our model results and those obtained with the conventional uniform stretching models can be up to 95 % for reasonable parameter choices. The models also predict up to 1 km of syn-rift uplift as a consequence of phase transitions. Mantle phase transitions, in particular the spinel-garnet-plagioclase-lherzolite transitions are responsible for the most significant effects on subsidence. Differences in mantle composition are shown to be a second-order effect. Parameterized density models are derived for crustal and mantle rocks, which reproduce the main effects of the phase transitions on subsidence.

¹This work has been submitted for publication in Earth and Planetary Science Letters (Kaus, B.J.P, Connolly, J.A.D., Podladchikov, Y.Y., Schmalholz, S.M., in review)

2.1 Introduction

One of the most widely used models for subsidence of sedimentary basins formed by extension is the uniform stretching model (USM), which assumes that subsidence is caused by crustal thinning and by thermal cooling [McKenzie, 1978]. An important feature of the USM formulation is that lithospheric density is assumed to depend only on temperature, a model we designate as the temperature-dependent-density (TDD) formulation. The TDD formulation has been applied successfully in many situations, but it cannot explain certain common observations [e.g. Ziegler and Cloetingh, 2004, and reference therein]. The most prominent difficulty is that many basins have relatively thin syn-rift sediments, but thick post-rift sediments [Walker et al., 1997, Royden and Keen, 1980, Beaumont et al., 1982, Sclater et al., 1980, Spadini et al., 1997]. To explain the thickness of the post-rift sediments with the TDD formulation, extensive stretching is required; this is at odds with the small thickness of syn-rift sediments. The widespread phenomenon of basin uplift during overall extension of the lithosphere [e.g. in the Vøring Basin; see Ren et al., 2003] creates a second difficulty for the TDD formulation. A thin crust is required ($< 1/7$ of the lithospheric thickness) to explain this phenomenon. Such crustal thicknesses contradict geophysical observations suggesting crustal thickness is typically 30-40 km. Uplift usually occurs preceding rifting or after a finite amount of extension. An additional problem with the TDD formulation is posed by the fact that many basins have a phase of accelerated subsidence rates during the post-rift thermal subsidence phase [Middleton, 1980, Hamdani et al., 1994]. The TDD formulation predicts that the subsidence rate decreases with time t as $t^{-0.5}$, and thus this can only be explained with non-thermal mechanisms [Podladchikov et al., 1994, Hamdani et al., 1994, 1991]. To rectify these problems refinements of the TDD formulation have been proposed that include depth-dependent stretching [Royden and Keen, 1980, Beaumont et al., 1982], active rifting [Huisman et al., 2001], interaction between lithospheric rheology and erosion [Burov and Poliakov, 2001] or mineral phase transitions [e.g. Podladchikov et al., 1994, Yamasaki and Nakada, 1997, O'Connell and Wasserburg, 1972, Artyushkov et al., 2000]. Here we focus on a refinement of this model in which lithospheric density is adjusted to account for phase transitions that occur in response to the geodynamic cycle.

Most of the models that have been proposed to explain shortcomings of the TDD

formulation involve complexity or rely on parameters such as the lithospheric rheology, which are poorly constrained. However the conditions and consequences of the metamorphic phase transitions that occur during lithospheric thinning are constrained from field observations, experimental studies and thermodynamic theory. We exploit the latter to construct a realistic lithospheric density model and to assess its consequences for basin subsidence.

That metamorphic phase transitions influence basin subsidence has been recognized for several decades. It has been suggested in Lovering [1958] and Kennedy [1959] that crustal phase transitions around the Moho could affect uplift and subsidence. Numerical and analytical studies that concentrated on phase transitions in crustal rocks [Hamdani et al., 1991, 1994, Middleton, 1980, Neugebauer and Spohn, 1978, Spohn and Neugebauer, 1978, Artyushkov et al., 2000], in mantle rocks [Podladchikov et al., 1994, Yamasaki and Nakada, 1997], or in both [Petrini et al., 2001] demonstrated that phase transitions cause syn-rift uplift preceding rifting, greater post-rift subsidence than in the TDD formulation and periods of accelerated subsidence. Lobkovsky and coworkers [Lobkovsky et al., 1993, 1996] proposed a model in which partial melt, emplaced and solidified in lenses below the rift center, is transformed into eclogite causing accelerated post-rift subsidence. Their model requires a nearly impermeable Moho and predicts that eclogite lenses remain present after the completion of extension, which may be seismically detectable.

The applicability of most of the models described above is limited, since they typically only consider a single discontinuous phase transition. Natural rocks have continuous reactions. Many of these reactions have only small density effects, but the cumulative effect of these reactions can be significant. The optimal approach is to consider all the reactions that may occur in the lithosphere. Such an analysis in combination with basin subsidence was done by Petrini et. al. [Petrini et al., 2001], who used a realistic density distribution for both mantle and crustal rocks and demonstrated that phase changes lead to more post-rift subsidence and less syn-rift subsidence. However, they restricted their analysis to small stretching factors ($\delta = 1.5$) and did not detect syn-rift uplift as observed in Podladchikov et al. [1994] and Yamasaki and Nakada [1997].

Here we follow the same approach by coupling realistic density distributions with a kinematic subsidence model. To estimate the sensitivity of the results to the chemical

composition of the lithosphere, we compute density models for a range of different mantle and crustal compositions. The results are then compared with the TDD formulation and parameterized density maps are derived that reproduce results of 'real' density maps up to reasonable accuracy and thus yield additional insight into the way phase transitions influence subsidence.

2.2 Representative phase diagrams and density distributions for crustal and upper mantle rocks.

Phase assemblages at the pressure (P) and temperature (T) conditions of interest were computed using free-energy minimization [Connolly, 1990, Connolly and Pettrini, 2002]. The minimization program requires thermodynamic data for end-member phase compositions together with solution models to compute relative proportions, compositions and densities of the stable mineralogy.

The bulk compositions considered here are listed in Table 2.1. Taylor and McLennan Taylor and McLennan [1985] proposed a model for the continental crust consisting of 75% Archean crust and 25% andesitic crust, to represent the different Archean and post-Archean crustal growth processes. A second crustal model assumes the crust to have a granodioritic composition [Wedepohl, 1995]. Three different mantle models have been employed. These are variations of the pyrolite model originally proposed by Ringwood [Ringwood, 1962] and represents the mean chemical composition of the upper mantle. The compositions that are used for calculations range from an incompatible-element-enriched mantle (Hawaiian Pyrolite) to a normal upper mantle (MOR-pyrolite), to a depleted mantle (Tinaquillo lherzolite). The crust is assumed to be saturated in water, whereas the mantle is anhydrous. Melting in the crust is not taken into account.

The model bulk rock compositions were simplified to include the oxides SiO_2 , TiO_2 , Al_2O_3 , FeO , MgO , CaO , Na_2O and MnO . Cr_2O_3 is not available in the database of Holland and Powell [Holland and Powell, 1998] and was therefore not considered. Chromium would effect the spinel-stability field, but due to its low volumetric content ($< 0.5\%$) its overall contribution to the mean bulk density is minor. The same reasoning justifies neglecting other minor components such as P_2O_5 and K_2O . The thermodynamic mineral database of Holland and Powell [Holland and Powell, 1998]

	Hawaiian Py- rolite ¹	MORB- harzburgite pyrolite ²	Depleted (Tinaquillo) pyrolite ¹	Total Crust ³	Granodiorite ⁴
SiO ₂	45.2	45	45	57.3	66.1
TiO ₂	0.71	0.17	0.1	0.9	0.54
Al ₂ O ₃	3.54	4.4	3.2	15.9	15.7
FeO	8.47	7.6	7.7	9.1	4.4
MgO	37.5	38.8	40	5.3	1.74
CaO	3.08	3.4	3	7.4	1.5
Na ₂ O	0.57	0.4	0.2	3.1	3.75
MnO	0	0.11	0	0	0
K ₂ O	0	0	0	1.1	2.78
H ₂ O	0	0	0	saturated	saturated

Table 2.1: Crustal and mantle compositions considered in this work. ¹Green and Falloon [1998], ²Green et al. [1979], ³Taylor and McLennan [1985], ⁴Wedepohl [1995]

was used in the calculations, together with solution models as listed in table 2.2.

The accuracy of the solution models was specified to resolve compositions with a maximum error of 2 mol%. Melting of mantle rocks is accounted for by the pMELTS model [Ghiorso et al., 2002]. Solution models compiled by Holland and Powell were used for crystalline solution phase. To test the effect of using solution models from other sources, calculations were done with a range of different solution models, and it was found that the effect on the density structure was relatively minor. Furthermore, some of the computations of Sobolev and Babeyko [Sobolev and Babeyko, 1994], who used independently developed software and databases, have been redone here. These computations showed good agreement, both for the density structure and for the phase diagram section topology, which gives additional confidence in the robustness of our approach.

The accuracy of the solution models was specified to resolve compositions with a maximum error of 2 mol%. Melting of mantle rocks is accounted for by the pMELTS model [Ghiorso et al., 2002]. Solution models compiled by Holland and Powell were used for crystalline solution phase. To test the effect of using solution models from other sources, calculations were done with a range of different solution models, and it was found that the effect on the density structure was relatively minor. Furthermore, some of the computations of Sobolev and Babeyko [Sobolev and Babeyko, 1994], who used independently developed software and databases, have been redone here. These

computations showed good agreement, both for the density structure and for the phase diagram section topology, which gives additional confidence in the robustness of our approach.

The most important mantle phase transitions are the spinel-lherzolite to garnet-lherzolite transition that occurs around at pressures from 12-20 kbar and the plagioclase-lherzolite to spinel/garnet-lherzolite that occurs at lower pressures (fig. 2.1). These results are in agreement with previous computational results and experimental data [see e.g. Wood and Yuen, 1983, Wood and Holloway, 1984]. Partial melting of mantle rocks commences at temperatures above 1200 °C, which is in agreement with experimental data [Green and Falloon, 1998].

The corresponding densities for these two compositions as well as for the other three compositions listed in table 2.1 are shown in figures 2.2 and 2.3. The most significant density change in the crust is the eclogite (plagioclase-out) transition (fig. 2.1). The 'total-crust' density distribution has densities that are slightly higher than the 'granodioritic' crustal composition. Mantle densities are discontinuous across the stability fields for garnet-, spinel-, and plagioclase-lherzolite, and at the onset of melting. The density change related to the spinel-garnet phase transition is around 30-40 kg/m³. This is the only phase transition that was considered in [Podladchikov et al., 1994] and , [Yamasaki and Nakada, 1997]. However, the density change that is related to the transformation of spinel-lherzolite to plagioclase-lherzolite (80-100 kg/m³) is larger than the density effect of the spinel-garnet transition. Therefore the plagioclase-in transition could potentially cause additional post-rift subsidence and syn-rift uplift.

2.3 One dimensional subsidence model

The influence of phase transitions on sedimentary basin subsidence was studied with a one-dimensional kinematic model. The lithosphere is assumed to consist of a crust and a mantle. During a rifting period of duration t_{rift} , the lithosphere is uniformly thinned in such a way that the ratio between the post-rift and the pre-rift thickness is

$$\delta = \frac{H_{lith}^{pre}}{H_{lith}^{post}} \quad (2.1)$$

here H_{lith}^{pre} is the pre-rift and H_{lith}^{post} the post-rift lithospheric thickness. During and after rifting, perfect isostasy is assumed, i.e. the pressure or lithostatic load at depth

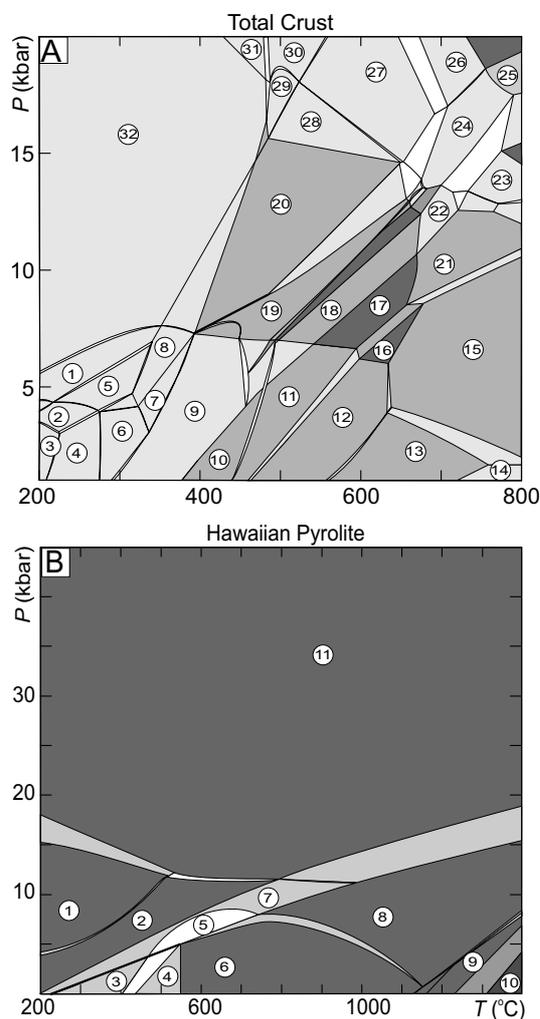


Figure 2.1: Phase diagram section of lithospheric rocks with A) a 'total crust' chemical bulk composition (see table 2.1). Numbers refer to the following phase relationships (see table 2.2 for abbreviations): 1=Chl+Phg+lws+sph, 2=Chl+Kfs+pmp+sph, 3=Chl+Kfs+sph+lmt, 4=Chl+Kfs+sph+prh, 5=Chl+Phg+pmp+sph, 6=Chl+Kfs+sph+lmt, 7=Bt+Chl+zso+sph, 8=Chl+Phg+zso+sph, 9=Bt+Chl+zso+sph, 10=Bt+Chl+sph, 11=Bt+Chl+rt, 12=Bt+Chl+ilm, 13=Bt+Crd+ilm, 14=Bt+Crd+fa+ilm, 15=Bt+Grt+ilm, 16=Bt+Amp+ilm, 17=Bio+rt, 18=Bt+Phg+rt, 19=Phg+zso+sph, 20=Phg+zso+sph, 21=Bt+Grt+rt, 22=Bt+Phg+Grt+rt, 23=Bt+Grt+rt, 24=Phg+Grt+ru, 25=Phg+Grt+rt, 26=Phg+Pa+Grt+rt, 27=Phg+Grt+zso+rt, 28=Phg+Grt+zso+sph+rt, 29=Phg+Grt+law+sph, 30=Phg+Grt+lws+rt, 31=Chl+Phg+lws+rt. 32=Chl+Phg+lws+sph. Phases 1-27, 29-31 contain qtz, phases 1-8,20,23-32 contain Cpx, phases 9-24,26-32 contain Amp and phases 1-19, 21-25 contain Pl B) a 'Hawaiian Pyrolite' mantle composition. Numbers refer to the phase relationships: 1=Cpx+Opx+Ol+Ilm+crn, 2=Cpx+Opx+Ol+Spl+Ilm, 3=Cpx+Opx+Pl₁+Pl₂+Ol+Ilm, 4=Cpx+Opx+Pl₁+Pl₂+Ol+rt, 5=Cpx+Opx+Pl₁+Ol+Spl+Ilm+rt, 6=Cpx+Opx+Pl+Ol+rt, 7=Cpx+Opx+Ol+Spl+Ilm+Pl, 8=Cpx+Opx+Pl+Ol+Ilm, 9=Cpx+melt+Pl₂+Ol+Ilm, 10=melt+Ol+Ilm, 11=Cpx+Opx+Ol+Ilm+Grt.

Symbol	Phase	Formula	Ref.
Amp	amphibole	$\text{Ca}_{2-2w}\text{Na}_{2w}\text{Mg}_{3+2y+z}x\text{Fe}_{(3+2y+z)(1-x)}\text{Al}_{3-3y-w}\text{Si}_{7+w+y}\text{O}_{22}(\text{OH})_2$	1
Bt	biotite	$\text{KMg}_{(3-y)x}\text{Fe}_{(3-y)(1-x)}\text{Al}_{1+2y}\text{Si}_{3-y}\text{O}_{10}(\text{OH})_2$	1
Car	carpolite	$\text{Mg}_x\text{Fe}_{(1-x)}\text{Al}_2\text{Si}_2\text{O}_6(\text{OH})_2$	ideal
Chl	chlorite	$\text{Mg}_{(5-y+z)x}\text{Fe}_{(5-y+z)(1-x)}\text{Al}_{2(1+y-z)}\text{Si}_{3-y+z}\text{O}_{10}(\text{OH})_8$	2
Cpx	clino- pyroxene	$\text{Na}_{1-y}\text{Ca}_y\text{Mg}_x\text{Fe}_{(1-x)y}\text{Al}_y\text{Si}_2\text{O}_6$	6
Crd	cordierite	$\text{Mg}_{2x}\text{Fe}_y\text{Mn}_{(1-x-y)}\text{Al}_2\text{SiO}_5(\text{OH})_2$	ideal
crn	corundum	Al_2O_3	1
Ctd	chloritoid	$\text{Mg}_x\text{Fe}_y\text{Mn}_{(1xy)}\text{Al}_2\text{SiO}_5(\text{OH})_2$	1
Grt	garnet	$\text{Fe}_{3x}\text{Ca}_{3y}\text{Mg}_{3(1-x-y)}\text{Al}_2\text{Si}_3\text{O}_{12}$	1
ilm	ilmenite	FeTiO_3	1
Ilm	ilmenite	$\text{Mg}_x\text{Mn}_y\text{Fe}_{1-x-y}\text{TiO}_3$	ideal
Kfs	alkali feldspar	$\text{Na}_x\text{K}_y\text{AlSi}_3\text{O}_8$	3
lws	lawsonite	$\text{CaAl}_2\text{Si}_2\text{O}_7(\text{OH})_2 \cdot (\text{H}_2\text{O})$	1
lmt	laumontite	$\text{CaAl}_2\text{Si}_4\text{O}_{12} \cdot (4\text{H}_2\text{O})$	1
Ol	olivine	$\text{Mg}_x\text{Fe}_{1-x}\text{SiO}_4$	1
Opx	ortho- pyroxene	$\text{Ca}_z\text{Mg}_{x(2-y)(1-z)}\text{Fe}_{(1-x)(2-y)(1-z)}\text{Al}_{2y}\text{Si}_{2-y}\text{O}_6$	1
Pl	plagioclase	$\text{Na}_x\text{Ca}_{1-x}\text{Al}_{2-x}\text{Si}_{2+x}\text{O}_8$	4
Phg	phengite	$\text{K}_x\text{Na}_{1-x}\text{Mg}_y\text{Fe}_z\text{Al}_{3-2(y+z)}\text{Si}_{3+y+z}\text{O}_{10}(\text{OH})_2$	1
Melt	melt	Na-Mg-Al-Si-K-Ca-Fe hydrous silicate melt	5
prh	prehnite	$\text{Ca}_2\text{Al}_2\text{Si}_3\text{O}_{10}(\text{OH})_2$	1
pmp	pumpellyite	$\text{Ca}_4\text{MgAl}_5\text{Si}_6\text{O}_{21}(\text{OH})_7$	1
qtz	quartz	SiO_2	1
rt	rutile	TiO_2	1
Sa	sanidine	$\text{Na}_x\text{K}_y\text{AlSi}_3\text{O}_8$	1
Spl	spinel	$\text{Mg}_x\text{Fe}_{1-x}\text{AlO}_3$	1
sph	sphene	CaTiSiO_5	1
St	staurolite	$\text{Mg}_{4x}\text{Fe}_{4y}\text{Mn}_{4(1xy)}\text{Al}_{18}\text{Si}_{7.5}\text{O}_{48}\text{H}_4$	1
zo	zoisite	$\text{Ca}_2\text{Al}_3\text{Si}_3\text{O}_{12}(\text{OH})$	1

Table 2.2: Mineral solution notation, formulas and model sources (1 = Holland and Powell [1998]; 2 = Holland et al. [1998]; 3 = Chatterjee and Froese [1975]; 4-Newton et al. [1980]; 5-Ghiorso et al. [2002]; 6-Holland and Powell [1996]). The compositional variables w , x , y , and z may vary between zero and unity and are determined as a function of pressure and temperature by free-energy minimization.

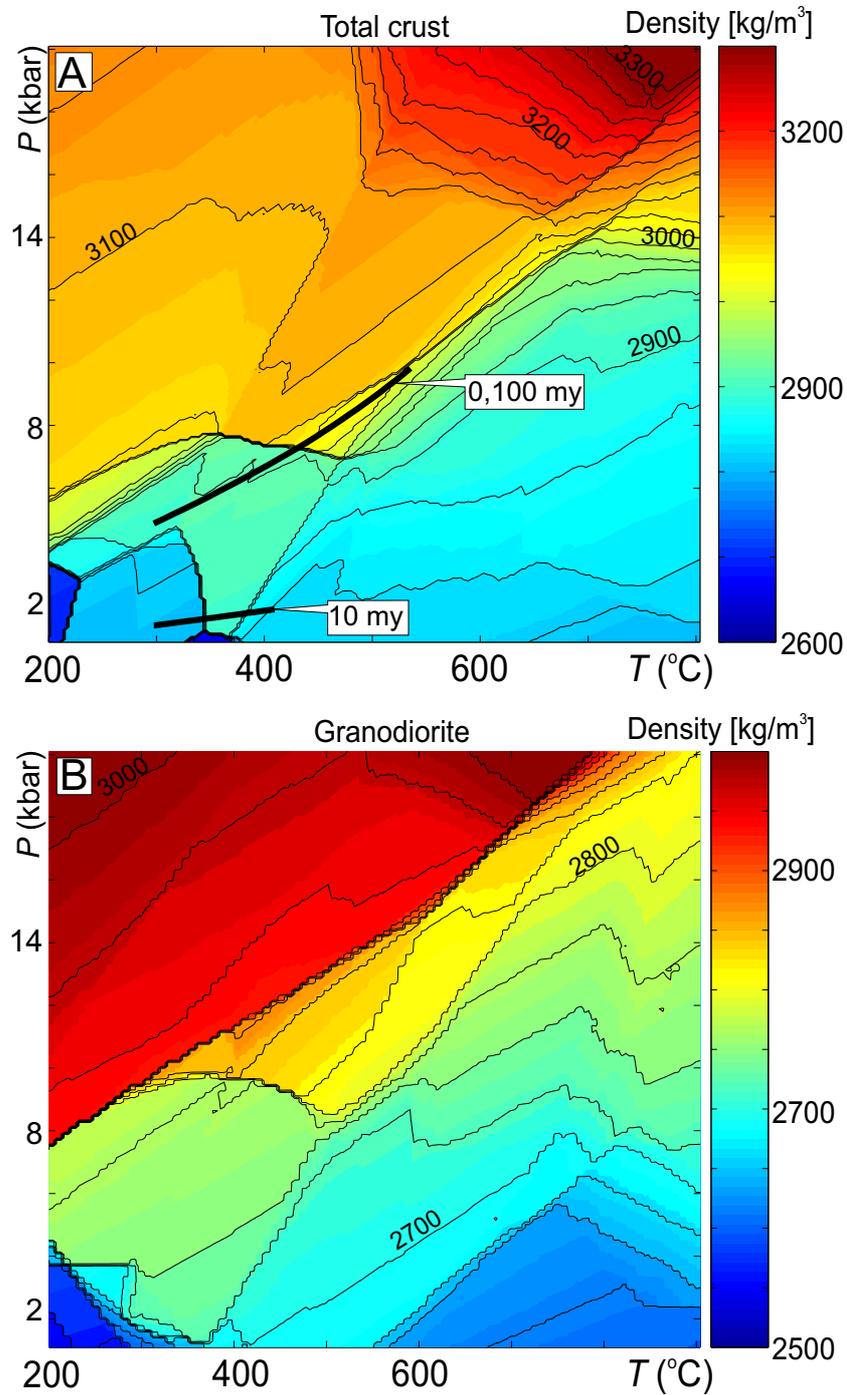


Figure 2.2: Density as a function of pressure and temperature for different crustal rocks, with a bulk chemical composition given in table 1. Superposed lines in A) shows the P-T distribution in a lithosphere with parameters of fig. 5 and $\delta = 6$ (see text).

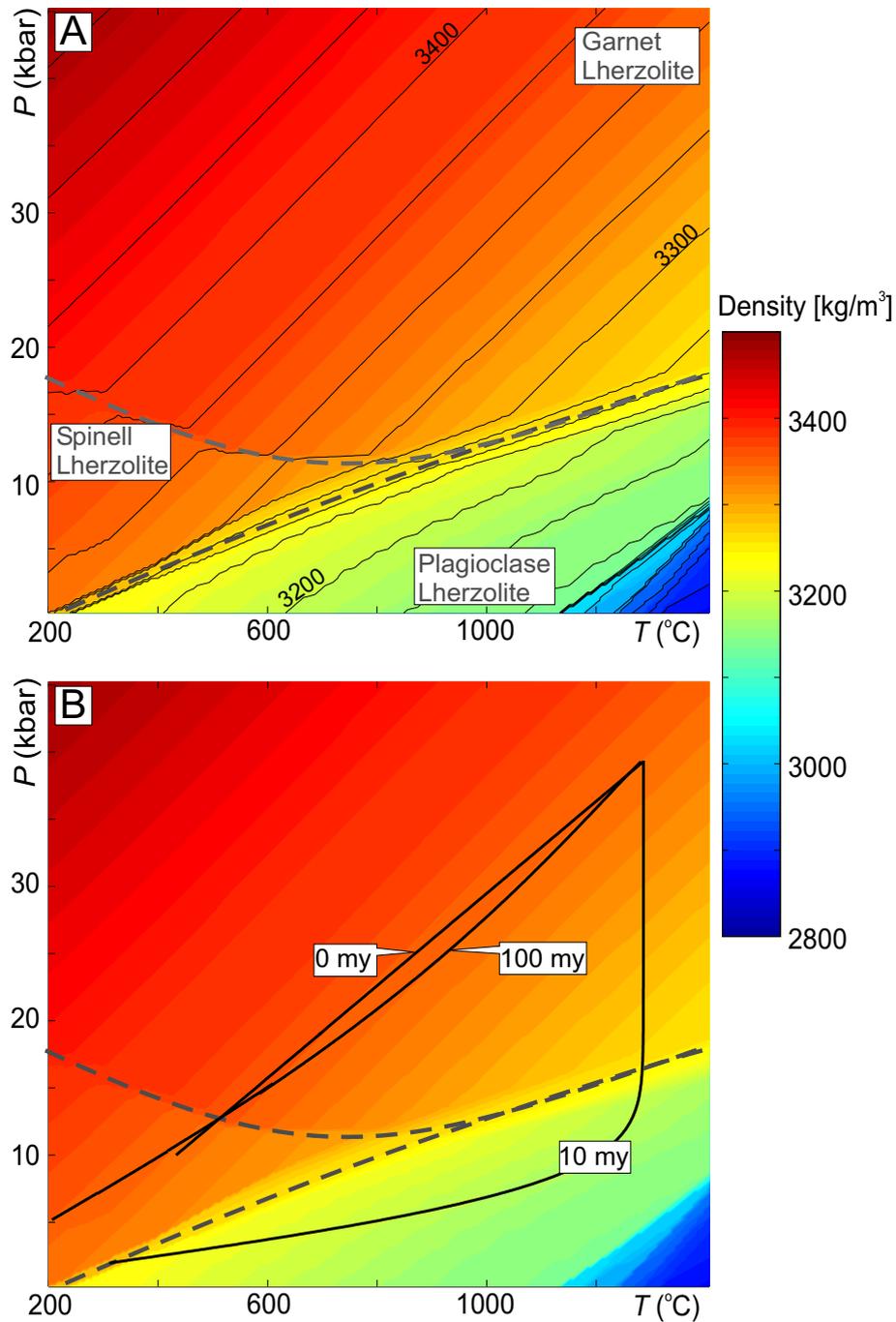


Figure 2.3: Density as a function of pressure and temperature for a Hawaiian Pyrolite mantle composition with a bulk chemical composition given in table 1. Superposed lines in B) shows the P-T distribution in a lithosphere with parameters of fig. 2.5 and $\delta = 6$.

H_{lith}^{pre} is constant. The subsiding basin is filled with sediments of constant density (2200 kg/m³), and thinning of the lithosphere is compensated by up-welling of mantle asthenosphere. Thinning of the lithosphere occurs incrementally between time steps to include effects of sediment blanketing and thermal cooling during rifting. Sediments are also thinned incrementally to conserve volume of the already deposited sediments. The effects of erosion are not incorporated and sediments can be uplifted above the original zero-level.

The thermal evolution during and after rifting is computed with:

$$\rho c_p \left(\frac{\partial T}{\partial t} + v_z \frac{\partial T}{\partial z} \right) = \frac{\partial}{\partial z} \left(k \frac{\partial T}{\partial z} \right) + A \quad (2.2)$$

where ρ is density, c_p heat capacity, t time, z vertical coordinate, v_z vertical velocity during thinning, T temperature, k thermal conductivity and A radioactive heat production. Radioactive elements are assumed to be only present in the crust. The boundary conditions are constant temperature of $T = 0^\circ\text{C}$ on top of the model and a constant temperature of $T = T_{base}$, at the base of the model. The initial condition is a steady-state temperature distribution. Density is computed with three approaches. The first approach is the classic approach in which density is temperature-dependent through the thermal expansivity α . The second approach takes the pressure dependence of density into account, which yields the following density-pressure-temperature relationship

$$\rho = \rho_0(1 - \alpha T + \beta P) \quad (2.3)$$

The third approach takes density from our computed density maps (figs. 2.2 and 2.3). At lower temperatures metamorphic reactions proceed slowly and metastable phase assemblages can occur [e.g. Austrheim et al., 1997] so that our assumption of equilibrium phase assemblages may not be valid. Therefore, we assume that the density structure at temperature lower than 200 °C, is the same as that 200 °C. Simulations in which this equilibrium temperature was increased to 300 °C did not show a significant difference. An additional effect that has been ignored in the current study is the effect of latent heat. This effect was considered by Yamasaki and Nakada [1997], who demonstrated that it causes maximum changes in temperature of around 10 °C, which is insignificant on a lithospheric scale [e.g. Connolly and Thompson, 1989].

Equation 2.2 is solved numerically with an implicit finite-difference scheme for the diffusion terms [Borse, 1997] and temperature advection is calculated using the method

Symbol	Meaning	Value
k	Thermal conductivity	$3 \text{ Wm}^{-1}\text{K}^{-1}$
A	Radioactive heat production	10^{-6} Wm^{-1}
c_p	Specific heat	$10^3 \text{ Jkg}^{-1}\text{K}^{-1}$
α	Thermal expansivity	$3 \times 10^{-5} \text{ K}^{-1}$
β	Compressibility	$1 \times 10^{-11} \text{ Pa}^{-1}$

Table 2.3: Model parameters used in this work.

of characteristics. The numerical code was benchmarked versus variable analytical solutions for diffusion and advection problems, versus the analytical solution of McKenzie [1978] and versus results of a numerical basin code independently developed by one of the co-authors (SMS). The nonlinearity in the problem, which is introduced through the dependency of density on temperature and pressure, in combination with isostatic equilibrium, is solved iteratively. For the calculations of density changes conservation of volume is assumed. Each one-dimensional element between two grid points is assigned a density and the size of this element does not change during the iterative density calculations. An alternative approach is to conserve mass so that the product of element length times its density is the same before and after the density changes. The numerical resolution was 200 grid points in vertical direction. Resolution tests have been performed to ensure that the temporal and spatial resolution is sufficient.

2.4 Results

2.4.1 Basin subsidence with and without phase transitions

A range of calculations have been performed for the solely temperature-dependent density (TDD) formulation, a model in which density is solely dependent on P and T (eq. 2.3), and a model in which 'realistic' densities are taken from thermodynamic calculations (fig. 2.4). Differences between the TDD formulation and the density formulation of equation 2.3 are insignificant, indicating that the term βP in equation 2.3 can be neglected. However, the incorporation of realistic densities does have a large influence on the subsidence behavior. In general, syn-rift subsidence is decreased and post-rift subsidence is increased. During the syn-rift phase most models are initially close to the subsidence behavior of TDD model. After ~ 5 million years deviations can be observed, which in some cases result in an uplift of several hundred meters. After rifting ceases, models with realistic densities give significantly more post-rift subsidence compared to TDD formulations. Moreover, realistic density models lead to episodes of both increased and decreased subsidence rate relative to TDD models.

Syn-rift subsidence in the TDD formulation is dependent on the density contrast between mantle and crust in such a way that a lower mean crustal density results in more syn-rift subsidence (for the same average density of the infilled sediments). An identical effect can be observed here in models with a realistic crustal density (fig. 2.4). Models with a 'granodioritic' density subside more during the syn-rift phase than models with a 'total-crust' density. This reflects the fact that the 'total-crust' model is largely based on inferred Archean crustal compositions that are denser than granodioritic compositions.

Post-rift subsidence in the TDD formulation is independent of the density contrast between crust and mantle. Thus the large differences that are observed during the post-rifting phase in all models with realistic densities (fig. 2.4c) are entirely due to the effect of phase transitions. Depending on the mineralogy, the additional post-rift subsidence ranges from $\sim 30\% - 100\%$.

To establish whether this additional subsidence is caused by mantle or by crustal phase transitions, two experiments were performed: one in which the crustal phase transitions and one in which the mantle phase transitions have been deactivated. Deactivating phase transitions has been done by replacing the realistic density model by

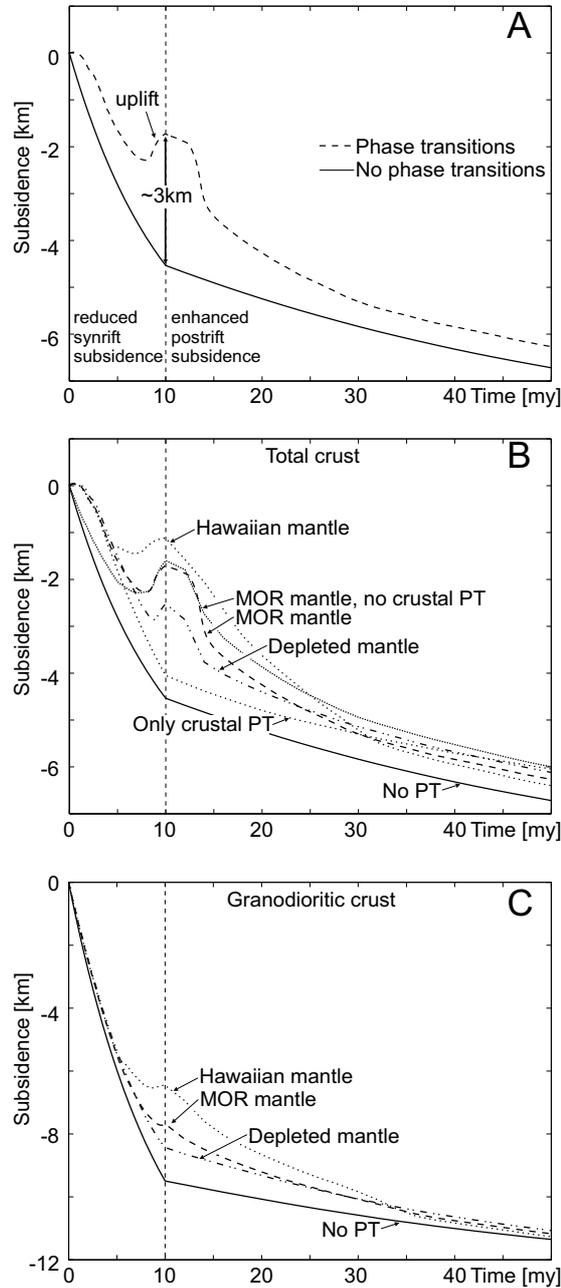


Figure 2.4: Effect of phase-transitions on basin subsidence for a model with $\delta = 3$, $H_{crust}^{pre} = 35$ km, $H_{mantle}^{pre} = 90$ km and $T_{base} = 1300^\circ\text{C}$. A) Comparison of a model with phase transitions and a model without phase transitions. The model with phase transitions has a 'total crust' crustal density and a MOR mantle density. The model without phase transitions has been computed with $\rho_0^c = 2900$ kgm^{-3} . B) The effect of mantle compositions on subsidence for 'total crust' crustal model. The model with crustal phase transitions deactivated assumes $\rho_0^c = 2995$ kgm^{-3} . C) The effect of mantle compositions on subsidence for 'granodioritic' crustal model. $\rho_0^c = 2730$ kgm^{-3} .

a TDD formulation (eq. 2.3 with $\beta = 0$). The results demonstrate that mantle phase transitions are, to a large extent, responsible for the additional post-rift subsidence that is observed (fig. 2.4B).

Syn-rift uplift is observed in almost all models with realistic densities (fig. 2.4), and occurs either at the onset of rifting or after a finite amount of thinning. The amount of syn-rift uplift at the onset of rifting is relatively small (maximum ≈ 50 meters) compared to the amount of uplift that is obtained after a finite amount of extension. However, it has an impact on the total amount of syn-rift subsidence, since it prevents the basin from subsiding for ~ 1 million years. In the next section we analyze the reasons for the syn-rift-uplift and study its dependence on various parameters.

2.4.2 Effect of stretching factor on basin subsidence

Increasing the stretching factor δ with a realistic density model (fig. 2.5) enhances the effect of phase transitions in that it results in less syn-rift subsidence and more post-rift subsidence than in simple density models. For example, a stretching factor of $\delta = 1.5$ causes additional post-rift subsidence of 20% – 40%, whereas stretching factors of $\delta = 4$ and $\delta = 6$ cause additional subsidence of 50% – 110%, depending on the composition of crustal and mantle. Up to 1000 m of syn-rift-uplift is observed in models with a 'total-crust' crustal composition combined with either a Hawaiian- or a MOR-pyrolitic mantle composition for stretching factors of 4 and 6. If syn-rift-uplift occurs, it is followed by a period of increased post-rift subsidence.

The phenomena that cause syn-rift uplift and extensive post-rift subsidence in the simulations with realistic density models can be understood by considering the variations in density caused by phase transformations during and after rifting (figs. 2.2A and 2.3B). At the onset of rifting, the mantle lithosphere is composed of spinel-lherzolite up to a pressure of 12 kbar (corresponding to a depth of 41 km) and of garnet-lherzolite at greater pressures. During rifting, spinel-lherzolite and part of the garnet-lherzolite transforms into plagioclase-lherzolite. The onset of syn-rift uplift (fig 2.5) corresponds to the introduction of plagioclase-lherzolite into the model mantle and the end of syn-rift-uplift, at ~ 8 myr, corresponds to the complete transformation of spinel-lherzolite into plagioclase-lherzolite. After rifting ceased, the mantle cools and plagioclase-lherzolite transforms back into garnet- and spinel-lherzolite. Simula-

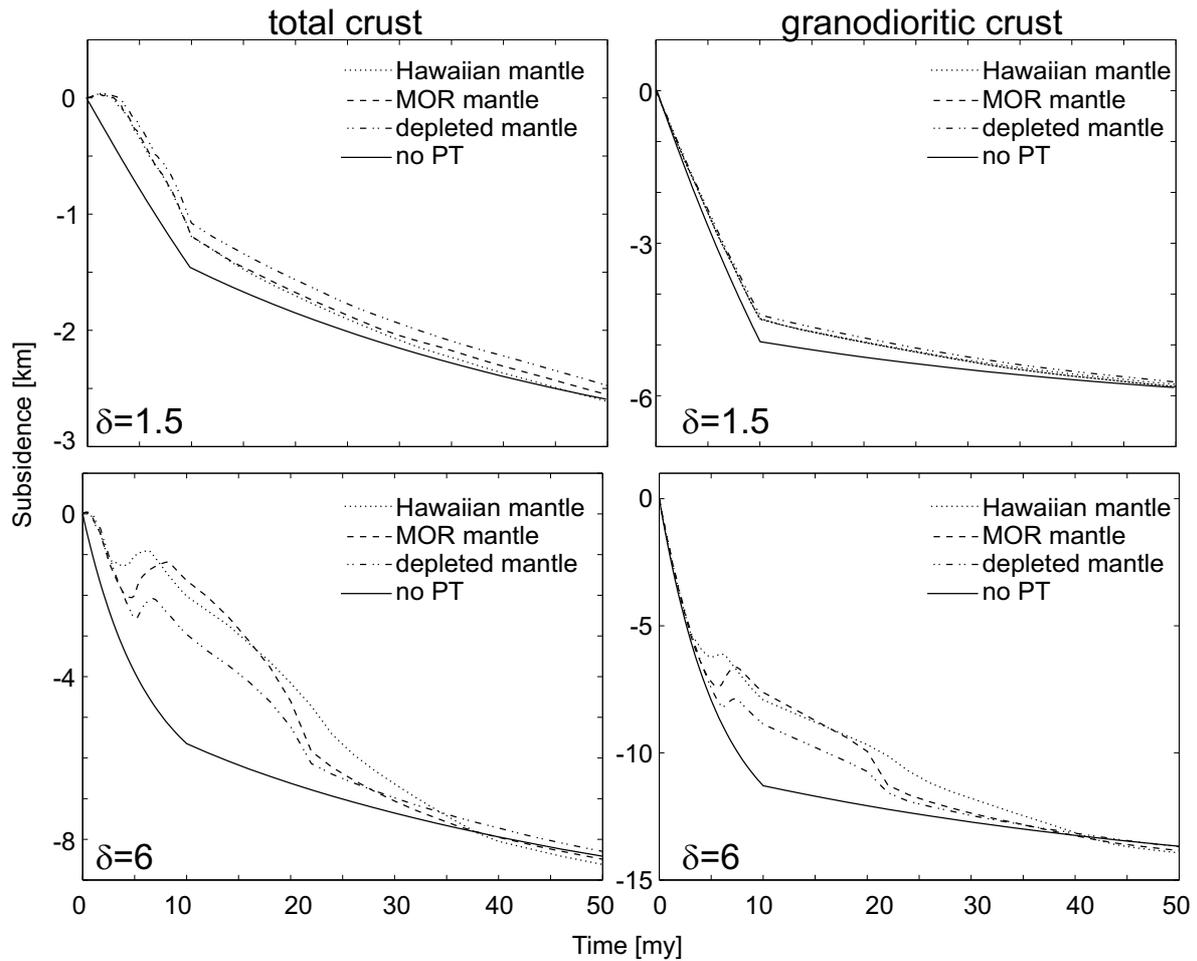


Figure 2.5: Effect of stretching factor δ on basin subsidence for different mantle and crustal compositions. ρ_c^0 in models without phase transitions is chosen to match the long-term total subsidence.

tions with lower stretching factors (e.g. $\delta = 1.5$), do not result in the formation of plagioclase-lherzolite and have no syn-rift uplift (see fig. 2.5). Thus the formation of plagioclase lherzolite is an influential reaction for basin subsidence.

Syn-rift uplift has a different magnitude for different mantle compositions. A depleted mantle composition leads to generally less syn-rift uplift (fig. 2.5). This can be attributed to the fact that a depleted mantle has a higher Mg-content, which favours the growth of olivine and orthopyroxene, and increases the density of the plagioclase-lherzolite by 10-25 kg/m³. This effect shows that syn-rift uplift is dependent on the mantle composition. The post-rift and total subsidence, however, are only weakly dependent on the mantle composition (fig. 2.5). We thus conclude that syn-rift uplift is a consequence of the generally lower density of the plagioclase-lherzolite compared to that of spinel- and garnet-lherzolite, and that this uplift will only occur if the stretching factor δ exceeds a minimum value δ_{crit} (fig. 2.5).

2.4.3 Parameterized density models

Our simulations with realistic density models indicate that the garnet-spinel- plagioclase-lherzolite transitions have the greatest influence on basin subsidence. The simulations also demonstrate that crustal phase transitions do not have a large impact on post-rift subsidence, but that the crustal rock assemblage does influence the syn-rift subsidence since it changes the *mean* density of the crust. The effect of the crustal density model is thus similar to a TDD formulation with a crustal density chosen to correspond to the mean density.

To verify whether the mantle transitions are indeed the most important ones, and to reduce the computed density maps to a simple parameterization, the computed 'real' density distributions have been fit by least-squares (see Appendix A). Partial melting has not been included in the current density-parameterization since the melt is assumed to be rapidly removed by dikes [Rubin, 1993] or a porous flow mechanism [Connolly and Podladchikov, 1998] which changes the bulk chemistry of the rocks and invalidates the model used here.

The 'real' density models are compared with the parameterized density models for different stretching factors, compositions and crustal/mantle thicknesses (fig. 2.6). The maximum error in total, syn-rift and post-rift subsidence between the real mod-

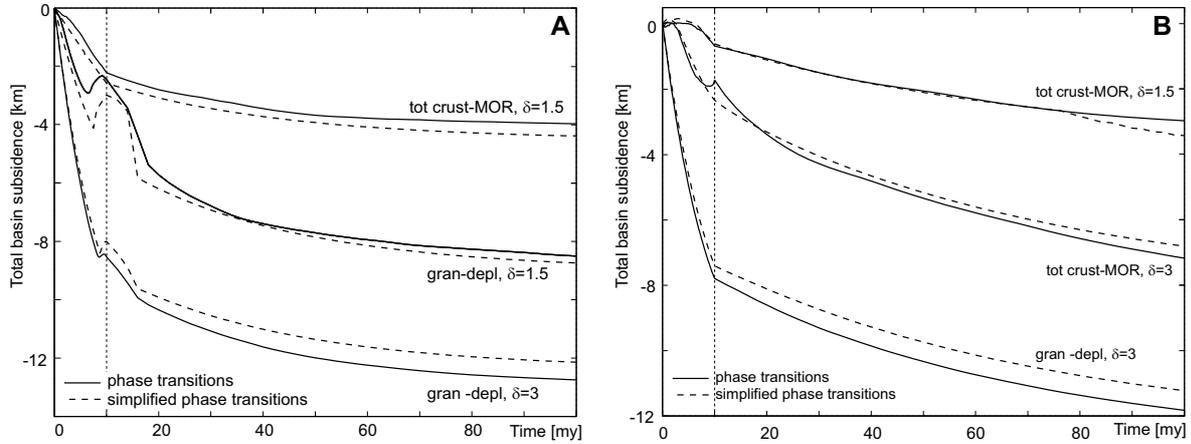


Figure 2.6: Comparison of basin subsidence with real and simplified density models, for a stretching factor of $\delta = 1.5$ and $\delta = 3$, and for a) a crustal thickness of 35 km and a mantle thickness of 70 km and b) a crustal thickness of 35 and a mantle thickness of 110 km.

els and the parameterized models is around 10 %. The parameterized models reproduce the most important features of the real models, including syn-rift-uplift and accelerated post-rift subsidence. The low density of plagioclase-lherzolite compared to spinel/garnet-lherzolite is responsible for syn-rift-uplift and a subsequent phase of accelerated post-rift subsidence (fig. 2.6). Increasing the density of plagioclase-lherzolite reduces uplift. The total subsidence after post-rift subsidence remains unaltered.

2.4.4 Deviations from the TDD pure-shear stretching model

In this section the parameterized density models are used to estimate the maximum syn-rift uplift and post-rift subsidence that can be produced by phase-transition models. The *maximum* effect of phase transitions in the models presented before was obtained for a combination of a 'total-crust' crust and a MOR-pyrolite mantle. The *minimum* effects occurred for a granodioritic crust and a depleted-pyrolite mantle. An estimate of the deviations that phase transitions give with respect to TDD formulation are thus be made by performing systematic computations with both the minimum and the maximum model.

Several thousand computations have been performed, in which the stretching factor and the crustal thickness were varied, but the total lithospheric thickness and the base

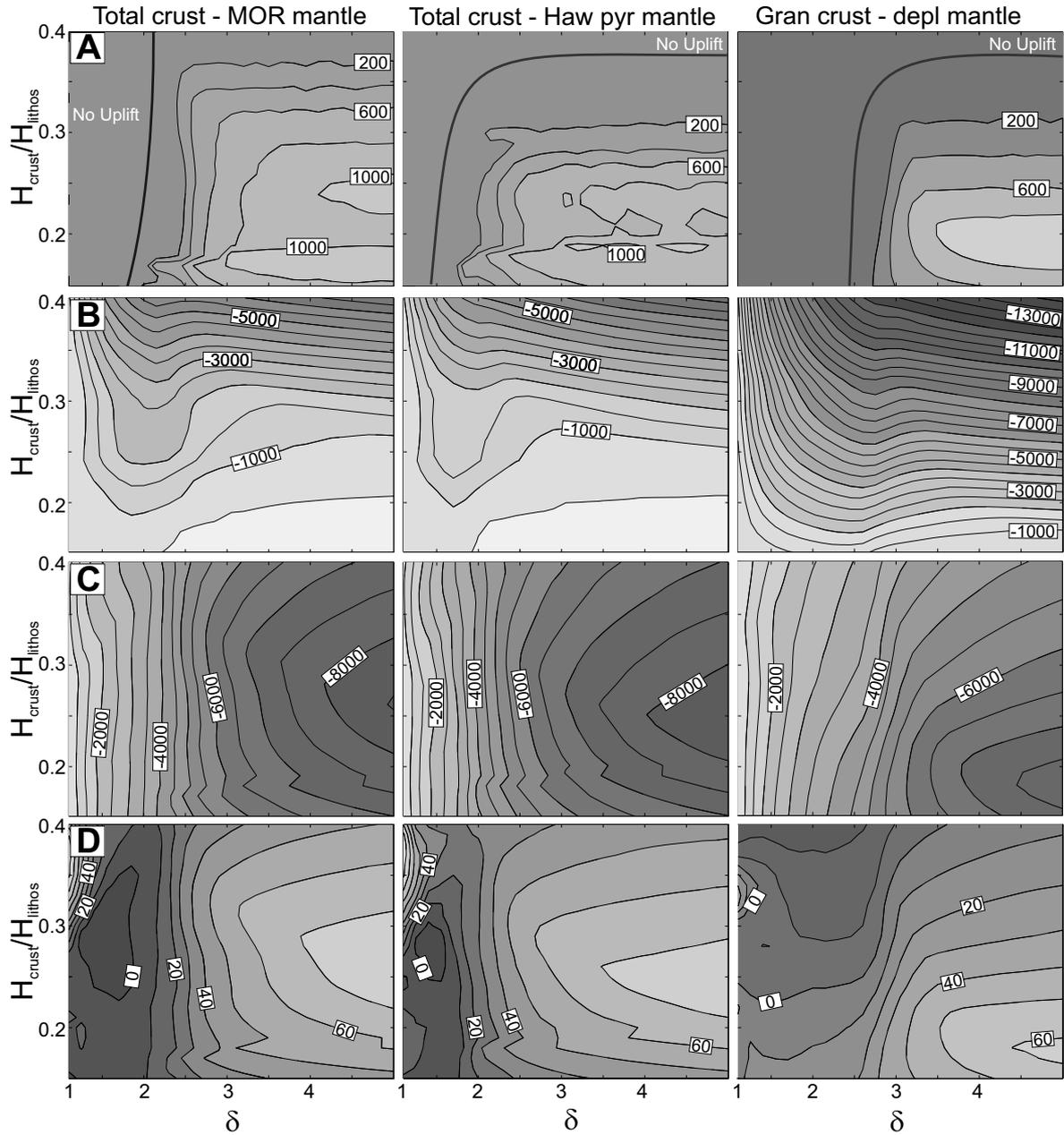


Figure 2.7: A) Contour plot of maximum syn-rift uplift at the end of the rifting phase as a function of stretching factor and crustal thickness normalized over the lithospheric thickness. The model on the left is with a density distribution ('MOR mantle' and 'total crust') which gives the maximum uplift. The model on the right gives the minimum uplift. B) Total syn-rift subsidence (after 10 Myrs of rifting). C) Total thermal subsidence (after 300 myrs) in meters. D) Percentage of additional post-rift subsidence compared to TDD models in which the density distribution is solely temperature-dependent.

temperature were maintained constant (fig. 2.7). Models with $H_{crust}/H_{lithos} < 0.2$ are characterized by a phase of uplift at the beginning of the rifting period, as predicted by previous work [Podladchikov et al., 1994]. This is followed by a phase of subsidence and, in some cases, by an additional uplifting phase (fig. 2.7a). Maximum obtainable syn-rift uplift ranges from 700 m in the minimum model to ~ 1000 m in the maximum model.

Decreasing T_{base} to 1000°C decreases the maximum syn-rift uplift by 30%. In addition, the critical stretching factor to initiate syn-rift uplift is increased from 2 – 2.5 to 3 – 3.5. Increasing T_{base} to 1400°C decreases the critical stretching factor from ~ 2 to 1.5. Increasing the total lithospheric thickness from 110 km to 160 km, increases the critical stretching factor required to initiate syn-rift uplift from 2 to 3, and narrows the range of H_{crust}/H_{lithos} for which syn-rift occurs. If lithospheric thickness is decreased, the H_{crust}/H_{lithos} -range widens from 0.15 to 0.4 but the maximum uplift is decreased by $\sim 20\%$.

Compared to the TDD formulation, phase transitions cause additional post-rift subsidence of at least 20%, and as much as 90% in the maximum model or 50% in the minimum model (fig. 2.7D). Increasing T_{base} to 1400°C increases the subsidence by 20%, whereas decreasing T_{base} to 1000°C halves the difference. Increasing the lithospheric thickness to 160 km increases the maximum deviations of the TDD formulation by 30% and decreasing the thickness to 80 km, decreases the deviations by 30%.

2.5 Discussion and conclusions

Realistic density models (appendix) that are reproduced by simple parameterizations have been implemented in kinematic basin simulations. The simulations demonstrate that, in general, phase transitions reduce syn-rift subsidence and increase post-rift subsidence compared to the uniform stretching (TDD) model, in which density is solely temperature-dependent. The deviations in post-rift subsidence from the TDD model resulting from realistic density models vary from $\sim 20\%$ at low stretching factors to $> 90\%$ at greater stretching factors. Differences in mineralogy may result in different amounts of subsidence, but all the cases considered generate trends that deviate significantly from that predicted by TDD models. In distinction from the TDD model that does not result in syn-rift uplift for reasonable parameters, use of realistic density

model can explain up to 1 km of syn-rift uplift. In contrast to simpler modelling efforts to account for the effect of phase transitions [e.g. Yamasaki and Nakada, 1997], which predicted syn-rift uplift preceding rifting, we obtain uplift during stretching. This syn-rift-uplift is caused by the formation of plagioclase-lherzolite from spinel- and garnet-lherzolite, which occurs during rifting (fig. 2.8). During the thermal post-rift stage, plagioclase-lherzolite is transformed back into spinel- and garnet-lherzolite, thereby causing an additional amount of postrift subsidence. A two-dimensional simulation points out that plagioclase-lherzolite mainly forms below the center of the basin (fig 2.8). Thus syn-rift uplift is to be expected in the center of the basin only. Plagioclase-lherzolite will only be detectable seismically during, or shortly after, an active extension phase.

In our simulations we assumed that phase assemblages are always at equilibrium. This assumption may require explanation. In general, metamorphic reactions are influenced by deformation, presence of volatiles, temperature, etcetera. In crustal rocks, volatiles strongly enhance metamorphic reaction kinetics [e.g. Rubie and Thompson, 1985, Austrheim et al., 1997]. If these volatiles are removed during prograde metamorphism, reaction rates drop and the prograde metamorphic assemblage are preserved [Connolly and Thompson, 1989, Rebay and Powell, 2002]. For anhydrous mantle assemblages, thermal activation is the dominant factor. It has been argued that anhydrous metamorphic reactions occur at geologically short timescales for temperatures above 600°C [Ahrens and Schubert, 1975]. We demonstrate that the mantle reactions have the largest impact on basin subsidence. Typical mantle temperatures are above 600°C (fig. 2.3a), at these temperatures both homogeneous and heterogeneous equilibrium are likely.

In all our models, the accommodation space created by subsidence is completely filled with sediments (i.e. water depth is zero). Once sediments are deposited, they will never be eroded. This implies that during phases of syn-rift uplift, sediments are uplifted above the initial zero-level. We have performed additional simulations in which these sediments were eroded infinitely fast above the zero-level. These simulations resulted in a factor two increase of the amount of syn-rift uplift. The total amount of subsidence, however, was nearly unaffected. Thus the simulations presented here can be regarded as being conservative with respect to the amounts of syn-rift uplift.

For more realistic two- and three-dimensional models the effect of flexural isostasy

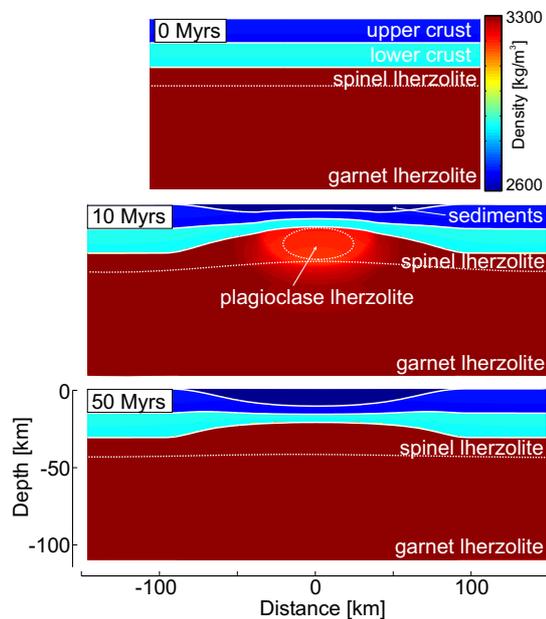


Figure 2.8: Results of 2-D computations with phase transitions activated in the mantle only (MOR-model). The maximum stretching factor in the center of the basin is 3. The upper crust has a T -dependent density with $\rho_0 = 2700 \text{ kgm}^{-3}$ and the lower a T -dependent density with $\rho_0 = 2900 \text{ kgm}^{-3}$. The initial crustal thickness is 35 km; stretching is active for 10 Myrs and the elastic thickness is 0 km. Plagioclase lherzolite appears in the mantle lithosphere during rifting but disappears after rifting has ceased. Modelling results have been obtained with TECMOD2D [www.geomodsol.com, Podladchikov et al., 2002].

has to be taken into account. A lithospheric flexural rigidity or effective elastic thickness [Watts, 2001, Burov and Diament, 1995] can decrease or damp the syn-rift uplift significantly only if the width of the area in which uplift occurs is considerably smaller than the flexural wavelength. Effective elastic thicknesses used for extensional basin modelling are usually small (smaller than around 5 km, e.g. Kusznir and Ziegler [1992], Bellingham and White [2002]) providing flexural wavelengths smaller than around 100 km. Hence, for such values of effective elastic thickness, syn-rift uplift will be only slightly decreased for most basin and passive margin dimensions. The EET is frequently assumed to be a function of the depth of a particular isotherm [Watts, 2001, and references therein]. Syn-rift uplift occurs at the end of rifting, when isotherms are most elevated and the temperature-dependent EET is smallest. Thus phase transitions

might have a considerable effect also in cases with large EET's. Furthermore, lateral heat conduction in two- and three-dimensional models will provide considerable different results only if the thinning factors vary strongly in the lateral direction causing strong lateral temperature variations (fig. 2.8). Rheology may potentially influence our results in two ways: (1) the viscosity of the asthenosphere determines to a large extent how well the lithosphere is isostatically compensated. Large asthenospheric viscosities may reduce and delay the uplift phase. (2) Increasing the effective viscosity of the upper lithosphere will result in larger Deborah numbers and therefore larger effective elastic thicknesses. This may decrease syn-rift uplift as discussed above. However more work in this direction is required to obtain insight in the interaction between phase transitions and rheologically controlled dynamical processes [see e.g. Artyushkov et al., 2000]. In this paper we have focussed on the effect of metamorphic reactions on basin subsidence by comparison with the simplest version of the stretching model. Subsequent modifications of this model such as depth-dependent stretching [Royden and Keen, 1980], depth of necking combined with flexural isostasy and erosion [e.g. Kooi et al., 1992, van Balen et al., 1995], ductile lower crustal flow [Burov and Cloetingh, 1997], magmatic underplating and intraplate stresses can also cause additional subsidence and uplift in e.g. the North Sea [Cloetingh et al., 1990, Kooi et al., 1991]. Most of these models have employed a simple TDD model. We have demonstrated here that the incorporation of more realistic density models may have a considerable effect in itself. Incorporating realistic densities in any of the models above will thus enhance the effects. Depth-dependent stretching caused by ductile crustal flow, for example, would, if combined with metamorphic reactions, require less differential stretching between crust and mantle lithosphere to obtain the same subsidence as depth-dependent stretching models without phase transitions. Phase transitions could in the same way enhance active rifting [Huismans et al., 2001], since it increases the buoyancy of the up-welling mantle dome. It seems unlikely that any of these mechanisms alone can fully explain the formation of sedimentary basins.

Models for extensional basin formation based on kinematic thinning are frequently applied to reconstruct particular observed basin stratigraphies and to estimate the basin's thermal history, which is essential for evaluating petroleum prospects. The thermal history is strongly controlled by the value of the thinning factors used to generate the observed basin subsidence. Our results indicate that models with and

without mineral phase transitions will provide significantly different thinning factors to generate the same amount of syn-rift or post-rift subsidence. Therefore, thermal histories provided by models with and without mineral phase transitions are likely to be considerably different. Reliable thermal history reconstructions should, therefore, compare the thermal histories obtained by models with and without mineral phase transitions.

2.6 Acknowledgements

We thank the Swiss National Fond, project number 21-61912.00, for financial support. YYP and SMS were supported by a Center of Excellence grant to PGP from the Norwegian research council. We thank Taras Gerya, Ed Ghent and Sierd Cloetingh for reviews of the manuscript.

2.7 Simplified phase diagrams

Simplified phase diagrams have been derived from the real phase diagrams by least-squares fitting. For this purpose, the mantle phase diagram has been divided into three parts: spinel-lherzolite, plagioclase-lherzolite and garnet lherzolite. The crustal phase diagram has been divided into two parts, and equations are given in table 2.4. The mantle phase diagram has been divided into three parts (table 2.5). The maximum error in density is around 1% for mantle compositions and around 6% for crustal compositions. The parameterized phase diagram is valid for the temperatures and pressures that are indicated. The density of rocks at lower temperatures than 200°C is assumed to be the density at a temperature of 200°C, since metamorphic reactions become too slow and our equilibrium phase diagram is no-longer applicable.

Composition	Low- P region			High- P region		
	ρ_0	α $\times 10^{-5}$	β $\times 10^{-11}$	ρ_0	α $\times 10^{-5}$	β $\times 10^{-11}$
total crust	2835	2.82	3.46	3060	1.96	1.34
granodioritic	2683	8.57	4.85	2918	4.11	2.09

Table 2.4: The simplified crustal phase diagram is composed of two regions. Density in each region can be computed with $\rho = \rho_0 (1 - \alpha T + \beta P)$, with P in Pa and T in °C. The boundary between the two regions is given by $P = 1.92 \times 10^6 T - 1.6 \times 10^7$ for the total crust and by $P = 2 \times 10^6 T - 3.2 \times 10^8$ for a granodioritic crust.

$\rho = \rho_0 (1 - \alpha T + \beta P)$ Pyrolite Composition	Spinel lherzolite			Plagioclase lherzolite			Garnet lherzolite		
	ρ_0	α $\times 10^{-5}$	β $\times 10^{-11}$	ρ_0	α $\times 10^{-5}$	β $\times 10^{-11}$	ρ_0	α $\times 10^{-5}$	β $\times 10^{-11}$
Hawaiian	3369	3.87	9.54	3388	3.54	1.34	3270	3.67	1.65
MOR	3363	3.87	8.62	3387	3.54	1.34	3249	3.39	1.45
Depleted	3353	3.88	9.25	3371	3.56	1.34	3275	3.51	1.50
$P = aT^2 + bT + c$ Composition	Plagioclase-Spinel boundary			Spinel-Garnet boundary					
	a	$b \times 10^6$	$c \times 10^8$	a	$b \times 10^6$	$c \times 10^9$			
Hawaiian	-345	1.96	-4.4	1825	-2.7	2.2			
MOR	-460	1.99	-4.4	1946	-3.1	2.4			
Depleted	-435	1.91	-4.1	1660	-2.9	2.4			

Table 2.5: The simplified mantle diagram is composed of three regions. Fitted functions for density in each region are indicated as well as the boundaries between the different regions.

Bibliography

- T. J. Ahrens and G. Schubert. Gabbro-eclogite reaction rate and its geophysical significance. *Review of Geophysics and Space Physics*, 13:383–400, 1975.
- E. V. Artyushkov, N.-A. Moerner, and D.H. Tarling. The cause of loss of lithospheric rigidity in areas far from plate tectonic activity. *Geophysical Journal International*, 143:752–775, 2000.
- H. Austrheim, M. Erambert, and A.K. Engvik. Processing of crust in the root of the caledonian continental collision zone: the role of eclogitization. *Tectonophysics*, 273: 129–153, 1997.
- C. Beaumont, C.E. Keen, and R. Boutilier. On the evolution of rifted continental margins: Comparison of models and observations for the Nova Scotia margin. *Geophysical Journal of the Royal Astronomical Society*, 70:667–715, 1982.
- P. Bellingham and N. White. A two-dimensional inverse model for extensional sedimentary basins - 2. application. *Journal of Geophysical Research-Solid Earth*, 107 (B10):art. no.–2260, 2002.
- G.J. Borse. *Numerical methods with MATLAB . A resource for scientists and engineers*. PWS Publishing Company, Boston, 1997.
- E. Burov and S. Cloetingh. Erosion and rift dynamics: new thermomechanical aspects of post-rift evolution of extensional basins. *Earth and Planetary Science Letters*, 150:7–26, 1997.
- E. Burov and A. Poliakov. Erosion and rheology controls on synrift and postrift evolution: Verifying old and new ideas using a fully coupled numerical model. *Journal of Geophysical Research-Solid Earth*, 106(B8):16461–16481, 2001.

- E. B. Burov and M. Diament. The effective elastic thickness (T_e) of continental lithosphere: What does it really mean? *Journal of Geophysical Research*, 100(B1): 3905–3927, 1995.
- N.D. Chatterjee and E. Froese. A thermodynamic study of the pseudo-binary join muscovite-paragonite in the system $KAlSi_3O_8$ - $NaAlSi_3O_8$ - Al_2O_3 - SiO_2 - H_2O . *American Mineralogist*, 60:985–993, 1975.
- S. Cloetingh, F. Gradstein, H. Kooi, A. C. Grant, and M. Kaminsky. Plate reorganisation: a cause of rapid late Neogene subsidence and sedimentation around the North Atlantic? *Journal of the Geological Society London*, 147:495–506, 1990.
- J. A. D. Connolly. Multivariable phase-diagrams - an algorithm based on generalized thermodynamics. *American Journal of Science*, 290(6):666–718, 1990.
- J. A. D. Connolly and K. Petrini. An automated strategy for calculation of phase diagram sections and retrieval of rock properties as a function of physical conditions. *Journal of Metamorphic Geology*, 20(7):697–708, 2002.
- J. A. D. Connolly and Y. Y. Podladchikov. Compaction-driven fluid flow in viscoelastic rock. *Geodinamica Acta*, 11(2-3):55–84, 1998.
- J. A. D. Connolly and A. B. Thompson. Fluid and enthalpy production during regional metamorphism. *Contributions to Mineralogy and Petrology*, 102(3):347–366, 1989.
- M.S. Ghiorso, M.M. Hirschmann, P.W. Reiners, and V.C. Kress. The pMELTS: A revision of MELTS for improved calculation of phase relations and major element partitioning related to partial melting of the mantle to 3GPa. *Geochemistry Geophysics Geosystems*, 3(5):doi. 10.1029/2001GC000217, 2002.
- D. H. Green, W.O. Hibberson, and A.L. Jaques. Petrogenesis of mid-oceanic ridge basalt. In M.W. Mc-Elhinny, editor, *The earth: it's origin, structure and evolution*, pages 265–299. Academic press, 1979.
- D.G. Green and T.J. Falloon. Pyrolite: A Ringwood concept and its current expression. In I. Jackson, editor, *The Earth's Mantle. Composition, structure and evolution*, page 566. Cambridge University Press, Cambridge, 1998.

- Y. Hamdani, J. C. Mareschal, and J. Arkani-Hamed. Phase changes and thermal subsidence in intracontinental sedimentary basins. *Geophysical Journal International*, 106:657–665, 1991.
- Y. Hamdani, J. C. Mareschal, and J. Arkani-Hamed. Phase change and thermal subsidence of the Williston basin. *Geophysical Journal International*, 116:585–597, 1994.
- T. Holland, J. Baker, and R. Powell. Mixing properties and activity-composition relationships of chlorites in the system $\text{MgO-FeO-Al}_2\text{O}_3\text{-SiO}_2\text{-H}_2\text{O}$. *European Journal of Mineralogy*, 10:395–406, 1998.
- T. Holland and R. Powell. Thermodynamics of order-disorder in minerals .2. symmetric formalism applied to solid solutions. *American Mineralogist*, 81:1425–1437, 1996.
- T. Holland and R. Powell. An internally consistent thermodynamic dataset for phases of petrological interest. *Journal of Metamorphic Geology*, 16:309–343, 1998.
- R. S. Huismans, Y. Y. Podladchikov, and S. Cloetingh. Transition from passive to active rifting: Relative importance of asthenospheric doming and passive extension of the lithosphere. *Journal of Geophysical Research-Solid Earth*, 106(B6):11271–11291, 2001.
- G.C. Kennedy. The origin of continents, mountain ranges and ocean basins. *American Scientist*, 97:491–504, 1959.
- H. Kooi, S. Cloetingh, and J. Burrus. Lithospheric necking and regional isostasy at extensional basins. 1. subsidence and gravity modelling with an application to the gulf of lions margin (se france). *Journal of Geophysical Research*, 97:17553–17571, 1992.
- H. Kooi, M. Hettema, and S. Cloetingh. Lithosphere dynamics and the rapid pliocene-quadernary subsidence phase in the southern north sea basin. *Tectonophysics*, 192: 245–259, 1991.
- N. J. Kusznir and P. A. Ziegler. The mechanics of continental extension and sedimentary basin formation - a simple-shear pure-shear flexural cantilever model. *Tectonophysics*, 215(1-2):117–131, 1992.

- L. I. Lobkovsky, S. Cloetingh, A. M. Nikishin, Y. Volozh, A. Lankreijer, A. Belyakov, V.G. Groshev, P. A. Fokin, E.E. Milanovsky, L.A. Pevzner, V. I. Gorbachev, and M.A. Korneev. Extensional basins of the former Soviet Union - structure, basin formation mechanisms and subsidence history. *Tectonophysics*, 266:251–285, 1996.
- L. I. Lobkovsky, A. Ismail-zadeh, B. M. Naimark, A. M. Nikishin, and S. Cloetingh. Mechanism of subsidence of the erath crust and sedimentary basin formation (in russian). *Doklady Akademii Nauk Sssr*, 330(2):256–260, 1993.
- J.R. Lovering. The nature of the Mohorovic discontinuity. *Transaction of the American Geophysical Union*, 39:947–955, 1958.
- D. McKenzie. Some remarks on the development of sedimentary basins. *Earth and Planetary Science Letters*, 40:25–32, 1978.
- M.F. Middleton. A model of intracratonic basin formation entailing deep crustal metamorphism. *Geophysical Journal of the Royal Astronomical Society*, 62:1–14, 1980.
- H. Neugebauer and T. Spohn. Late stage development of mature Atlantic-type continental margins. *Tectonophysics*, 50:275–305, 1978.
- R.C. Newton, T.V. Charlu, and O.J. Kleppa. Thermochemistry of the high structural state plagioclases. *Geochemica et Cosmochimica Acta*, 44:933–941, 1980.
- R. J. O’Connell and G.J. Wasserburg. Dynamics of submergence and uplift of a sedimentary basin underlain by a phase change boundary. *Review of Geophysics*, 10: 335–368, 1972.
- K. Petrini, J. A. D. Connolly, and Y. Y. Podladchikov. A coupled petrological-tectonic model for sedimentary basin evolution: the influence of metamorphic reactions on basin subsidence. *Terra Nova*, 13(5):354–359, 2001.
- Y. Podladchikov, S. M. Schmalholz, and D. W. Schmid. Inverse modelling of sedimentary basins. In *Progress in industrial mathematics at ECMI 2000*, pages 625–629. Springer-Verlag, Berlin, 2002.

- Y. Y. Podladchikov, A. N. B. Poliakov, and D. A. Yuen. The effect of lithospheric phase-transitions on subsidence of extending continental lithosphere. *Earth and Planetary Science Letters*, 124(1-4):95–103, 1994.
- G. Rebay and R. Powell. The formation of eclogite facies metatroctolites and a general petrogenetic grid in $\text{Na}_2\text{O}-\text{CaO}-\text{FeO}-\text{MgO}-\text{Al}_2\text{O}_3-\text{SiO}_2-\text{H}_2\text{O}$ (NCFMASH). *Journal of Metamorphic Geology*, 20(9):813–826, 2002.
- S. Ren, J.I. Faleide, O. Eldholm, J. Skogseid, and F. Gradstein. Late Cretaceous-Paleocene tectonic development of the NW vøring basin. *Marine and Petroleum Geology*, 20:177–206, 2003.
- A.E. Ringwood. A model for the upper mantle. *Journal of Geophysical Research*, 67:857–866, 1962.
- L. Royden and C.E. Keen. Rifting process and thermal evolution of the continental margin of eastern Canada determined from subsidence curves. *Earth and Planetary Science Letters*, 51:343–361, 1980.
- D. C. Rubie and A. B. Thompson. Kinetics of metamorphic reactions at elevated temperatures and pressures. In A. B. Thompson and D. C. Rubie, editors, *Metamorphic Reactions: Kinetics, Textures and Deformation*, pages 27–79. Springer Verlag, New York, 1985.
- A. M. Rubin. Dikes vs diapirs in viscoelastic rock. *Earth and Planetary Science Letters*, 119(4):641–659, 1993.
- J. G. Sclater, L. Royden, F. Horvath, B. Burchfield, S. Semken, and L. Stegena. The formation of intra-Carpathian basins as determined from subsidence data. *Earth and Planetary Science Letters*, 51:139–162, 1980.
- S. V. Sobolev and A. Y. Babeyko. Modeling of mineralogical composition, density and elastic-wave velocities in anhydrous magmatic rocks. *Surveys in Geophysics*, 15(5):515–544, 1994.
- G. Spadini, A. Robinson, and S. Cloetingh. Thermo-mechanical modelling of Black Sea formation, subsidence and sedimentation. In A. Robinson, editor, *Regional and petroleum geology of the Black Sea and surrounding areas*, pages 19–38. 1997.

- T. Spohn and H. Neugebauer. Metastable phase transition models and their bearings on the development of Atlantic-type geosynclines. *Tectonophysics*, 50:387–412, 1978.
- S.R. Taylor and S.M. McLennan. *The continental crust: its composition and evolution*. Blackwell Scientific Publications, 1985.
- R. T. van Balen, P.A. van der Beek, and S.A.P.L. Cloetingh. The effect of rift shoulder erosion on stratal patterns at passive margins: Implications for sequence stratigraphy. *Earth and Planetary Science Letters*, 134:527–544, 1995.
- I.M. Walker, K.A. Berry, J.R. Bruce, L. Bystol, and J.H. Snow. Structural modelling of regional depth profiles in the Vøring Basin: implications for the structural and stratigraphic development of the Norwegian passive margin. *Journal of the Geological Society*, 154:537–544, 1997.
- A. B. Watts. *Isostasy and flexure of the lithosphere*. Cambridge University Press, 1 edition, 2001.
- K. H. Wedepohl. The composition of the continental crust. *Geochemica et Cosmochimica Acta*, 59(7):1217–1232, 1995.
- B. J. Wood and J. R. Holloway. A thermodynamic model for subsolidus equilibria in the system CaO-MgO-Al₂O₃-SiO₂. *Geochemica et Cosmochimica Acta*, 48:159–176, 1984.
- B. J. Wood and D. Yuen. The role of lithospheric phase transitions on seafloor flattening at old ages. *Earth and Planetary Science Letters*, 66:303–314, 1983.
- T. Yamasaki and M. Nakada. The effects of the spinell-garnet phase transition on the formation of rifted sedimentary basins. *Geophysical Journal International*, 130: 681–692, 1997.
- P. A. Ziegler and N. Cloetingh. Dynamic processes controlling evolution of rifted basins. *Earth-Science Reviews*, 64:1–50, 2004.

Chapter 3

The interplay between folding, diapirism and erosion in compressional settings¹

Abstract

Domes and basins are evidence for vertical movements in both compression and extension tectonic environments. They are thus evidence for interplay between gravity and tectonic forces in structuring the continental crust. We employ analytical and numerical techniques to investigate the respective roles of gravity and compression during the growth of crustal-scale buckle anticlines and diapirs submitted to instantaneous erosion. The analytical perturbation method, which explores the onset of both types of instability, yields a 'phase-diagram' discriminating eight folding-diapirism modes, five of which are geologically relevant. Numerical simulations show that the phase diagram is applicable to evolved, finite amplitude stages. Calculated strain fields in both diapirs and folds show normal sense of shear at the interface, if the upper layer is thick compared to the lower layer. We conclude that the present-day structural techniques applied for distinguishing diapiric domes and folds are ambiguous if detachment folding and intense erosion take place during deformation, and that domes displaying extensional structures at their periphery do not necessarily reflect extension.

¹Part of this work has been published in : Burg, J.P., Kaus, B.J.P, Podladchikov, Y.Y. (2004) Dome structures in collision orogens. Mechanical investigation of the gravity/compression interplay, in Whitney, D.L., Teyssier, C., and Siddoway, C.S., *Gneiss domes in orogeny: Geological Society of America Special Paper* 380, p. 47-66.)

3.1 Introduction

Collision mountain systems are long, linear to arcuate belts at the Earth's surface. In these mountains, abundant folds and thrusts reflect regional shortening. Isostatic considerations, gravimetric studies and seismic information show that horizontal shortening is intrinsically related to crustal thickening and it requires 5 to 7 km of crustal root to balance each km of mountain range above sea level. In other words, a mountain grows 5 to 7 times more downward than upward. Consequently, collision-mountains are sites where the continental crust is buried, thus subjected to intense metamorphism and igneous activity. Thickening of the buoyant crust and subsequent uplift create a high topography. The mountain belt becomes also a region of erosion, which digs out deep crustal levels and supplies sedimentary basins. The long-term process results in ancient orogens being levelled to flatlands that expose metamorphic and magmatic rock associations, those that were part of the mountain roots. Typically, these high-grade metamorphic regions display large closed structures termed domes and basins. Application of plate tectonics at understanding collision orogens has focused on horizontal relative movements because convergence is one to two orders of magnitude larger than orogenic vertical movements. Horizontal transport is classically inferred from recumbent folds and thrust systems. Domes and basins are evidence for vertical movements (e.g. Brun [1983]). Five distinct origins have been, postulated:

- Folding: Many domes may be double plunging anticlines and/ or culminations of crossing anticlines of two separate generations and different trends [Ramsay, 1989, Snowden and Bickle, 1976].
- Diapirism of igneous intrusions: A usual explanation of blob-like geological map patterns in high-grade metamorphic terranes postulates that mobilized rock masses rose buoyantly in the core of domes in response to the gravity instability resulting from the low density and viscosity of the granitic core below the denser and stronger wall rocks [e.g. Ramberg, 1972, 1981, Brun et al., 1981, Ramsay, 1989]. This mechanism has been modelled mathematically [Ramberg, 1972, Fletcher, 1972] and experimentally [e.g. Dixon, 1975].
- Reactivation of basement plutons: The development of the type mantle gneiss domes, in Finland [Eskola, 1949] invokes two orogenic events. During the first

orogeny, granite plutons were emplaced in metasediments and metavolcanites. Erosion exposed the plutons and country rock, which were covered by a younger sequence of sediments. Injection of new magma during the second orogeny reactivated the old plutons, causing them to expand upwards and thereby fold the overlying strata into domes and basins. We consider this interpretation as a restrictive case of polyorogenic yet superposed diapirism and/or folding.

- Extensional culminations: rocks of mid-crustal levels are brought to shallower levels by tectonic denudation and erosion, horizontal extension along major, shallow-dipping detachments scraping away the overlying cover [e.g. Coney and Harms, 1984]
- Upward impingement: A strong or rigid basement block forces bending of its plastic cover [Gzovsky et al., 1973].

These five mechanisms are not mutually exclusive. For example, fold interference is not incompatible with gravity instability since both mechanisms could operate synchronously, in particular where density contrast is invariably present between core and surrounding rocks [Snowden and Snowden, 1981]. These five mechanisms produce upward movement of lower and mid-crustal levels during orogeny, but refer to different force systems since compression and extension are predominantly horizontal forces of opposite sign and diapirism versus impingement involve predominantly vertical forces acting with and against gravity, respectively. By chance for geologists, they develop symptomatic structural features that allow identifying which mechanism was dominant [Brun, 1983]. In particular, extensional core complexes display a marked asymmetry of metamorphic grade and ages contrasting with the symmetry in folds and diapirs (Fig. 3.1). Our aim is to discuss the growth and the mechanical characteristics of two types of domes in the light of analytical solutions and two-dimensional numerical codes: (1) large upright folds for which upward amplification is fundamentally a response to horizontal compression; (2) magmatic bodies for which diapiric (i.e. piercing) rise controlled by the vertical gravity seems to play a significant role. Extensional metamorphic core complexes in which rocks of middle crustal levels are uplifted and exposed by a process dominated by large offsets along low-angle normal faulting are not discussed in this work because mechanical insight has been given by Lavier and Buck [2002].

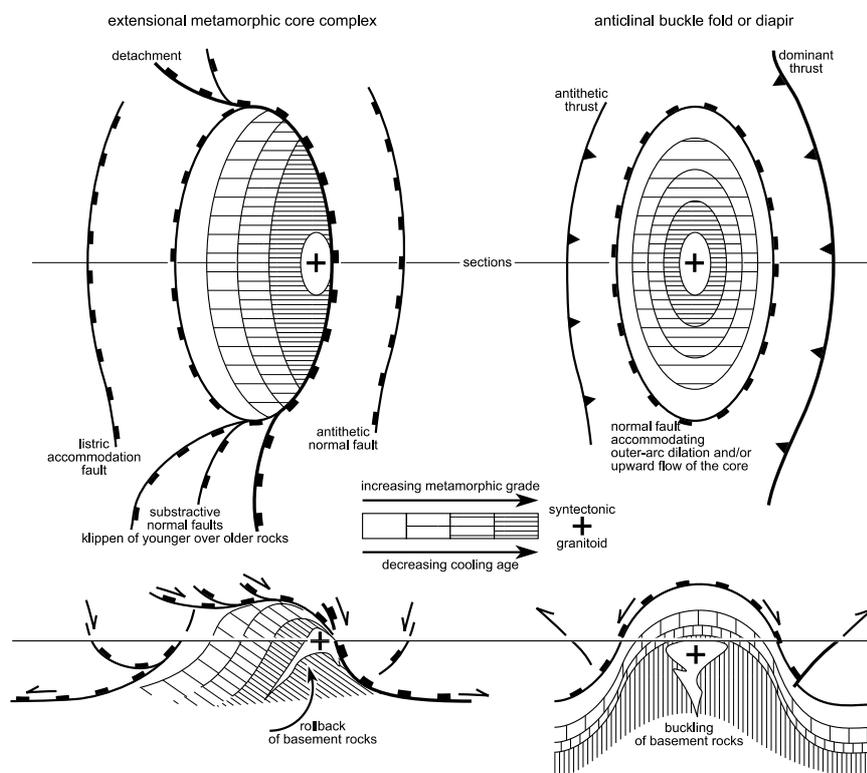


Figure 3.1: Synthetic maps (top) and sections (below) showing the symmetry and asymmetry that characterize folds and diapirs on the one hand, and extensional metamorphic core complexes on the other hand.

3.2 Crustal-scale folds

3.2.1 Geological information

Although geologists have accepted the existence of large recumbent folds (fold nappes) with several tens of kilometers long inverted limbs [e.g. Arthaud, 1970, Ramsay, 1981], they resisted the concept of big buckle folds, with the intuitive belief that high amplitude buckles could not stand against gravity [e.g. Ramberg, 1970]. However, the concept becomes valid in places where erosion can behead crustal-scale anticlines during their growth, thus eliminating the height and relief problem. This is particularly the case in the Himalayan syntaxes for which the mechanical consistency of the erosion condition has been tested numerically [Burg and Podladchikov, 1999, 2000]. In these

very active Himalayan regions, buckle folding has allowed exhumation of 30 km deep rocks within less than 5 Myr [Zeitler et al., 1993, Burg et al., 1997, 1998]. These neotectonics examples bring support to crustal fold interpretation of older domes such as in the Variscides [Stipska et al., 2000] and some mantled gneiss domes [Snowden and Bickle, 1976], although the folding perception is gone out of fashion. The apparent lack of periodically spaced anticlines neighboring domes interpreted as folds is a common criticism of the buckle interpretation since smaller scale simulations emphasize periodicity in buckling [e.g. Currie et al., 1962, Smith, 1977]. The relative isolation of large-scale folds is a puzzling singularity on which we will comment in the light of the numerical modelling presented here.

3.2.2 Mechanical background

A considerable body of work has shown, both theoretically and experimentally, that if a thin layer undergoing layer-parallel shortening is more competent (i.e. stiffer) than the surrounding material, this condition is unstable and buckling as an instability of the stiff layer will occur, while the entire system is deforming in pure shear [e.g. Price and Cosgrove, 1990]. Early work focused on the analysis of buckling of a layer, either elastic or viscous, in an infinite viscous matrix taking into account a simple linear relationship between stress and strain or stress and strain-rate [Smoluchowski, 1910, Biot, 1961, Ramberg, 1964, Ramberg and Stephansson, 1964]. Modelling relevant to lithospheric-scale deformation, assumes that a powerlaw viscous layer of thickness H (representing the crust) floats on a viscous halfspace (representing the mantle) (Fig. 3.2A). We used a setup similar to that of Schmalholz et al. [2002], and we additionally implemented fast erosion at the top surface and considered an inverse density contrast between the layer and the underlying halfspace. If the viscous lower halfspace does not exert any shear stress on the layer, the thin-plate theory applies [e.g. Reddy, 1999].

$$\frac{\mu_1}{3n} \frac{\partial^2}{\partial x^2} \left(H_l(x)^3 \frac{\partial^3 W(x, t)}{\partial x^2 \partial t} \right) + \frac{\partial}{\partial x} \left(P H_l(x) \frac{\partial W^m(x, t)}{\partial x} \right) - (\sigma_{Ntop} - \sigma_{Nbot}) = 0 \quad (3.1)$$

where μ_1 is the effective viscosity of the upper layer, n its power-law exponent and P is the mean viscous layer-parallel stress, given by [e.g. Turcotte and Schubert, 2002, Schmalholz et al., 2002]:

$$P = 4\mu_1 \dot{\epsilon}_{BG} \quad (3.2)$$

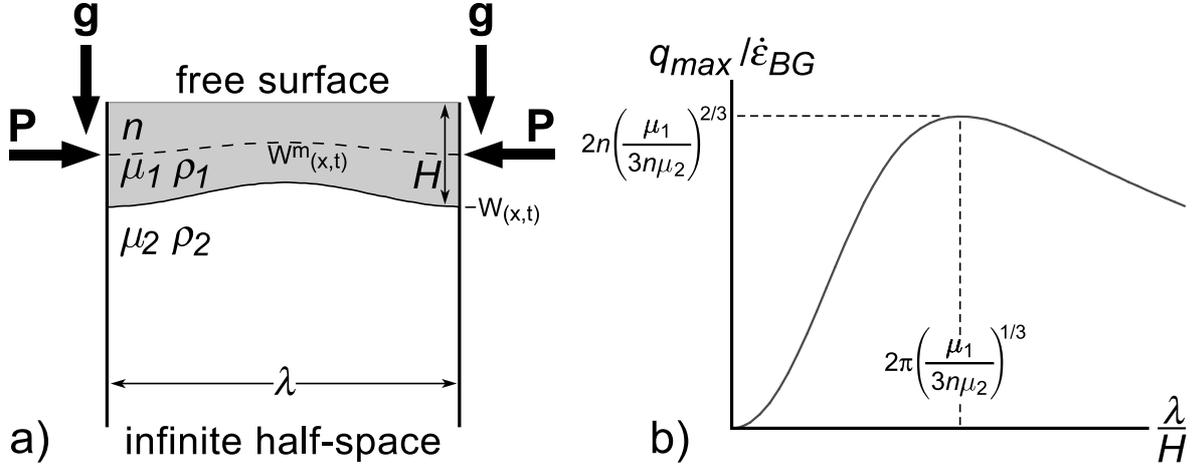


Figure 3.2: A) Setup for analytical investigation of a powerlaw viscous plate floating on a linearly viscous halfspace. B) Growth rate plotted versus normalized wavelength for the case without gravity. Symbols as in the text.

where $\dot{\epsilon}_{BG}$ is the background pure shear shortening rate.

$W(x, t)$ describes the deflection of the lower boundary of the layer, which is assumed to be sinusoidal with a time-dependent amplitude $A(t)$:

$$W(x, t) = A(t)\sin(\omega x) \quad (3.3)$$

where $\omega = 2\pi/\lambda$ is the wavenumber.

$W^m(x, t) = W(x, t)/2$ is the deflection of the middle line of the layer. $H_l(x)$ is the thickness of the layer:

$$H_l(x) = H - W(x, t) = H - A(t)\sin(\omega x) \quad (3.4)$$

where H is the mean thickness of the layer.

Finally σ_{Ntop} and σ_{Nbot} are the vertical forces exerted at the top and bottom of the layer, respectively. The top is kept flat, so $\sigma_{Ntop} = 0$. The layer bottom is deflected and has a density contrast, so the vertical forces that act on the layer are due to the gravitational load and the viscous drag of the underlying halfspace with viscosity μ_2 [e.g. Turcotte and Schubert, 2002, equation 6-165, page 251]:

$$\sigma_{Nbot} = 2\mu_2 \omega \frac{\partial W(x, t)}{\partial t} - (\rho_1 - \rho_2)gW(x, t) \quad (3.5)$$

Substituting (3.3), (3.4) and (3.5) into (3.1), and keeping only linear terms (omitting $A(t)^2$ terms, since $A(t)$ is assumed to be small) gives a linear ordinary differential equation (ODE) for $A(t)$:

$$\frac{\mu_1 H^3 \omega^4}{3n} \frac{\partial A(t)}{\partial t} - \frac{P}{2} H \omega^2 A(t) + 2\mu_2 \omega \frac{\partial A(t)}{\partial t} - (\rho_1 - \rho_2) g A(t) = 0 \quad (3.6)$$

The solution of this ODE has the form $A(t) = A_0 e^{(qt)}$ where q is the growth rate, given by:

$$q = \frac{3n \left(\frac{PH\omega^2}{2} + (\rho_1 - \rho_2)g \right)}{\omega (\mu_1 H^3 \omega^3 + 6\mu_2 n)} \quad (3.7)$$

Substituting P from equation (3.2) into equation (3.7) yields:

$$\frac{q}{\dot{\epsilon}_{BG}} = \frac{3n \left(2\mu_1 H \omega^2 + \frac{(\rho_1 - \rho_2)g}{\dot{\epsilon}_{BG}} \right)}{\omega (\mu_1 H^3 \omega^3 + 6\mu_2 n)} \quad (3.8)$$

Setting the gravity term g to zero yields a solution for folding without gravity. Plotting $q/\dot{\epsilon}_{BG}$ versus λ/H for this case shows that the growth rate has a single maximum (Fig. 3.2B) that can be found by setting the derivative of equation (3.7) versus ω to zero and solving for ω . The only positive solution for ω yields the so-called dominant wavelength [Biot, 1961, Ramberg and Stephansson, 1964]:

$$\lambda_{dom} = 2\pi H \left(\frac{\mu_1}{3n\mu_2} \right)^{\frac{1}{3}} \quad (3.9)$$

with the corresponding growth rate:

$$q_{max} = 2n \left(\frac{\mu_1}{3n\mu_2} \right)^{\frac{2}{3}} \dot{\epsilon}_{BG} \quad (3.10)$$

Equation (3.9) shows that the dominant wavelength is dependent on the thickness of the layer and on the viscosity contrast between the upper and lower layers, a parameter that has a large amount of uncertainty in nature. Equation (3.10) shows that the growth rate is essentially dependent on the viscosity contrast. Note that the expressions for dominant wavelength and growthrate have a factor 3 in the denominator, which is appropriate for folding of one interface only.

It is interesting to compare these expressions with the solutions obtained for a folded upper surface without erosion, [Schmalholz et al., 2002]. The expressions for dominant

wavelength are the same, but the dominant growth rate is twice smaller under fast erosion. This counter-intuitive result, which was confirmed by more complete thick-plate analysis, can be understood by noting that the folding instability is driven through deflections on interfaces between different layers. Since there is only one deflected interface in this set-up, the rate at which the instability grows is twice slower than when two interfaces are present. Another point worth noting is the effect of gravity. It can be seen from equation 3.7 that a stable density configuration ($\rho_1 < \rho_2$) will decrease the growth rate, and a rather large density contrast will eliminate the folding instability. If the density configuration is unstable ($\rho_1 > \rho_2$), calculations show that the dominant growth rate and wavelength go both to infinity.

One should note that the above derivation has the implicit assumption that folds are periodic by inserting a sinusoidal interface deflection (Eq. 3.3). However, the introduction of nonlinear effects for example in the matrix can result in a more localized type of folding [e.g. Hunt et al., 1996]. In addition, both laboratory [Abbassi and Mancktelow, 1992] and numerical [Zhang et al., 2000] experiments, pointed out the strong localizing effect that finite amplitude initial perturbations can have on the distribution and morphology of buckle folds.

The analysis presented above has been extended to large-scale problems by including the effects of a finite, low-viscous bottom layer called "matrix", a denomination we will keep in this work to stay in line with earlier literature. Three different folding modes exist depending on both the thickness of the lower viscous layer and the efficiency of gravitational versus compressional forces [Schmalholz et al., 2002]. For example, gravity plays the dominant role in the lithospheric buckling of Central Asia, whereas the thickness of the underlying 'soft' layer controlled buckling in the Zagros and Jura mountains. In the following part of this work, we further employ the setup described above [see also Goff et al., 1996].

3.3 Magmatic/diapiric domes

3.3.1 Geological Information

Diapirism and intrusion are processes involved when a geological formation (the source layer) has come under sufficient stress (including gravity driven components) to flow,

pierce and break through overlying strata of higher density and strength. Magmas commonly have densities lower than those of the overlying rocks and consequently tend to ascend through passageways or zones of weakness. Most magma does not reach the Earth's surface but crystallizes at depth to form plutonic bodies of igneous rocks. If plutons are diapirs (a concept still disputed) magma stops rising where surrounding rocks have lower density and/or at the temperature-equivalent depth where magma cools and solidifies [e.g. Vigneresse and Clemens, 2000, Burov et al., 2003]. In fact, the coeval emplacement at similar depth of magmas with different composition indicates that there is no neutral buoyancy level in the crust. Accordingly, purely gravity-driven igneous diapirism should not exist [e.g. Vigneresse and Clemens, 2000]. The abundance of pluton-cored domes in orogens, in particular in the European Variscides [e.g. Zwart, 1967] points to the participation of any tectonic deformation during magma ascent [Brun et al., 1981, Vigneresse and Clemens, 2000]. A typical history comprises a deformation-controlled mechanical instability that becomes the location from which buoyant upwelling of relatively light magma starts while "ballooning" characterizes final emplacement [Pitcher, 1979, Pons et al., 1992]. A similar history is invoked for salt tectonics [e.g. Jackson and Talbot, 1989, Poliakov et al., 1996] and many migmatite domes [see discussion in Teyssier and Whitney, 2002]. Diapiric ascent thus most often occurs in a regional tectonic setting whose actual role requires new investigations.

3.3.2 Mechanical background

Here, we only summarize statements from previous work. The driving force of diapirism is the density inversion [Biot and Odé, 1965]. The spontaneous rise of buoyant domes into a denser overburden in the Earth's gravity field is strongly inspired by salt tectonics [e.g. Talbot and Jackson, 1987]. In its simplest form, the rock system consists of two horizontal layers, each of which has uniform thickness, density and Newtonian viscosity [e.g. Woidt, 1978]. Whether viscosities are equal or different, the setting is unstable if the overlying layer is denser than the lower one and the interface separating the two layers is not perfectly flat. The gravitational instability of a heavy fluid overlying a lighter fluid is named Rayleigh-Taylor instability and is able to grow to finite amplitude, independent of background shortening or extension.

Dimensional analysis allows gaining first insight in the basic parameters that control

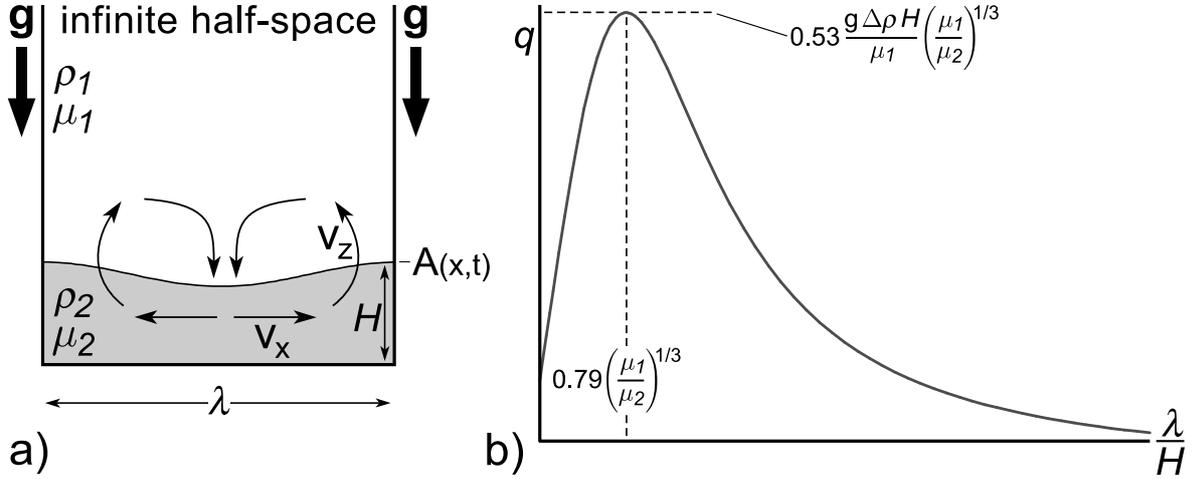


Figure 3.3: A) Setup for analytical investigation of a low density viscous fluid underlying a higher density viscous halfspace. B) Growth rate plotted versus normalized wavelength.

the Rayleigh-Taylor instability [Whitehead, 1988]. A layer of thickness H , density ρ_2 and viscosity μ_2 underlies an infinite viscous halfspace (see Fig. 3.3A). The horizontal velocity v_x in the layer is much larger than the vertical velocity v_z . Force balance in the layer is given by the lateral pressure gradient p :

$$\frac{p}{\lambda} = \mu_2 \frac{v_x}{H^2} \quad (3.11)$$

where λ is the wavelength of the sinusoidal perturbation on the interface separating the two fluids. The upper halfspace with a viscosity μ_1 and density ρ_1 will have a force balance between buoyancy, vertical velocity and viscosity:

$$\frac{p}{\lambda} = \mu_1 \frac{v_z}{\lambda^2} + \frac{(\rho_1 - \rho_2)gA(t)}{\lambda} \quad (3.12)$$

where the density difference between the two layers is assumed to be > 0 , g the gravitation acceleration and $A(t)$ the time-dependent amplitude of the perturbation. Combining equations (3.11) and (3.12) gives:

$$\mu_2 \frac{v_x}{H^2} - \mu_1 \frac{v_z}{\lambda^2} = \frac{(\rho_1 - \rho_2)gA(t)}{\lambda} \quad (3.13)$$

Assuming further that the fluids are incompressible, [see equation 6-53 in Turcotte and

Schubert, 2002, p. 235]:

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} = 0$$

gives:

$$\frac{v_x}{\lambda} + \frac{v_z}{H} = 0 \quad (3.14)$$

Like for folding, the interface instability grows exponentially with time ($A(t) = A_0 e^{qt}$) at a growth rate q while the vertical velocity is related to amplitude as:

$$v_z = \frac{\partial A(t)}{\partial t} \quad (3.15)$$

where $\partial/\partial t$ expresses the time derivative. Substituting Equations (3.14) and (3.15) in equation (3.13) gives:

$$\left(\frac{\mu_2 \lambda}{H^3} + \frac{\mu_1}{\lambda^2} \right) q A_0 e^{qt} = - \frac{(\rho_1 - \rho_2) g A_0 e^{qt}}{\lambda} \quad (3.16)$$

and solving for q :

$$q = - \frac{(\rho_1 - \rho_2) g}{\lambda \left(\frac{\mu_2 \lambda}{H^3} + \frac{\mu_1}{\lambda^2} \right)} \quad (3.17)$$

This simple derivation shows that the growth rate depends on the wavelength of the perturbation [Whitehead, 1988]. Figure 3.3B plots the growth rate as a function of normalized wavelength (λ/H). For the specific wavelength/thickness ratio:

$$\frac{\lambda_{dom}}{H} = 0.79 \left(\frac{\mu_1}{\mu_2} \right)^{\frac{1}{3}} \quad (3.18)$$

the growth rate is maximum:

$$q = 0.53 \frac{g(\rho_1 - \rho_2)H}{\mu_1} \left(\frac{\mu_1}{\mu_2} \right)^{\frac{1}{3}} \quad (3.19)$$

This equation predicts the most favorable condition for the development of Rayleigh-Taylor instabilities (pure diapirism). More complete derivations using a two-dimensional perturbation analysis also reach this basic result [Whitehead, 1988], albeit with different coefficients. The notable conclusions gained from these simple mechanical considerations are:

- The viscosity contrast is the chief parameter controlling the wavelength of a diapiric instability. As such, the dependence of the dominant wavelength on viscosity is the same as for folding of a viscous layer in a viscous matrix (see equations 3.9 and 3.10).
- Density influences the growth rate (and thus velocity) only and has a negligible effect on the wavelength.

3.4 Analytical perturbation method

Several authors have considered the case of a Rayleigh-Taylor instability under compression and demonstrated that above a certain strain rate the folding instability becomes dominant [e.g. Conrad and Molnar, 1997, Ismail-Zadeh et al., 2002]. To solve our geological preoccupation, which is how to know fold-domes from diapir-domes, we decided to explore the transition between both instability modes in more detail. We addressed the problem by using three methods for the simplest two-layer system. First, the analytical perturbation method, which is valid for the onset of both types of instability, is used to derive a 'phase-diagram' discriminating different modes of deformation. Second, numerical simulations were performed to test the applicability of the phase diagram and to study the geometries that develop during the nonlinear finite amplitude stages. Finally, we calculated and compared the patterns of finite strain, which geologists could hopefully use to distinguish the different deformation modes from field observations.

3.4.1 Method

In order to get insight in the transition between diapirism and buckling modes, the simplest model consists in a layer of high viscosity and density over a layer with lower viscosity and density (see Fig. 3.4). The system is subjected to layer-parallel, pure-shear background deformation at a constant strain rate. Gravity is present. The bottom boundary is rigid (no-slip condition) and the top is a fast redistribution boundary (no stress condition on the top boundary kept flat throughout experiments). Rheology is assumed to be linearly viscous. With this setup, a standard perturbation method [e.g. Smith, 1977, Fletcher, 1977] was employed to derive an analytical solution that

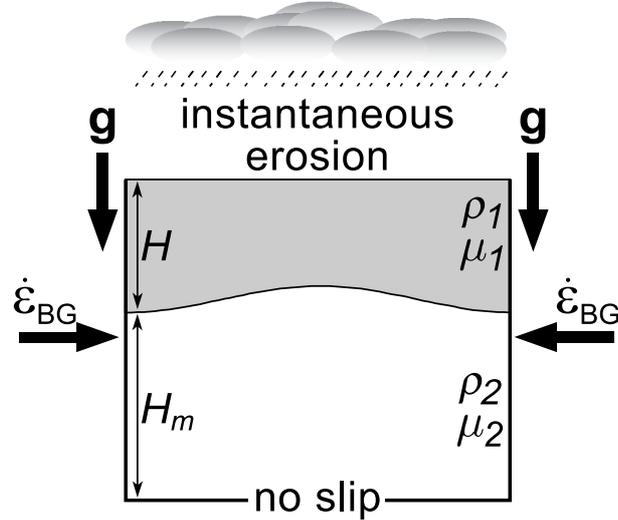


Figure 3.4: Setup used in this work. A layer of high viscosity and density lies over a layer with lower viscosity and density. The system is subjected to layer-parallel, pure-shear background deformation at a constant strain rate. Gravity is present. The bottom boundary is rigid and the top is kept flat. Symbols as in the text.

describes the growth rate of the layer interface as a function of its wavelength, density- and viscosity-structure and the thickness of both layers (see Appendix). The complete analytical solution is rather complicate and grants only limited insight on the controlling parameters. Therefore, analytical solutions regarding the dominant wavelength and growth rate were obtained for several end-member cases. For instance, an expression for dominant growth rate and wavelength was derived in the case of thin-layer diapirism by setting the background strain rate to zero and making the lower layer much thinner than the upper layer. The result obtained in this way is similar as the result that was derived in section 3 on the basis of dimensional analysis (Equations 3.18 and 3.19).

3.4.2 Phase diagram: folding versus diapirism

With the technique summarized above, eight deformation modes could be distinguished, each with a different expression of dominant growth rate and wavelength (see Table 3.1 for scaling laws). These deformation modes are displayed in a two-

dimensional phase diagram (Fig. 3.5) using two non-dimensional numbers, B and B_{det} , constructed on the basis of rheological and geometrical parameters. They are:

$$\begin{aligned} B_{det} &= \left(\frac{H_m}{H} \right) R^{-\frac{1}{3}} \\ B &= 0.39 Ar R^{-\frac{2}{3}} \end{aligned} \quad (3.20)$$

where R is the viscosity contrast (μ_1/μ_2) between the upper and the lower (matrix) layer, respectively, H and H_m the thickness of the upper and lower layer, respectively, and Ar the Argand number expressing the importance of gravity over the background shortening rate [adapted from England and McKenzie, 1982]:

$$Ar = \frac{(\rho_1 - \rho_2)gH_m}{2\mu_2\dot{\epsilon}_{BG}} \quad (3.21)$$

Note that this definition of the Argand number differs from the original definition [England and McKenzie, 1982]: $Ar = \frac{\rho_1}{\rho_2} \left(1 - \frac{\rho_1}{\rho_2} \right) \frac{\rho_1 g H}{\mu_1 v_0 / H}$, where v_0 is the indentation velocity. Equation 3.21 also differs from the definition of [Schmalholz et al., 2002], who define $Ar = \frac{\Delta\rho g H}{2\mu_1 \dot{\epsilon}_{BG}}$, where $\Delta\rho$ is the density difference between the material above (air or water) the high-viscous plate and the material below the plate. This point emphasises that the definition of the Argand number is problem related. With equation (3.21), gravity dominates over compression if the Argand number (and therefore the B number) is large and one expects diapiric-type structures. On the other hand, low Argand numbers indicate folding or homogeneous thickening modes.

Contours of the dominant growth rate and dominant wavelength versus B and B_{det} are shown in Figure 3.5. The curves of iso-dominant wavelength and/or iso-growth rate have different slopes in different regions of the diagram, thus defining eight fields. The mechanical 'phase-boundaries' are the locus of the inflexion points where slopes change. Inside each field, the growth rate of the given mode is larger than the growth rate of any of the other seven modes.

It can be readily seen that diapiric modes (gravity dominated) roughly occur for $B > 1$ and folding modes (where compression dominates) exist when $B < 1$. Somewhat expectedly the transition depends largely on the Argand number. Three fields cover the diapiric modes:

1. Thin-layer diapirism, when the interlayer boundary is very close to the bottom, no-slip boundary.

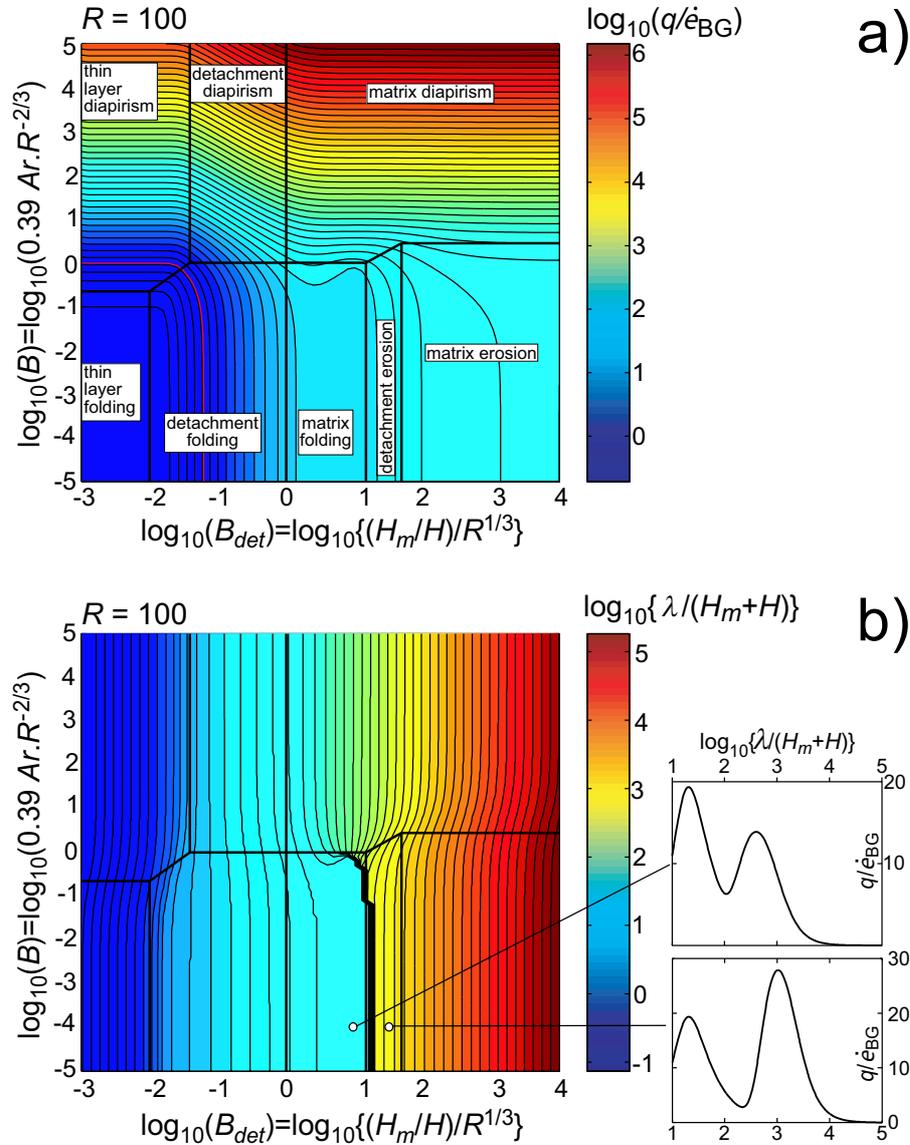


Figure 3.5: a) Growth rates normalized over the background strain rate as a function of B and B_{det} plotted for an R -value of 100. Thick lines indicate phase boundaries that were calculated on the basis of simplified dominant wavelength and growth rate expressions. b) Contours of dominant wavelengths normalized over the total thickness of the system as a function of B and B_{det} , plotted for an R -value of 100. Switching from a detachment folding-mode to a detachment-diapirism mode, by changing B but keeping B_{det} at a constant level, the dominant wavelength remains almost unchanged. The insets show that the growth rate curve for folding with erosion has two peaks. The second peak becomes larger than the first one in the detachment-erosion field.

Deformation Mode	Dominant wavelength	Dominant growthrate
Thin-layer folding	$\lambda_{tlf} = 4.46H \left(\frac{H_m}{H}\right) R^{\frac{1}{3}}$	$q_{tlf} = 0.86R^{-\frac{1}{3}}\dot{\epsilon}_{BG}$
Detachment folding	$\lambda_{df} = 3.3H \left(\frac{H_m}{H}\right)^{\frac{1}{2}} R^{\frac{1}{6}}$	$q_{df} = 0.86 \left(\frac{H_m}{H}\right) R^{\frac{1}{3}}\dot{\epsilon}_{BG}$
Matrix folding	$\lambda_{mf} = 4.45HR^{\frac{1}{3}}$	$q_{mf} = 0.86R^{\frac{2}{3}}\dot{\epsilon}_{BG}$
Detachment erosion	$\lambda_{de} = 6.30H \left(\frac{H_m}{H}\right)^{\frac{3}{4}} R^{\frac{1}{4}}$	$q_{de} = 0.18 \left(\frac{H_m}{H}\right) R^{\frac{1}{12}}\dot{\epsilon}_{BG}$
Matrix erosion	$\lambda_{me} = 5.24H \left(\frac{H_m}{H}\right)$	$q_{me} = 0.60R\dot{\epsilon}_{BG}$
Thin-layer diapirism	$\lambda_{tld} = 2.795H \left(\frac{H_m}{H}\right) R^{\frac{1}{3}}$	$q_{tld} = 0.28ArR^{-\frac{2}{3}}\dot{\epsilon}_{BG}$
Detachment diapirism	$\lambda_{dd} = 3.86H \left(\frac{H_m}{H}\right)^{\frac{1}{2}} R^{\frac{1}{6}}$	$q_{dd} = \frac{1}{3}Ar \left(\frac{H_m}{H}\right) R^{-\frac{1}{3}}\dot{\epsilon}_{BG}$
Matrix diapirism	$\lambda_{md} = 2.75H \left(\frac{H_m}{H}\right)$	$q_{md} = \frac{1}{3}Ar\dot{\epsilon}_{BG}$

Table 3.1: Scaling laws for the different deformation modes in Fig. 3.5, derived using the perturbation method described in the Appendix. The growthrate of two deformation modes is equal at their boundary (e.g. $q_{dd} = q_{md}$). Thus the scaling laws given here can be used to compute the position of phase boundaries.

2. Detachment diapirism, when the interlayer boundary is further up from the bottom, which still influences the dynamics of the system.
3. Matrix diapirism, when presence of the stiff but thin top layer exerts no influence.

Five fields cover the folding modes. With increasing B_{det} numbers they are:

1. Thin layer folding, when the interlayer boundary is very close to the bottom boundary. In this case, growth rates are slower than the background shortening-rate. Folds will still grow, but at such a slow rate that they will not become discernable during the later stages.
2. Detachment folding, when the interface is further up from the bottom, but the influence of the bottom is still felt. The low-viscosity lower layer acts like a thin channel and the growth rate of the interlayer boundary is limited by the rate with which viscous fluids can flow through this channel. Therefore growth rates increase with increasing channel thickness (ie. increasing B_{det} number)
3. Matrix folding, when the lower layer is thick enough to act like an infinite matrix. Growth rates depend solely on the viscosity contrast and background strain rates.

4. Detachment erosion, which has significantly longer dominant wavelengths than the matrix folding field. The growth rates are sensitive to viscosity contrast, matrix thickness and background strain rate.
5. Matrix erosion, which has no dependency on the thickness of the top stiff but thin layer.

In light of previous work [Schmalholz et al., 2002], the folding modes 1, 2 and 3 were expected. The new modes 4 and 5 are due to the action of erosion. The growth rate curves of these two modes draw attention because they have two maxima. For small B_{det} numbers (modes 1-3), the first maximum always dominates. However, if B_{det} becomes larger than $10^{1.2}$ (for $R=100$), the second peak overtakes the first one (insets of Fig. 3.5B). This is accompanied by a sudden increase in dominant wavelength.

Not all of the eight fields on the phase diagram are expected to occur in nature. The thin-layer folding field is one example. The growth rate being always smaller than background shortening implies that the system will deform homogeneously rather than produce an instability. The erosion-modes that require very large thickness ratios is another example. Consider for example a 50 km thick crust. If a viscosity contrast of 100 is assumed between a strong upper crust and a weak lower crust, the thickness of the strong layer should be around 350 m in order to fall within the detachment-erosion mode ($\log_{10}(B_{det}) = 1.5$). Estimates of effective thickness of the upper crust for Central Asia vary between 6 and 12 km [Schmalholz et al., 2002]. Owing to these limits, only 5 fields are further considered in the current study (Fig. 3.6A). Changing the viscosity contrast affects the topology of the phase diagram (Fig. 3.6B). The boundaries between the diapiric modes 2 and 3 and between the folding modes 2 and 3 are independent of the viscosity contrast, a fact due to the definition of B and B_{det} . The other boundaries move across the diagram so that the other fields widen with increasing viscosity contrast.

3.4.3 Discussion

The first point extractable from the phase-diagram (Fig. 3.5) is that the detachment folding and detachment diapirism modes have very similar dominant wavelengths. For example, consider a model with given layer thicknesses and viscosity contrast resulting in $\log_{10}(B_{det}) = -1$. If compression is very slow (i.e. $\dot{\epsilon}_{BG} \rightarrow 0$) the Argand number

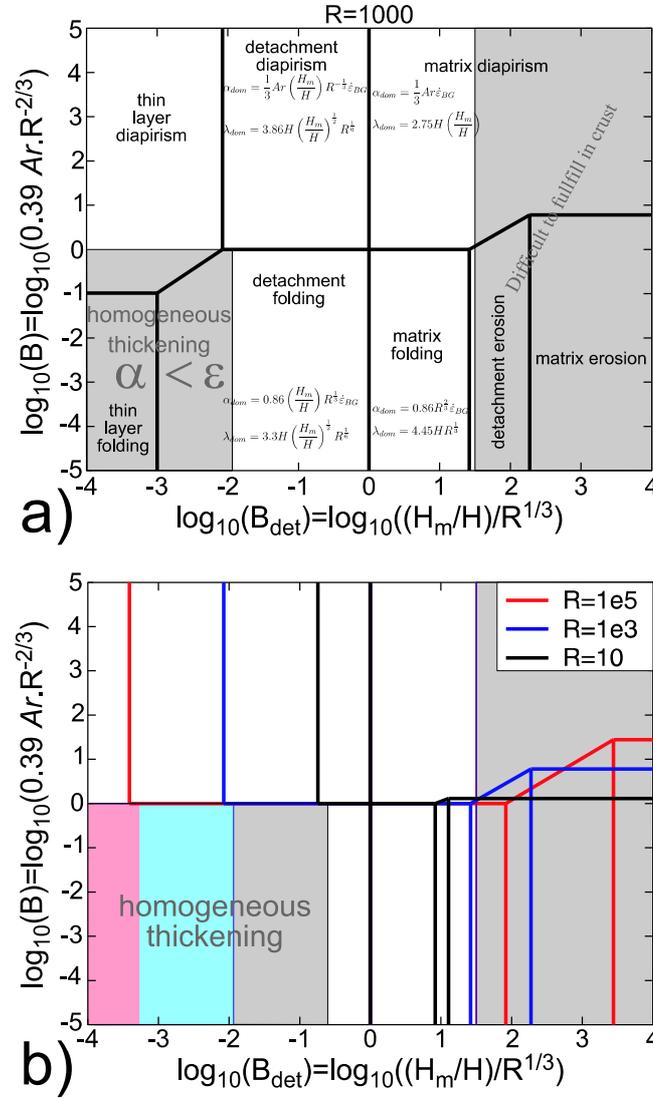


Figure 3.6: a) Phase diagram for crustal deformation for a viscosity contrast of $R = 1000$ as a function of B and B_{det} . Not all phase fields are likely to occur at crustal scale, because some require a very thin layer on a very thick matrix (detachment erosion and matrix erosion fields) and others have a growth rate smaller than the background shortening rate (thin-layer folding) and will therefore not produce finite amplitude domes. b) Topology of the phase diagram as a function of viscosity contrasts. The non-dimensional numbers B and B_{det} were chosen in such a way that the boundary between the four fields that are thought to be relevant for the crust is independent on the viscosity contrast. The other fields increase in size with increasing viscosity contrast.

is very large ($Ar \rightarrow \infty$) and, consequently, the B number is large. A diapiric, purely gravity driven deformation mode will develop with a given wavelength (Fig 3.5B). If the system now suddenly undergoes compression, the Argand number and therefore the B number decrease and detachment folding becomes favored. However, the dominant wavelength remains unchanged. Instabilities that started forming in the diapiric field are likely to continue amplifying in the folding mode without discernable consequences on the wavelength. This observation may help understanding site and shapes of salt diapirs that pierce periodical folds of the Zagros Mountains [Alavi, 1994]. A second point is that the shapes of both diapirs and folds developing while being instantaneously eroded have two-dimensional profiles that look alike; both have chimney (cusp) shapes with very steep walls and narrow heads. The three-dimensional shape of the resulting structures is currently only partially understood for diapirs. 3D tubular diapirs amplify slightly faster than two-dimensional (linear, wall-like) diapirs at finite amplitudes [Kaus and Podladchikov, 2001]. Therefore, diapirs tend to form finger-like (or in map-view circular) intrusions. Three-dimensional folding, on the other hand, has so far only been studied in the linear initial stages, for a layer embedded in an infinite matrix [Fletcher, 1991, 1995]. The effects of both finite amplitude and a detachment layer remain to be examined. At the moment, we can only speculate that existing large-amplitude three-dimensional folds will continue amplifying in the same detachment folding mode.

3.5 Numerical simulations

The boundaries separating the different deformation modes are relatively sharp (Fig. 3.5), which points to rapid changes in deformation mechanism from one mode to the next. However, the analytical perturbation method used to derive the diagram has an important shortcoming in that it implies infinitely small perturbations of the interface and, therefore is valid only for starting conditions. Linearization is too strong an assumption for understanding geological structures that accumulate large amounts of strain. We performed numerical simulations to study the importance of nonlinear finite amplitude effects and, therefore, to estimate the validity of the analytical predictions.

3.5.1 Numerical technique

We used the code GANGO, which is a two-dimensional Eulerian finite-difference/spectral method that builds on the technique described in [Schmalholz et al., 2001]. It solves the momentum equations for the slow motion of rocks on geological timescales. Incompressibility is assumed. The governing partial differential equations are solved on a two-dimensional domain that uses a spectral approximation in the horizontal direction and a conservative finite difference approximation in the vertical direction. Time stepping is done with an implicit algorithm in order to accurately track the highly unstable stages of the initiating instabilities. The interface between layers of different, sharply-varying material parameters is described by a marker line that eliminates numerical diffusion. The two-layer initial setup is the same as used for the analytical method (Fig. 3.4). It mimics a high-viscosity, high-density layer, analogue to the upper crust, resting on a low-viscosity, low-density layer representing the deeper crust. The density inversion could represent, for example, magma-rich migmatites or salt at depth. Rheology is linearly viscous and constant within each layer. Horizontal boundary conditions are assumed to be periodic on top of a background pure-shear velocity field. A no-slip boundary condition was set at the bottom boundary, fast redistribution is allowed on the top boundary. This means that the topography created at the surface is instantaneously flattened, depressions being filled with a high-viscosity, high-density material similar to that of the upper layer. Thus the modelling particularly concerns regions of very fast erosion and sedimentation, as it is suggested in the Himalayan syntaxes [Burg and Podladchikov, 2000].

3.5.2 Comparison between numerics and analytics

In order to check the accuracy of the code we numerically calculated growth rates by imposing different sinusoidal perturbations of small amplitude at the interlayer boundary and recording the growth rate under four different deformation modes: detachment and matrix diapirism and detachment and matrix folding. The analytical and numerical results show a good agreement (Fig. 3.7). To study the nonlinear finite amplitude effects, growth rates were also calculated for sinusoidal perturbations with larger initial amplitudes. In this case, the growth rate curves are similar in shape, but the magnitudes are generally smaller. Therefore, the dominant wavelength of finite amplitude instabilities is close to the wavelength predicted by the analytical solution. Accordingly, the phase diagram can be applied to large strain cases.

3.5.3 Geometries developing at finite amplitude stages

Previous numerical simulations have demonstrated that an initially horizontal layer perturbed with random noise develops structures whose wavelength is close to the dominant wavelength [e.g. Schmalholz and Podladchikov, 1999, Kaus and Podladchikov, 2001]. Exceptions to this general rule can occur if the amplitude of the noise is large compared to the thickness of the layer [e.g. Mancktelow, 1999], if a large non-dominant initial perturbation is present [e.g. Schmeling, 1987, Kaus and Podladchikov, 2001], or if the dominant growth rate is smaller than the background strain rate. Simulations starting with an initial low-amplitude random noise performed for four deformation modes (not shown here) indeed developed into structures with a wavelength close to the dominant one. To study the geometries developing during finite amplitude stages, it is therefore sufficient to start with a sinusoidal perturbation of dominant wavelength. Results of such calculations are displayed in Figure 3.8 for which, in addition to the B and B_{det} numbers, the initial perturbation amplitude was varied. Figure 3.9 shows simulations for the same B and B_{det} parameters with different viscosity contrasts. The resulting structural characteristics can be summarized as follows:

- *Thin-layer diapirism to detachment diapirism*: a tendency to form balloon-on-a-string diapirs [Podladchikov et al., 1993] is observed, both at small and large initial amplitudes and viscosity contrasts. After the diapir reaches the surface, its form changes to a stock or chimney-like shape.

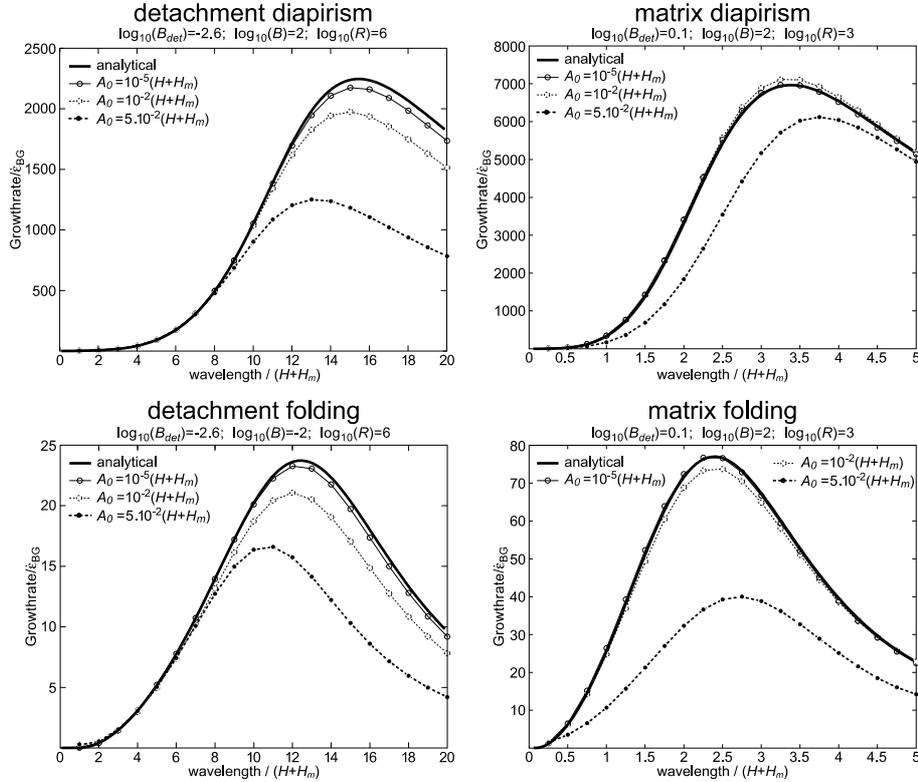


Figure 3.7: Comparison of analytical growth rates with numerically determined growth rates for 4 different deformation modes. In all cases the numerically determined growth rates with a small amplitude perturbation are very close to analytical results. Cases where the initial perturbation has a larger amplitude generally result in smaller growth rates, while the dominant wavelength remains almost unchanged.

- *Detachment diapirism:* Simulations with a viscosity contrast of 3000 formed always chimney-like geometries after the low-viscosity low-density material reached the surface. The simulation with a lower viscosity contrast formed a balloon-on-a-string diapir. The smaller dominant wavelength and thickness ratio in this simulation compared to other cases can explain this difference. The interface is far below the eroding top surface during the initial stages and the shape of the diapir evolves as if the top were a no-slip or free-slip boundary [e.g. Woitd, 1978]. Additional simulations showed that, as a rule-of-a-thumb, chimney-like diapirs form if the dominant wavelength is at least 4-5 times larger than the thickness of overburden.

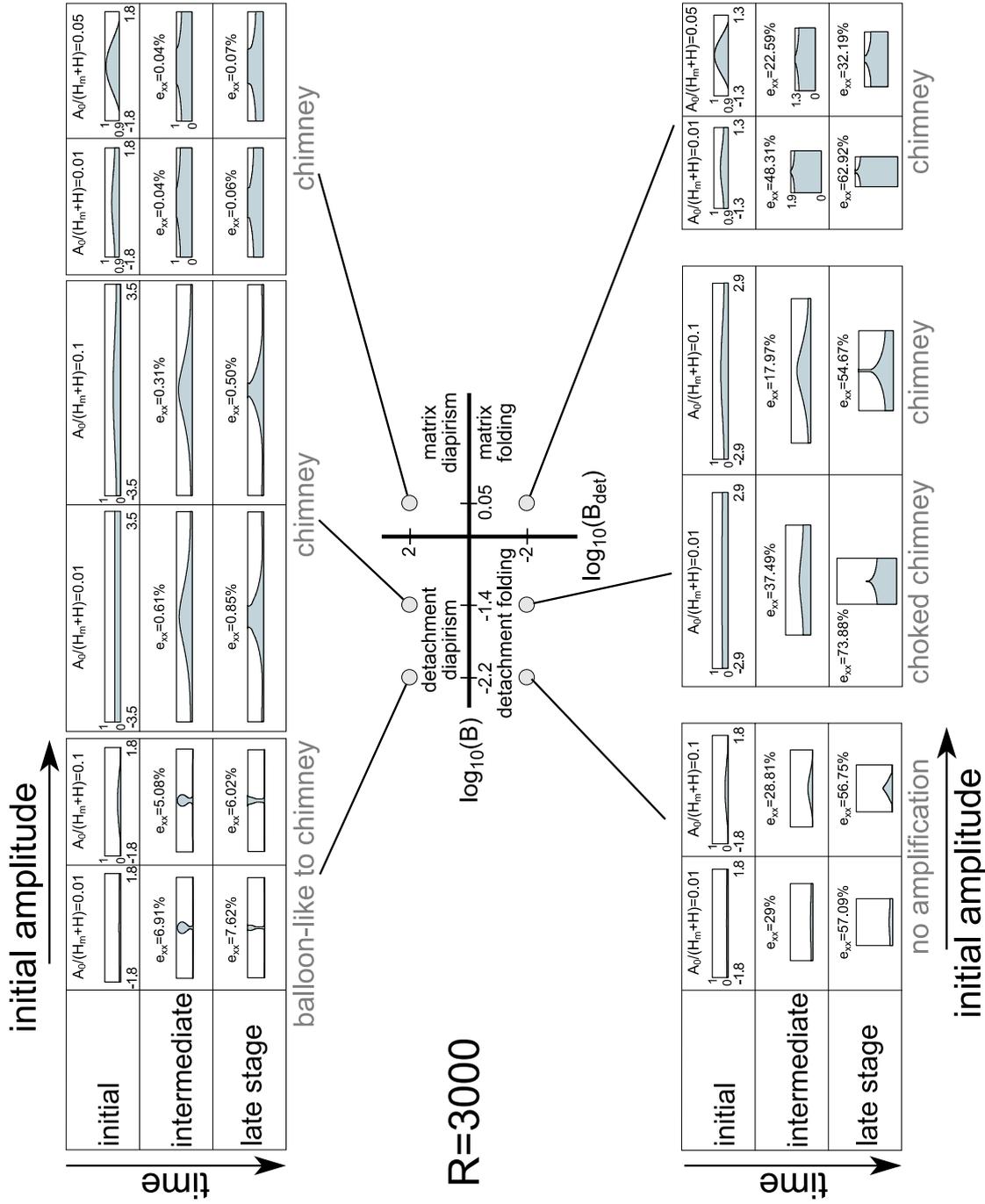


Figure 3.8: Geometrical results of numerical simulations for the four deformation modes relevant for crustal-scale deformation. The viscosity contrast R is 3000 in all simulations. The initial setup has a sinusoidal perturbation whose wavelength is the dominant wavelength. Two simulations were performed at specific (B, B_{det}) points: one with an initial amplitude A_0 of 0.01 times the height of the initial box and one with a larger initial amplitude. Varying the initial amplitude has relatively little effects on the final structures that form in the detachment-diapirism, the matrix diapirism and the matrix folding fields, whereas it has significant effects in the detachment-folding field where the antinodal crest does not always reach the surface.

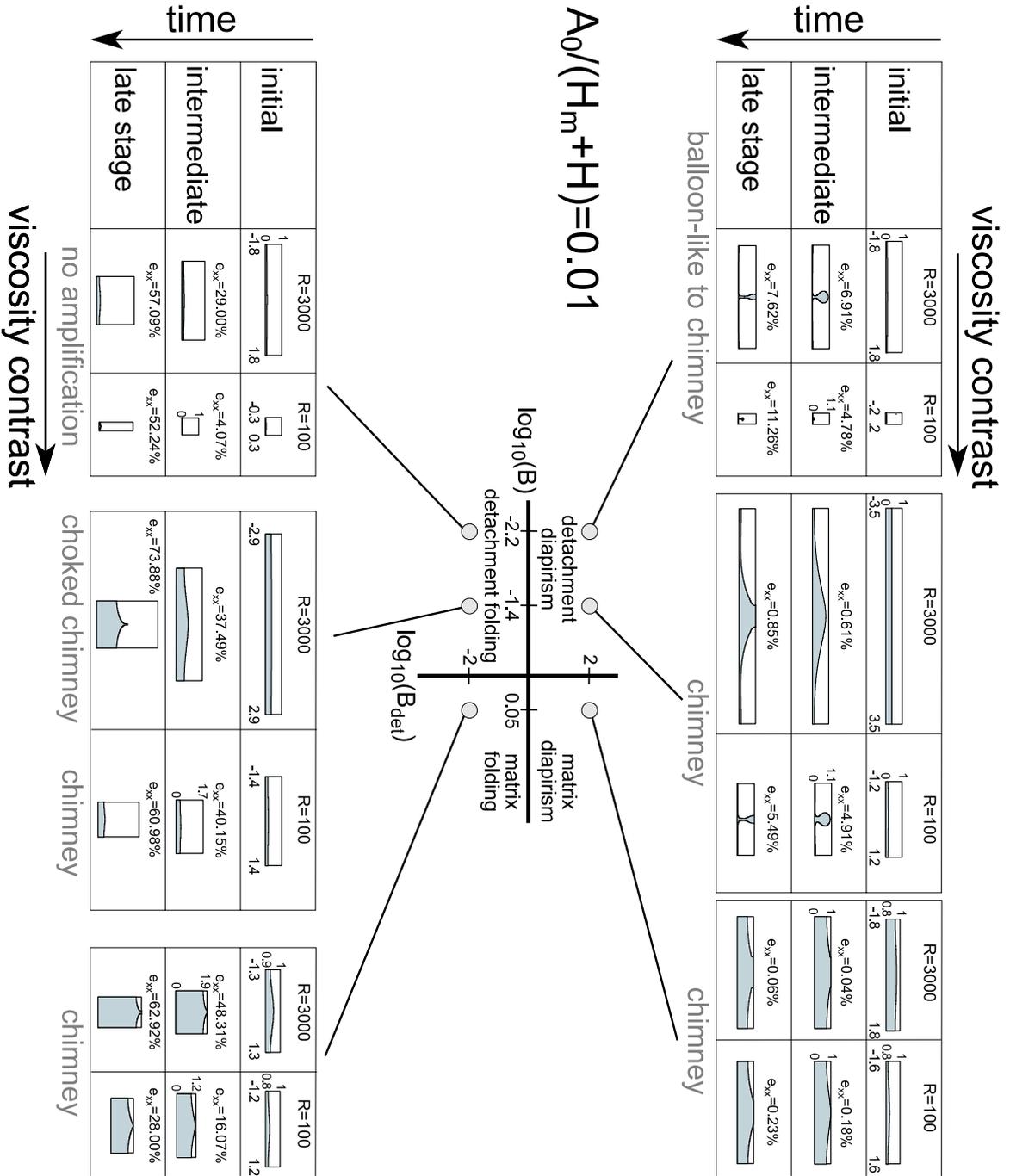


Figure 3.9: Results of numerical simulations where the viscosity ratio between the dense, high viscosity layer (white) and a lighter, low viscosity layer (grey) is varied. The initial amplitude is set to 0.01 the box thickness. All other parameters are the same as in figure 3.8.

- *Matrix diapirism*: All simulations evolve towards chimney-like structures.
- *Thin-layer folding to detachment folding*: Dynamic growth rates are too small to have active amplification of the interlayer surface whose change in shape is due to the overall pure-shear shortening.
- *Detachment folding*: The simulation with a small initial amplitude and large viscosity contrast has active amplification, but insufficient to reach the surface. The thereby developed 'hidden' chimney-like structure would be difficult to observe in orogenic belts. Simulations that started with a larger initial amplitude, on the other hand, produced structures able to reach the surface, a stage after which the geometry evolves towards chimney-like structures. Simulations with a smaller viscosity contrast have smaller growth rates and no active amplification.
- *Matrix folding*: All simulations reach the surface and form chimney shapes similar to those in the matrix-diapirism field.

In conclusion two different dome-like geometries can be experimented: (1) the classical diapir structure, which forms if very thin layers are present, and (2) chimney-like structures, which form both in folding- and diapiric deformation modes.

3.5.4 Dynamics of dome-formation

Numerical results show that the final stage morphology of both diapiric and folding deformation modes are very similar, at least regarding the shape of the interface between different material properties. The aim of this section is to investigate differences in the dynamic evolution of buckling and diapiric deformations. Figure 3.10 shows snapshots in the evolution of a simulation in the detachment diapirism field and in the detachment folding field. Comparisons are made for the stream function and effective stress [second invariant of the deviatoric stress tensor, see e.g. Ranalli, 1995]. During the first two stages, before the growing structures reach the surface, the velocity patterns are very similar. Both diapirs and folds behave as if the low-viscosity material were flowing through a thin channel towards the centre of the dome. However, differences exist in the magnitude and distribution of effective stress: much larger values are built up in the folding mode, which is due to larger background strain rates set in this simulation.

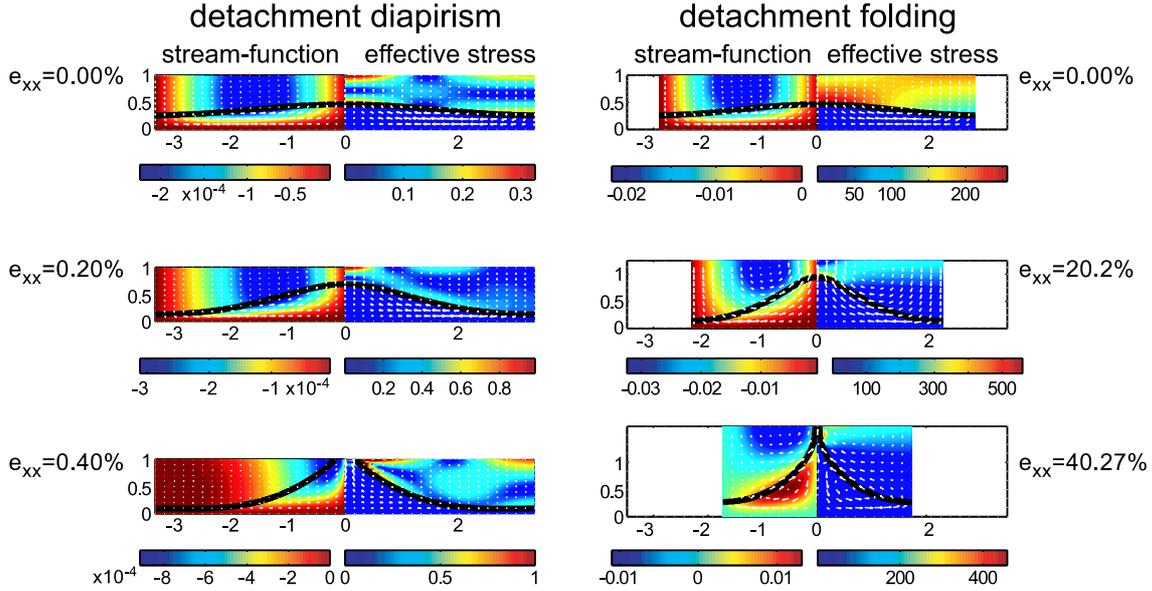


Figure 3.10: Results of numerical calculations, showing different snapshots of a) a simulation in the detachment diapirism field ($B_{det} = 10^{-1.4}$, $B = 10^2$) and b) a simulation in the detachment folding field ($B_{det} = 10^{-1.4}$, $B = 10^{-2}$). Colours indicate stream function (left) and effective stress (right). Active velocity (i.e. velocity minus background pure-shear velocity) is indicated by white arrows. The viscosity contrast R is set to 3000 and the initial amplitude A_0 is 0.1 times the height of the box. During the initial stages, the main difference between folding and diapirism is the magnitude of effective stress in the high viscous layer, which is due to the fact that the background strain rate is larger in the folding case. Remarkably the active velocity field and the geometry of the interface are similar for both cases. Only after the low viscosity/low density material has reached the surface and is being eroded, the deformation patterns changes significantly: whereas the diapir continues to move upward, with a dramatically increased rate, the material in the core of the fold mainly follows a downward movement (gets squeezed away). e_{xx} -shortening in the horizontal (x) direction.

During the last stage, after the interlayer boundary reached the surface, considerable differences occur in the velocity field. Upward flow in the centre of the eroding dome continues in the diapiric mode, with acceleration due to the fact that velocities are no longer controlled by the highest but by the lowest viscosity region. Folds, however, change the velocity pattern from upwards movement to downward squeezing of low viscosity material near the centre of the dome. This is accompanied with a decrease in absolute velocity, and can be explained since the dominant mechanism changes from a folding instability to flow-between-rigid-plates. If this channel has a wedge-shaped form, the material gets squeezed outwards rather than through the channel. Simulations within the matrix fields confirm these results: initial velocity patterns are similar, but patterns change once erosion of the dome material starts.

Chimney-like structures are appealing features to explain the formation of isolated domes because the so-called dominant wavelength becomes very large under fast-erosion conditions. These two aspects, along with the occurrence of steep limbs, are reminiscent of the Himalayan Syntaxes. However, the bulk form and distribution are scantily discriminating. Therefore, we tried to establish further discrimination between diapirs and folds from the strain field in and around domes.

3.6 Strain in and around domes

In order to predict the strain distribution in and around domes, a feature observable to geologists, a passive grid was inserted in the numerical simulations and moved with the calculated velocity distributions. Results are displayed in the form of finite strain ellipsoids for an intermediate stage in figure 3.11 and for a late stage in figure 3.12, which are chosen for their likeliness in the shape of the interface. Expectedly, the largest strains are recorded within the low viscosity materials. Diapir simulations show a normal sense of shear towards the diapir-overburden interface. The matrix-folding simulations show a reverse sense of shear towards this interface, which is in agreement with fold vergence and senses of shear geologists have been using (Fig. 3.14). However the detachment folding simulations behave opposite to expectations with a normal sense of shear at the interface. This is explained by the fact that the thin viscous channel largely controls deformation and, indeed, the strain pattern resembles that of viscous pipe flow. Only after extrusion of low-viscosity material the finite strain

pattern starts switching from normal to reverse sense of shear near the centre of the dome, whereas in diapirs the pattern remains relatively unchanged. We conclude that the present-day structural techniques applied for distinguishing diapiric domes and folds are ambiguous if detachment folding takes place during fast erosion, and that domes displaying extensional structures at their periphery do not necessarily reflect extension.

3.7 Discussion

We presented results for a two-layer system with simple Newtonian rheology, which is a fair approximation of the creep behavior of salt and salt and sediments. However, rocks generally have a brittle rheology and/or a power-law rheology known to be depth-dependent through the strong temperature effect on the constitutive laws [e.g. Ranalli, 1995]. In the light of previous studies on folding and diapirism, we will first comment on the influence of power-law and brittle rheologies on our results. The main effect of power-law rheology on the folding instability is a marked increase in growth rates and a small decrease in dominant wavelength [e.g. Schmalholz et al., 2002, Smith, 1977, see also equations 3.9 and 3.10]. The same effect was also recognised for diapirs, around which the deformation aureole becomes smaller with increasing powerlaw-exponent [Weinberg and Podladchikov, 1995].

Temperature dependency of viscosity may have an influence on the results presented here which are, strictly spoken, only valid for quasi-adiabatic conditions. Thermal diffusion is known to be more effective if (1) growth rates of the instabilities are low and (2) the instability wavelength is small [e.g. Conrad and Molnar, 1997]. Both conditions may exist in the thin-layer folding and thin-layer detachment modes (Fig. 3.5), and would slow down the growth of domes. Since in crustal cases, the viscosity differences are related to composition differences, the overall thermal effect are likely to be minor at initial stages. During evolved stages however, cooling may stop domes from rising, an effect which will be more pronounced for short-wavelength structures. Nonlinear feedback mechanisms such as those due to shear heating [e.g. Regenauer-Lieb and Yuen, 2003] may also modify the finite amplitude structures presented here. Inspection of figures 3.11 and 3.12 indicates that shear-heating may be most efficient for the detachment-folding mode, where it would potentially transfer the two-layer system

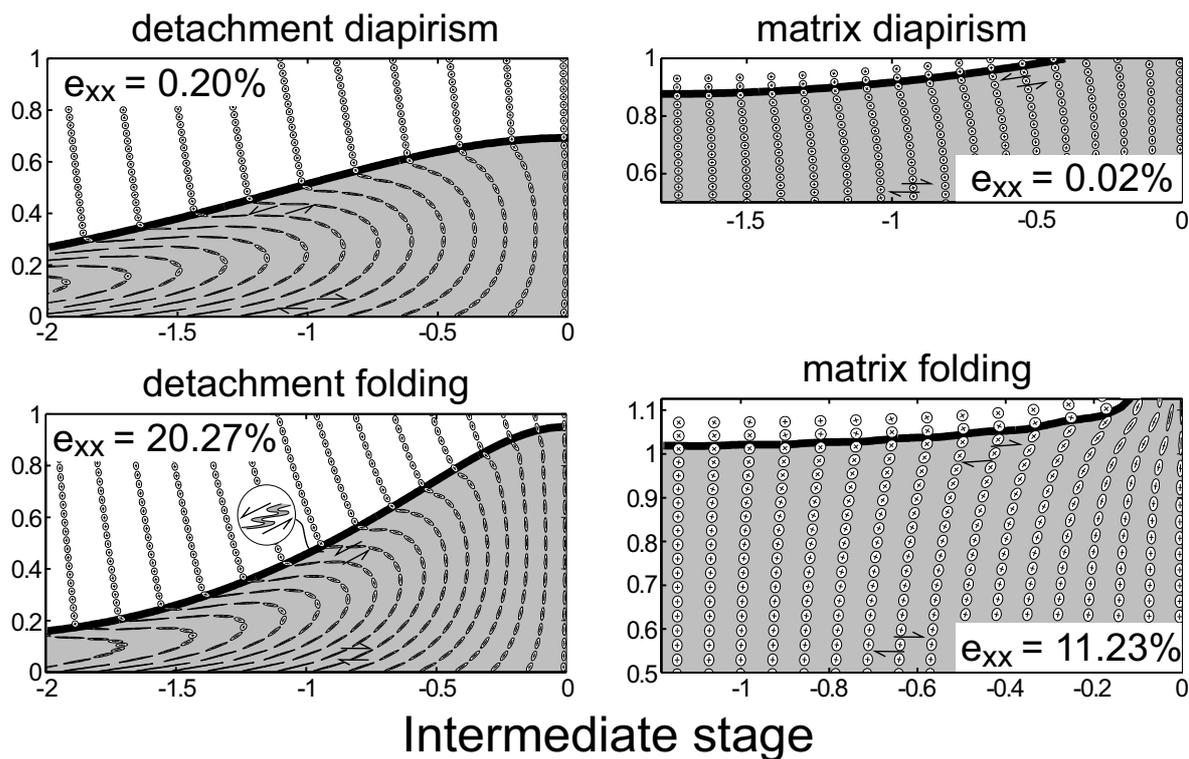


Figure 3.11: Strain patterns around diapirs (left) and folds (right) during intermediate stages, when the low viscosity material has not reached the surface yet. The following parameters were used: a) $\log_{10}(B_{det}) = -1.4, \log_{10}(B) = 2, R = 3000, A_0/(H + H_m) = 0.1$, b) $\log_{10}(B_{det}) = -0.05, \log_{10}(B) = 2, R = 3000, A_0/(H + H_m) = 0.1$ c) $\log_{10}(B_{det}) = -1.4, \log_{10}(B) = -2, R = 3000, A_0/(H + H_m) = 0.1$ and d) $\log_{10}(B_{det}) = 0.05, \log_{10}(B) = 2, R = 3000, A_0/(H + H_m) = 0.01$. Both detachment folding and detachment diapirism show a very similar strain pattern, which resembles the pattern classically attributed to diapirs only.

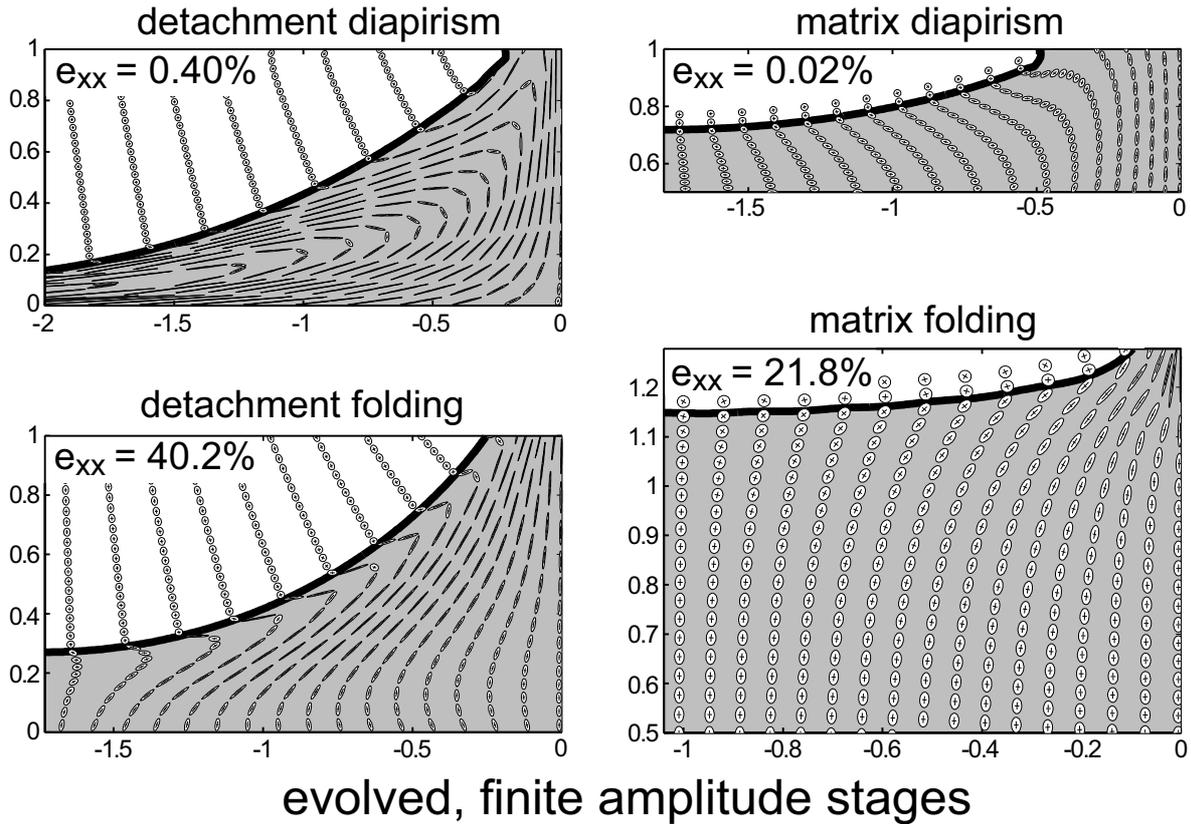


Figure 3.12: Strain patterns around diapiric structures (left) and folding structures (right) during late stages, when the low viscosity material has reached the surface and is being eroded. All parameters are the same as in figure 3.11. Matrix folding and matrix diapirism still have opposite senses of shear close to the boundary. Detachment diapirism and detachment folding have the same sense of shear close to the interface. However the detachment folding sense-of shear flips towards the centre of the anticline.

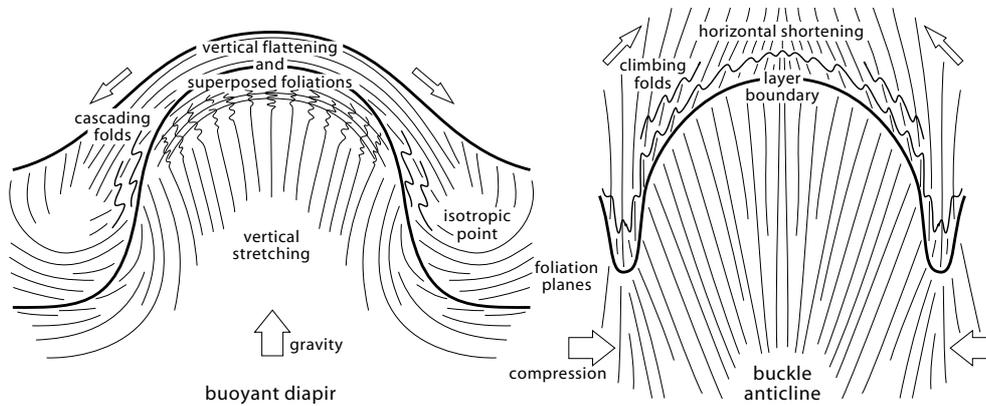


Figure 3.13: Structural characteristics of a cylindrical diapir and a cylindrical buckle-anticline; adapted from [Dixon, 1987]. This classical morphology is consistent with matrix diapirism and folding (Fig. 3.12).

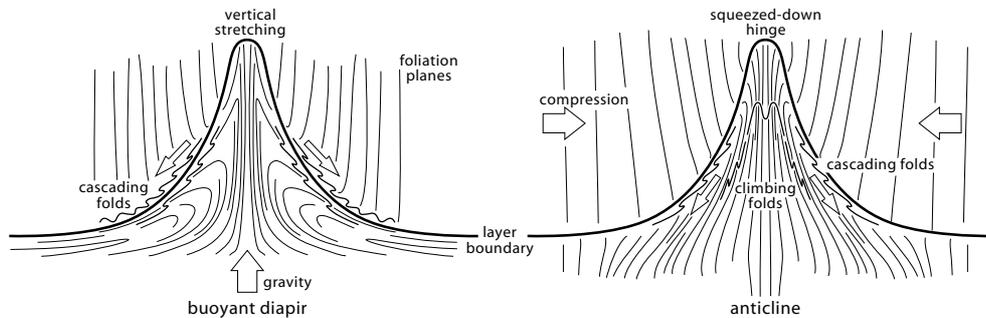


Figure 3.14: Structural characteristics of a diapir and a buckle-anticline submitted to instantaneous erosion as modeled in this work. This morphology refers to detachment diapirism and folding (Fig. 3.12).

into a three layer system with a weak shearzone intercalated between the low-viscous matrix and the highly-viscous overlying layer. The consequences of such a zone on the geometry and dynamics of dome formation are currently unknown, and should be addressed in future studies.

A brittle rheology has only minor effects on salt diapirism [Poliakov et al., 1996] in the absence of compression or extension. Gerbault and coauthors [Gerbault et al., 1999] also argued that the brittle rheology has a relatively minor effect on lithospheric-scale folding. Zuber [1987] and Schmalholz et al. [2002] have considered the effect of a depth-dependent rheology. For this purpose, they redefined the thicknesses (H and H_m) of the layers, yet the results remained similar to those obtained with depth-independent rheologies. Accordingly, one may expect that power-law and brittle rheologies will change the definitions of the two non-dimensional numbers B and B_{det} but the main results of this paper, as first order strain patterns and phase diagram, will not change significantly.

A point worth discussing again is the fact that natural folds are often non-periodic on a large scale, whereas they appear to be rather periodic on the outcrop scale. The simulations presented in this paper shed some light on this process, by showing that instabilities starting from a larger initial perturbation grow faster. Recent work on viscoelastic folding pointed out that elasticity triggers more localized folding [e.g. Schmalholz and Podladchikov, 1999]. Localization is also obtained if geometrical (with a non-dominant wavelength and finite amplitude) and/or rheological heterogeneities are present. It is obvious that a mountain chain like the Himalayas contains such heterogeneities at the onset of continental shortening (e.g. sutures, igneous intrusions etc.). Therefore, large-scale folding, especially during continental collision, may naturally be localized. Folding of the oceanic lithosphere, on the other hand, is expected to be much more periodic, since the oceanic lithosphere is more homogeneous. It is thus logical that regular undulations in topography, gravity signal and Moho-depth observed in the Central Indian Basin over hundreds of kilometers [e.g. Zuber, 1987, Gerbault, 2000, and references therein], are periodic buckle folds. It is also evident that three-dimensional constraints linked to corner effects in original bends of continental boundaries [e.g. Brun and Burg, 1982] will force fold localization.

3.8 Conclusions

Simplified analytical calculations result in a phase diagram separating different modes of deformation as a function of two non-dimensional numbers: B and B_{det} . Five modes of deformation are important in non-extensional, crustal-scale doming, ranging from diapiric to folding modes. Numerical simulations further demonstrated that the evolved diapirs and folds are similarly looking chimney-like, rather isolated, domes and erosion plays a key role in this isolation. The strain distribution around these domes allows the discrimination of folds versus diapirs only in the case where $B_{det} > 1$, implying that the low-viscosity lower layer is relatively thick. In the other cases, the use of asymmetric structures can erroneously point to a diapiric origin even if the dome has formed in compression only. If shortening continues after the core of the dome is eroded, important differences will be observed between diapirs and folds: the core of diapirs continues rising whereas folds develop a core syncline reflecting downward movement in the center of the dome.

3.9 Acknowledgements

We thank the Swiss National Fond, project number 21-61912.00, for financial support. We particularly thank Stefan Schmalholz for enlightening discussions. Dave Yuen and Peter Huddleston have provided very helpful reviews.

3.10 Perturbation analysis

For geological processes inertial terms can be ignored and force equilibrium in a two-dimensional case, ignoring the effect of gravity, is given by:

$$\begin{aligned}\frac{\partial\sigma_{xx}}{\partial x} + \frac{\partial\sigma_{xz}}{\partial z} &= 0 \\ \frac{\partial\sigma_{xz}}{\partial x} + \frac{\partial\sigma_{zz}}{\partial z} &= 0\end{aligned}\tag{3.22}$$

In addition one assumes that material is incompressible:

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} = 0\tag{3.23}$$

Rheology is linearly viscous (with a constant viscosity μ inside each layer) and given by:

$$\begin{aligned}\sigma_{xx} &= -p + 2\mu \frac{\partial v_x}{\partial x} \\ \sigma_{zz} &= -p + 2\mu \frac{\partial v_z}{\partial z} \\ \sigma_{xz} &= \mu \left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right)\end{aligned}\quad (3.24)$$

where p is pressure. Inserting equations (3.24) into the force-balance equations (3.22) gives:

$$\begin{aligned}-\frac{\partial p}{\partial x} + 2\mu \frac{\partial^2 v_x}{\partial x^2} + \mu \frac{\partial}{\partial z} \left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right) &= 0 \\ -\frac{\partial p}{\partial z} + 2\mu \frac{\partial^2 v_z}{\partial z^2} + \mu \frac{\partial}{\partial x} \left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right) &= 0\end{aligned}\quad (3.25)$$

Pressure can be eliminated by taking the derivative of the first equation versus z and subtracting the derivative of the second equation versus x :

$$2\mu \left(\frac{\partial^3 v_x}{\partial z \partial x^2} - \frac{\partial^3 v_z}{\partial x \partial z^2} \right) + \mu \left(\frac{\partial^2}{\partial z} - \frac{\partial^2}{\partial x^2} \right) \left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right) = 0 \quad (3.26)$$

We split the velocity in a sinusoidal perturbed part and a background part, due to the pure-shear shortening.

$$\begin{aligned}v_x &= \tilde{v}_x(z) \exp(I\omega x) - \dot{\epsilon}_{BG} x \\ v_z &= \tilde{v}_z(z) \exp(I\omega x) + \dot{\epsilon}_{BG} z\end{aligned}\quad (3.27)$$

where $I = \sqrt{-1}$ and $\dot{\epsilon}_{BG}$ is the background pure-shear strain rate assumed to be constant over the model. Inserting equations (3.27) into equation (3.26), using the incompressibility constraint (3.24) and dividing by $\mu I \exp(I\omega x) / \omega$ yields a 4th order ODE for $\tilde{v}_z(z)$:

$$\frac{\partial^4 \tilde{v}_z(z)}{\partial z^4} - 2\omega^2 \frac{\partial^2 \tilde{v}_z(z)}{\partial z^2} + \omega^2 \tilde{v}_z(z) = 0 \quad (3.28)$$

A general solution of equation (3.28) has the form:

$$\tilde{v}_z(z) = A e^{(\omega z)} + B z e^{(\omega z)} + C e^{(-\omega z)} + D z e^{(-\omega z)} \quad (3.29)$$

where A, B, C and D are four constants. Having expressed $\tilde{v}_z(z)$, we can determine all other velocities, stresses and the pressure from equations (3.23), (3.24) and (3.27), respectively.

The studied case is a 2-layer system (see Fig. 3.4). Within each layer equation (3.29) applies. Thus there are 8 unknown constants that need 8 equations. Two of these equations apply at the no-slip lower boundary ($z = -H_m$):

$$\begin{aligned}\tilde{v}_z^{lo}(-H_m) &= 0 \\ \tilde{v}_x^{lo}(-H_m) &= 0\end{aligned}\tag{3.30}$$

Two equations apply at the upper boundary ($z = H$), where an infinitely fast erosion boundary condition is set:

$$\begin{aligned}\tilde{\sigma}_{zz}^{up}(H) &= 0 \\ \tilde{\sigma}_{xz}^{up}(H) &= 0\end{aligned}\tag{3.31}$$

At the interface between the layers ($z = 0$), continuity of velocity is required. After making a first order Taylor expansion around $z = 0$, the condition can be written

$$\begin{aligned}\tilde{v}_x^{lo}(0) - \tilde{v}_x^{up}(0) &= 0 \\ \tilde{v}_z^{lo}(0) - \tilde{v}_z^{up}(0) &= 0\end{aligned}\tag{3.32}$$

Stresses across the interface should also be continuous:

$$\begin{aligned}\tilde{\sigma}_{xz}^{lo}(0) - \tilde{\sigma}_{xz}^{up}(0) &= 4\dot{\epsilon}_{BG}(\mu^1 - \mu^2) \frac{\partial \eta}{\partial x} \\ \tilde{\sigma}_{zz}^{lo}(0) - \tilde{\sigma}_{zz}^{up}(0) &= (\rho^1 - \rho^2) g \eta\end{aligned}\tag{3.33}$$

where $\eta = A(t)\cos(\omega x)$ is the sinusoidally perturbed interface and $A(t)$ the perturbation amplitude which grows exponentially with time. The growth on top of the background pure-shear thickening is given by:

$$A(t) = A_0 e^{qt}\tag{3.34}$$

where q is the growth rate of the interface. After noting that $\frac{\partial \eta}{\partial t} = \tilde{v}_z(0)$ one can derive that $\eta = \frac{\tilde{v}_z(0)}{q}$. Substituting this relationship in equations (3.33) yields:

$$\begin{aligned}\tilde{\sigma}_{xz}^{lo}(0) - \tilde{\sigma}_{xz}^{up}(0) - 4\dot{\epsilon}_{BG}(\mu^1 - \mu^2) \omega \frac{\tilde{v}_z^{up}(0)}{q} &= 0 \\ \tilde{\sigma}_{zz}^{lo}(0) - \tilde{\sigma}_{zz}^{up}(0) - (\rho^1 - \rho^2) g \frac{\tilde{v}_z^{up}(0)}{q} &= 0\end{aligned}\tag{3.35}$$

equations (3.30), (3.31), (3.32) and (3.35) thus are 8 equations for 8 unknowns (A-F). They can be written in matrix form:

$$\mathbf{AC} = \mathbf{R} \quad (3.36)$$

where \mathbf{A} is an 8 by 8 matrix, containing coefficients in front of the unknowns (A-F), \mathbf{C} is an 8 by 1 vector containing the 8 unknowns, and \mathbf{R} is an 8 by 1 vector, which is filled with zeros in the current case.

The system (3.36) has a nontrivial solution only if $\det(\mathbf{A})=0$. In practice this tedious task is done by using the symbolic manipulation package MAPLE, but even in this case the analytical solutions are complicated. An alternative method was described in Smith [1977], who determined the growth rate q numerically by inserting a random value for q in \mathbf{A} and iterating until the $\det(\mathbf{A})$ is zero. This gives the growth rate for a given wave number and for given physical parameters. This task can be done using the linear algebra package MATLAB. By scanning ranges of ω 's, the fastest growth rate is found. Analytical expressions of the maximum growth rate versus physical parameters were found by calculating growth rates for limiting cases (e.g. setting the background strain rate to zero gives pure diapiric modes). These analytical expressions were then compared to analytical expressions that were derived for simplified cases (e.g. the formula for dominant wavelength and growth rate for matrix folding was compared to an analytical solution that was derived for the case of folding if a layer resting on an infinite matrix). Once expressions for growth rates are known for every deformation mode, the boundaries between deformation modes can be calculated by requiring that the growth rate of two adjacent phases are equal at the phase boundary.

Bibliography

- M. R. Abbassi and N. S. Mancktelow. Single layer buckle folding in nonlinear materials .1. experimental-study of fold development from an isolated initial perturbation. *Journal of Structural Geology*, 14(1):85–104, 1992.
- M. Alavi. Tectonics of the Zagros orogenic belt of Iran: new data and interpretations. *Tectonophysics*, 229(3-4):211–238, 1994.
- F. Arthaud. *Etude tectonique et microtectonique comparée de deux domaines hercyniens: les nappes de la Montagne Noire (France) et l'anticlinorium de l'Iglesiente (Sardaigne)*. Université des Sciences et Techniques du Languedoc, Série Géologie Structurale, 1, Montpellier, 1970.
- M.A. Biot. Theory of folding of stratified viscoelastic media and its implications in tectonics and orogenesis. *Geological Society of America Bulletin*, 72:1595–1620, 1961.
- M.A. Biot and H. Odé. Theory of gravity instability with variable overburden and compaction. *Geophysics*, 30(2):215–227, 1965.
- J.-P. Brun. L'origine des dômes gneissiques: Modèles et tests. *Bulletin de la Société géologique de France*, 25:219–228, 1983.
- J. P. Brun and J. P. Burg. Combined thrusting and wrenching in the Ibero-Armorican arc: A corner effect during continental collision. *Earth and Planetary Science Letters*, 61:319–332, 1982.
- J.-P. Brun, D. Gapais, and B. Le Theoff. The mantled gneiss domes of Kuopio (Finland): interfering diapirs. *Tectonophysics*, 74:283–304, 1981.

- J.-P. Burg, P. Davy, P. Nievergelt, F. Oberli, D. Seward, Z. Diao, and M. Meier. Exhumation during crustal folding in the Namche-Barwa syntaxis. *Terra Nova*, 9 (2):53–56, 1997.
- J.-P. Burg, P. Nievergelt, F. Oberli, D. Seward, P. Davy, J.-C. Maurin, Z.Z. Diao, and M. Meier. The Namche Barwa syntaxis: evidence for exhumation related to compressional crustal folding. *Journal of Asian Earth Sciences*, 16(2-3):239–252, 1998.
- J.-P. Burg and Y. Podladchikov. Lithospheric scale folding: numerical modelling and application to the Himalayan syntaxes. *International Journal of Earth Sciences*, 88 (2):190–200, 1999.
- J.-P. Burg and Yu. Podladchikov. From buckling to asymmetric folding of the continental lithosphere: Numerical modelling and application to the Himalayan syntaxes. In M.A. Khan, P.J. Treloar, M.P. Searle, and M.Q. Jan, editors, *Tectonics of the Western Himalayas and Karakoram*, volume 170, pages 219–236. Geological Society Special Publications, London, 2000.
- E. Burov, C. Jaupart, and L. Guillou-Frottier. Ascent and emplacement of buoyant magma bodies in brittle-ductile upper crust. *Journal of Geophysical Research-Solid Earth*, 108(B4):art. no.-2177, 2003.
- P.J. Coney and T.A. Harms. Cordilleran metamorphic core complexes: Cenozoic extensional relics of Mesozoic compression. *Geology*, 12:550–554, 1984.
- C.P. Conrad and P. Molnar. The growth of Rayleigh-Taylor-type instabilities in the lithosphere for various rheological and density structures. *Geophysical Journal International*, 129:95–112, 1997.
- J.B. Currie, H.W. Patnode, and R.P. Trump. Development of folds in sedimentary strata. *Geological Society of America Bulletin*, 73(6):655–674, 1962.
- J. M. Dixon. Mantled gneiss domes. In C.K. Seyfert, editor, *Encyclopedia of structural geology and plate tectonics*, volume 10, pages 398–412. Nostrand Reinhold Company Inc, 1987.

- J.M. Dixon. Finite strain and progressive deformation in models of diapiric structures. *Tectonophysics*, 28:89–124, 1975.
- P. England and D.P. McKenzie. A thin viscous sheet model for continental deformation. *Geophysical Journal of the Royal Astronomical Society*, 70:295–321, 1982.
- P.E. Eskola. The problem of mantled gneiss domes. *Quarterly Journal of the Geological Society of London*, 104:461–476, 1949.
- R. C. Fletcher. Folding of a single viscous layer - exact infinitesimal amplitude solution. *Tectonophysics*, 39(4):593–606, 1977.
- R. C. Fletcher. 3-dimensional folding of an embedded viscous layer in pure shear. *Journal of Structural Geology*, 13(1):87–96, 1991.
- R. C. Fletcher. 3-dimensional folding and necking of a power-law layer - are folds cylindrical, and, if so, do we understand why. *Tectonophysics*, 247(1-4):65–83, 1995.
- R.C. Fletcher. Application of a mathematical model to the emplacement of mantled gneiss domes. *American Journal of Science*, 272:197–216, 1972.
- M. Gerbault. At what stress level is the central Indian Ocean lithosphere buckling? *Earth and Planetary Science Letters*, 178(1-2):165–181, 2000.
- M. Gerbault, E. B. Burov, A. N. B. Poliakov, and M. Daignieres. Do faults trigger folding in the lithosphere? *Geophysical Research Letters*, 26(2):271–274, 1999.
- D. Goff, D.V. Wilschko, and R.C. Fletcher. Decollement folding as a mechanism for thrust-ramp spacing. *Journal of Geophysical Research*, 101(B5):11341–11352, 1996.
- M.V. Gzovsky, A.S. Grigoryev, O.I. Gushchenko, A.V. Mikhailova, A.A. Nikonov, and D.N. Osokina. Problems of the tectonophysical characteristics of stresses, deformations, fractures and deformation mechanisms of the earth's crust. *Tectonophysics*, 18(1/2):167–205, 1973.
- G. Hunt, H. Muhlhaus, B. E. Hobbs, and A. Ord. Localized folding of viscoelastic layers. *Geologische Rundschau*, 85:58–64, 1996.

- A. T. Ismail-Zadeh, H. E. Huppert, and J. R. Lister. Gravitational and buckling instabilities of a rheologically layered structure: implications for salt diapirism. *Geophysical Journal International*, 148(2):288–302, 2002.
- M.P.A. Jackson and C.J. Talbot. Anatomy of mushroom-shaped diapirs. *Journal of Structural Geology*, 11(1/2):211–230, 1989.
- B. J. P. Kaus and Y. Y. Podladchikov. Forward and reverse modeling of the three-dimensional viscous Rayleigh-Taylor instability. *Geophysical Research Letters*, 28(6):1095–1098, 2001.
- L.L. Lavier and W.R. Buck. Half graben versus large-offset low-angle normal fault: Importance of keeping cool during normal faulting. *Journal of Geophysical Research*, 107(B6):10.1029/2001JB000513, 2002.
- N. S. Mancktelow. Finite-element modelling of single-layer folding in elasto- viscous materials: the effect of initial perturbation geometry. *Journal of Structural Geology*, 21(2):161–177, 1999.
- W.S. Pitcher. The nature, ascent and emplacement of granitic magma. *Journal of the Geological Society of London*, 136:627–662, 1979.
- Y. Podladchikov, C. Talbot, and A. N. B. Poliakov. Numerical models of complex diapirs. *Tectonophysics*, 228(3-4):189–198, 1993.
- A. Poliakov, Y.Y. Podladchikov, E. Dawson, and C.J. Talbot. Salt diapirism with simultaneous brittle faulting and viscous flow. In I. Alsop, D. Blundell, and I. Davison, editors, *Salt Tectonics*, volume 100, pages 291–302. Geological Society Special Publication, London, 1996.
- J. Pons, C. Oudin, and J. Valero. Kinematics of large syn-orogenic intrusions: example of the Lower Proterozoic Saraya batholith (Eastern Senegal). *Geologische Rundschau*, 81(2):473–486, 1992.
- N.J. Price and J.W. Cosgrove. *Analysis of geological structures*. Cambridge, 1990.

- H. Ramberg. Selective buckling of composite layers with contrasted rheological properties, a theory for simultaneous formation of several orders of folds. *Tectonophysics*, 1(4):307–341, 1964.
- H. Ramberg. Folding of laterally compressed multilayers in the field of gravity, ii, numerical examples. *Physics of the Earth and Planetary Interiors*, 4:83–120, 1970.
- H. Ramberg. Theoretical models of density stratification and diapirism in the earth. *Journal of Geophysical Research*, 77:877–889, 1972.
- H. Ramberg. The role of gravity in orogenic belts. In K.R. McKlay and N.J. Price, editors, *Thrust and nappe tectonics*, volume 9, pages 125–140. Geological Society Special Publication, London, 1981.
- H. Ramberg and O. Stephansson. Compression of floating elastic and viscous plates affected by gravity: a basis for discussing crustal buckling. *Tectonophysics*, 1:101–120, 1964.
- J.G. Ramsay. Tectonics of the Helvetic nappes. *Geological Society Special Publication*, 9:293–309, 1981.
- J.G. Ramsay. Emplacement kinematics of a granite diapir: the Chindamora batholith, Zimbabwe. *Journal of Structural Geology*, 11(1/2):191–209, 1989.
- G. Ranalli. *Rheology of the earth*. Chapman and Hall, London, 2nd edition, 1995.
- J.N. Reddy. *Theory and analysis of elastic plates*. Taylor and Francis, Philadelphia Pa, 1999.
- Regenauer-Lieb and D. Yuen. Modeling shear zones in geological and planetary sciences: solid- and fluid-thermal-mechanical approaches. *Earth-Science Reviews*, 63: 295–349, 2003.
- S. M. Schmalholz and Y. Podladchikov. Buckling versus folding: Importance of viscoelasticity. *Geophysical Research Letters*, 26(17):2641–2644, 1999.
- S. M. Schmalholz, Y. Y. Podladchikov, and D. W. Schmid. A spectral/finite difference method for simulating large deformations of heterogeneous, viscoelastic materials. *Geophysical Journal International*, 145(1):199–208, 2001.

- S.M. Schmalholz, Y. Podlachidkov, and J.-P. Burg. Control of folding by gravity and matrix thickness: Implications for large-scale folding. *Journal of Geophysical Research*, 107(B1):10.1029/2001JB000355, 2002.
- H. Schmeling. On the relation between initial conditions and late stages of the Rayleigh-Taylor instability. *Tectonophysics*, 133:65–80, 1987.
- R.B. Smith. Formation of folds, boudinage, and mullions in non-Newtonian materials. *Geological Society of America Bulletin*, 88(2):312–320, 1977.
- M. Smoluchowski. Versuche ueber Faltungerscheinungen schwimmender elastischen Platten. *Akad. Wiss. Krakau. Math. Kl.*, pages 727–734, 1910.
- P.A. Snowden and M.J. Bickle. The Chindamora batholith: diapiric intrusion or interference fold? *Journal of the Geological Society of London*, 132:131137, 1976.
- P.A. Snowden and D.V. Snowden. Petrochemistry of the late Archaean granites of the Chindamora batholith, Zimbabwe. *Precambrian Research*, 16(103-129), 1981.
- P. Stipska, K. Schulmann, and V. Hck. Complex metamorphic zonation of the Thaya dome: result of buckling and gravitational collapse of an imbricated nappe sequence. In J.W. Cosgrove and M.S. Ameen, editors, *Forced folds and fractures*, volume 169, pages 197–211. Geological Society Special Publication, London, 2000.
- C.J. Talbot and M.P.A. Jackson. Internal kinematics of salt diapirs. *Bulletin of the American Association of Petroleum Geologists*, 71(9):1068–1093, 1987.
- C. Teyssier and D. L. Whitney. Gneiss domes and orogeny. *Geology*, 30(12):1139–1142, 2002.
- D.L. Turcotte and G. Schubert. *Geodynamics*. Cambridge University Press, Cambridge, second edition edition, 2002.
- J.-L. Vigneresse and J.D. Clemens. Granitic magma ascent and emplacement: neither diapirism nor neutral buoyancy. In B. Vendeville, Y. Mart, and J.-L. Vigneresse, editors, *Salt, shale and igneous diapirs in and around Europe*, volume 174, pages 1–19. Geological Society Special Publication, London, bureau de recherches géologiques et minières edition, 2000.

-
- R. F. Weinberg and Y. Y. Podladchikov. The rise of solid-state diapirs. *Journal of Structural Geology*, 17(8):1183–1195, 1995.
- J.A. Whitehead. Fluid models of geological hotspots. *Annual Reviews of Fluid Mechanics*, 20:61–87, 1988.
- W.D. Wooldt. Finite element calculations applied to salt dome analysis. *Tectonophysics*, 50:369–386, 1978.
- P.K. Zeitler, C.P. Chamberlain, and H.A. Smith. Synchronous anatexis, metamorphism, and rapid denudation at Nanga Parbat (Pakistan Himalaya). *Geology*, 21:347–350, 1993.
- Y. Zhang, N. S. Mancktelow, B. E. Hobbs, A. Ord, and H.B. Muhlhaus. Numerical modelling of single-layer folding: clarification of an issue regarding the possible effect of computer codes and the influence of initial irregularities. *Journal of Structural Geology*, 22(10):1511–1522, 2000.
- M. T. Zuber. Compression of oceanic lithosphere: an analysis of intraplate deformation in the Central Indian basin. *Journal of Geophysical Research*, 92(B6):4817–4825, 1987.
- H.J. Zwart. The duality of orogenic belts. *Geologie en Mijnbouw*, 46(8):283–309, 1967.

Chapter 4

Initiation of localized shear in visco-elasto-plastic rocks.

Abstract Shear-localization is a process of primary importance for the onset of subduction and the evolution of plate-tectonics on Earth. In this chapter we focus on a model in which shear-localization is initiated through shear-heating. The rheology employed is linear Maxwell viscoelastic with von Mises plasticity and an exponential dependence of viscosity on temperature. Dimensional analysis reveals that four non-dimensional parameters control the initiation of shear-zones. The onset of shear-localization is systematically studied with 0-D, 1-D and 2-D numerical models, both under constant stress and under constant velocity boundary conditions. Mechanical phase-diagrams demonstrate that six deformation modes exist under constant velocity boundary conditions. A constant stress boundary condition, on the other hand, exhibits only two deformation modes (localization versus no-localization). Scaling laws for the growth rate of temperature are computed for all deformation modes. Numerical and analytical solutions demonstrate that diffusion of heat may inhibit localization. Initial heterogeneities are required to initiate localization. The derived scaling laws are applied to Earth-like parameters. For a given heterogeneity-size, stable (non-seismic) localization only occurs for a range of effective viscosities. Localization is inhibited if viscosity is smaller than a minimum threshold, which is a function of the heterogeneity size. Two-dimensional models are presented for a lithosphere subjected to homogeneous extension with an initially circular inclusion. The models demonstrate that both plasticity and elasticity are required to form lithospheric-scale shear-zones. The simplified rheological model is compared with a more realistic, and more complex

model of olivine that takes diffusion-, powerlaw and Peierls creep into account. Good agreement exists between the models. The simplified model proposed in this study thus reproduces the main physics of ductile faulting. Two-dimensional late-stage simulations of lithospheric-scale shear-localization are presented that confirm the findings of the initial stage analysis.

4.1 Introduction

Strain-localization is one of the long-lasting research topics in theoretical geodynamics. There are several motivations for this. First, earthquakes are frequently related to large-scale, plate bounding fault-zones. The formation of faults in the upper 10-15 km of the lithosphere is most likely a result of brittle fracture [Scholz, 2002]. This process is often modeled with simple Mohr-Coulomb friction or Byerlees law [Brace and Kohlstedt, 1980] combined with elastic or viscoelastic, distributed deformation [e.g. Poliakov et al., 1994, Lavier et al., 2000]. Once a brittle fault-zone has formed, continuing deformation on the fault is governed by the rheology of the fault gouge. Repeating earthquake cycles can occur if the fault gouge has the ability to both weaken and strengthen depending on the applied velocity and/or stress. Rheological models that have this ability include the rate and state-dependent friction model [Dieterich, 1979, Ruina, 1983], viscoelastic damage mechanics [e.g. Lyakhovskiy et al., 1997, submitted 2004], or dynamic grain-growth models [Montesi and Hirth, 2003].

Whereas the occurrence of earthquakes in the upper brittle part of the lithosphere appears to be relatively well understood, the presence of earthquakes deeper in the lithosphere is more puzzling. Laboratory experiments indicate that the brittle strength of rocks should only apply for the upper 10-20 km. Deeper parts of the oceanic and continental lithosphere, which are able to sustain significant differential stresses, are most likely dominated by ductile deformation mechanisms [Kohlstedt et al., 1995]. Experiments on olivine indicated that under large confining stresses, the deformation mechanism of olivine changes from diffusion creep or dislocation creep to the so-called Peierls creep mechanisms, which effectively limits the maximum differential stress of the rock sample [e.g. Goetze and Evans, 1979]. Qualitatively, the Peierls mechanism can be regarded as von Mises plasticity, in which the yield stress is weakly dependent on

strainrate and temperature. It is well-known from engineering studies on metals that von Mises plasticity does not lead to spontaneous localization of strain in small bands; it rather tends to form a broad band of plastic deformation (so-called metal plasticity). Thus, additional mechanisms are required to explain the formation of shear-zones (and earthquakes) in the ductile, high strength part of the lithosphere. One possibility that has been explored is the self-lubrication model in which effective viscosity drops both with stress and strainrate [e.g. Bercovici et al., 2001]. Another possibility is the shear-heating model in which viscosity is reduced due to increased temperature [e.g. Yuen et al., 1978].

The problem of shear-localization in the non-brittle part of the lithosphere has gained significant attention from the mantle-convection community over the last decade. This attention is mainly driven by the fact that early models of mantle convection with a realistic temperature-dependent viscosity [Moresi and Solomatov, 1995], did not develop plate-tectonic behavior but instead lead to the formation of a one-plate planet. While this result may be applicable for Venus, it is obviously not relevant to the Earth. Subsequent work showed that weak plate boundaries, initially prescribed in the models, are required to develop plate tectonic behavior. Self-consistent models of plate-like behavior, in which fault zones form spontaneously from the rheology, have mainly concentrated on damage-like self-lubricating models [e.g. Bercovici, 1993, 1996, Tackley, 1998, 2000c, Bercovici, 2003, Auth et al., 2003, Ogawa, 2003] or on pseudo-plastic yield stress rheologies [Trompert and Hansen, 1998, Tackley, 2000a,b]. The pseudo-plastic rheologies seem to perform slightly better in generating plate-like behavior with mid-oceanic ridges [Tackley, 2000a] and subduction in 3D geometries. However, these models still face some drawbacks. First, they should be regarded as being ad-hoc (although more recent developments have derived the governing equations from first-principles see e.g. Bercovici et al. [2001]). Second, plate-like behavior occurs for a very small parameter range only, and for yield-strengths that are up to an order of magnitude lower than laboratory-derived measurements [Tackley, 2000b]. Asymmetric subduction does not form spontaneously and the amount of strike-slip motion (e.g. like on the San Andreas fault) in these models is lower than on the Earth. Third, the pseudo-plastic rheology is instantaneous and has no memory effect; once the applied strainrate is reduced the faults disappear. In Auth et al. [2003] 2D numerical results have been reported which employed a more sophisticated damage rheology formulation, and that did result in

asymmetric subduction. However the applicability of their models to 3-D cases is still to be tested.

A different line of research has focussed on the generation of lithospheric-scale shear-zones by induced thermal weakening due to shear-heating. Early work [Turcotte and Oxburgh, 1968, Yuen et al., 1978] on this instability demonstrated that it is possible to generate a weak zone with this mechanism, but only under the presence of very large stresses and under somewhat restrictive conditions of a constant thickness shear-zone, whose boundaries are maintained at a constant temperature. If, on the other hand, a heterogeneity is initially present shear-localization occurs more easily [Fleitout and Froidevaux, 1980, Ogawa, 1987, Montesi and Zuber, 2002] but still requires relatively large stresses (typically > 1000 MPa). In nature, the stress in the lithosphere is limited by, for example, the Peierls creep mechanism. One-dimensional models of shear-localization that incorporate laboratory-derived rheological laws demonstrate that the Peierls mechanism generally inhibits shear-localization, unless a constant stress boundary condition or large initial stresses are present [Kameyama et al., 1999]. Essentially all models cited above have been developed for 0-D or 1-D cases and are thus applicable to study the evolution of temperature and strain rate on preexisting zones of weakness. However, a different question, relevant in the context of self-consistent plate tectonics models, is how these zones are created. Advances in this direction have been made by Regenauer-Lieb and Yuen [1998], who demonstrated that quasi-adiabatic shearzones develop in a pre-stressed lithosphere under extension if the shear-zone is initiated by a notch-like initial heterogeneity [Regenauer-Lieb and Yuen, 1998]. These studies have been followed by work that concentrated on the effect of thermo-mechanical coupling with plasticity and more complex rheologies, including water-dependent diffusion creep, powerlaw creep and Peierls creep [Regenauer-Lieb et al., 1999, Regenauer-Lieb and Yuen, 2000a,b, Regenauer-Lieb et al., 2001, Regenauer-Lieb and Yuen, 2004]. Regenauer-Lieb et al. [2001] demonstrated that it is possible to generate an asymmetric lithospheric-scale shearzone by adding ~ 10 km's of sediments on top of the oceanic plate, provided the material is governed by a 'wet'-rheology. This conclusion has been confirmed in more recent work [Regenauer-Lieb and Yuen, 2004] in which several thermal heterogeneities have been introduced in a lithosphere extended at constant velocity.

These studies thus highlight the possibility of creating lithospheric-scale shear-zones

with a mechanism consistent with current knowledge on rheology. However, drawbacks of the mentioned studies are that they are numerically very expensive, require a complex rheology with many parameters, and have ignored the mantle flow. Moreover, a quantitative understanding of the parameters that control the initiation and the evolution of ductile shearzones is not available. The focus of the current paper is to study the *initiation* of a shear-zone in the framework of a somewhat simpler rheology than that employed by Regenauer-Lieb and Yuen [1998]. The initiation phase is defined here as the stage during which the temperature rises by several degrees. We use a simple linear visco-elasto-plastic model, in which the viscosity is related to temperature by a simple exponential relationship. Whereas our model is not entirely consistent with state-of-the art rheological laws, it has the advantage of having fewer parameters. This allows for a systematic study covering the whole parameter space. It is demonstrated that four nondimensional parameters control the initiation of shear-zones, under both constant stress and constant velocity boundary conditions. At the end of this chapter, and in future work, we will focus on the late-stage evolution of shear-zones.

4.2 Mathematical model and non-dimensionalization

4.2.1 Governing equations

We employ two different rheologies: (i) a visco-elasto-plastic rheology with an exponential temperature-dependent (linear) viscosity (the 'simple' model) and (ii) a rheology which is more applicable to olivine and considers the effects of diffusion creep, powerlaw creep and Peierls low-temperature plasticity (appendix A). In this section the governing equations for the "simple" model are described and suitable non-dimensionalization schemes are proposed for different boundary conditions. The rheology of dry and wet olivine is described in more detail in appendix A, where rules are given to transform model (ii) into model (i).

It is assumed that deformation-induced volume changes are negligible compared to the total shear deformation of rocks, which leads to the incompressibility constraint:

$$\frac{\partial v_i}{\partial x_i} = 0 \quad (4.1)$$

here v_i is velocity and x_i the spatial coordinates. Assuming that the effects of inertia and gravity can be ignored, force equilibrium gives:

$$\frac{\partial \sigma_{ij}}{\partial x_j} = 0 \quad (4.2)$$

where σ_{ij} are stresses. Furthermore:

$$\sigma_{ij} = -P\delta_{ij} + \tau_{ij} \quad (4.3)$$

where τ_{ij} are deviatoric stresses and pressure P is given by:

$$P = -\frac{\sigma_{ii}}{3} \quad (4.4)$$

The strainrate is defined as

$$\dot{\epsilon}_{ij} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \quad (4.5)$$

The rheology is visco-elasto-plastic:

$$\dot{\epsilon}_{ij} = \dot{\epsilon}_{ij}^{vis} + \dot{\epsilon}_{ij}^{el} + \dot{\epsilon}_{ij}^{pl} = \frac{1}{2\mu} \tau_{ij} + \frac{1}{2G} \frac{D\tau_{ij}}{Dt} + \dot{\lambda} \tau_{ij} \quad (4.6)$$

where D/Dt denotes the objective derivative of the stress tensor versus time t , μ is the shear viscosity, G the elastic shear modulus and $\dot{\lambda}$ a to-be-determined variable to ensure that the yield criterion is not violated. Laboratory experiments [e.g. Goetze and Evans, 1979] indicate that the stress of rocks in the upper mantle is limited by the Peierls creep mechanism (see Appendix A). In our simplified rheological model we approximate this mechanism by the pressure-independent von Mises yield criteria [Simo and Hughes, 2000], which can be written as

$$F = |\tau| - \sigma_y$$

where $|\tau| = (0.5\tau_{ij}\tau_{ij})^{0.5}$ is the effective stress (or the radius of the Mohr-circle) and σ_y is the yield stress. Viscosity is assumed to be temperature-dependent according to the Frank-Kamenetzky approximation [Moresi and Solomatov, 1995, McKenzie, 1977, Frank-Kamenetzky, 1969]:

$$\mu = \mu_0 e^{(-\gamma T^{di})} \quad (4.7)$$

where $T^{di} = T - T_0$ is the temperature difference compared the to initial temperature, T temperature, μ_0 initial viscosity, T_0 initial temperature and γ can be computed from the rheology according to $\gamma = Q/(RT_0^2)$ (with Q being the activation energy and R the gas-constant). Conservation of energy is given by

$$\left(\frac{\partial T^{di}}{\partial t} + v_i \frac{\partial T^{di}}{\partial x_i} \right) = \kappa \frac{\partial^2 T^{di}}{\partial x_j^2} + \frac{\tau_{ij}(\dot{\epsilon}_{ij}^{vis} + \dot{\epsilon}_{ij}^{pl})}{\rho c_p} \quad (4.8)$$

where κ is thermal diffusivity, ρ density and c_p heat capacity and the last term denotes shear-heating due to dissipative, non-recoverable processes.

4.2.2 Non-dimensionalization

We will consider both cases in which the far-field stress is kept constant and cases in which the far-field strainrate is maintained at a constant level $\dot{\epsilon}_{BG}$. The number of non-dimensional parameters that occur in the present setup can be minimized by choosing as characteristic values σ_0 , μ_0 , κ and γ . Characteristic values for stress, time, temperature, length and strainrate are given by:

$$\begin{aligned} \sigma^* &= \sigma_0 \\ t^* &= \frac{\mu_0}{G} \end{aligned}$$

$$T^* = \frac{1}{\gamma}$$

$$L^* = \sqrt{\frac{\kappa\mu_0}{G}} \quad (4.9)$$

In the case of a constant strainrate boundary condition, the rheological equation (eq. 7.3) in nondimensional form (with $\tilde{\cdot}$ denoting nondimensional variables) can be written as:

$$\tilde{\dot{\epsilon}}_{ij} = \frac{\sigma_0 e^{\tilde{T}^{di}}}{2\mu_0} \tilde{\tau}_{ij} + \left(\frac{\sigma_0 G}{\mu_0} \right) \frac{1}{2G} \frac{D\tilde{\tau}_{ij}}{D\tilde{t}} + \frac{\sigma_0}{2\mu_0} \dot{\lambda} \tilde{\tau}_{ij} \quad (4.10)$$

which may be simplified to

$$\frac{2\mu_0 \tilde{\dot{\epsilon}}_{ij}}{\sigma_0} = e^{\tilde{T}^{di}} \tilde{\tau}_{ij} + \frac{D\tilde{\tau}_{ij}}{D\tilde{t}} + \dot{\lambda} \tilde{\tau}_{ij} \quad (4.11)$$

Since the background strain rate is maintained at a constant value, the terms on the left-hand-side can be collected into a nondimensional parameter $B_{vis} = \frac{2\mu_0 \dot{\epsilon}_{BG}}{\sigma_0}$, where $(0.5\dot{\epsilon}_{ij}\dot{\epsilon}_{ij})^{0.5} = \dot{\epsilon}_{BG}$, indicates the magnitude of the steady-state viscous stress versus the initial stress. The energy equation in nondimensional form is given by

$$\left(\frac{\partial \tilde{T}^{di}}{\partial \tilde{t}} + \tilde{v}_i \frac{\partial \tilde{T}^{di}}{\partial \tilde{x}_i} \right) = \frac{\partial^2 \tilde{T}^{di}}{\partial \tilde{x}_j^2} + Br \tilde{\tau}_{ij} (\tilde{\epsilon}_{ij}^{vis} + \tilde{\epsilon}_{ij}^{pl}) \quad (4.12)$$

where $Br = \frac{\sigma_0^2 \gamma}{\rho c_p G}$ is the modified Brinkman number that indicates the efficiency of shear-heating. If, additionally, a heterogeneity length scale R is introduced, four nondimensional numbers are present:

$$B_{pl} = \frac{\sigma_0}{\sigma_y}$$

$$Br = \frac{\sigma_0^2 \gamma}{\rho c_p G}$$

$$Pe = \frac{R}{\sqrt{\frac{\kappa\mu_0}{G}}}$$

$$B_{vis} = \frac{2\mu_0 \dot{\epsilon}_{BG}}{\sigma_0} \quad (4.13)$$

Physically, B_{pl} denotes the importance of the initial stress versus the yield stress, Br the efficiency of shear-heating, Pe the ratio of an heterogeneity length scale over the diffusion length scale, and B_{vis} the ratio of steady-state viscous stress over initial stress.

Depending on the type of boundary and initial conditions, several combinations of nondimensional parameters are valid:

1. Constant stress boundary condition. This is a type of condition that may be applicable inside a subducting slab, where the stress is approximately constant with time. $B_{vis} = 1$ in this case.

2. Constant strain rate boundary condition with a relaxed initial stress state ($B_{vis} \gg 1$). The condition may apply to the stress state of a lithosphere at the onset of rifting.

3. Constant strain rate boundary condition with viscous initial stress distribution ($B_{vis} = 1$). An example would be a lithosphere that has been deforming for several Maxwell times and starts deforming in a different direction.

4. Constant strainrate boundary condition with little movement and viscous initial stress ($B_{vis} \ll 1$). This represents the condition when a lithosphere with a large initial stress is *held* almost undeformed. Most of the effects will be due to the release of elastic energy.

Upper bounds for the non-dimensional parameters given above can be estimated by assuming that $G \simeq 10^{10} - 10^{11}$ Pa, $\kappa = 10^{-6}$ m²s⁻¹, $\gamma = 0.01 - 0.1$, $\rho c_p = 3 \times 10^6$ [see for example Turcotte and Schubert, 1982]. Viscosity is assumed to vary from $\mu_0 = 10^{18} - 10^{28}$ Pas. Moreover, the theoretical yield strength of rocks (given by the point at which bonds between atoms start to break) is around $\frac{1}{10}$ of the shear modulus, giving a value of $\sigma_y \leq 10^9 - 10^{10}$ Pa [e.g. Scholz, 2002]. Typical yield strengths of rocks vary between 500 – 1000 MPa [Regenauer-Lieb and Yuen, 2003]. Physical meaningful results can only be obtained if the initial stress, σ_0 , is less than or equal to the yield stress. The stress state of the lithosphere may be estimated from stress released from earthquakes [typically around 10 MPa, Lachenbruch and Sass, 1991] and from rheology-independent force balance considerations about the average lithospheric stress required to support mountain belts [around 100 MPa, Jeffrey, 1959]. Maximum stress values occur in a preloaded lithosphere and will have values close to the yield stress (1000 MPa). From these considerations, the initial stress is assumed to vary from $\sigma_0 = 1 - 1000$ MPa. Strain rate estimates for geodynamic processes give $\dot{\epsilon}_0 = 10^{-12} - 10^{-18}$ s⁻¹. An upper bound for the size of the heterogeneity is obviously the thickness of the lithosphere $\sim O(100)$ km, so $R = 0.1 - 10^5$ m. Using these values, the following parameter-range can be estimated to be realistic for the non-dimensional

numbers given in eq. (4.13):

$$\begin{aligned}
10^{-4} &\leq B_{pl} \leq 0.4 \\
3 \times 10^{-8} &\leq B_r \leq 3 \\
1 \times 10^{-9} &\leq Pe \leq 3 \times 10^3 \\
2 \times 10^{-9} &\leq B_{vis} \leq 2 \times 10^{10}
\end{aligned} \tag{4.14}$$

The characteristic length scale L^* varies between 30 m and 10^5 km.

4.3 Numerical methods

The initiation of shear-localization is studied with 0-D, 1-D, and 2-D models. In this section the numerical techniques employed for the different models are described, whereas results obtained with these methods are given in section 4.4. All values are in non-dimensional units; we have dropped \sim for readability. Transformation of the non-dimensional into dimensional units can be done by multiplying them with the characteristic values (eq. 4.9).

4.3.1 0-D model

In order to study the process of 1-D and 2-D shear-localization, one should first study the effects purely due to the employed rheology. A 0-D model serves this purpose by ignoring any spatial derivatives. In the case of a constant stress boundary condition ($\frac{d\tau}{dt} = 0$ and $\dot{\lambda} = 0$), the model is describe by a single ordinary differential equation (ODE) :

$$\frac{\partial T^{di}}{\partial t} = f_s Br \tau^2 \frac{e^{T^{di}}}{2} \tag{4.15}$$

where f_s is a parameter that describes the type of the applied shear; $f_s = 1$ for simple shear and $f_s = 2$ for pure shear. Moreover, $\tau = 1$ due to the chosen non-dimensionalization. An analytical solution of eq. 4.15 (assuming that $T^{di}(0) = 0$) is:

$$T^{di}(t) = \ln \left(-\frac{1}{f_s Br \left(t - \frac{2}{f_s Br} \right)} \right) \tag{4.16}$$

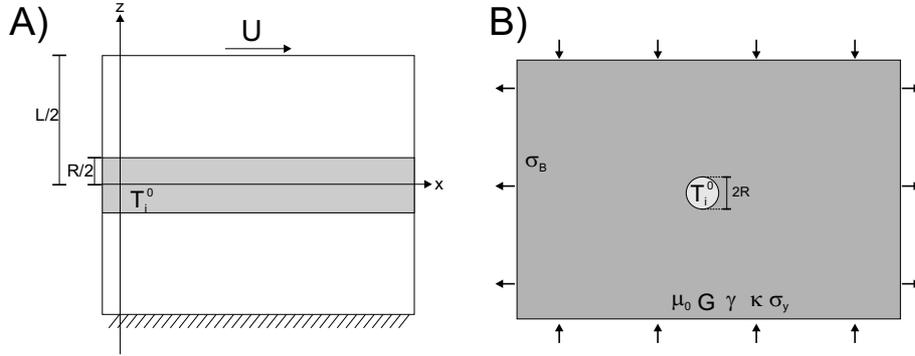


Figure 4.1: A) Setup of the one-dimensional model. B) Setup of the two-dimensional model.

The solution of eq. (4.16) always tends to 'blow-up' after a blow-up time $t \simeq \frac{2}{f_s B r}$ (the point where the solution becomes imaginary). A constant stress boundary condition will thus lead to thermal runaway, a well-established fact [e.g. Turcotte and Schubert, 1982, Kameyama et al., 1999].

In the case of a constant strainrate boundary condition, two coupled ODE's apply:

$$\begin{aligned} \frac{\partial T^{di}}{\partial t} &= f_s B r \tau (\dot{\epsilon}_{BG} - \dot{\epsilon}^{el}) \\ \frac{d\tau}{dt} &= 2\dot{\epsilon}_{BG} - e^{T^{di}} \tau - 2\dot{\lambda} \tau \end{aligned} \quad (4.17)$$

here $\dot{\epsilon}_{BG} = B_{vis}/2$ and $\dot{\epsilon}^{el} = \frac{1}{2} \frac{d\tau}{dt}$. If $\frac{d\tau}{dt}$ is discretized as $\frac{\tau - \tau^{old}}{dt}$, an expression for $\dot{\lambda}$ is given by:

$$\dot{\lambda} = \begin{cases} 0, & \text{if } \tau \leq \sigma_y \\ -\frac{1}{2dt} - \frac{e^{T^{di}}}{2} + \frac{\dot{\epsilon}_{BG}}{\sigma_y} + \frac{1}{2dt\sigma_y} \tau^{old}, & \text{if } \tau > \sigma_y \end{cases} \quad (4.18)$$

where $\sigma_y = B_{pl}^{-1}$. A general analytical solution for eqs. (4.17) does not exist. Therefore we have chosen to solve the system of equations numerically (see Appendix B for MATLAB source codes).

4.3.2 1-D model

The 1-D model (Fig. 4.1a) assumes that a slab of thickness L is subjected to simple shear, with either a constant shear stress or a constant velocity (which is equivalent to

a constant background strain rate). Localization is initiated by increasing the initial temperature to T_i^0 at $z = [-R/2..R/2]$. The commonly-made assumption of small strains and small rotations [Cherukuri and Shawki, 1995] is employed. In this case, eqs. (4.2) can be simplified to:

$$\frac{\partial \tau_{xz}}{\partial z} = 0 \quad (4.19)$$

and equation 4.12 can be simplified to

$$\frac{\partial T^{di}}{\partial t} = \frac{\partial^2 T^{di}}{\partial z^2} + Br\tau_{xz}(\dot{\epsilon}_{xz} - \dot{\epsilon}_{xz}^{el}) \quad (4.20)$$

Note that eq. (4.19) implies constant τ_{xz} over the model domain. Rheology is given by:

$$\frac{d\tau_{xz}}{dt} = 2\dot{\epsilon}_{xz} - e^{T^{di}}\tau_{xz} - 2\dot{\lambda}\tau_{xz} \quad (4.21)$$

Shear stress can be obtained by integrating eq. (4.21) with respect to z :

$$L \frac{d\tau_{xz}}{dt} = U - \tau_{xz} \int_{-L/2}^{L/2} (e^{T^{di}} + \dot{\lambda}) dz \quad (4.22)$$

From this expression it is apparent that elasticity is not important under a constant stress boundary condition, since $\frac{d\tau_{xz}}{dt} = 0$. In this case, the velocity at the boundary U will change (typically increase) with time. Time discretization of equation 4.22 gives:

$$L \frac{\tau_{xz} - \tau_{xz}^{old}}{dt} = U - \tau_{xz} \int_{-L/2}^{L/2} (e^{T^{di}} + \dot{\lambda}) dz \quad (4.23)$$

which can be rewritten as:

$$\tau_{xz} = \frac{U dt}{L + \eta dt} + \frac{L}{L + \eta dt} \tau_{xz}^{old} \quad (4.24)$$

with

$$\eta = \int_{-L/2}^{L/2} (e^{T^{di}} + \dot{\lambda}) dz \quad (4.25)$$

In the case of constant strainrate boundary conditions, the problem is determined by equations (4.20), (4.24) and (4.25). In the case of constant stress boundary conditions,

the elastic terms in equation (4.21) disappear, and the problem is reduced to one with quasi-viscous rheology governed by equations (4.20) and (4.21) (with $\frac{d\tau_{xz}}{dt} = 0$).

Boundary conditions for temperature are adiabatic both at the upper and the lower boundaries ($\frac{\partial T^{di}}{\partial z} = 0$). The initial condition of temperature is uniform ($T^{di} = 0$), except in the center of the domain between $-R \leq z \leq R$, where the temperature is raised to T_i^0 . The stress at $t = 0$ is given by $\tau(t = 0) = \sigma_0/\sigma_0 = 1$. The governing equation for temperature is discretized with a standard second order implicit finite difference method on a non-uniform grid with a numerical resolution of at least 2001 gridpoints. The governing equation for momentum is advanced in time by numerical integration of the temperature field [Fleitout and Froidevaux, 1980, Kameyama et al., 1999]. The timestep is variable and similar to the one employed in the ODE-solver (Appendix B). Resolution tests have been performed to ensure that the temporal and spatial resolution are sufficient. A total of more than 100'000 1D simulations have been performed to study the onset of shear localization in a systematic manner.

4.3.3 2-D model

In order to test whether the one-dimensional simulations are applicable in two-dimensional settings, we have performed over 500 two-dimensional simulations. For this we have used two numerical codes, recently developed in the frame of this thesis. The first code, GANGO, is a finite-difference/spectral method which solves the governing equations for visco-elasto-plasticity in an Eulerian/Lagrangian framework, using a staggered finite-difference discretization in the vertical direction and a spectral approach in the horizontal direction. Lateral variations in viscosity are treated iteratively. Stress advection (and rotation), which is required for the treatment of viscoelasticity, is done by using a semi-Lagrangian advection scheme. The energy equation is solved on a grid that has a resolution two times higher than the mechanical part of the code. Time movement is done by an implicit approach, which additionally checks that the temperature increase per timestep is not larger than a specified value. The code has been extensively benchmarked versus a range of different setups (thermal advection, diffusion, shear heating, stress around a cylindrical inclusion, viscoelastic buckling, stress rotation). Further details of the numerical code can be found in [Kaus et al., 2004], and a summary of benchmark results can be obtained from the first author's web-

page (<http://www.geology.ethz.ch/sgt/staff/boris/>). Relevant for the current work are comparisons of stress distribution around a 2D elliptical viscous inclusion in a viscous matrix, for which 2D analytical solutions were described in Schmid and Podladchikov [2003]. Moreover, comparisons have been made with the 0-D ODE solution of Appendix B.

The second numerical code, SloMo, employs the finite element technique to solve the incompressible Stokes equations. The velocity-pressure formulation is employed on quadrilateral elements with quadratic shape functions for velocity and discontinuous linear shape functions for pressure. The rheology is temperature-dependent visco-elasto-plastic. In the present study, the code has been used in a purely Lagrangian manner, which has the advantage that stress advection problems do not arise but the disadvantage that simulations are terminated when the elements become too distorted. However this drawback was not found to be very severe for the current study since (i) we focus on the onset of the instability and (ii) the finite element computations serve as a comparison tool with the simulations performed with GANGO, since this code is faster for the present setup. Excellent agreement exists between the results obtained with SloMo and those obtained with GANGO.

Two different 2D models are employed. The first setup consists of a circular inclusion of radius $R/2$ where the initial temperature is increased to a value T_i^0 (see fig. 4.1). Two different boundary conditions are employed: (1) a pure-shear background strainrate condition, during which the model is extended at a constant strainrate $\dot{\epsilon}_{bg}$. (2) A constant stress boundary condition, which is numerically treated by iteratively changing $\dot{\epsilon}_{bg}$ until the second invariant of the stress tensor at the lateral boundaries is within 0.1% of the required boundary stress. The thermal boundary conditions are periodic in the lateral directions and zero-flux on the lower and upper boundary. The size of the domain is 30-50 times larger than the radius of the inclusion, to ensure that the boundaries do not influence the results. Our 2-D simulations have been performed under extension. If a compressional boundary condition is applied, variations in viscosity between the lithosphere and the underlying mantle may induce additional instabilities like buckling, which makes it more difficult to separate the different effects. The initial heterogeneity is chosen to be circular, since this represents the worst-case scenario for shear-localization: differently shaped inclusions generate larger stress perturbations [e.g. Schmid and Podladchikov, 2003], and hence simplify initiation of localization.

The second 2D setup is almost identical to the first setup, with the difference that a layer of 'air' is added on top, and a layer of asthenosphere at the bottom of the model. Again a circular heterogeneity is introduced in order to initiate localization.

4.4 Initiation stage

4.4.1 0-D model

Stress and temperature evolutions computed for both the simplified rheological model and the dry-olivine model are presented in figure 4.2. High initial temperatures result in relatively small stresses and thus relatively small amounts of shear-heating (fig. 4.2a). Lower initial temperatures result in an increase of stress up to the yield stress after which the stress remains constant or decreases due to the decrease of the temperature-dependent effective viscosity. During the viscoelastic stress-buildup stage, relatively little shear-heating occurs (most of the deformation goes into reversible elastic stress-increase). Once the yield stress has been reached, however, $\frac{d\tau}{dt} = 0$ and all further deformation is done by dissipative plastic and viscous mechanisms. There is thus a sudden increase in shear-heating, reflected by a kink in the temperature-time evolution (fig. 4.2b). If plasticity is deactivated (fig. 4.2c) stress-buildup continues in the simplified model and saturates at the viscous steady-state level in the dry-olivine model (which is unrealistically large). Once the steady-state stress is reached, elastic strainrates are negligible and shear-heating is at its maximum, which results in a faster heating-rate for the olivine model than for the simplified model. In the (unrealistic) case without plasticity, the simplified model does not describe the evolution very well. This is caused by the fact that the rheology of the complex model switches from diffusion creep to powerlaw creep at large stresses. The simplified rheology is estimated with only one of these two deformation mechanisms, and thus overestimates the strength of the material. If the initial stress of the model is (unrealistically) large (fig. 4.2d) and plasticity is not considered, temperature may at some stage increase in an almost explosive way (this has previously been called a runaway instability). In future work we demonstrate that strictly speaking a runaway process does not occur under constant strainrate/velocity boundary conditions, but we will keep the term since the process may indeed result in large temperatures and eventually melting. It

is interesting that the results obtained here with a 0-D model are in good agreement with results obtained with the 1-D, simple-shear, model of Kameyama et al. [1999] in terms of the time required to start the runaway instability.

A comparison of the simplified 0-D model (having 3 parameters) with the olivine model (having 12 parameters; see appendix A) shows that the results are in fairly good agreement concerning the overall temperature and stress evolutions. The largest differences occur when plasticity is deactivated (fig. 4.2c), in which case the simplified model underestimates the growthrates of temperature. Given the uncertainties in rheological constants, initial temperatures and initial stresses of the lithosphere, we regard the simplified model to be acceptable to study the onset of shear-localization. The remaining part of this chapter therefore concentrates on this model.

The previous results thus point out that diverse responses (e.g. runaway-like/no runaway, plastic yielding/no plastic yielding) occur for different initial temperatures or different non-dimensional parameters. In order to compare the various models, we define a growth rate q_{0D} which is given by:

$$q_{0D} = \frac{\Delta T}{\Delta t} \quad (4.26)$$

where Δt is the time interval required to raise the temperature by a given unit ΔT . Throughout most of this work, we will concentrate on the initiation stage of shear-localization. All numerical simulations have been integrated until an increase ΔT in temperature was detected.

In the nondimensional 0-D model, $T^{di}(t = 0) = 0$. The non-dimensional parameters in this setup are B_{vis} , Br and B_{pl} . A contour plot of q_{0D} versus Br and B_{vis} for two different values of B_{pl} (Fig. 4.3) demonstrates that five different domains occur, each with its own dependence of q_{0D} on the non-dimensional parameters (see table 4.1 for scaling relationships). The boundaries between these different domains are relatively sharp. The domains represent different modes of deformation and have significantly different stress-time evolutions. We can distinguish:

1. *Viscous mode.* Stress changes from the initial value to the viscous stress ($\tau = 2\mu\dot{\epsilon}_{BG}$) with little overall increase in temperature. After this, stress drops due to the rise in temperature. The growthrate ($q_{0D} \propto BrB_{vis}^2$) is dependent on the elastic shear modulus, as well as on the background strainrate and effective viscosity.

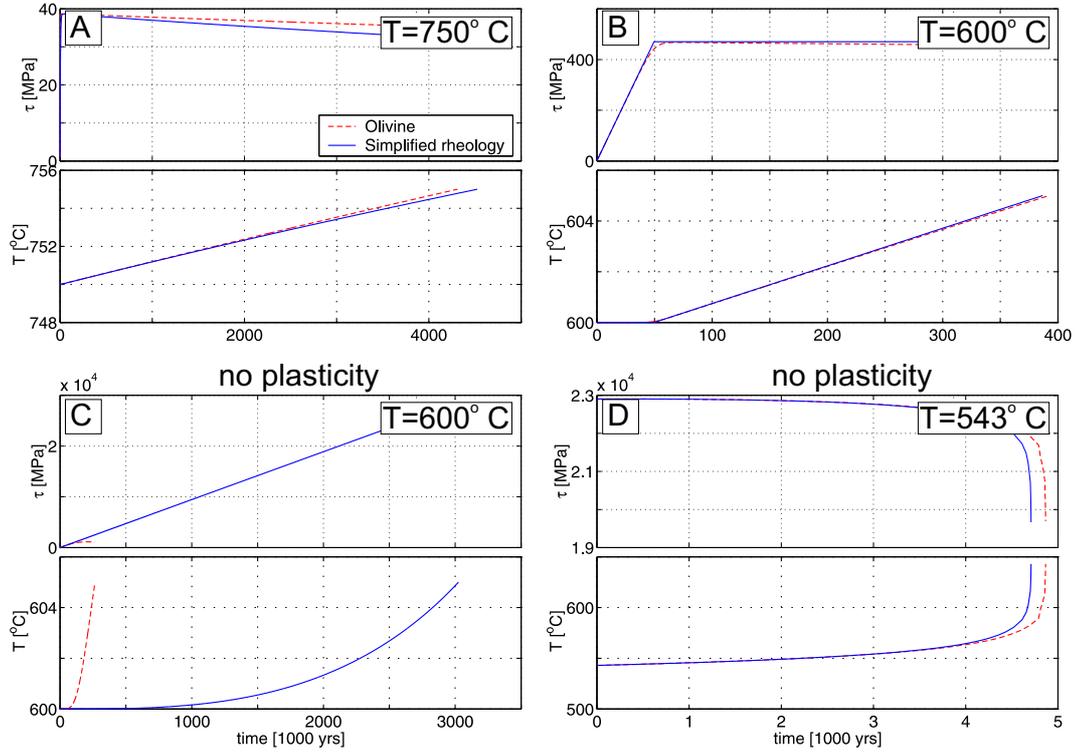


Figure 4.2: General behavior of the 0-D model. Models A-C are computed for dry olivine using $G = 5 \times 10^{10}$ Pa, $\dot{\epsilon} = 3 \times 10^{-15}$ s $^{-1}$, $\rho c_p = 3 \times 10^6$ Jkg $^{-1}$ m $^{-3}$, $\sigma_0 = 1$ MPa and pure-shear deformation. Models A and B include the effect of plasticity (or Peierls creep), whereas this has been disabled in models C and D. Model D is shown as a comparison with a 1D-model of [Kameyama et al., 1999, see their fig. 3] using $G = 8 \times 10^{10}$ Pa, $\dot{\epsilon} = 1.6 \times 10^{-14}$ s $^{-1}$, $\rho c_p = 2.4 \times 10^6$ Jkg $^{-1}$ m $^{-3}$, $\sigma_0 = 23 \times 10^9$ Pa under overall simple-shear conditions. The following parameters apply for the different models: A) $\mu_0 = 6.5 \times 10^{21}$ Pas, $\gamma = 0.039$ K $^{-1}$, $\sigma_y = 104$ MPa, $B_{vis} = 39$, $Br = 1.3 \times 10^{-7}$ and $B_{pl} = 1 \times 10^{-3}$. B) $\mu_0 = 9.7 \times 10^{26}$ Pas, $\gamma = 0.05$ K $^{-1}$, $\sigma_y = 470$ MPa, $B_{vis} = 5.8 \times 10^6$, $Br = 1.5 \times 10^{-7}$ and $B_{pl} = 2.1 \times 10^{-3}$. C) as B) but with $\sigma_y = 10^{10}$ Pa and $B_{pl} = 10^{-10}$. D) $\mu_0 = 7.5 \times 10^{23}$ Pas, $\gamma = 0.097$ K $^{-1}$, $\sigma_y = 10^{16}$ Pa, $B_{vis} = 1.0$, $Br = 133$ and $B_{pl} = 10^{-10}$. MATLAB scripts described in Appendix A and B have been used to obtain A-C. Model D has been computed for the dry-olivine rheology of [Kameyama et al., 1999, see their fig. 3].

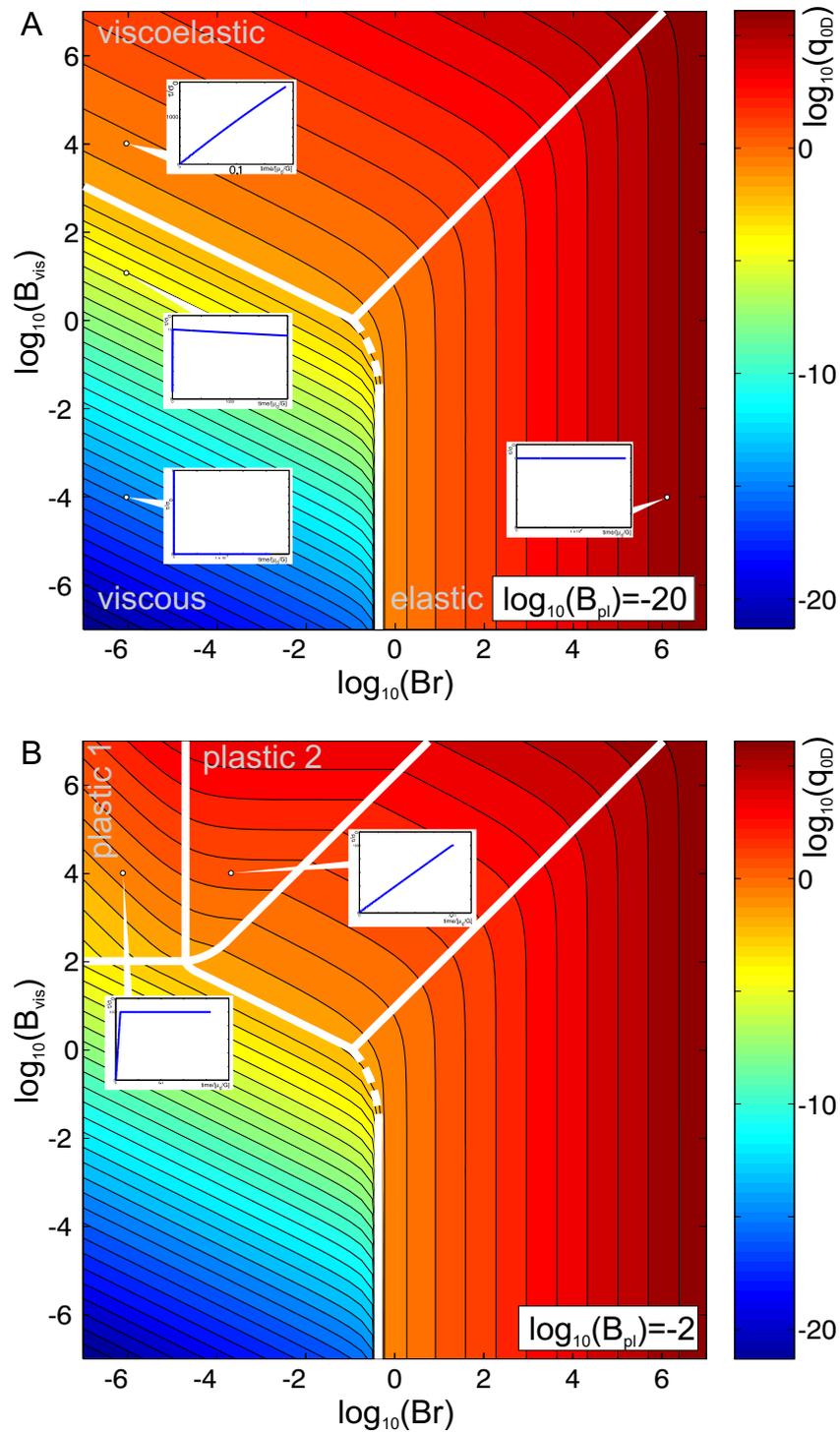


Figure 4.3: Numerically computed growthrate q_{0D} as a function of B_{vis} , Br and B_{pl} for a constant strainrate boundary condition and a 0-D simple-shear setup. Computations are performed for $\Delta T = 0.1$. Insets show stress evolution versus time at various locations. A) $B_{pl} = 10^{-20}$ (no plasticity), B) $B_{pl} = 10^{-2}$. Scaling relationships for q_{0D} in the various domains are summarized in table 4.1. Expressions for the boundaries between different deformation modes (white lines) are given in table 4.2.

Domain	Equation	$\Delta T = 0.01$	$\Delta T = 0.05$	$\Delta T = 0.1$
Viscous	$q_{0D} = f_V Br B_{vis}^2$	$f_V = 0.50$	$f_V = 0.49$	$f_V = 0.48$
Elastic	$q_{0D} = f_E Br$	$f_E = 0.50$	$f_E = 0.51$	$f_E = 0.53$
Viscoelastic	$q_{0D} = f_{VE} Br^{\frac{1}{3}} B_{vis}^{\frac{2}{3}}$	$f_{VE} = 0.03$	$f_{VE} = 0.08$	$f_{VE} = 0.12$
Plastic 1	$q_{0D} = f_{P1} Br B_{vis} B_{pl}^{-1}$	$f_{P1} = 0.50$	$f_{P1} = 0.50$	$f_{P1} = 0.50$
Plastic 2	$q_{0D} = f_{P2} B_{vis} B_{pl}$	$f_{P2} = 0.01$	$f_{P2} = 0.05$	$f_{P2} = 0.10$

Table 4.1: Scaling laws for growthrate q_{0D} in the 0-D setup with constant strainrate boundary conditions for different values of ΔT . Growthrates with a constant stress boundary condition are identical to the expressions in the elastic domain.

If the initial stress equals the viscous stress, the growth rate is approximately equal to the growth rate in the case of a constant stress boundary condition. If $B_{vis} < 1$, the viscous steady-state stress is smaller than the initial stress and stress relaxation occurs.

2. *Viscoelastic mode.* Stress increases almost linearly with time. During the initial stage, little thermally-induced stress reduction is observed in this mode. The growthrate ($q \propto Br^{\frac{1}{3}} B_{vis}^{\frac{2}{3}}$) depends on the amount of elasticity, viscosity and strainrate, but not on the initial stress.
3. *Elastic mode.* Stress remains close to its initial value during the onset of shear-localization. This mode only occurs for large Br -numbers, and thus requires large initial stresses. The initial growthrates of this mode are almost identical to the growthrates under constant stress boundary condition.
4. *Plastic 1 mode.* Stress increases linearly from the initial stress to the yield stress ($\sigma_y = B_{pl}^{-1}$). Yielding occurs after only a small amount of bulk heating. Once the material yields, elastic strainrates are zero (since $\dot{\epsilon}_{el} = \frac{d\tau}{dt}$) and all further deformation occurs by dissipative elastic and viscous processes. The amount of shear-heating thus sharply rises once yielding occurs (see also fig. 4.2b).
5. *Plastic 2 mode.* The overall behavior is similar to the plastic 1 mode, with the difference that yielding now occurs at the end of the numerical simulation, when the overall temperature is almost ΔT .

The growthrates in most domains (with exception of the viscoelastic domain) are relatively insensitive to the total temperature increase ΔT used to determine q_{0D} (table 4.1), meaning that eq. 4.26 is appropriate. In the viscoelastic domain, the temperature evolution versus time is exponential rather than linear, and hence the true growthrate is $< q_{0D}$ in the beginning of the simulation and $> q_{0D}$ at the end.

In the case of a constant stress boundary condition, only a single domain exists with growthrate $q_{0D} \propto f_{cs}Br$. Interestingly, this expression is identical to the elastic case with a constant strainrate boundary condition ($f_{cs} = f_E$, see table 4.1). This can be understood by realizing the the elastic case has an almost constant stress evolution during the initial stages (inset in fig. 4.3a).

The boundaries between different modes of deformation are well defined and relatively sharp. Most boundaries are continuous and in their location can be computed analytically by requesting that the growthrates of two adjacent domains are equal at the phase boundary (see table 4.2). A discontinuous boundary occurs between the viscous and the elastic domain, which is accompanied with a sharp increase in growthrate.

All growthrates determined in this section are positive. Any deformation will thus result in a temperature increase. In order to understand whether this will result in shear-localization, a 1-D model is required.

4.4.2 1-D model

In the 1-D setup, a small perturbation either in temperature or in viscosity is introduced at the onset of the simulations (see Fig. 4.1a). If the matrix viscosity is μ_0 , its temperature will increase as $T_{bg}(t) = q_{0D}t \frac{G}{\mu_0}$. The temperature of a weak zone with

Boundary	Equation	$\Delta T = 0.01$	$\Delta T = 0.05$	$\Delta T = 0.1$
Viscous- Elastic	$Br = f_{ve}$	$f_{ve} = 0.04$	$f_{ve} = 0.22$	$f_{ve} = 0.43$
Viscous- Viscoelastic	$B_{vis} = f_{vve}Br^{-\frac{1}{2}}$	$f_{vve} = 0.12$	$f_{vve} = 0.26$	$f_{vve} = 0.35$
Elastic- Viscoelastic	$B_{vis} = f_{eve}Br$	$f_{eve} = 68.0$	$f_{eve} = 16.1$	$f_{eve} = 9.28$
Viscous- Plastic 1	$B_{vis} = f_{vp1}B_{pl}^{-1}$	$f_{vp1} = 1.0$	$f_{vp1} = 1.0$	$f_{vp1} = 1.0$
Plastic 1-Plastic 2	$Br = f_{p1p2}B_{pl}^2$	$f_{p1p2} = 0.02$	$f_{p1p2} = 0.1$	$f_{p1p2} = 0.20$
Plastic 2-Viscoelastic	$B_{vis} = f_{p2ve}BrB_{pl}^{-3}$	$f_{p2ve} = 27$	$f_{p2ve} = 4.1$	$f_{p2ve} = 1.73$

Table 4.2: Expression for the phase boundaries in fig. 4.3.

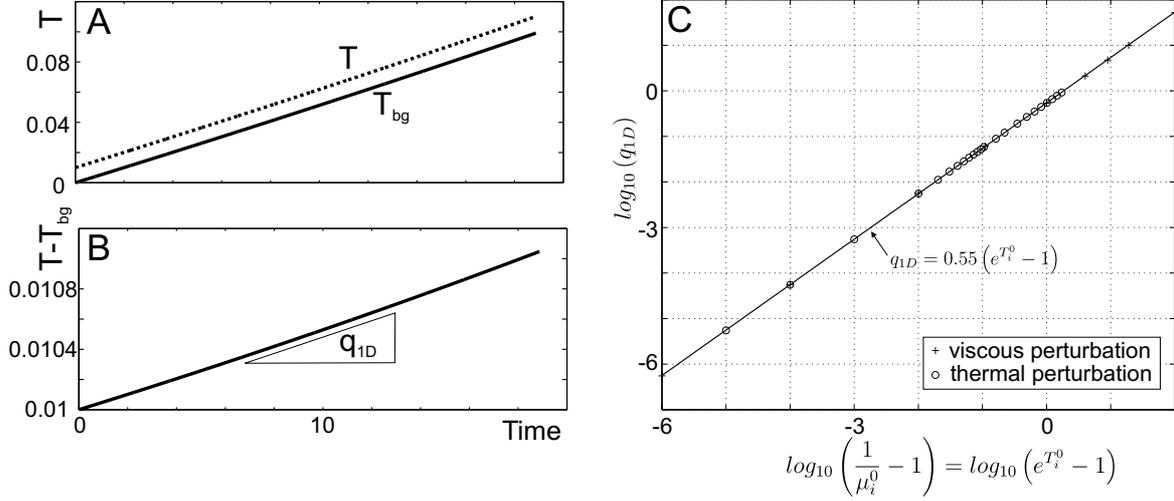


Figure 4.4: A) Temperature evolution versus time for a 1-D simulation with a constant stress boundary condition and $Pe = 1000$, $Br = 10^{-2}$, $L/R = 100$. B) T (maximum temperature) minus T_{bg} (background temperature) of the same simulation. The slope of the curve is q_{1D} in equation 4.29. C) Numerically determined growth rates versus initial perturbation temperature or viscosity for a constant stress boundary condition with $Br = 1$, $Pe = 1000$ and $\Delta T = 0.1$.

viscosity μ_i^0 will be fairly similar, but with a different timescale: $T(t) = q_{0D}t \frac{G}{\mu_i^0}$. The difference in temperature between the two domains is thus (see also Fig. 4.4):

$$T(t) - T_{bg}(t) = \left(\frac{\mu_0}{\mu_i^0} - 1 \right) \frac{G}{\mu_0} q_{0D}t = \left(\frac{1}{\mu_i^0} - 1 \right) q_{1D}t \quad (4.27)$$

where μ_i^0 is the initial viscosity of the shearzone ($\mu_i^0 < \mu_0$), $T_{bg}(t)$ is the temperature evolution of the matrix far away from the shear-zone (which can essentially be described with a 0-D model; see eq. 4.26), q_{1D} is the 1-D growth rate, and the viscosity of the matrix was chosen to be the scaling viscosity. In the case of a thermal perturbation, the temperature evolution can thus be expressed as:

$$T(t) - T_{bg}(t) = T_i^0 + \left(e^{T_i^0} - 1 \right) q_{1D}t \quad (4.28)$$

which yields for a small initial thermal perturbation ($T_i^0 \ll 1$):

$$T(t) - T_{bg}(t) = T_i^0 (1 + q_{1D}t) \quad (4.29)$$

Note that if $q_{1D} > 0$, the heating rate in the center of the domain is larger than the background heating rate. Temperature in the center will thus increase faster, which results in a low viscosity zone and shear-localization. If, on the other hand, $q_{1D} < 0$, the initial temperature perturbation diffuses away, until the whole domain has approximately the same temperature and heats with the same rate.

The growthrate q_{1D} has been computed numerically by integrating the governing equations 4.20 and 4.21 until $T(t) > \Delta T + T_i^0$, with $\Delta T = 0.1$ (reflecting an increase in temperature of 1 – 10 K). In order to obtain insight on the influence of the nondimensional parameters on the initiation of shear-localization, more than 100'000 numerical simulations have been performed. Convergence tests showed that the spatial and time resolution employed was sufficient.

In addition to the non-dimensional parameters B_{pl} , B_{vis} and Br , two new parameters occur in the 1-D setup namely $Pe = R/L^*$ and R/L . Pe describes the ratio between the size of the initial heterogeneity R and the diffusion lengthscale (Eq. 4.9). L/R is the ratio between the size of the domain and the size of the heterogeneity. If $L > R$ is sufficiently large, this ratio does not influence the dynamics of the initial stages. Otherwise, the boundaries will influence the dynamics at the center of the deforming domain and the growthrate is reduced. The critical L/R is dependent on L^* and was found to be around 2 to 3 for the setup considered here. This is rather small compared to typical shear-zones, and therefore we assume this effect to be negligible.

Constant stress boundary condition

In the case of a constant stress boundary condition, $B_{vis} = 1$ and B_{pl} plays no role since $\sigma_0 \leq \sigma_y$. Therefore, the only non-dimensional parameters that may influence the problem are Br and Pe . An expression of q_{1D} , for sufficiently large Pe -numbers, is given by:

$$q_{1D} = 0.55Br \quad (4.30)$$

The dependence of q_{1D} on Pe is shown on figure 4.5A. Diffusion starts to reduce the growthrate if $Pe < 0.12$ and inhibits localization (q_{1D} becomes negative) if $Pe < 0.06$. The boundary between localization and no-localization (i.e. diffusion dominates) is given by (see also fig. 4.5B):

$$Pe = 10^c Br^{-0.5} \quad (4.31)$$

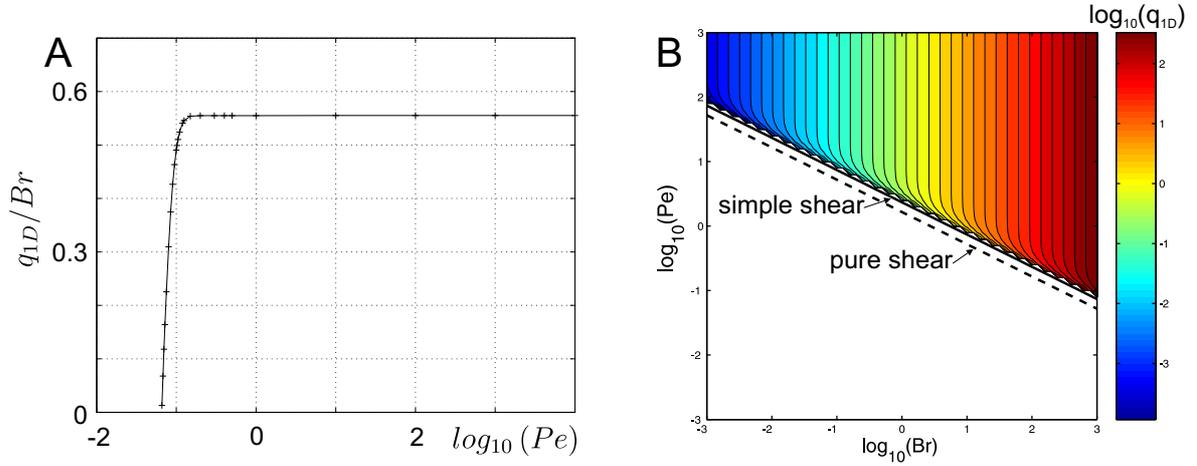


Figure 4.5: A) Normalized growthrate versus Pe for a constant stress boundary condition with $Br = 1000$, $\Delta T = 0.1$ and $T_i^0 = 0.01$. Diffusion inhibits localization if $Pe < 0.06$. B) Numerically computed growthrate q_{1D} versus Pe and Br with $\Delta T = 0.1$. Diffusion dominates in the white area. Expressions of the black lines are given in the text.

where $c = 2.3$ under simple-shear and $c = 1.7$ under pure-shear boundary conditions. It will be shown later that this expression can be derived analytically.

Constant velocity boundary condition

The 1-D growthrate in simulations with a constant velocity boundary condition are shown for different Pe , B_{pl} , B_{vis} and Br -numbers in figure 4.6. The 0-D and the 1-D growthrates are similar (see also Table 4.3). Differences, however, are related to the effect of diffusion, which may inhibit localization if it is faster than the rate of heat production.

Diffusion is proportional to Pe^{-2} . The equations for the phase boundaries that involve diffusion demonstrate that this boundary can be predicted under the assumption that the diffusion growthrate is given by $q_{diff} \approx \frac{2}{Pe^2}$ (the factor can be predicted analytically as will be shown in the next section). The requirement $q_{diff} = q_{1D,vis}$ then predicts the phase boundary.

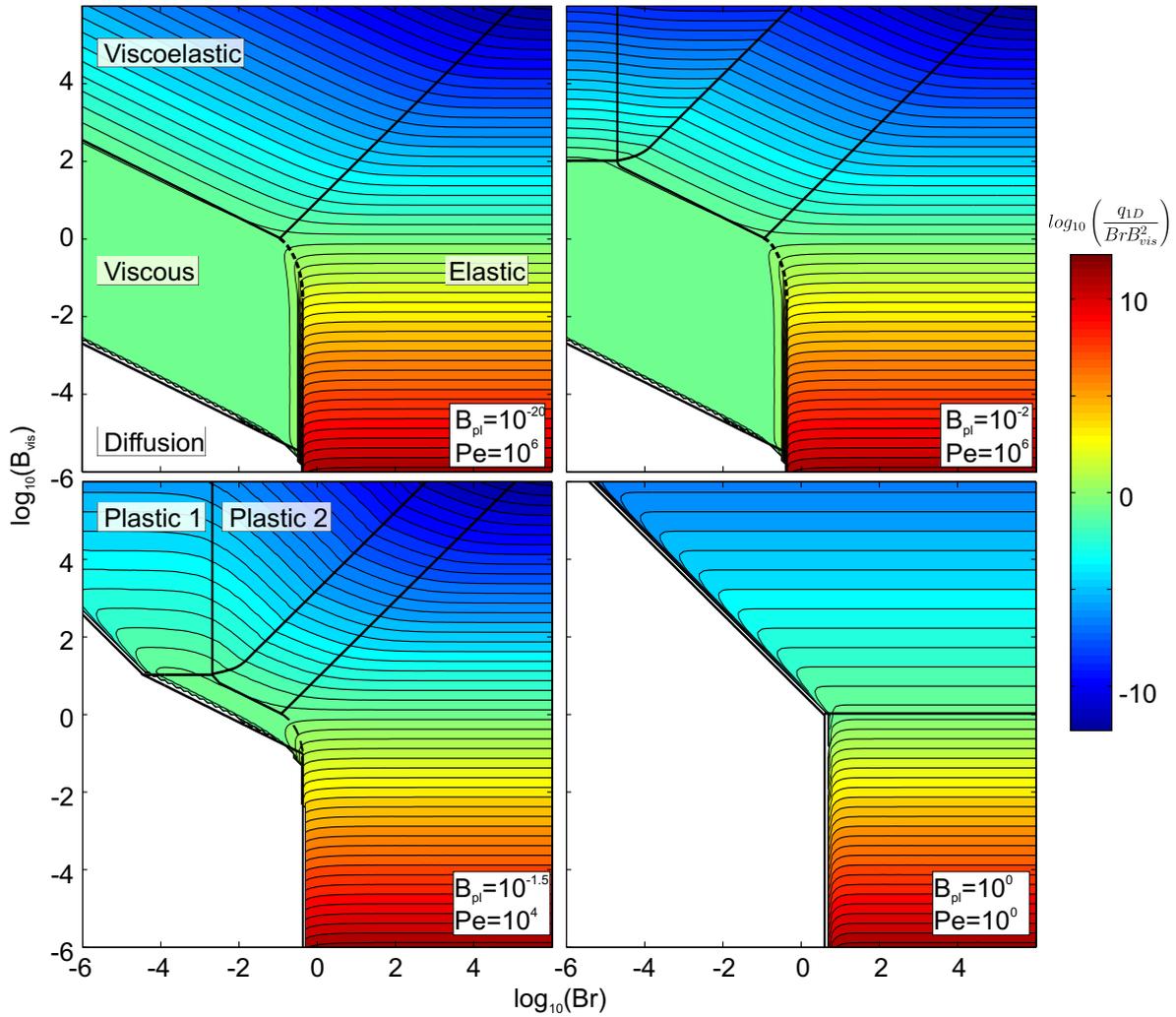


Figure 4.6: Growthrate versus Br and B_{vis} for various B_{pl} and Pe numbers in a 1-D constant velocity model ($\Delta T = 0.1, T_i^0 = 0.01$). Scaling laws for q_{1D} are indicated in table 4.3 for simulations with $Pe > 1$. The white areas are areas with negative growthrates (diffusion-dominated). The lower-right plot is a case where diffusion also influences the quasi-elastic domain (see text for discussion).

Localization and diffusion

Analytically, the effect of diffusion on inhibiting shear-localization can be understood by noting that the time t_{rise} , required to increase the temperature of the inclusion from T_i^0 to $T_i^0 + \Delta T$, in the absence of diffusion, is dependent on the background growthrate of temperature as:

$$t_{rise} = \frac{\Delta T}{q_{0D} + T_i^0 q_{1D}} \approx \frac{\Delta T}{q_{0D}} \quad (4.32)$$

where $T_i^0 \ll 1$. An analytical solution for the temperature evolution with time at the center of a 1D shearzone of width R , with initial temperature T_i^0 and with a constant heat production $Q = q_{1D} T_i^0$, is given by [e.g. Carslaw and Jaeger, 1959, Cardwell et al., 1978]:

$$\frac{T(t^n) - T_{bg}(t^n)}{T_i^0} = \left(q_{1D} P e^2 \left(t^n + \frac{1}{8} \right) + 1 \right) \operatorname{erf} \left(\frac{1}{4\sqrt{t^n}} \right) + \frac{q_{1D} P e^2 \sqrt{t^n}}{2\sqrt{\pi}} \exp \left(- \left(\frac{1}{4\sqrt{t^n}} \right)^2 \right) - \frac{q_{1D} P e^2}{8} \quad (4.33)$$

where $t^n = \frac{t}{P e^2}$ is the rescaled time. Diffusion will be more efficient than heating if at a given time t^n , the left-hand-side of 4.33 becomes < 1 . An analytical solution for the transition between diffusion-dominated and heating-dominated temperature evolution is thus:

$$q_{1D} P e^2 = \frac{1 - \operatorname{erf} \left(\frac{1}{4\sqrt{t^n}} \right)}{\left(t^n + \frac{1}{8} \right) \operatorname{erf} \left(\frac{1}{4\sqrt{t^n}} \right) + \frac{\sqrt{t^n}}{2\sqrt{\pi}} \exp \left(- \left(\frac{1}{4\sqrt{t^n}} \right)^2 \right) - \frac{1}{8}} \quad (4.34)$$

Domain	Equation	$\Delta T = 0.01$	$\Delta T = 0.05$	$\Delta T = 0.1$
Viscous	$q_{1D} = f_V B r B_{vis}^2$	$f_V = 0.50$	$f_V = 0.50$	$f_V = 0.50$
Elastic	$q_{1D} = f_E B r$	$f_E = 0.51$	$f_E = 0.53$	$f_E = 0.55$
Viscoelastic	$q_{1D} = f_{VE} B r^{\frac{1}{3}} B_{vis}^{\frac{2}{3}}$	$f_{VE} = 0.03$	$f_{VE} = 0.08$	$f_{VE} = 0.13$
Plastic 1	$q_{1D} = f_{P1} B r B_{vis} B_{pl}^{-1}$	$f_{P1} = 0.48$	$f_{P1} = 0.51$	$f_{P1} = 0.53$
Plastic 2	$q_{1D} = f_{P2} B_{vis} B_{pl}^{-1}$	$f_{P2} = 0.01$	$f_{P2} = 0.06$	$f_{P2} = 0.11$

Table 4.3: Scaling laws for growthrate q_{1D} in the 1-D setup with constant strainrate boundary conditions for different values of ΔT . Growthrates with a constant stress boundary condition are identical to the expressions in the elastic domain. Expressions have been derived for $Pe \gg 1$.

The solution is plotted in figure 4.7A. At small times, diffusion is not very efficient. At later stages (larger t^n) diffusion becomes more efficient, and may eliminate the initial increase in temperature in the center of the shearzone. If, however, $q_{1D} > \frac{3}{Pe^2}$ diffusion will not be able to inhibit shear-localization.

The analytical solution can be compared with the previously obtained numerical results (which have been integrated until $T(t) = T_i^0 + \Delta T$) by noting that $q_{1D} \approx q_{0D}$, and that $t_{rise} \simeq \frac{\Delta T}{q_{1D}}$. Substituting this expression into equation 4.34 gives an implicit equation for q_{1D} . The critical growthrate required to overcome the effect of diffusion is given by

$$q_{1D} = \frac{f(\Delta T)}{Pe^2} \quad (4.35)$$

where $f(\Delta T)$ is a constant that is solely dependent on ΔT (which can be obtained numerically from eq. 4.34; see fig. 4.7B). For small ΔT

$$\Delta T \simeq 10^{-2} q_{1D} Pe^2 \quad (4.36)$$

The prediction of eq. 4.33 was verified by 1-D numerical simulations, which have been integrated until $t = t_{rise}$. An excellent agreement exist between the numerically determined time evolution of the maximum temperature in the shearzone and eq. 4.33.

Equation 4.35 can be used to predict the boundary between the viscous deformation mode and the diffusion mode shown in figure 4.6. In this figure $\Delta T = 0.1$, giving $f(0.1) \approx 10^{0.3}$. The condition for inhibiting localization thus gives $0.55BrB_{vis}^2 = 10^{0.3}Pe^{-2}$ or $B_{vis} = 2Pe^{-1}Br^{-0.5}$, which is in excellent agreement with the numerical

Boundary	Equation	$\Delta T = 0.1$
Viscous- Elastic	$Br = f_{ve}$	$f_{ve} = 0.43$
Viscous- Viscoelastic	$B_{vis} = f_{vve}Br^{-\frac{1}{2}}$	$f_{vve} = 0.36$
Elastic- Viscoelastic	$B_{vis} = f_{eve}Br$	$f_{eve} = 8.7$
Viscous- Plastic 1	$B_{vis} = f_{vp1}B_{pl}^{-1}$	$f_{vp1} = 1.0$
Plastic 1-Plastic 2	$Br = f_{p1p2}B_{pl}^2$	$f_{p1p2} = 0.20$
Plastic 2-Viscoelastic	$B_{vis} = f_{p2ve}BrB_{pl}^{-3}$	$f_{p2ve} = 1.77$
Diffusion-Viscous	$B_{vis} = f_{dv}Pe^{-1}Br^{-0.5}$	$f_{dv} = 2$
Diffusion-Plastic 1	$B_{vis} = f_{dp1}B_{pl}Pe^{-2}Br^{-1}$	$f_{dp1} = 3.8$

Table 4.4: Expression for the phase boundaries in fig. 4.6.

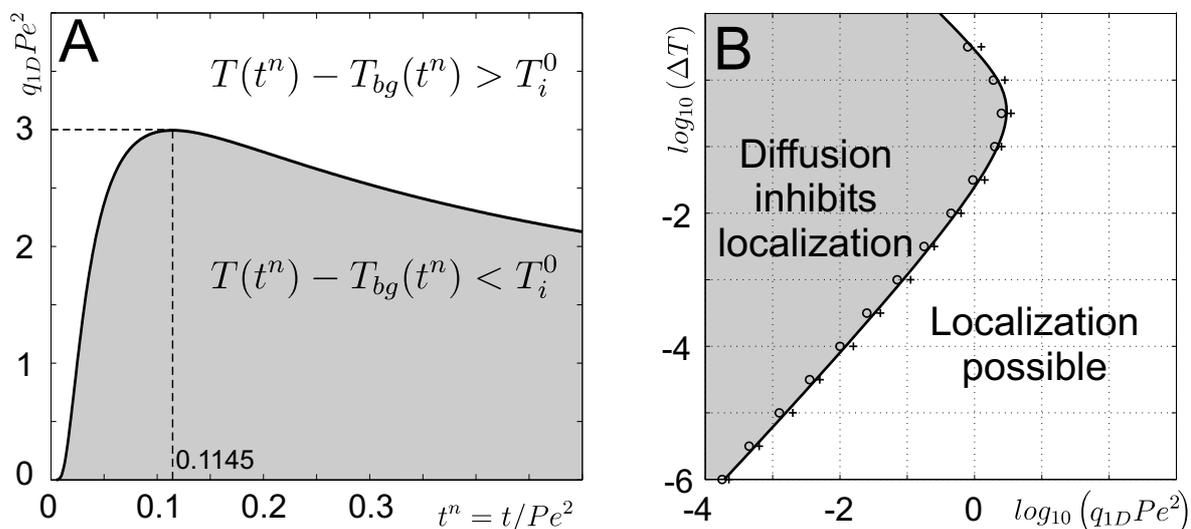


Figure 4.7: A) Growthrate versus t^n (eq. 4.34). Grey area is diffusion-dominated. B) Growthrate versus ΔT . The black line is computed analytically from equations 4.32 and 4.34. 1-D numerical simulations have been performed to verify the analytical solution and are indicated by crosses ($T(t_{rise}) > T_i^0$) and circles ($T(t_{rise}) < T_i^0$). Maximum occurs at $\Delta T \simeq 0.34$, $q_{1D}Pe^2 \simeq 3$.

simulations (see table 4.4). Similarly one can predict that diffusion will inhibit localization if $Pe < 0.06$, in the case of a constant stress boundary condition and with the parameters of fig. 4.5.

4.4.3 Link between current and classical analysis of shear-localization

At this stage, it is interesting to compare the results obtained sofar with the more classical analysis of shear-localization, such as the one presented in Turcotte and Schubert [1982]. Their analysis is done for steady-state conditions, constant stress boundary conditions, and assumes that a layer of constant thickness H is deformed under simple shear. The lower boundary is thermally insulating and the upper boundary is of constant temperature. Rheology is viscous. They demonstrated that the maximum temperature in the sheared layer is a function of the Brinkman number Br^* . Turcotte and Schubert [1982] derive Br^* for an Arrhenius-type rheology. One can show that in

the case of a Frank-Kamenetzky rheology (eq. 4.7), Br^* is defined as:

$$Br^* = \frac{\sigma_0^2 H^2 \gamma}{k \mu_0} \quad (4.37)$$

where $k = \kappa \rho c_p$ denotes thermal conductivity. The maximum dimensionless temperature ΔT in the sheared layer depends on Br^* as [Turcotte and Schubert, 1982]:

$$Br^* = \frac{2 \left(\frac{Br \ln(\Delta T)}{2} \right)}{\left(\cosh \left(\frac{Br \ln(\Delta T)}{2} \right)^{0.5} \right)^2} \quad (4.38)$$

For small temperature increase (subcritical branch), the following relationship holds:

$$\Delta T \simeq 0.5 Br^* \quad (4.39)$$

There are two important differences between the 'classical' analysis outline above and the one studied in this chapter. They are:

1. The classical analysis assumes a constant thickness shearzone, whereas the work presented here only prescribes the thickness of the initial thermal perturbation. Due to the effects of shearheating and diffusion, the shearzone thickness can evolve with time.

2. The classical analysis assumes steady-state conditions ($\frac{\partial T}{\partial t} = 0$), whereas in the present work $\frac{\partial T}{\partial t} > 0$.

In the following we show that it is possible use the results of the 'classical' analysis, in order to predict the expression for growthrate q_{0D} . If the shearzone is thin, and of constant temperature ΔT , the energy equation can be simplified to:

$$\frac{\partial \Delta T}{\partial t} = -\kappa \frac{\Delta T}{H^2} + q_{0D} \quad (4.40)$$

At steady-state conditions, $\frac{\partial \Delta T}{\partial t} = 0$. With the use of eqns. (4.39) and (4.38), one obtains

$$\begin{aligned} q_{0D} &= 0.5 \kappa \frac{Br^*}{H^2} \\ q_{0D} &= 0.5 \frac{\sigma_0^2 \gamma}{\rho c_p \mu_0} = 0.5 Br \end{aligned} \quad (4.41)$$

This result is identical to the result found previously by numerical analysis (table 4.1).

4.4.4 2-D model

The results presented in the last section were valid for the initiation stage in a 1-D simple shear setup, or in other words for deformation along an existing weak-zone. The purpose of the current section is to understand how such weak-zones can be created in 2-D pure-shear extensional setting (see fig. 4.1 for the setup). Shear localization is initiated by increasing the initial temperature to T_i^0 , or decreasing the viscosity to μ_i^0 in a circular heterogeneity with radius R . Two cases will be considered: (1) the background strain rate is maintained constant and (2) the far-field stress is maintained at a constant level.

Three examples of shear-localization under constant strainrate boundary conditions are shown in figure 4.8. In these examples, plasticity is not activated. As in 0-D and 1-D, we can distinguish the following deformation modes:

- The viscous simulation reaches the viscous steady-state stress ($\sigma^{2nd} = B_{vis}$) very soon after the onset of extension. A 'flower-like' shear-heating distribution in and around the inclusion is obtained at this stage, which is in agreement with recently obtained 2-D analytical results of Schmid and Podladchikov [2003] and Schmid [2002]. Shear-heating is largest in the lobes around the inclusion.
- The viscoelastic simulation shows a distinctly different behavior. In this case shear-heating mainly occurs in the center of the heterogeneity. This can be understood by recalling that shear-heating is proportional to $\tau_{ij} (\dot{\epsilon}_{ij} - \dot{\epsilon}_{ij}^{el})$. Increasing temperature reduces the viscosity and makes the material more 'viscous' (since it lowers the Deborah-number; $De = \frac{2\mu\dot{\epsilon}}{G}$). The more viscous a material, the lower the elastic strainrate and thus the more heat is dissipated.
- The elastic simulation behaves like the visco-elastic simulation with the difference that stress, and therefore shear-heating, stays at an almost constant level. This is in agreement with the 0-D and 1-D models.

The viscoelastic mode ultimately switches into a viscous mode (heating around instead of inside the inclusion) but only when the viscous steady-state stress has been reached (after 2-3 Maxwell times). As will be shown later, this does not occur for realistic parameters, since the stresses are too large. Once this switch occurs, however, linear low viscosity shearzones are initiated from the inclusion.

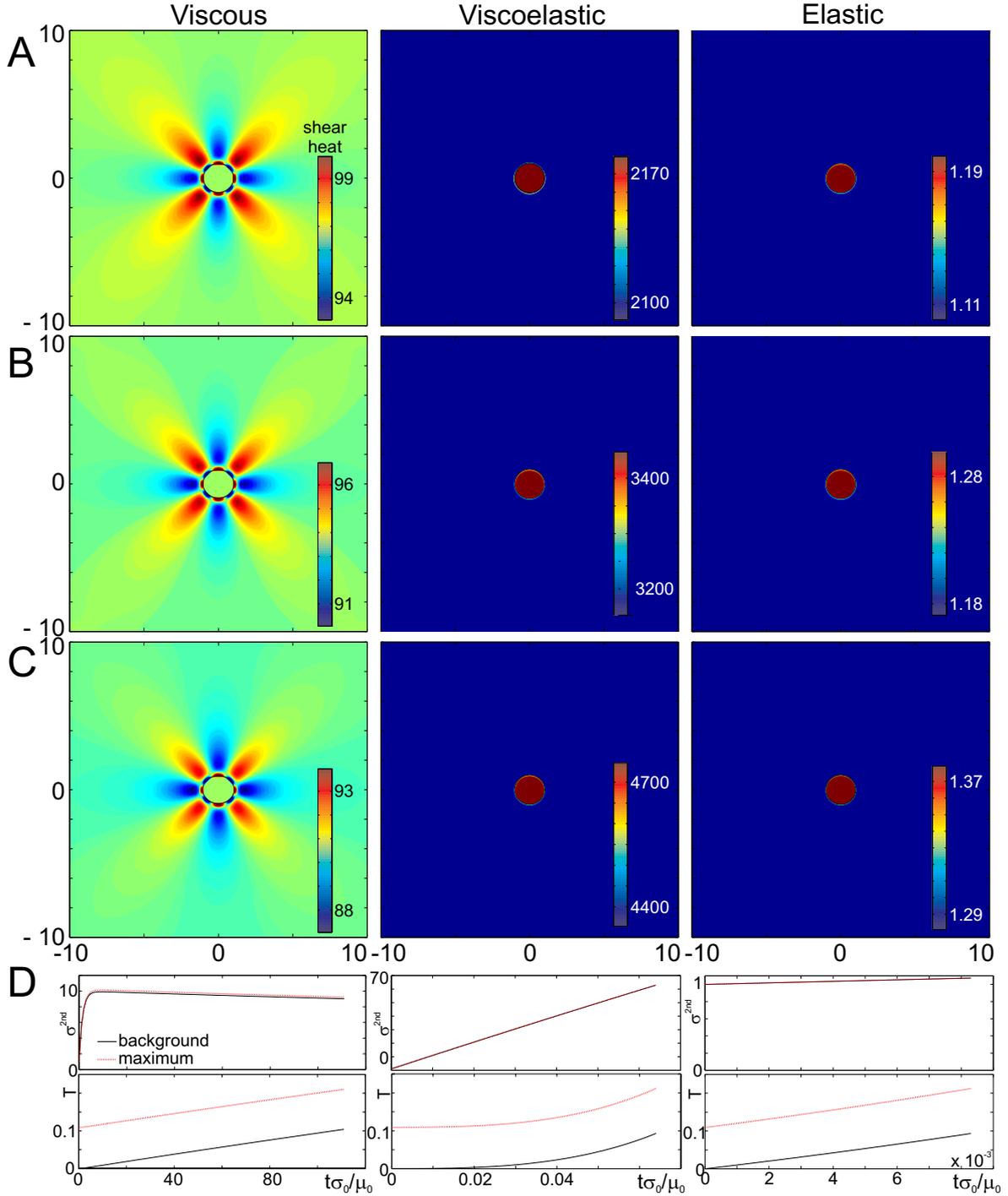


Figure 4.8: Examples of the initiation stage in a 2D pure-shear setup (fig. 4.1) for 3 different, non-plastic, deformation modes. Colors indicate the amount of shear-heating. A) evolution at the time when $T_{max} - T_i^0 \simeq 0.033$, B) $T_{max} - T_i^0 \simeq 0.067$, C) $T_{max} - T_i^0 \simeq 0.1$. D) Time-dependent parameters (2^{nd} invariant of the stress tensor and temperature). $B_{pl} = 10^{-20}$, $Pe = 10^3$ in all simulations. Other parameters are: viscous: $Br = 10^{-5}$, $B_{vis} = 10$; viscoelastic: $Br = 10^{-3}$, $B_{vis} = 10^3$; elastic: $Br = 10^1$, $B_{vis} = 10^1$.

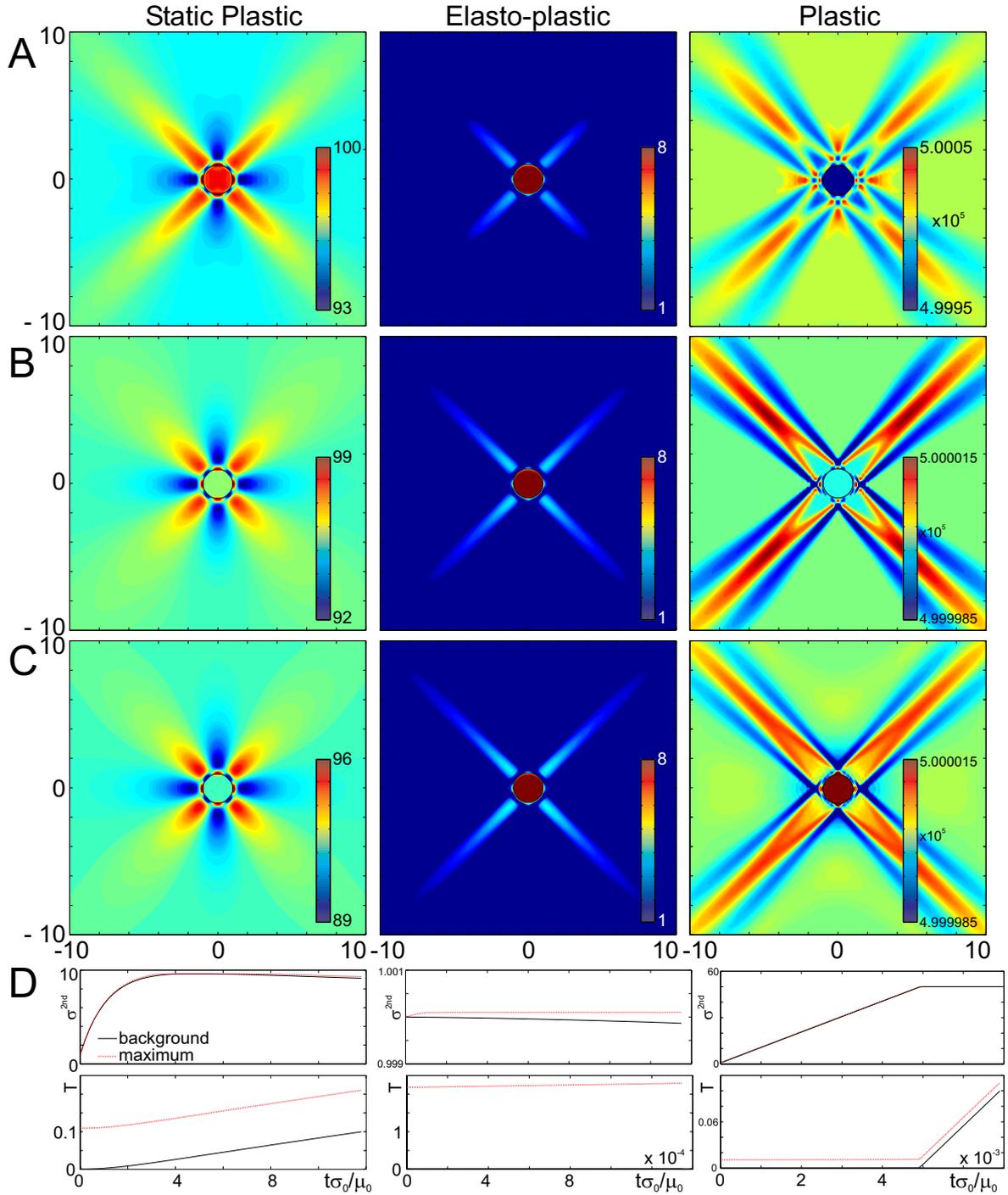


Figure 4.9: Snapshots of shear-heating at three different times (A-C) for the different plastic modes that occur in the 2-D model. $Pe = 10^3$ in all simulations. Other parameters are: static-plastic: $Br = 10^{-4}$, $B_{vis} = 10$, $B_{pl} = \frac{1}{9.6}$; elasto-plastic: $Br = 10^1$, $B_{vis} = 1$, $B_{pl} = \frac{1}{1.0001}$; plastic: $Br = 10^{-4}$, $B_{vis} = 10^4$, $B_{pl} = \frac{1}{50}$. The static-plastic mode yields only in part of the domain. With ongoing deformation, viscosity decreases and the stress drops below the yield stress (switches to the viscous mode). The elasto-plastic mode also yields in part of the domain only but otherwise behaves elastically. The plastic mode results in whole domain yielding.

If plasticity is considered in the simulations, the pattern of deformation changes. We have observed three different types of plastic deformation modes (fig. 4.9):

- The plastic mode corresponds to the plastic-1 and plastic-2 modes in the 1-D setup. During the pre-yielding stages of the simulation, shear-heating is maximum in the inclusion (as in the viscoelastic or elastic modes). At yielding, brittle shear-zones initiate at an angle of 45° from the inclusion and 'cut' the model domain. Initiation occurs in the four lobes around the inclusion, since the stress at these points is slightly larger than the background value. With continuous deformation, the rest of the model domain yields. At this stage, large, linear zones with a slightly reduced viscosity have been created. The setup is now similar to the 1-D setup with a pure-shear rather than a simple-shear background deformation. Continuing deformation may result in weak shearzones.
- The static-plastic mode occurs when the yield stress is slightly larger than viscous steady-state stress of the matrix close to the viscous-plastic boundary. In this deformation mode, yielding occurs locally and for a finite amount of time only. The maximum stress around a weak inclusion is larger than the stress away from the inclusion. The amount of stress-enhancement depends on the aspect-ratio of the inclusion and on the viscosity contrast between inclusion and matrix. Circular inclusions have a maximum stress-increase of $\sim 30\%$ [see e.g. Schmid, 2002], whereas infinitely long elliptical inclusions may have an infinite stress-increase at their tips. In the static-plastic mode, yielding occurs for a finite-amount of time only since the overall temperature increase results in a viscosity reduction, which will ultimately drop the stress below the yield level. After this, deformation proceeds as in the viscous mode. No long-living stable shearzone is thus expected in this domain.
- The elasto-plastic mode occurs when the yield stress is very close to the initial stress. Whereas shear-heating in the inclusion is largest, zones of local yielding exist and propagate outwards. Ongoing deformation results in thermal runaway.

Simulations performed with a constant stress boundary condition are similar to the elastic mode under constant strainrate conditions.

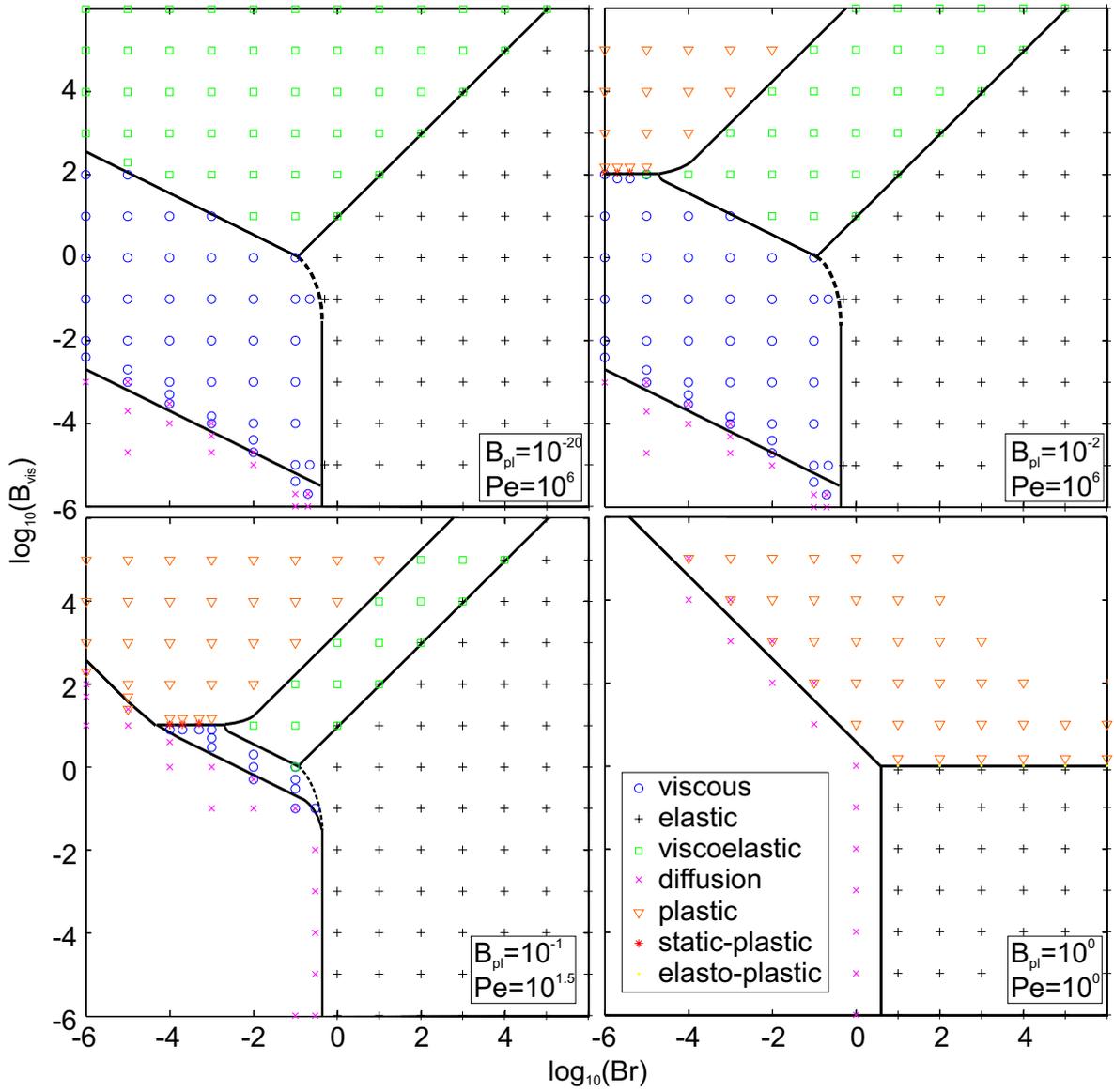


Figure 4.10: 2D simulations of shear-localization around circular inclusions. $T_i^0 = 0.01, \Delta T = 0.1$. Black lines are phase boundaries derived with 1-D simulations (fig. 4.6).

Domain	Growthrate in center	Growthrate outside
Viscous	$q_{2D}^{cen} \approx 0$	$q_{2D}^{out} = 0.25B_{vis}^2Br$
Elastic	$q_{2D}^{cen} = 1.1Br$	$q_{2D}^{out} \approx 0$
Viscoelastic	$q_{2D}^{cen} = 0.15Br^{\frac{1}{3}}B_{vis}^{\frac{2}{3}}$	$q_{2D}^{out} \approx 0$
Plastic	$q_{2D}^{cen} = 0.05BrB_{vis}B_{pl}^{-1}$	$q_{2D}^{out} \approx 0$

Table 4.5: Scaling laws for growthrate in the 2-D setup with constant strainrate boundary conditions. Growthrates have been determined both in the center of the inclusion and outside the inclusion (stars in fig. 4.8) for $Pe \gg 1$ and $\Delta T = 0.1$.

In order to understand whether the different deformation modes described above form for the same range of parameters as the corresponding 0-D and 1-D modes, we have performed systematic 2-D simulations. The results (fig. 4.10) indicate that this is the case. Diffusion in 2-D is slightly more efficient than in 1-D but differences are small and the diffusion boundaries derived in the previous section can be applied to 2-D cases. It should be noted that the timescale in our model is given by the Maxwell relaxation time ($= \mu_0/G$). The total amount of *strain*, required to raise the temperature to ΔT , is dependent on the relaxation time. The inclusion remains circular during deformation only for sufficiently small relaxation times, or sufficiently large growthrates. In other cases, the inclusion becomes elliptical and Pe decreases, enhancing the effect of diffusion.

In 2-D, shear-heating is maximum in either the center of the inclusion (e.g. fig. 4.8A), or in one of the lobes around the inclusion (e.g. fig. 4.8B,C). To further compare the 2-D with the 1-D results, growthrates have determined both inside and outside the inclusion (table 4.5). In general, either the inside or the outside of the inclusion heats with similar rates as in the 1-D case. The viscous cases produce more heat outside, and the viscoelastic and elastic cases more heat inside the inclusion (in agreement with fig. 4.8).

The dependence of the growthrate on inclusion temperature or viscosity is similar to the 1-D case for small viscosity contrasts (eq. 4.27). For viscosity contrasts larger than 10-100, the growthrate in the viscous domain saturates, whereas the viscoelastic undergoes a switch from heating inside to heating outside the inclusion.

The 2-D simulations thus point out the importance of plasticity in creating linear zones of enhanced shear-heating. Plasticity has two effects: Firstly, it limits the stress

level to reasonable values. Secondly, the onset of plastic yielding is accompanied by the creation of linear zones in which shear heating is slightly larger than the background level. Whereas this does not immediately result in large-scale shear localization since the absolute increases in temperature are small (generally $< 1\text{K}$), it does help to localize deformation at later stages in the deformation. Once these linear zones have been created, ongoing deformation and localization is essentially a one-dimensional process.

4.5 Possible implications for shear-localization on planets

We have demonstrated that shear-localization may occur by different deformation modes. In the current section we provide insight into the possible relevance of these modes for earth-like conditions.

For a given inclusion size, initial stress and yield stress, a phase diagram of deformation modes versus effective viscosity μ_0 and background strainrate $\dot{\epsilon}_{BG}$ is shown in figure 4.11. For the creation of this diagram, we assumed small amounts of heating and small initial thermal perturbations. The results demonstrate that shear-localization is possible for earth-like parameters. However, the transition between localization and no-localization is critically dependent on the background strainrate, the effective viscosity and the size of the initial heterogeneity. Small background strainrates and/or small effective viscosities require large pre-existing heterogeneities for shear-localization to initiate. In order to verify whether the scaling laws, derived for the initial stages, also describe the behavior during later stages, we have performed additional late-stage 1-D numerical simulations integrated until $\Delta T \geq 10$ (corresponding to 200-1000 K). The results indicate that the diffusion-viscous and the viscous-viscoelastic boundaries are indeed correctly predicted from the initial stage analysis. The plastic-diffusion boundary, however, does not seem to occur during later stages. An analysis of the numerical simulations showed that significant stress drop occurs during late stages of simulations within the plastic-viscous mode. Thus, this mode becomes very similar to the viscous mode, which explains why the transition to no-localization is described by the diffusion-viscous boundary. The boundary between diffusion and localization is independent of the initial stress and the elastic shear-modulus of the lithosphere, but

is strong function of the heterogeneity-size (fig. 4.11):

$$\dot{\epsilon}_{BG} = \frac{1.4}{R} \sqrt{\frac{\kappa \rho C_p}{\mu_0 \gamma}} \quad (4.42)$$

The initial stress and the elastic shear-modulus, however, influence the Br -number (indicating the efficiency of shear-heating), which in turn influences the growthrate and is thus proportional to the amount of strain required to raise the temperature by a certain value. Large initial stresses help to form shear-zones.

A new effect occurs during late stages, namely the transition from aseismic creep to "seismic" creep, the latter being characterized by a rapid increase in temperature and strainrate in the center of the shear-zone (similar to the effect described in figure 4.2D). For the parameters of figure 4.11, both the viscoelastic and the elastic modes are not present. The elastic mode is absent since the initial stress of $\sigma_0 = 10$ MPa yields $Br \simeq 3 \times 10^{-5}$, which is too small for the elastic mode. The viscoelastic mode is absent since B_{pl} is sufficiently large ($B_{pl} \simeq 0.01$).

Two-dimensional simulations, performed for lithospheric parameters, generally confirm the 1-D results (fig. 4.12). For a given strainrate of $3 \times 10^{-15} \text{ s}^{-1}$, no localization occurs for lithosphere viscosities which are around an order of magnitude lower than the critical value of $O(10^{23})$ Pas determined from the 1D scaling laws. The larger the viscosity, the stronger the localization process. It is important to note that the main reduction in viscosity (and therefore the main localization process) does *not* occur when the lithosphere first yields, but later. The simulation of fig. 4.12C, for example, starts yielding after ~ 1 % extension. The main localization, however, occurs between 6-7 % extension. Before this stage the lithosphere is extended in a pure-shear mode. Afterwards strain localizes in thin zones (which have been formed during initial yielding of the lithosphere; see also fig. 4.9). The larger the initial viscosity of the lithosphere, the more elastic the lithosphere (the Deborah-number = $2\mu_0\dot{\epsilon}_0/G$ is larger) and the more vigorously the localization stage occurs. Simulations, performed for viscosities much larger than those shown in figure 4.12, are numerically difficult due to the very high strain rates that develop together with a fast drop in stress and large increase in temperature. This phenomenon is physically most likely related to the development of earthquakes and was also observed by other authors [e.g. Regenauer-Lieb et al., 1999, Regenauer-Lieb and Yuen, 2000a, 2004]. It is important to note that plasticity helps creating lithospheric-scale shear-zones (fig. 4.12A and D). For realistic strain

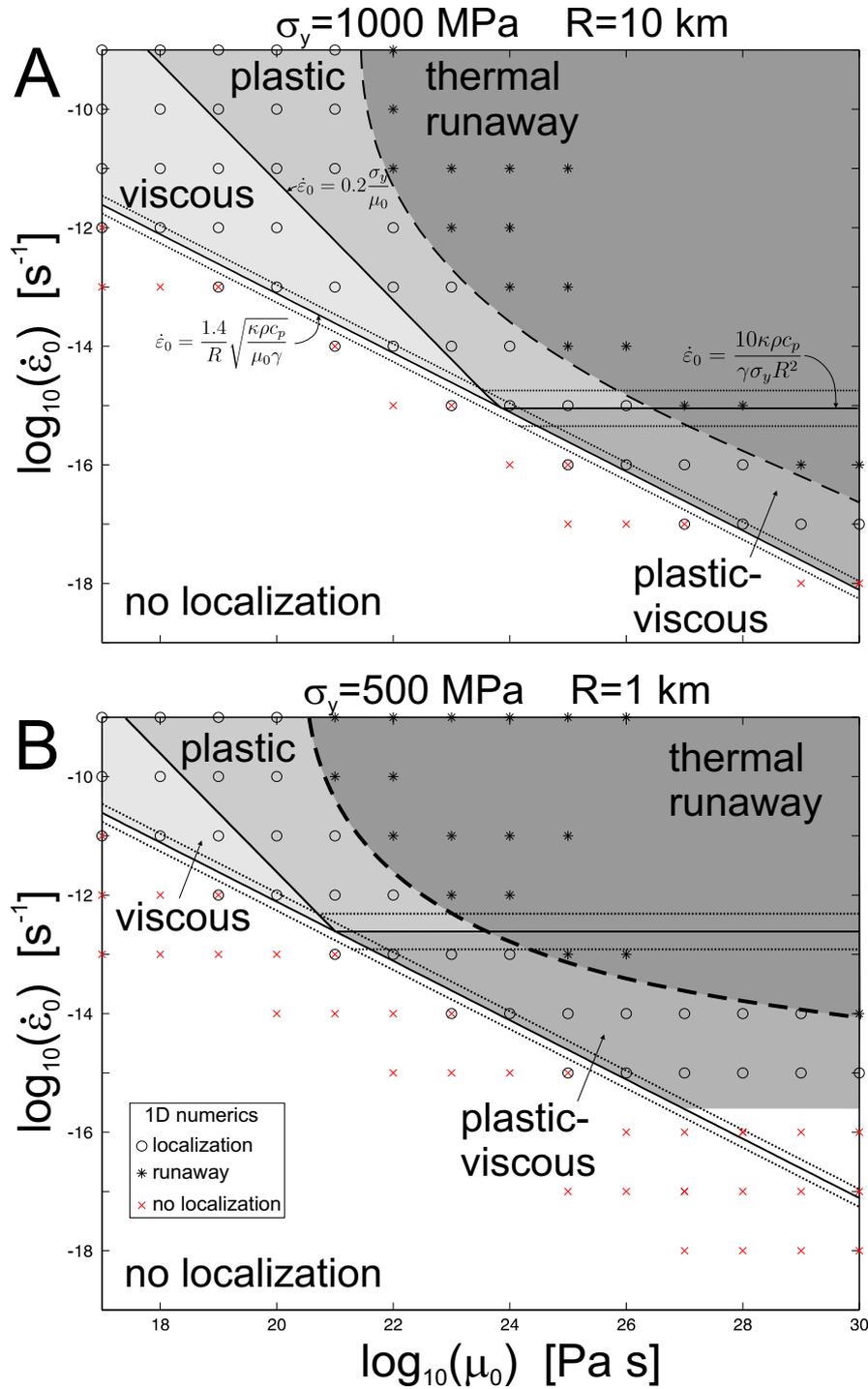


Figure 4.11: Deformation modes as a function of strainrate and effective viscosity for an inclusion size of A) $R = 10 \text{ km}$ and B) $R = 1 \text{ km}$ (with $\gamma = 0.1 \text{ K}^{-1}$, $\sigma_0 = 10 \text{ MPa}$, $\rho c_p = 3 \times 10^6 \text{ Jm}^2\text{Ks}^{-1}\text{kg}^{-1}$, $G = 10^{11} \text{ Pa}$). Boundaries between modes are derived from the 1-D initial stage scaling laws for a simple-shear, constant velocity boundary condition (dotted lines are boundaries with a twice larger or smaller growthrate, and thus take e.g. the effect of pure-shear into account) and have been verified by numerical simulations. One-dimensional late-stage numerical simulations (integrated until $\Delta T \geq 10$ and with $R/L = 0.01$, $T_i^0 = 0.01$) are indicated by symbols. The condition for thermal runaway is $\max(\dot{\epsilon}) \geq 10^8 \dot{\epsilon}_0$.

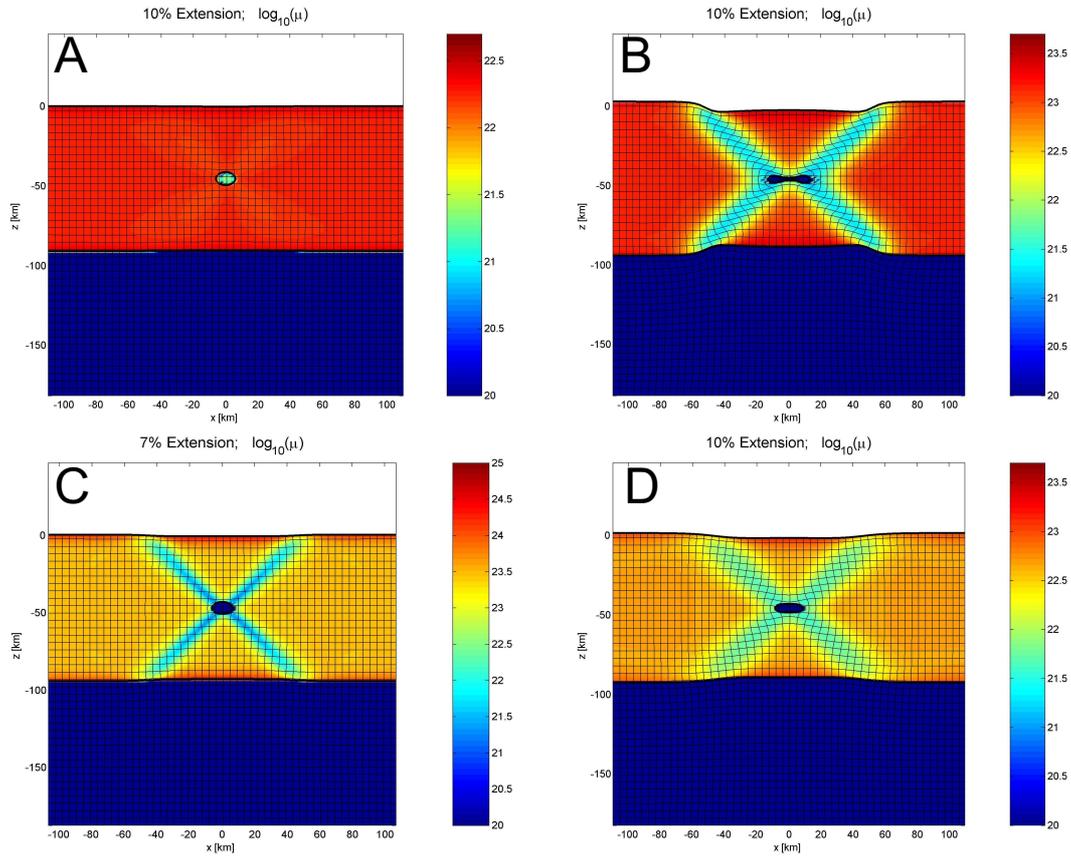


Figure 4.12: 2-D results for a lithosphere with an initially circular weak inclusion of 5 km radius that is extended with a pure-shear constant strainrate of $3 \times 10^{-15} \text{s}^{-1}$. The initial viscosity is A) $\mu_0 = 5 \times 10^{22}$, B) $\mu_0 = 5 \times 10^{23}$, C) $\mu_0 = 10^{25}$. Model D) has the same parameters as model B), but with plasticity deactivated. Other parameters are $G = 10^{11}$, $\gamma = 0.1\rho c_p = 3 \times 10^6$, $\kappa = 10^{-6}$, $\sigma_0 = 50 \text{ MPa}$, $\sigma_y = 1000 \text{ MPa}$. The free surface is approximated by a layer of $\mu = 10^{19} \text{ Pas}$ in the top 50 km's of the model (with a no-stress boundary condition on top of the layer). A moderate resolution of 301×64 is employed. Black lines represent a passive marker grid. Gravity is not present.

rates and viscosities, plasticity is required to limit maximum stress levels, which would otherwise be unrealistically large. Plasticity also helps in creating linear zones where localization occurs during later stages. Localization is also possible in the non-plastic modes. However localized zones in the viscous mode are relatively thick and require large heterogeneities, whereas deformation in the viscoelastic mode develops unrealistically large stresses. That the creation of shearzones around weak inclusions is difficult was also found by Mancktelow [2002] for a strain-weakening viscous rheology rather than a thermal weakening rheology.

It is interesting to compare the results obtained here with other works. Regenauer-Lieb et al. [2001] presented a model in which subduction was initiated by adding ~ 10 km of sediments on top of an oceanic plate. They obtained lithospheric-scale shearzones for wet olivine rheologies only. Whereas their model setup and rheology differs from the one employed in this study, we can still obtain some insight in the process by employing the scaling relationships derived here. The initial growthrate of the shear-localization instability under a constant stress simple shear boundary condition is $q_{1D} = 0.5Br$. Diffusion may inhibit shear localization if $q_{1D} < 3/Pe^2$. By using the definitions of the non-dimensional parameters (eq. 4.13), it can be shown that diffusion will influence shear-localization if $\mu_0 > \frac{0.5\gamma\sigma_0^2 R^2}{3\kappa\rho c_p}$. We can estimate $\sigma_0 \approx 250$ MPa (10 km thick pile of sediments with a density of 2500 kg/m³) and $R \approx 15$ km (their sediment pile has a width of around 30 km's). Employing common values for the other parameters ($\gamma = 0.1, \rho c_p \kappa = 3$), gives a minimum viscosity of $\mu_0 \approx 10^{23}$ Pas, above which diffusion will influence shear-localization. If the constant stress boundary condition is applicable, the models should ultimately result in thermal runaway after $t \frac{G}{\mu_0} \approx \frac{2}{Br}$ or $t \approx \frac{2\rho c_p \mu_0}{\sigma_0^2 \gamma}$. This time increases with increasing viscosity, from $t \approx 3$ Myrs for $\mu_0 = 10^{23}$ Pas to $t \approx 300$ Myrs for $\mu_0 = 10^{25}$ Pas (using the parameters above). Thus, for large effective viscosities, diffusion influences localization and the time required to produce shear-localization increases significantly. Both arguments seem to indicate that low effective viscosities are required to initiate shearzones under constant stress boundary conditions. Wet rheologies generally have smaller effective viscosities for the same ambient conditions than dry rheologies. Thus, the scaling laws derived here seem to give a physical explanation of the observation that water is required to initiate subduction [Regenauer-Lieb et al., 2001]. One-dimensional numerical simulations, performed under pure-shear boundary conditions and for different olivine rheologies and inclusion

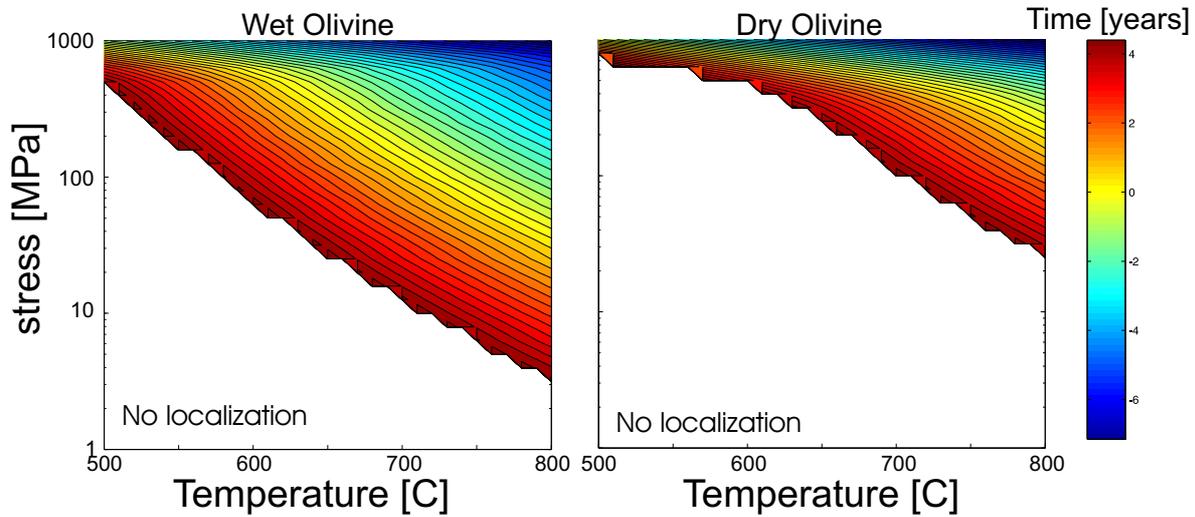


Figure 4.13: Results of 1-D simulations for the initial stages under a constant stress boundary condition for wet and dry olivine rheology. The size of the heterogeneity is 10 km. Colors indicate the time required to raise the temperature 1 K. Wet rheology allows localization to occur at smaller stresses compared to the dry rheology.

sizes confirm this (see fig. 4.13).

Regenauer-Lieb and Yuen [2004] studied the development of thermal shear-zones in a lithosphere with a water-dependent rheology. They employed a constant velocity boundary condition and showed that water is important in creating aseismic lithospheric-scale shearzones. Under dry conditions they obtained seismic shearzones, preventing further computation. A similar effect can be observed here (fig. 4.11). For a given strainrate, thermal runaway occurs at large effective viscosities (i.e. dry rheology), whereas stable localization occurs at smaller viscosities (wet rheology).

We have ignored the strainrate-dependence of viscosity in 1-D and 2-D simulations. The powerlaw-dependence of viscosity on strainrate $\mu \propto \dot{\epsilon}^{\frac{1}{n}-1}$ will produce additional weakening of viscosity with increasing strainrate. During later stages, this effect may be significant and would increase both the shear-localization and the thermal runaway domain on figure 4.11.

4.6 Conclusions

We have identified the parameters that control the initiation of thermal shearzones in a visco-elasto-plastic material. The feedback mechanism responsible for localization is a reduction of viscosity due to an increase in temperature through shear-heating. It was demonstrated that up to five deformation modes occur, as a result of the rheological constitutive equations. These deformation modes are a function of the non-dimensional numbers $B_{vis} = \frac{2\mu_0\dot{\epsilon}_0}{\sigma_0}$, $Br = \frac{\sigma_0^2\gamma}{\rho c_p G}$ and $B_{pl} = \frac{\sigma_0}{\sigma_y}$. Each deformation mode has a characteristic stress-evolution. Comparison of the simplified rheology with a more complex rheology which takes into account the effect of diffusion creep, powerlaw creep and Peierls plasticity, shows that the simplified model is an acceptable first order approximation of the more complex models.

Systematic one-dimensional simulations show that shear-localization, if initiated by a reduction in viscosity or an increase in temperature, occur with the same deformation modes as in the 0-D case. Diffusion may inhibit localization. A growthrate of temperature increase was derived and scaling laws are computed for the dependency of this growthrate on the various non-dimensional parameters. An analytical solution is derived for the effect of diffusion on shear-localization.

Two-dimensional simulations generally behave like the 1-D simulations. If a circular, rather than linear, initial heterogeneity is present and a pure shear boundary condition is applied, plasticity is required to create large-scale linear shear-zones. For earth-like parameters, shear-localization through the mechanism presented here is possible, but requires sufficiently large effective viscosities and heterogeneity sizes. Heterogeneities of $O(1)$ km, for example, require an effective viscosity $O(10^{24})$ Pas, to initiate localization. For viscosities in excess of 10^{26} Pas, thermal runaway occurs. Smaller heterogeneities will also produce heat, but this heat is rapidly removed by the effects of thermal diffusion, and hence no localized shearzones form.

4.7 Appendix A: Rheology of dry olivine and conversion to the simplified model

Viscous deformation of olivine in the ductile regime occurs mainly by diffusion creep, power-law creep and the Peierls mechanism also called low-temperature plasticity [e.g. Kameyama et al., 1999, Regenauer-Lieb and Yuen, 2003]. If, in addition, the effect of elasticity in a Maxwell material is considered, then the combined rheological equation becomes:

$$\dot{\epsilon}_{ij} = \frac{1}{2G} \frac{d\tau_{ij}}{dt} + \dot{\epsilon}_{ij}^{vis} \quad (4.43)$$

with

$$\dot{\epsilon}_{ij}^{vis} = \dot{\epsilon}_{ij}^{pc} + \dot{\epsilon}_{ij}^{dc} + \dot{\epsilon}_{ij}^{pp} \quad (4.44)$$

where the subscripts *pc*, *dc* and *pp* denote the effects of powerlaw creep, diffusion creep, and Peierls plasticity respectively. These effects are described in more detail in the following sections.

4.7.1 Powerlaw creep

The flow-law for powerlaw creep is generally determined with triaxial compression experiments and has the form [Karato and Jung, 2003]

$$\dot{\epsilon}_{11} = A_p C_{OH}^r (\tau_{11} - \tau_{33})^n \exp\left(-\frac{Q_p + PV}{RT}\right) \quad (4.45)$$

where A_p , Q_p , V are experimentally determined values, P is pressure, $R = 8.31 \text{ J mol}^{-1} \text{ K}^{-1}$ the universal gas constant, T temperature, C_{OH} hydroxile concentration, r a parameter ($=0$ in dry cases) and 1 indicates the direction of maximum compression. Unfortunately, equation 4.45 is not in a general tensorial form, which is required for numerical computations. Molnar et al. [1998], among others, demonstrate how the tensorial form can be obtained. The first assumption is that the material is incompressible. Triaxial compression then gives:

$$\dot{\epsilon}_{11} = -(\dot{\epsilon}_{22} + \dot{\epsilon}_{33}) = -2\dot{\epsilon}_{22} = -2\dot{\epsilon}_{33} \quad (4.46)$$

If the material is also isotropic:

$$\tau_{11} = -2\tau_{22} = -2\tau_{33} \quad (4.47)$$

a general tensorial form of equation 4.45 is:

$$\dot{\epsilon}_{ij} = AC_{OH}^r A_p \sigma_{2nd}^{n-1} \exp\left(-\frac{Q_p + PV}{RT}\right) \tau_{ij} \quad (4.48)$$

where A is a constant and $\sigma_{2nd} = \sqrt{\frac{1}{2}\tau_{ij}\tau_{ij}}$ is the second invariant of the deviatoric stress tensor. In a (ideal) triaxial test, shear stresses are not present and the second invariant simplifies to:

$$\sigma_{2nd} = \sqrt{\frac{1}{2}(\tau_{11}^2 + \tau_{22}^2 + \tau_{33}^2)} = \left(\frac{\sqrt{3}}{2}\right) \tau_{11} \quad (4.49)$$

spelling out equation 4.48 for $\dot{\epsilon}_{11}$ gives:

$$\begin{aligned} \dot{\epsilon}_{11} &= AA_p C_{OH}^r \left(\left(\frac{\sqrt{3}}{2}\right) \tau_{11}\right)^{n-1} \exp\left(-\frac{Q_p + PV}{RT}\right) \tau_{11} \\ &= AA_p C_{OH}^r \left(\frac{\sqrt{3}}{2}\right)^{n-1} \tau_{11}^n \exp\left(-\frac{Q_p + PV}{RT}\right) \end{aligned} \quad (4.50)$$

Equation 4.45 can be rewritten as (using the fact that $\tau_{33} = -\frac{1}{2}\tau_{11}$):

$$\dot{\epsilon}_{11} = A_p C_{OH}^r \left(\frac{3}{2}\right)^n \tau_{11}^n \exp\left(-\frac{Q_p + PV}{RT}\right) \quad (4.51)$$

Equation 4.51 and equation 4.50 should be the same. This can only be the case if $A = \frac{3^{(n+1)/2}}{2}$. The general tensor expression for powerlaw creep is thus:

$$\dot{\epsilon}_{ij}^{pc} = \frac{3^{(n+1)/2}}{2} A_p C_{OH}^r \sigma_{2nd}^{n-1} \exp\left(-\frac{Q_p + PV}{RT}\right) \tau_{ij} \quad (4.52)$$

The pressure term in equation 4.52 introduces is maximum $\sim 10 - 20\%$ of Q_p (for depths of up to 200 km) and has therefore been ignored in this study.

4.7.2 Diffusion creep

Diffusion creep is linearly viscous (i.e. $n = 1$), but has a grainsize-dependency. The tensorial form can be derived as described above and is given by [e.g. Karato et al., 1986]

$$\dot{\epsilon}_{ij}^{dc} = \frac{3}{2} A_d d^{-p} \tau_{ij} \exp\left(-\frac{Q_d}{RT}\right) \quad (4.53)$$

where d is the grainsize and p an experimentally determined constant.

4.7.3 Peierls low temperature plasticity

At large stresses, the diffusion creep and powerlaw creep flow laws break down and are replaced by Peierls plasticity. The experimentally determined flow law for Peierls plasticity is given by [e.g. Goetze, 1978, Goetze and Evans, 1979]:

$$\dot{\epsilon}_{11} = \dot{\epsilon}_0 \exp \left(-\frac{H_a}{RT} \left(1 - \frac{\tau_{11} - \tau_{33}}{\sigma_0} \right)^2 \right) \quad (4.54)$$

Note that a different form of equation 4.54 was used in Regenauer-Lieb and Kohl [2002], Regenauer-Lieb and Yuen [2004] and Karato et al. [2001]. They additionally introduced a dependency of H_a on the melting temperature (as a function of pressure). This correction increases H_a by a maximum amount of $\sim 7\%$ at a depth of 100 km, and for the sake of simplicity, this effect has been ignored here.

A more general form of equation 4.54 is clearly given by:

$$\dot{\epsilon}_{ij} = A \frac{\tau_{ij}}{\sigma_{2nd}} \dot{\epsilon}_0 \exp \left(-\frac{H_a}{RT} \left(1 - \frac{b\sigma_{2nd}}{\sigma_0} \right)^2 \right) \quad (4.55)$$

here A and b are geometrical factors that should be derived by matching the tensorial form with the triaxial expression. Making again the approximation of incompressibility and isotropy (eqs. 4.47 and 4.49), allows to rewrite eq. 4.55 as:

$$\dot{\epsilon}_{11} = A \frac{2\tau_{11}}{\sqrt{3}\tau_{11}} \dot{\epsilon}_0 \exp \left(-\frac{H_a}{RT} \left(1 - \frac{b\sqrt{3}\tau_{11}}{2\sigma_0} \right)^2 \right) \quad (4.56)$$

and equation 4.54 as:

$$\dot{\epsilon}_{11} = \dot{\epsilon}_0 \exp \left(-\frac{H_a}{RT} \left(1 - \frac{3\tau_{11}}{2\sigma_0} \right)^2 \right) \quad (4.57)$$

comparing 4.56 with 4.57 implies that $A = \frac{\sqrt{3}}{2}$ and $b = \sqrt{3}$. Thus 4.55 becomes:

$$\dot{\epsilon}_{ij}^{pp} = \frac{\sqrt{3}}{2\sigma_{2nd}} \dot{\epsilon}_0 \exp \left(-\frac{H_a}{RT} \left(1 - \sqrt{3} \frac{\sigma_{2nd}}{\sigma_0} \right)^2 \right) \tau_{ij} \quad (4.58)$$

It should be noted that Peierls plasticity is only valid for low temperatures $T \leq 1000\text{K}$ and stresses $O(100)$ MPa. At lower stresses the particular expression of Peierls plasticity may interfere with the other flowlaws.

Table 4.6 summarizes values for dry and wet olivine employed in this study.

Variable	Value	Units	Wet/Dry	Source
Q_d	290×10^3	J/mol	dry	1
Q_d	250×10^3	J/mol	wet	1
A_d	7.7×10^{-8}	$\text{Pa}^{-1}\text{s}^{-1}\text{mm}^3$	dry	1
A_d	1.5×10^{-9}	$\text{Pa}^{-1}\text{s}^{-1}\text{mm}^3$	wet	1
p	3	-	both	1
d	0.1	mm	both	
A_p	2.4×10^5	$\text{MPa}^{-3.5}\text{s}^{-1}$	dry	2
Q_p	510×10^3	J/mol	dry	2
n	3.5	-	dry	2
A_p	1.9×10^3	$\text{MPa}^{-3}\text{s}^{-1}$	wet	2
Q_p	420×10^3	J/mol	wet	2
n	3	-	wet	2
C_{OH}	4000	ppm H/Si	wet	2
r	1.3	-	wet	2
r	0	-	dry	2
$\dot{\epsilon}_0$	5.7×10^{11}	s^{-1}	dry	3
H_a	536×10^3	J/mol	dry	3
σ_0	8.5×10^9	Pa	dry	3
$\dot{\epsilon}_0$	1.2×10^{12}	s^{-1}	wet	3
H_a	498×10^3	J/mol	wet	3
σ_0	9.1×10^9	Pa	wet	3

Table 4.6: Rheological parameters for olivine employed in this study. References: 1-Karato et al. [1986], 2-Karato and Jung [2003], 3-Regenauer-Lieb and Kohl [2002]

4.7.4 Conversion to simplified model

Equations 4.58, 4.53 and 4.52 can be used to estimate the values of γ and μ_0 in eq. 4.7. In the case when diffusion creep dominates the initial stress evolution, viscosity can be approximated by

$$\mu_0 = \frac{|\dot{\epsilon}^{vis}(T, |\tau|)|}{2|\tau|} \quad (4.59)$$

If powerlaw or dislocation creep dominates the steady-state stress evolution, μ_0 can be obtained from equation 4.52. Finally, γ can be computed numerically by computing

μ_0 effective viscosity at T and at $T + \Delta T$. This gives:

$$\gamma = -\frac{\ln\left(\frac{\mu_0(T+\Delta T)}{\mu_0(T)}\right)}{\Delta T} \quad (4.60)$$

We have employed $\Delta T = 1^\circ \text{K}$.

The simplified rheology consists thus of a viscous part (estimated either from powerlaw or from diffusion creep), a plastic part (estimated from Peierls plasticity) and an elastic part. Good agreement exist between the simplified and the complex models, except if a transition from diffusion-creep to powerlaw creep occurs during a model run. In this case, the simplified model typically underestimates stresses and growth rates of temperature (fig. 4.2). A MATLAB script that implements the rheology described above, with the parameters listed in table 4.6, is shown in figure (4.14).

4.8 Appendix B: MATLAB codes for the 0-D model

The MATLAB script ODE_VEP.m (www.mathworks.com) which integrates the coupled ODE's in equation 4.17 is shown on figure 4.15. A variable, adaptive timestep is employed together with an implicit time-integration. A MATLAB script which calls ODE_VEP.m and compares the dry olivine model with the simplified model is shown on figure 4.16.

4.9 Acknowledgements

Neil Mancktelow, Dani Schmid, Klaus Regenauer-Lieb, Chris Hieronymous, Jeroen van Hunen and Ritske Huismans have provided valuable input during various stages of this work. James Connolly helped to improve the English. We thank the Swiss National Science Foundation for financial support (project 21-61912.00).

```

function [str_rate_creep, str_rate_diff, str_rate_pow, str_rate_pei, mu_eff, gamma, sigma_y, mu_eff_best] = ...
    Olivine(varargin)
%=====
% OLIVINE.M
%
T      = varargin{1};
stress = varargin{2};
flowlaw_type = varargin{3};
if nargin>3
    str_rate_bg = varargin{4};
end
R      = 8.3145;    % [J/mol/K      ]
switch flowlaw_type
case 'Olivine_dry'
    % Diffusion creep parameters for dry olivine (Karato 1986)
    Qd = 290e3;    % [J/mol      ]
    Ad = 7.7e-8;   % [Pa-1/s-1/mm^3 ]
    d  = 0.1;     % [mm      ]
    % Powerlaw creep parameters for dry olivine (Karato and Jung 2003)
    Qp = 510e3;    % [J/mol      ]
    Ap = 1.26e-12; % [MPa-3/s-1    ]
    n  = 3;       % [      ]
    % Peierls mechanism parameters (Regenauer-Lieb & Kohl 2003)
    str_rate0 = 5.7e11; % [s^-1      ]
    Ha = 536e3;    % [J/mol      ]
    sigma_0 = 8.5e9; % [Pa      ]
case 'Olivine_wet'
    % Diffusion creep parameters for wet olivine (Karato 1986)
    Qd = 290e3;    % [J/mol      ]
    Ad = 1.5e-9;   % [MPa-1/s-1/mm^3 ]
    d  = 0.1;     % [mm      ]
    % Powerlaw creep parameters for wet olivine (Karato and Jung 2003)
    Qp = 470e3;    % [J/mol      ]
    r  = 1.3;     % [      ]
    COH = 4000;   % [ppm H/Si    ]
    Ap = 3.63e-18*COH^r; % [Pa-3/s-1    ]
    n  = 3;       % [      ]
    % Peierls mechanism parameters (Regenauer-Lieb & Kohl 2003)
    str_rate0 = 1.2e12; % [s^-1      ]
    Ha = 498e3;    % [J/mol      ]
    sigma_0 = 9.1e9; % [MPa      ]
otherwise
    error('Flowlaw type is unknown.')
end

% Compute parameters with T+1 (needed to compute gamma):
[str_rate_creep_T1, str_rate_diff, str_rate_pow, str_rate_pei] = ...
    Flowlaws(stress, T+1, R, Qd, Ad, d, Qp, Ap, n, str_rate0, Ha, sigma_0);
% Compute parameters with given T:
[str_rate_creep, str_rate_diff, str_rate_pow, str_rate_pei] = ...
    Flowlaws(stress, T, R, Qd, Ad, d, Qp, Ap, n, str_rate0, Ha, sigma_0);

% Compute parameters like gamma and initial effective viscosity.
mu_eff_T1 = stress./str_rate_creep_T1./2;
mu_eff    = stress./(str_rate_creep*1e10)./2; mu_eff=mu_eff*1e10;
gamma     = -log(mu_eff_T1./mu_eff)./1;

if nargin>3
    % Yield stress with the given background strainrate (this expression is only valid for temperatures <1000):
    Fac = (R*T./Ha.*log(sqrt(3))./2.*str_rate0./str_rate_bg).^0.5;
    sigma_y = sigma_0./sqrt(3).*(1 - Fac );
    sigma_y(find(Fac>=.99)) = 10000e6;
end

% Estimate the effective viscosity which best matches the model.
mu_eff_best = mu_eff;
if nargin>3
    stress = [str_rate_bg./[(3.^((n+1)./2))./2.*Ap.*exp(-Qp./R./T)]]^(1/n);
    if stress<sigma_y
        mu_eff_best = stress./str_rate_bg./2; %effective viscosity for powerlaw rheology
    else
        mu_eff_best = mu_eff;
    end
end

%=====
function [str_rate_creep, str_rate_diff, str_rate_pow, str_rate_pei] = ...
    Flowlaws(stress, T, R, Qd, Ad, d, Qp, Ap, n, str_rate0, Ha, sigma_0);
% Computes the flowlaw for olivine parameters

str_rate_diff = 3./2.*Ad.*d^(3).*exp(-Qd./R./T).*stress; % Diffusion creep strainrate
str_rate_pow  = (3.^((n+1)./2))./2.*Ap.*stress.^n.*exp(-Qp./R./T); % Powerlaw creep strainrate
str_rate_pei  = sqrt(3)./2./stress.*str_rate0.*exp(-Ha./R./T.*(1 - sqrt(3).*stress./sigma_0).^2).*stress; % Peierls

str_rate_pei(find(stress<100e6))= 0; %Peierls plasticity is only applicable at large stresses

% Total strainrate
str_rate_creep = str_rate_diff + str_rate_pow + str_rate_pei;

```

Figure 4.14: The MATLAB script Olivine.m which can be used to compute the rheology of wet and dry olivine at given temperature and pressure conditions. The script also outputs μ_0 , γ and σ_y for use in the simplified rheological model.

```

function [T_vec,Tau_vec,time] = ...
    ODE_VEP(Sigma_yield, str_rate, Br, T, Tmax,eps, Rheology, ShearType, Temp_init, Stress_char, Temp_char)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% ODE_VEP.M
%
% OD time evolution of a visco-elasto-plastic body deformed under constant strainrate.
%
% 2004, Boris Kaus and Yuri Podladchikov
%
% Initialization
dTdt_vec = [0,0];
Tau       = 1;
Tau_vec   = [1];
ShearHeat_vec = [0];
T_vec     = T;
time      = 0;
dT_dt    = Br*Tau*str_rate;
dt        = min(eps*(abs(T)+eps)/(abs(dT_dt)+eps),eps);
keepgoing = 1;
Tau_new   = Tau;
T_new     = T;

switch ShearType
case 'Simple Shear'
fs       = 1;
case 'Pure Shear'
fs       = 2;
otherwise
error('Unknown Shear Type')
end

while keepgoing == 1

    switch Rheology
    case 'Frank_Kamenetzky'
        %Linear exponential case
        mu_eff = exp(-T);
    otherwise %Fully nonlinear case for olivine
        [str_rate_creep, str_diff, str_pow, str_pei, mu_eff0, gamma] = ...
            Olivine((Temp_init)*Temp_char, 1 *Stress_char,Rheology);
        [str_rate_creep, str_diff, str_pow, str_pei, mu_eff, gamma] = ...
            Olivine((T_new +Temp_init)*Temp_char, Tau_new*Stress_char,Rheology);
        mu_eff = (mu_eff/mu_eff0);
    end
    Tau_new = min(mu_eff*(Tau + 2*str_rate*dt)/(mu_eff+dt), Sigma_yield); %Stress evolution
    str_el  = (Tau_new-Tau)/dt/2; %Elastic strainrate
    dT_dt   = fs*Br*Tau_new*(str_rate-str_el); %Thermal evolution
    ShearHeat = Tau_new*(str_rate-str_el);
    T_new    = T + dt*dT_dt;

    if abs(T_new - T) < eps*(1+abs(T_new)+abs(T)) ...
        & abs(Tau_new - Tau) < eps*(1+abs(Tau_new)+abs(Tau))
        % Accept current timestep
        T = T_new;
        Tau = Tau_new;

        % Exit criteria
        if T > Tmax
            dtc = (T-Tmax)/(T-T_vec(end));
            T = T - dtc*(T - T_vec(end));
            Tau = Tau - dtc*(Tau - Tau_vec(end));
            dt = dt*(1 - dtc);
            keepgoing = 0;
        end
        dTdt_vec = [dTdt_vec,dT_dt];
        Tau_vec = [Tau_vec,Tau];
        T_vec = [T_vec,T];
        time = [time,time(end)+dt];
        ShearHeat_vec = [ShearHeat_vec, ShearHeat];

        if abs(T_new - T) < eps/10*(1+abs(T_new)+abs(T))
            % Increase the timestep
            dt = dt*1.01;
        end
    else
        % Do not accept the current step and decrease the timestep
        dt = dt/2;
    end

    % plotting
    if mod(length(time),1000)==0
        figure(1), clf
        subplot(211),plot(time,Tau_vec,'-x'),ylabel('\tau\sigma_0'),xlabel('t/[mu_0/G]')
        subplot(212),plot(time,T_vec),ylabel('T'), xlabel('t/[mu_0/G]')
        drawnow, pause(1/10)
    end
end
end

```

Figure 4.15: The MATLAB script ODE_VEP.m which integrates eq. (4.17) versus time under constant strainrate boundary conditions.

```

%ODE_VEP_real_start
%compute stress evolution in a lithosphere with 'real' parameters
clear;
% Insert values in dimensional units
Sigma0_dim = 1e6; % Starting stress [Pa ]
T0_dim = 600; % Starting temperature [C ]
G_dim = 5e10; % elastic shear module [Pa ]
str_rate_dim = 3e-15; % Background strainrate [1/s ]
kappa_dim = 1e-6; % Diffusivity [m2/s ]
rho_cp_dim = 3e6; % density*heat capacity [Jm2K/s ]
T0_dim = T0_dim+273; % Temperature in Celcius
ShearType = {'Simple Shear','Pure Shear'}
ShearType = ShearType{2}; % Type of shear
eps = 1e-2; % Accuracy of the solver
Tmax = 5; % T-increase to terminate solver [K ]
R = 3000; % Size of heterogeneity [m ]
Rheology = 'Olivine_dry'; % Dry Olivine model

% Compute characteristic values at given stress, temperature and strainrate levels=====
% Compute characteristic values at temperature=T0_dim and Stress=Sigma0_dim
[str_rate_creep, str_diff, str_pow, str_pei, mu0, gamma, Sigma_y,mu_best] = ...
    Olivine(T0_dim, Sigma0_dim, Rheology,str_rate_dim);

% Compute characteristic values and non-dimensional numbers
Stress_char = Sigma0_dim;
Time_char = mu0/G_dim;
Temp_char = 1/gamma;
Length_char = sqrt(kappa_dim*Time_char);

Str_rate_char = 1./Time_char;
Visc_char = Stress_char*Time_char;
Br = 0.5*Sigma0_dim^2*gamma/rho_cp_dim/G_dim;
Bvis = 2*mu0*str_rate_dim/Sigma0_dim;
Bpl = Sigma0_dim/Sigma_y;
Pe = R/Length_char;
%=====

% % Compute some solver-specific values
Temp_init = T0_dim/Temp_char;
T = 0; % Initial temperature [ ]
Tau = 1; % Initial stress [ ]
Sigma_yield = 1e10; % Not used with real data [ ]
str_rate = Bvis/2; % Background strainrate [ ]
Tmax = Tmax/Temp_char; % Solver termination [ ]
% Compute the time-evolution for the dry-olivine rheological model=====
[T_vec1,Tau_vec1,time1] = ODE_VEP(Sigma_yield, str_rate, Br, T, Tmax, eps, ...
    Rheology, ShearType, Temp_init, Stress_char, Temp_char);

% Transform non-dimensional values to dimensional ones
Stress_vec_dim1 = Tau_vec1.*Stress_char;
Temp_vec_dim1 = (T_vec1+Temp_init).*Temp_char-273;
Time_vec_dim1 = time1.*Time_char;
Myrs = 1e6*365*24*3600;
%=====

% Compute the time-evolution for the simplified rheological model=====
mu0 = mu_best; %Take the 'best estimate viscosity' as scale
Time_char = mu0/G_dim;
Bvis = 2*mu0*str_rate_dim/Sigma0_dim
Temp_init = T0_dim/Temp_char;
T = 0; % Initial temperature [ ]
Tau = 1; % Initial stress [ ]
str_rate = Bvis/2; % Background strainrate [ ]
Rheology = 'Frank_Kamenetzky'; % Simplified model
Sigma_yield = Sigma_y/Stress_char; % Yield stress must be specified for simplified model
[T_vec2,Tau_vec2,time2] = ODE_VEP(Sigma_yield, str_rate, Br, T, Tmax, eps, ...
    Rheology, ShearType, Temp_init, Stress_char, Temp_char);

Stress_vec_dim2 = Tau_vec2.*Stress_char;
Temp_vec_dim2 = (T_vec2+Temp_init).*Temp_char-273;
Time_vec_dim2 = time2.*Time_char;
Myrs = 1e6*365*24*3600;
%=====

% Visualization
figure(1),clf
subplot(211),
subplot(211),
plot(Time_vec_dim1/Myrs*1000,Stress_vec_dim1/1e6,'r--',Time_vec_dim2/Myrs*1000,Stress_vec_dim2/1e6,'b-');
ylabel('\tau [MPa]')
title([' log10(B_{pl})=',num2str(log10(Bpl)), ' log10(Br)=',num2str(log10(Br)), ' log10(B_{vis})=',num2str(log10(Bvis))])
legend('Olivine','Simplified rheology')
grid on

subplot(212)
plot(Time_vec_dim1/Myrs*1000,Temp_vec_dim1,'r--',Time_vec_dim2/Myrs*1000,Temp_vec_dim2,'b-')
ylabel(' T [°C]')
xlabel('time [1000 yrs]')
grid on

```

Figure 4.16: The MATLAB script ODE_VEP_real_start.m which should be used together with ODE_VEP.m.

Bibliography

- C. Auth, D. Bercovici, and U.R. Christensen. Two-dimensional convection with a self-lubricating, simple-damage rheology. *Geophysical Journal International*, 154: 783–800, 2003.
- D. Bercovici. A simple model of plate generation from mantle flow. *Geophysical Journal International*, 114:635–650, 1993.
- D. Bercovici. Plate generation in a simple model of lithosphere-mantle flow with dynamic self-lubrication. *Earth and Planetary Science Letters*, 144:41–51, 1996.
- D. Bercovici. The generation of plate tectonics from mantle convection. *Earth and Planetary Science Letters*, 205:107–121, 2003.
- D. Bercovici, Y. Ricard, and G. Schubert. A two-phase model for compaction and damage 1. general theory. *Journal of Geophysical Research-Solid Earth*, 106(B5): 8887–8906, 2001.
- F.W. Brace and D.L. Kohlstedt. Limits on lithospheric stress imposed by laboratory experiments. *Journal of Geophysical Research*, 50:6248–6252, 1980.
- R.K. Cardwell, D.S. Chinn, G.F. Moore, and D. L. Turcotte. Frictional heating on a fault zone with finite thickness. *Geophysical Journal of the Royal Astronomical Society*, 52:525–530, 1978.
- H.S. Carslaw and J. C. Jaeger. *Conduction of heat in solids*. Oxford University Press, New York, 2nd edition, 1959.
- H. P. Cherukuri and T. G. Shawki. Energy-based localization theory: I. Basic framework. *International Journal of Plasticity*, 11:15–40, 1995.

- J. H. Dieterich. Modeling of rock friction. 1. Experimental results and constitutive equations. *Journal of Geophysical Research*, 84(NB5):2161–2168, 1979.
- L. Fleitout and C. Froidevaux. Thermal and mechanical evolution of shear zones. *Journal of Structural Geology*, 2(1-2):159–164, 1980.
- D.A. Frank-Kamenetzky. *Diffusion and heat transfer in chemical kinetics*. Plenum Press, New York, 1969.
- C. Goetze. The mechanisms of creep in olivine. *Philosophical Transactions of the Royal Society of London*, 288:99–119, 1978.
- C. Goetze and B. Evans. Stress and temperature in the bending lithosphere as constrained by experimental rock mechanics. *Geophysical Journal of the Royal Astronomical Society*, 59:463–478, 1979.
- H. Jeffrey. *The Earth*. Cambridge University Press, Cambridge, 1959.
- C. Kameyama, D. A. Yuen, and S. Karato. Thermal-mechanical effects of low temperature plasticity (the Peierls mechanism) on the deformation of a viscoelastic shear zone. *Earth and Planetary Science Letters*, 168:159–162, 1999.
- S. Karato and H. Jung. Effects of pressure on high-temperature dislocation creep in olivine. *Philosophical Magazine*, 83(3):401–414, 2003.
- S. Karato, M. S. Paterson, and J. D. Fitzgerald. Rheology of synthetic olivine aggregates: influence of grain size and water. *Journal of Geophysical Research*, 91(B): 8151–8176, 1986.
- S. Karato, D. Yuen, and M. R. Riedel. Rheological structure and deformation of subducted slabs in the mantle transition zone: Implications for mantle circulation and deep earthquakes. *Physics of the Earth and Planetary Interiors*, 127:83–108, 2001.
- B.J.P. Kaus, Y.Y. Podladchikov, and D. W. Schmid. Eulerian spectral/finite difference method for large deformation modelling of visco-elasto-plastic geomaterials. *Bolletino Geofisico*, 45:346–349, 2004.

- D. L. Kohlstedt, B. Evans, and S. J. Mackwell. Strength of the lithosphere: Constraints imposed by laboratory measurements. *Journal of Geophysical Research*, 100(B9):17587–17602, 1995.
- A. H. Lachenbruch and J. H. Sass. Heat flow from the Cajon Pass, fault strength, and tectonic implications. *Journal of Geophysical Research*, 97(B4):4995–5015, 1991.
- L. Lavier, W.R. Buck, and A.N.B. Poliakov. Factors controlling normal fault offset in an ideal brittle layer. *Journal of Geophysical Research*, 105(B10):23431–23442, 2000.
- V. Lyakhovsky, Y. Ben-Zion, and A. Agnon. Distributed damage, faulting, and friction. *Journal of Geophysical Research-Solid Earth*, 102(B12):27635–27649, 1997.
- V. Lyakhovsky, Y. Ben-Zion, and A. Agnon. A viscoelastic damage rheology and rate- and state-dependent friction. *Geophysical Journal International*, submitted 2004.
- N. S. Mancktelow. Finite-element modelling of shear zone development in viscoelastic materials and its implications for localisation of partial melting. *Journal of Structural Geology*, 24:1045–1053, 2002.
- D. McKenzie. Surface deformation, gravity anomalies, and convection. *Geophysical Journal of the Royal Astronomical Society*, 48:211–238, 1977.
- P. Molnar, G. Houseman, and C.P. Conrad. Rayleigh-Taylor instability and convective thinning of mechanically thickened lithosphere: effects of non-linear viscosity decreasing exponentially with depth and of horizontal shortening of the layer. *Geophysical Journal International*, 133:568–584, 1998.
- L.G.J. Montesi and G. Hirth. Grain size evolution and the rheology of ductile shear zones: from laboratory experiments to postseismic creep. *Earth and Planetary Science Letters*, 211:97–110, 2003.
- L.G.J. Montesi and M. T. Zuber. A unified description of localization for application to large scale tectonics. *Journal of Geophysical Research*, 107:10.1029/2001JB000465, 2002.
- L. Moresi and V. Solomatov. Numerical investigations of 2D convection in a fluid with extremely large viscosity variations. *Physics of Fluids*, 7:2154–2162, 1995.

- M. Ogawa. Shear instability in a viscoelastic material as the cause of deep focus earthquakes. *Journal of Geophysical Research*, 92(B1):13801–13810, 1987.
- M. Ogawa. Plate-like regime of a numerically modeled thermal convection in a fluid with temperature-, pressure-, and stress-history-dependent viscosity. *Journal of Geophysical Research*, 108(B2):2067, 2003.
- A. N. B. Poliakov, H. J. Herrmann, and Y. Y. Podladchikov. Fractal plastic shear bands. *Fractals*, 2:567–581, 1994.
- Regenauer-Lieb and D. Yuen. Modeling shear zones in geological and planetary sciences: solid- and fluid-thermal-mechanical approaches. *Earth-Science Reviews*, 63:295–349, 2003.
- K. Regenauer-Lieb and T. Kohl. Water solubility and diffusivity in olivine: Its role for planetary tectonics. *Mineralogical Magazine*, 2002.
- K. Regenauer-Lieb, J. P. Petit, and D. Yuen. Adiabatic shear bands in the lithosphere: Numerical and experimental approaches. *Electronic Geosciences*, 4:2, 1999.
- K. Regenauer-Lieb and D. Yuen. Rapid conversion of elastic energy into shear heating during incipient necking of the lithosphere. *Geophysical Research Letters*, 25(14):2737–2740, 1998.
- K. Regenauer-Lieb and D. Yuen. Quasi-adiabatic instabilities associated with necking processes of an elasto-viscoplastic lithosphere. *Physics of the Earth and Planetary Interiors*, 118:89–102, 2000a.
- K. Regenauer-Lieb and D. Yuen. Positive feedback of interacting ductile faults from coupling of equations of state, rheology and thermal-mechanics. *Physics of the Earth and Planetary Interiors*, in press, 2004.
- K. Regenauer-Lieb, D. Yuen, and J. Branlund. The initiation of subduction: Criticality by addition of water? *Science*, 294:578–580, 2001.
- K. Regenauer-Lieb and D. A. Yuen. Fast mechanisms for the formation of new plate boundaries. *Tectonophysics*, 322:53–67, 2000b.

- A. Ruina. Slip instability and state variable friction laws. *Journal of Geophysical Research*, 88:10359–10370, 1983.
- D. W. Schmid. *Finite and infinite heterogeneities under pure and simple shear*. Phd, Swiss Federal Institute of Technology, 2002.
- D. W. Schmid and Y. Podladchikov. Analytical solutions for deformable elliptical inclusions in general shear. *Geophysical Journal International*, 155:269–288, 2003.
- C. H. Scholz. *The mechanics of earthquakes and faulting*. Cambridge University Press, 2nd edition, 2002.
- J.C. Simo and T. J. R. Hughes. *Computational Inelasticity*. Springer, New York, 2nd edition, 2000.
- P. Tackley. Self-consistent generation of tectonic plates in three-dimensional mantle convection. *Earth and Planetary Science Letters*, 157:9–22, 1998.
- P. Tackley. The quest fore self-consistent generation of plate tectonics in mantle convection models. volume 121, pages 47–72. American Geophysical Union, Washington D.C., 2000a.
- P. Tackley. Self-consistent generation of tectonic plates in time- dependent, three-dimensional mantle convection simulations, 1. pseudoplastic yielding. *G3*, 01(23), 2000b.
- P. Tackley. Self-consistent generation of tectonic plates in time- dependent, three-dimensional mantle convection simulations. 2. strain weakening and asthenosphere. *G3*, 01(25), 2000c.
- R. Trompert and A. Hansen. Mantle convection simulations with rheologies that generate plate-like behavior. *Nature*, 395:686–689, 1998.
- D. L. Turcotte and E. R. Oxburgh. A fluid theory for the deep structure of dip-slip fault zones. *Physics of the Earth and Planetary Interiors*, 1:1113–1138, 1968.
- D. L. Turcotte and G. Schubert. *Geodynamics*. Wiley, 1982.

- D. A. Yuen, L. Fleitout, G. Schubert, and C. Froidevaux. Shear deformation zones along major transform faults and subducting slabs. *Geophysical Journal of the Royal Astronomical Society*, 54(1):93–119, 1978.

Chapter 5

Forward and reverse modelling of the three-dimensional viscous Rayleigh-Taylor instability¹

Abstract A combined finite-difference/spectral method is used to model the 3D viscous Rayleigh-Taylor instability. Numerically calculated growth rate spectra are presented for an initial sinusoidal perturbation of the interface separating two fluids with amplitude $10^{-3}H$ and $0.2H$, where H is the height of the system. At small initial amplitude, growth rate spectra closely follow linear theory, whereas the calculation with higher initial amplitude shows wavelength selection towards 3D perturbations. Numerical simulations and analytical theory are used to evaluate the applicability of previous 2D numerical models, which is shown to depend on (1) the wavelength and amplitude of an initially 2D sinusoidal perturbation and (2) the amplitude of background noise. It is also shown that reverse (backward) modeling is capable of restoring the initial geometry as long as overhangs are not developed. If overhangs are present, the possibility of restoring the initial conditions is largely dependent on the stage of overhang development.

¹This work has been published in slightly modified form in: Kaus, B.J.P, Podladchikov, Y.Y. (2001) Forward and reverse modeling of the three-dimensional viscous Rayleigh-Taylor instability, *Geophysical Research Letters* Vol.28, No.6, p.1095-1098.

5.1 Introduction

The Rayleigh-Taylor (RT) instability arising when a heavier fluid overlies a fluid with lower density has attracted attention of the Earth science community for some time. There are a numerous situations, where a RT-type model is applicable to nature. Examples are batholiths [Pons et al., 1992], salt tectonics [Podladchikov et al., 1993], and convective thinning of the lithosphere [Houseman and Molnar, 1997]. The RT instability has been intensively studied by laboratory experiments [e.g. Talbot et al., 1991], analytical methods [linear and nonlinear stability analysis, e.g. Ribe, 1998, Conrad and Molnar, 1997] and numerical simulations [e.g. Schmeling, 1987, Podladchikov et al., 1993]. Numerical calculations of the viscous RT instability, however, have been restricted to the 2D case, mainly because of limited computational power. Three-dimensional numerical simulations have been reported, but were done for viscous fluids with inertial forces [e.g. He et al., 1999]. In this paper, forward and reverse numerical simulations of the 3D viscous RT instability in absence of inertial forces are presented.

5.2 Mathematical model and numerical method

The Rayleigh-Taylor instability for the slowly creeping flow of viscous incompressible Newtonian fluids with constant viscosity is described by the Stokes system of equations, which are given by:

$$\frac{\partial V_i}{\partial x_i} = 0 \quad (5.1)$$

$$-\frac{\partial P}{\partial x_i} + \mu \Delta V_i + \rho g_i = 0 \quad (5.2)$$

where P is pressure, $V_i = (u, v, w)$ is velocity, $x_i = (x, y, z)$ are coordinates, μ =viscosity, $g_i = (0, 0, -g)$ is the gravitational acceleration and ρ is the density of the fluid. Non-dimensionalization was done taking H , $H^2 g \Delta \rho / \mu$, $H g \Delta \rho$ and $\mu / H g \Delta \rho$ as characteristic length, velocity, pressure and time respectively, where $\Delta \rho$ is the density difference between the upper and lower fluid. The equations (5.1-5.2) are solved here for a Cartesian box with height H and an aspect ratio of 5:5:1. The boundary conditions

are periodic for the lateral directions and no-slip on the upper and lower boundaries. The numerical method used for solving Eqns. 5.1-5.2 is a 3-D extension of the method used by [Schmalholz and Podladchikov, 1999] to model folding instabilities. It uses a spectral method for the horizontal directions and a conservative finite difference method on a regular grid for the vertical direction. The interface-tracking algorithm is a 3-D extension of the particle-line method described in Ten et al. [1998]. The interface is moved through time using an implicit time-marching scheme with adaptive time-step. Lines are added if the distance between two adjacent lines exceeds a given threshold. A resolution of 128×128 harmonics in horizontal direction and 513 grid points in vertical direction was used for the simulations shown in figures 2 and 4, and a resolution of $64 \times 64 \times 257$ for all other simulations. The numerical code runs on a single Pentium II, 400 MHz processor and needs approximately 200MB of RAM and 8-10 hours of CPU time for the lower resolution simulations and 1GB of RAM and 7 days of CPU time for the high resolution simulations presented here.

5.3 Infinitesimal and finite growth rate calculations

An unstable system consisting of two superposed immiscible fluid layers each of thickness 0.5 and same viscosity is considered. Infinitesimally initial perturbations on the interface separating the two fluids grow exponentially with time according to the relation $A(t) = A_0 \exp(qt)$, where A_0 is the initial amplitude, q is the growth rate and t is time, which can be calculated using a linear stability analysis [e.g. Chandrasekhar, 1961, Conrad and Molnar, 1997, Turcotte and Schubert, 1982] The initial spatial perturbations on the interface are split into a series of “normal modes”:

$$z_{int}(x, y) = 0.5 + dh_{int} \cos(k_x x) \cos(k_y y) \quad (5.3)$$

where $k_x = 2\pi/\lambda_x$ and $k_y = 2\pi/\lambda_y$ are wavenumbers in the x - and the y -direction respectively, λ_x and λ_y are wavelengths in x - and y -direction, and dh_{int} is an infinitesimally amplitude ($dh_{int} \ll \lambda_x, \lambda_y$). Each of these normal modes can be analyzed separately and has a non-dimensional growth rate q :

$$q = \frac{(k^2 + 2)e^{-k} - e^{-2k} - 1}{4k[-2ke^{-k} + e^{-2k} - 1]} \quad (5.4)$$

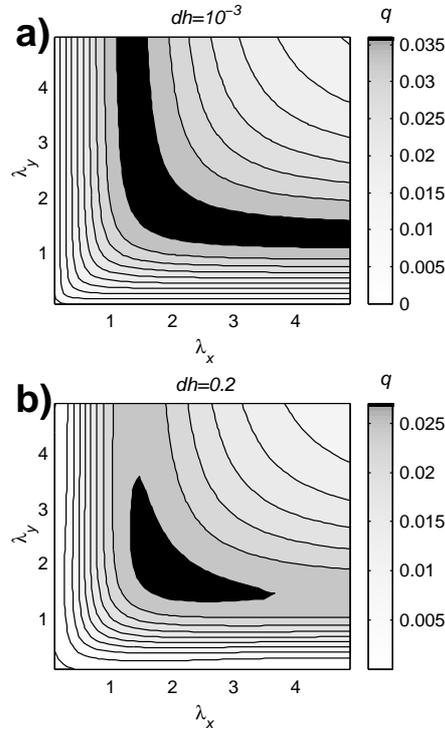


Figure 5.1: Contour plot of numerically calculated dimensionless growthrate q of a sinusoidal perturbation of the interface with an initial amplitude dh of A) 10^{-1} and B) 0.2. See text for explanation.

where $k = \sqrt{k_x^2 + k_y^2}$. The growth rate has a maximum of $q \approx 0.03835$ if $k \approx 4.895$. Note that Eq. (5.4) includes the 2D solution [e.g. Turcotte and Schubert, 1982] as a special case, i.e. $k_y = 0$. According to the linear stability analysis, which is only valid for very small perturbations of the interface, purely 2D waveforms have the same growth rate as an infinite number of 3D waveforms (composed by linear superposition of normal modes having wave number vectors of same length but different orientation). Thus a weakly nonlinear analysis is needed to constrain the pattern selection [Ribe, 1998], a situation similar to the Rayleigh-Bénard instability [see e.g. Godreche and Manneville, 1998, for discussion]. The growth rate q can also be calculated numerically by assuming a similar initial sinusoidal perturbation but of finite amplitude dh . The result of such a calculation is shown in figure 5.1. For initial amplitude of $dh = 10^{-3}$, the numerically calculated growth rate approaches the analytical growth rate (Eqn. 5.4) with an

accuracy of 1% (Fig. 5.1A). However, calculations using larger initial amplitude (e.g. $dh=0.2$, Fig. 5.1B) show a clear selection towards more 3D ($\lambda_x=\lambda_y$) normal mode perturbations. This result is in general agreement with the analytical calculations of [Ribe, 1998].

5.4 Forward modeling results

To study the competition between 2-D and 3-D initial perturbations, and thus to test the validity of previous 2-D numerical simulations, we performed a forward simulation of an initial (non-dominant) 2-D sinusoidal perturbation of the form $z_{int}(x, y) = 0.5 - 10^{-2} \cos(2\pi x/5)$ with normally distributed (white) noise with a variance of 5×10^{-4} . Two-dimensional numerical simulations by Schmeling [1987] showed already that such a configuration is unstable and leads to the breakup of the initial perturbation. This was also observed in our three-dimensional simulation, with the difference that the initial 2D perturbation decomposes into irregular 3-D structures (fig. 5.2). A two-dimensional Fourier transform of the interface revealed that the simultaneous growth and superposition of several dominant 3D normal modes is responsible for the development of irregular 3D structures. A rough estimate of the survival of initial sinusoidal 2D perturbations vs. dominant normal modes growing out of the background noise can be made by using linear stability growth rates. The growth in amplitude of an initial 2-D perturbation can be expressed $A^{2D}(t) = A_0^{2D} e^{(qt)}$, whereas that of a dominant mode, growing out of noise, is $A^{dom}(t) = A_0^{dom} e^{(q_{dom}t)}$. We define the characteristic time t^* as the time needed for an initial perturbation to reach an amplitude of 0.5. The initial perturbation survives if $A^{2D}(t^*) > A^{dom}(t^*)$. This condition can be written as:

$$\frac{\ln(A_0^{2D}/0.5)}{\ln(A_0^{dom}/0.5)} < \frac{q}{q_{dom}} \quad (5.5)$$

where $q_{dom} \approx 0.03835$ and q is calculated from linear stability analysis (Eqn. 5.4). Full numerical simulations are in good agreement with the prediction of equation 5.5 and show that breaking up of the initial 2D perturbation leads in most cases to a 3D geometry (see fig. 5.4). Equation 5.5 can thus be used to predict if an initial 2-D perturbation will survive or breaks up into 3-D structures.

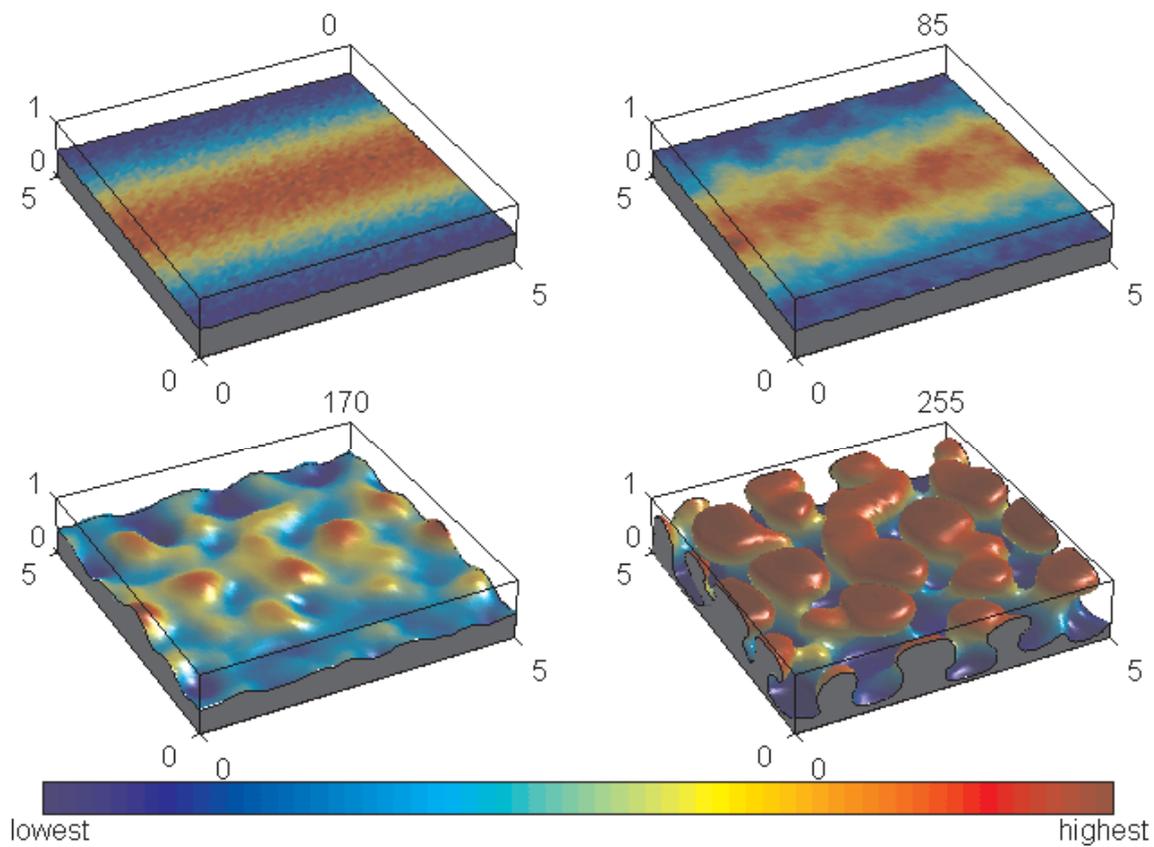


Figure 5.2: Forward simulation of a RT instability for an initial two-dimensional perturbation of the interface with an amplitude of 10^{-2} with addition of normally distributed random noise with a variance 5×10^{-4} . Numbers at the top are nondimensional times, colors show relative heights of the interface.

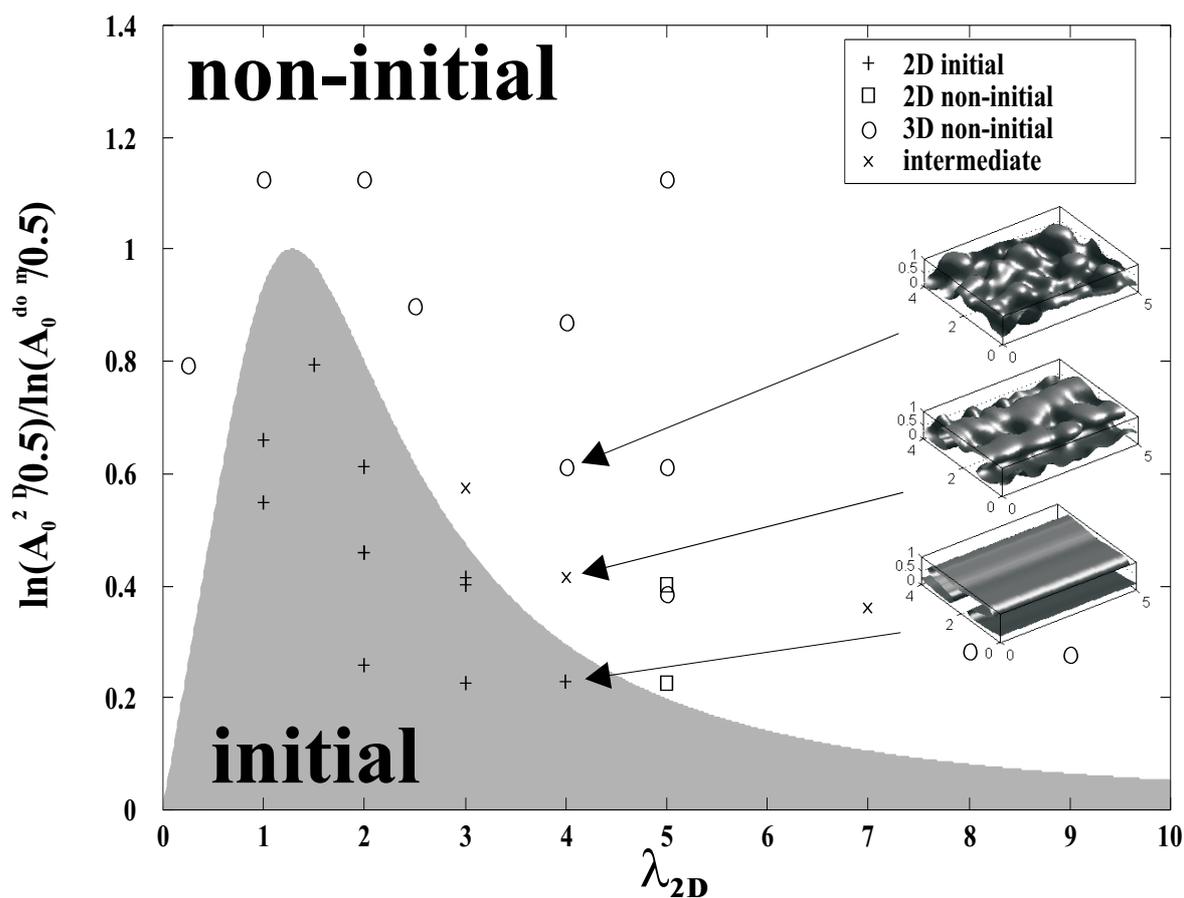


Figure 5.3: Phase diagram predicting the survival of an initial sinusoidal 2D perturbation over background noise depending on its initial amplitude (A_0), initial wavelength (λ_{2D}) and the initial amplitude of the largest dominant normal mode present in the noise (A_0^{dom}). Crosses show results of numerical simulations, where the initial perturbation survived, whereas open circles, x-marks and squares show results where the initial perturbation broke up into 2D, intermediate or 3D structures. Insets show the interface at the end of numerical simulations that resulted in 3D, intermediate and 2D structures respectively.

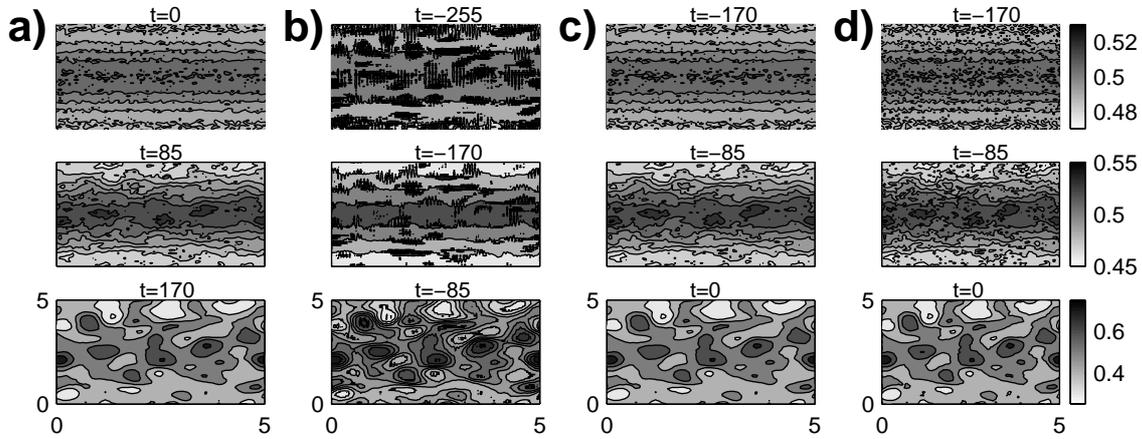


Figure 5.4: Reverse simulation of the 3D RT instability shown in Figure 2. (A) Contour plot of the interface height in the forward simulation at times $t=0$, 85 and 170 respectively. (B) Reverse simulation starting from $t=255$ in the forward simulation. (C) Reverse simulation starting from $t=170$ in the forward simulation. (D) Reverse simulation which started from $t=170$, but with added noise of maximum amplitude 5×10^{-3} .

5.5 Reverse modeling results

Inverse modeling is of major practical importance [Bennett, 1992, Marchuk, 1982] and reverse modeling of 3-D diapiric structures is of special interest for earth scientists (e.g. 3D restoration of salt domes). Reverse modeling was done, using the same numerical code as for forward simulations, but with negative timesteps. Four reverse simulations, started at different stages, recovered the 2-D initial perturbation (Fig. 5.5). Similar simulations showed that 1) reversing with a lower resolution ($64 \times 64 \times 257$) than that of the forward model gives approximately the same results, 2) the adding of lines to the interface, done in forward simulations, produces interpolation errors, which cumulate during reverse modeling and 3) these errors are even larger if overhangs are present. We thus speculate that restoration errors are partly due to cumulative growth of numerical errors and partly of physical origin. The physical origin of the noise effects can be explained by noting that the growth rate of the numerical errors during forward simulations is given by Eqn. 5.4. The amplitudes of the errors grow exponentially ($q > 0$ for all wavelengths), but slower than the true physical (dominant) modes. Conversely, during forward modeling of stably stratified fluids or reverse modeling of initial stages of the RT instability (when overhangs are not developed yet), the numerical errors

decay exponentially with a rate governed by eqn. 5.4, but $q < 0$. Therefore, it is to be expected that the reverse modeling of the RT instability is numerically more stable than the forward modeling (see also Fig. 5.5). However, development of overhangs drastically changes the situation. The overturned layers are stably stratified and their modeling is better posed for the forward simulations than for the reverse ones. Small perturbations at the lower side of an overhang, which are due to the remeshing of the interface, start amplifying during reverse simulations. Reversing of the RT instability is thus difficult if overhangs are developed and the success of recovering the initial conditions depends on the stage of overturn development. The most extreme case would be a fully developed overturn (complete stable stratification), from which reversing will no longer recover the initial perturbation.

5.6 Conclusions

We present fully dynamical numerical simulations of the 3D viscous RT instability. Forward simulations show that an initial 2D perturbation may decompose into 3D structures if the amplitude of background noise is high compared to the amplitude of the 2D perturbation. We quantified the 2D-3D transition, using linear stability theory (Eq. 5.5). Numerical simulations show good agreement with this prediction (Fig. 5.4). Reverse (backward) modeling of the RT instability is capable of restoring the initial 2D conditions from intensively deformed 3D structures. However, the accuracy of the reverse model deteriorates if overhangs are prominent in the 3D structures because the overhangs result in a stable configuration with little memory of the initial conditions.

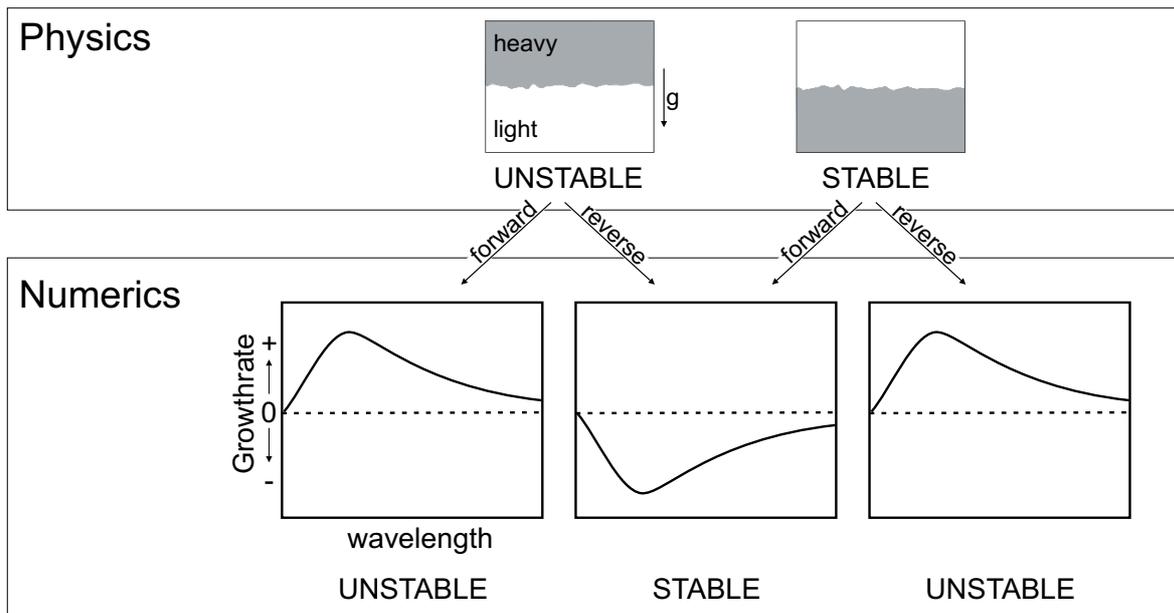


Figure 5.5: Cartoon explaining the numerics and physics of the RT instability. Starting with a physically unstable situation, two cases can occur numerically: a forward run (positive timestep) will result in the RT instability which has -in 2D- one wavelength that grows with a maximum, but finite, speed. This wavelength will be 'filtered' and survives during late stages. All growthrates are positive and thus numerically unstable. Reverse modeling, on the other hand, results in negative growthrates. Small perturbations decay exponentially fast and this case is numerically stable. The diffusion equation, on the other hand, does not have a maximum or minimum growthrate (infinitely small wavelengths decay infinitely fast during forward runs). Reverse modeling of the diffusion equation will *amplify* these small wavelengths infinitely fast and will thus lead to numerical blowup. Therefore the RT instability is numerically reversible but the diffusion equation is not.

Bibliography

- A.F. Bennett. *Inverse methods in physical oceanography*. Cambridge University Press, 1992.
- S. Chandrasekhar. *Hydrodynamic and hydromagnetic instabilities*. Oxford University Press, 1961.
- C.P. Conrad and P. Molnar. The growth of Rayleigh-Taylor-type instabilities in the lithosphere for various rheological and density structures. *Geophysical Journal International*, 129:95–112, 1997.
- Claude Godreche and P. Manneville, editors. *Hydrodynamics and Nonlinear Instabilities*. Hydrodynamics and Nonlinear Instabilities. Cambridge University Press, 1998.
- X. He, Zhang. R., C. H. Chen, and G.D. Doolen. On the three-dimensional Rayleigh-Taylor instability. *Physics of Fluids*, 11(5):1143–1152, 1999.
- G. Houseman and P. Molnar. Gravitational (Rayleigh-Taylor) instability of a layer with non-linear viscosity and convective thinning of continental lithosphere. *Geophysical Journal International*, 128:125–150, 1997.
- G.I. Marchuk. *Methods of numerical mathematics*. Springer Verlag, 1982.
- Y. Podladchikov, C. Talbot, and A. Poliakov. Numerical models of complex diapirs. *Tectonophysics*, 228:349–362, 1993.
- J. Pons, C. Oudin, and J. Valero. Kinematics of large syn-orogenic intrusions: example of the lower proterozoic saraya batholith (eastern senegal). *Geologische Rundschau*, 81(2):473–486, 1992.

- N.M. Ribe. Spouting and planform selection in the Rayleigh-Taylor instability of miscible viscous fluids. *Journal of Fluid Mechanics*, 377:27–45, 1998.
- S. M. Schmalholz and Y. Podladchikov. Buckling versus folding: Importance of viscoelasticity. *Geophysical Research Letters*, 26(17):2641–2644, 1999.
- H. Schmeling. On the relation between initial conditions and late stages of the Rayleigh-Taylor instability. *Tectonophysics*, 133:65–80, 1987.
- C. Talbot, P. Ronnlund, H. Schmeling, H.A. Koyi, and M.P.A. Jackson. Diapiric spoke patterns. *Tectonophysics*, 188:187–201, 1991.
- A. A. Ten, Y. Y. Podladchikov, D. A. Yuen, T. B. Larsen, and A. V. Malevsky. Comparison of mixing properties in convection with the particle-line method. *Geophysical Research Letters*, 25(16):3205–3208, 1998.
- D. L. Turcotte and G. Schubert. *Geodynamics*. Wiley, 1982.

Chapter 6

Transition from exponential to buoyancy-controlled diapirism.

Abstract

The transition from the initial to the finite amplitude stages of a 3D diapir are studied numerically. During initiation, the amplitude growth of the high/low density fluid interface versus time is exponential or slightly super-exponential, in agreement with (non-) linear stability theory. At some stage, however, growth becomes linear with time and is described by the Stokes equation for a sphere rising through a viscous media. A simple theory is proposed to link the initial stages to the later stages of diapir development. The difference between 2D and 3D perturbations is studied, which shows that 2D perturbations rise slower than 3D perturbations during the later stages. For a 3D perturbation to behave like a 2D structure, dominant mode aspect ratios in excess of 100 are required.

6.1 Introduction

Diapirism is a process relevant to several geodynamic processes, for example salt tectonics and mantle plumes. Therefore, a large number of experimental, numerical and analytical work has been devoted to study the instability of a low-density fluid underlying a higher-density fluid. These studies can be essentially divided into two groups: (1) those that concentrate on the initiation stage, which can be described by the linear or non-linear Rayleigh-Taylor instability theory [e.g. Chandrasekhar, 1961, Danes,

1964, Whitehead and Luther, 1975, Canright and Morris, 1993, Ribe, 1998], (2) those that concentrate on late stages, once a plume has been fully developed with a spherical, inflating head [e.g. Whitehead and Luther, 1975, van Keken, 1997]. Little work has been dedicated to the intermediate stages, i.e. linking initial with fully developed plume stages. An exception is the work of Bercovici and Kelly [1997], who employed a simplified lubrication theory valid for large viscosity contrasts.

The linear stability analysis predicts that the amplitude of a perturbation should increase exponentially with time. Nonlinear effects may even induce super-exponential (faster) growth [e.g. Ribe, 1998]. According to these theories, the velocity of the perturbation (derivative of amplitude versus time) should thus increase exponentially with time, which would result in supersonic speeds if scaled to the Earth's mantle. Whereas such a situation may occur if one of the fluids is inviscid, the finite viscosity of rocks and materials used in laboratory experiments causes a drag on the plume-head and limits its velocity. The falling or rising of spherical objects through viscous or power-law fluids is known to be mainly dependent on the viscosity of the surrounding fluid and the radius of the sphere [e.g. Stokes, 1851, Happel and Brenner, 1975, Weinberg and Podladchikov, 1994]: its velocity is constant and known as the Stokes velocity. The Stokes velocity is thus the upper limit for the velocity of an object rising through viscous fluids. It is however not clear how to link the initial (exponential) stages to the fully-developed stages. The goal of this chapter is to make this link. For the sake of simplicity, we concentrate on linear viscous rheologies.

6.2 Model and methods

We assume the fluids to be incompressible and slowly moving (negligible inertial terms). The rheology is linearly viscous. In this case the governing equations are

$$\frac{\partial v_i}{\partial x_j} = 0 \quad (6.1)$$

$$\frac{\partial \sigma_{ij}}{\partial x_j} = \rho g \vec{e}_z \quad (6.2)$$

$$\sigma_{ij} = -P\delta_{ij} + \mu \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \quad (6.3)$$

where v_i is velocity, x_j spatial coordinates, σ_{ij} are stresses, g the gravitational acceleration acting in the vertical direction, ρ is density and $P = -\frac{\sigma_{ii}}{3}$ is pressure. Our analysis is isothermal, 3D, and restricted to an isoviscous system with two fluids of different density. Time evolution of the interface $\Gamma(x, y)$ between the two fluids is computed according to:

$$\frac{\partial \Gamma(x, y)}{\partial t} + v_i \nabla \Gamma(x, y) = 0 \quad (6.4)$$

The numerical method is a finite-difference/spectral method [Kaus, 2000, Kaus and Podladchikov, 2001] which employs a Fourier spectral approximation in the horizontal direction and a conservative finite-difference method in the vertical direction. The interface is described by a triangular surface, which allows an accurate setting of initial perturbation and monitoring of the late-stage behavior. The method is essentially identical to that described in Schmalzl and Loddoch [2003]. An adaptive, implicit time stepping strategy is employed. The code reproduces linear-stability growth rates with an accuracy $> 0.5\%$ and has been verified versus Stokes flow. A no-slip ($v_z = v_x = v_y = 0$) lower boundary condition and a flat stress-free upper boundary ($\tau_{zz} = \tau_{xz} = \tau_{yz} = 0$) condition are employed. Non-dimensionalization is done by using H (height of computational domain) as characteristic length, μ_2 (viscosity of low density material) as characteristic viscosity, which gives $H^2 g(\rho_2 - \rho_1)/\mu_2$, $Hg(\rho_2 - \rho_1)$, $\mu_2/Hg(\rho_2 - \rho_1)$ as characteristic velocity, pressure and time, respectively.

6.3 Initial stages

The classical linear stability analysis assumes that the interface between two fluids of different density is horizontal and perturbed with a sinusoidal perturbation of infinitesimally small amplitude. A perturbation will grow in amplitude according to $A(t) = A_0 e^{qt}$, where q is the growth rate, which depends on factors like the wavelength of the 2-D perturbation, viscosity contrast, density difference, layer thickness etc. In a 2-D case one wavelength exists, called the dominant wavelength, for which the growth rate is maximum. Numerical simulations, starting from low-amplitude random noise, indicate that this wavelength will indeed be most pronounced during later stages. Figure 6.1 shows the dominant wavelength and growth rate for different viscosity contrasts and layer thicknesses. Three domains can be distinguished for the given parameters. Here we focus on thin layers, where two limiting cases can be

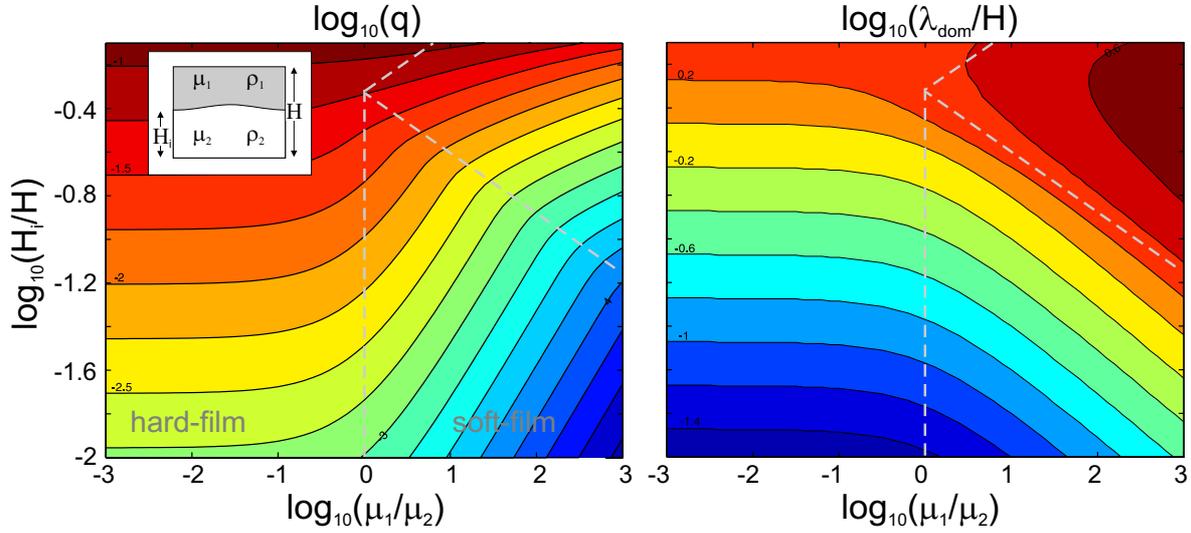


Figure 6.1: Dominant growthrate and dominant wavelength versus layer thickness and viscosity contrast for velocity-free lower boundary and a stress-free upper boundary.

distinguished. The hard-film limit (highly viscous thin layer) is characterized by the following dominant wavelength and growthrate dependencies:

$$q_{max} = 0.16 \frac{H(\rho_1 - \rho_2)g}{\mu_2} \frac{H_i}{H} \quad (6.5)$$

$$\frac{\lambda_{dom}}{H} = 2.95 \frac{H_i}{H} \quad (6.6)$$

The soft-film limit (thin, low viscosity layer) has asymptotic expressions

$$q_{max} = 0.14 \frac{H(\rho_1 - \rho_2)g}{\mu_2} \frac{H_i}{H} \left(\frac{\mu_1}{\mu_2} \right)^{-\frac{2}{3}} \quad (6.7)$$

$$\frac{\lambda_{dom}}{H} = 2.80 \frac{H_i}{H} \left(\frac{\mu_1}{\mu_2} \right)^{\frac{1}{3}} \quad (6.8)$$

These dominant wavelength expression have been derived for a two-dimensional setup with a perturbation of the form $\Gamma = H_i + A_0 \cos(2\pi/\lambda_x x)$. It has been shown [see e.g. Biot, 1966, Ribe, 1998] that the dominant growthrate is identical in a 3D setup with a perturbation of the form $\Gamma = H_i + A_0 \cos(2\pi/\lambda_x x) \cos(2\pi/\lambda_y y)$ as long as no background compression or extension is exerted. Instead of one dominant wavelength, however, the 3D case has an infinite number of dominant modes. These modes can be

computed from the 2D dominant wavelength as

$$\left(\frac{2\pi}{\lambda_{dom}}\right)^2 = \left(\frac{2\pi}{\lambda_x}\right)^2 + \left(\frac{2\pi}{\lambda_y}\right)^2 \quad (6.9)$$

Nonlinear perturbation analysis [Whitehead and Luther, 1975, Ribe, 1998] and numerical simulations [Kaus and Podladchikov, 2001], show that the infinite number of dominant wavelengths are a consequence of the assumption of infinitesimally small perturbations. In fact, purely 3D modes (with $\lambda_x = \lambda_y$) amplify slightly faster than modes with large aspect ratios.

6.4 Finite amplitude effects

Figure 6.2 shows the evolution of the maximum amplitude of the interface versus time. Thin layers undergo a transition from initially exponential growth to linear growth during the later stages. This linear growth, when velocity reaches a maximum, suggests that late stages may be described by the Stokes equation, which is given by [e.g. Stokes, 1851, Happel and Brenner, 1975, Weinberg and Podladchikov, 1994]:

$$V_{stokes} = \frac{1}{3} \frac{(\rho_1 - \rho_2)gR^2}{\mu_2} \left(\frac{\mu_2 + \mu_1}{\mu_2 + \frac{3}{2}\mu_1} \right) \quad (6.10)$$

where R is the radius of the sphere. However a plume is not a perfect sphere, at least for the constant viscosity case shown on figure 6.2. Thus what is the appropriate radius R matching the observed velocity? Two different methods are available to estimate R . The first method one consists in matching the observed velocity with V_{stokes} . The second method is based on a volume argument. If the initial sinusoidal perturbation has the form $\Gamma(x, y) = H_i + A_0 \cos(\frac{2\pi}{\lambda_x}x) \cos(\frac{2\pi}{\lambda_y}y)$, the volume of material that will form the head and tail of the plume is roughly given by $V_{geom} = \lambda_y \lambda_x H_i$. Under the assumption that all this material enters a spherical plume-head, its radius can be estimated by conservation of mass

$$R_{geom} = \left(\frac{3\lambda_y \lambda_x H_i}{4\pi H^3} \right)^{\frac{1}{3}} \quad (6.11)$$

Since 'real' plumes do not always have a perfectly spherical shape, there is a difference between the geometrically obtained radius, R_{geom} and the radius obtained from velocity computations:

$$R = CR_{geom} \quad (6.12)$$

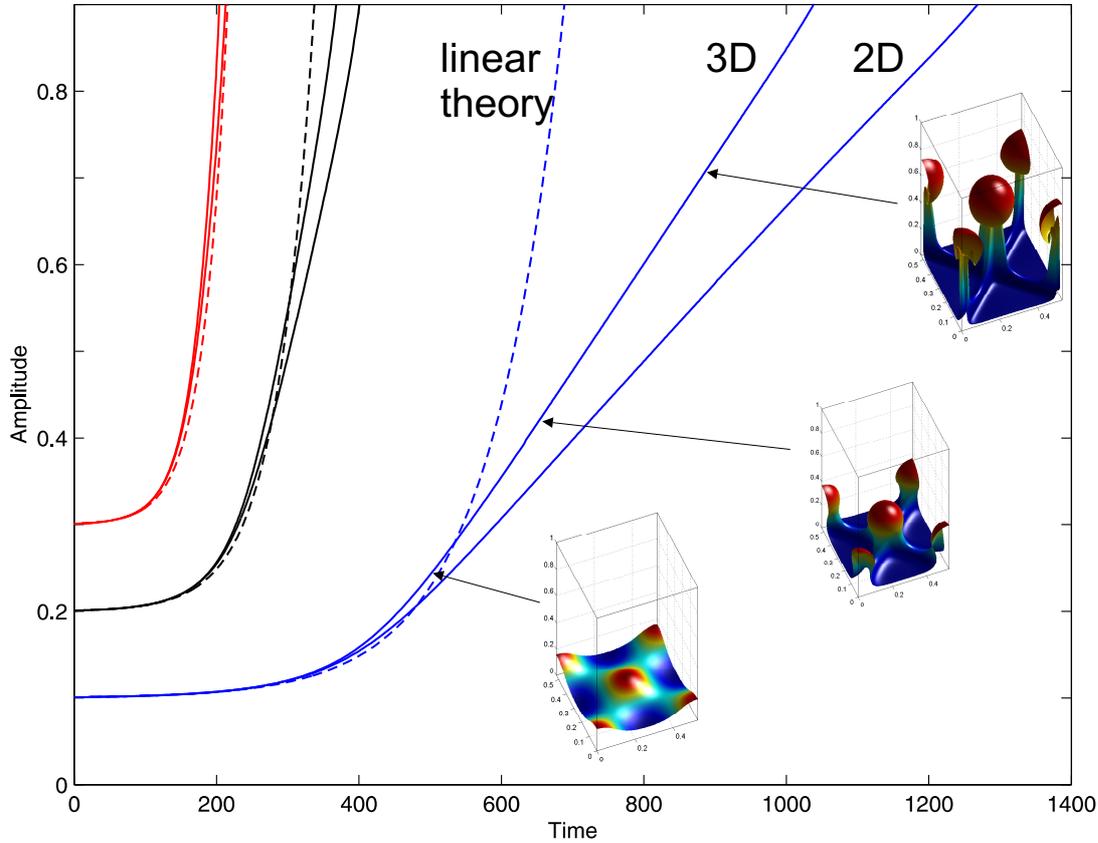


Figure 6.2: Late stage behavior of the amplitude versus time. During the initial stages, growth is exponential. Thin layers, however, exhibit linear growth during later stages. 2D perturbations amplify slower than 3D perturbations in all cases.

here C is a correction factor, which depends on viscosity contrast, shape of the diapir, layer thickness etc. We have numerically determined the correction factor C , and found that for the given boundary conditions and isoviscous material properties, $C \simeq 0.356$. This value is appropriate for 3D dominant modes of aspect ratio $\lambda_x/\lambda_y \leq 10$, and for thin layers of maximum thickness $H_i \leq 0.2H$. Thicker layers do not exhibit a clear transition from exponential to linear growth (e.g. fig. 6.2). This fact allows us to extend the linear stability theory by taking the finite, late-stage velocity into account. The increase of amplitude versus time is given by

$$A(t) = A_0 + \int_0^t V(t)dt \quad (6.13)$$

The velocity is the minimum of the Stokes velocity and the velocity predicted by linear

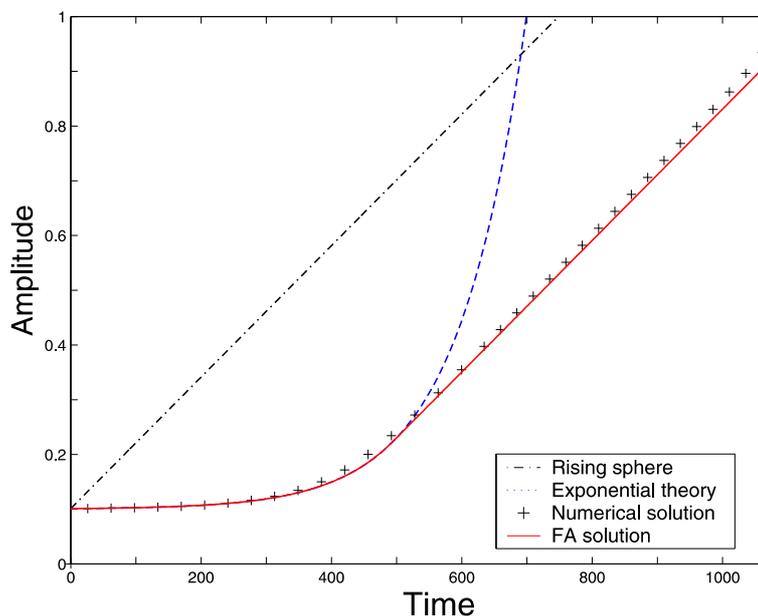


Figure 6.3: Comparison of the finite amplitude formulation with exponential theory and 3D numerical solution (with $\lambda_x = \lambda_y = 0.53H$) with a no-slip lower boundary and no-stress upper boundary condition.

stability theory:

$$V(t) = \min (V_{exp}(t), V_{stokes}) \quad (6.14)$$

here $V_{exp}(t)$ is the velocity obtained by the exponential linear stability theory

$$V_{exp}(t) = A_0 q e^{qt} \quad (6.15)$$

and V_{stokes} is computed by combining eqs. (6.10),(6.11) and (6.12). A comparison between the simple finite amplitude theory and numerical results shows good agreement (fig. 6.3). A similar agreement exists for dominant modes with an aspect ratio of up to 10.

6.5 2D vs 3D

The simple theory outlined above is only valid for 3D structures. Two-dimensional structures of the form $\Gamma = H_i + A_0 \cos(2\pi/\lambda_x x)$ rise slower than 3D ones, which may

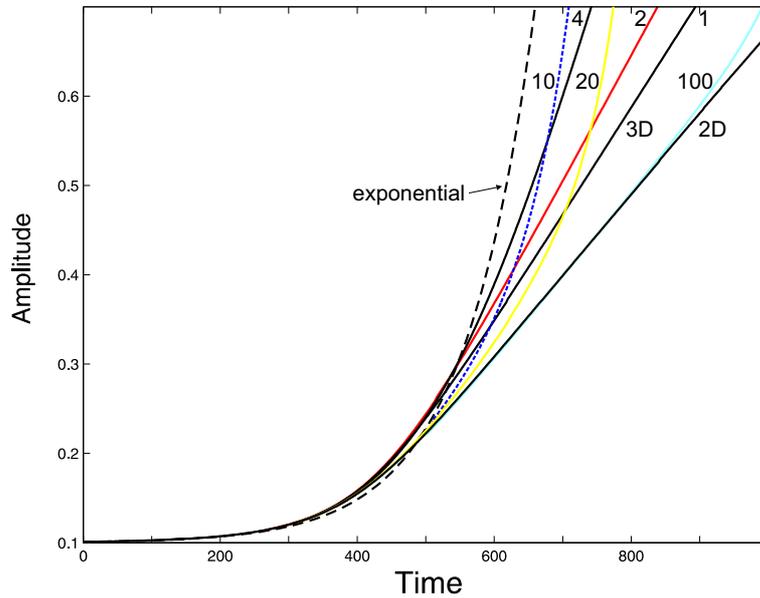


Figure 6.4: Effect of aspect ratio (defined as $R_a = \lambda_x/\lambda_y$) of initial sinusoidal perturbation growth of amplitude. 3D perturbations with aspect ratio $1 \leq R_a \leq 10$ amplify faster than 3D perturbations with $R_a = 1$. 2D behavior requires $R_a \geq 100$. $\mu_1/\mu_2 = 1$.

be attributed to the fact that cylinders have a larger surface area than spheres (fig. 6.2). But when can a 3D structure effectively be regarded as a 2D one in terms of its rising velocity? In order to understand this, a number of 3D simulations have been formed with dominant modes of increasing aspect ratio $R_a = \lambda_y/\lambda_x$. Results show that very large aspect-ratios ($R_a > 100$) are required to approximate 2D behavior (fig. 6.4).

Unfortunately, a simple equation of the form (6.10) does not exist for the rising velocity of cylinders in low-Reynolds number fluids (this is known as the Stokes paradox, see e.g. Happel and Brenner [1975]), precluding the derivation of a simple 2D finite amplitude theory.

6.6 Discussion and conclusions

Three-dimensional numerical simulations show that the finite amplitude behavior of diapirs rising from a thin source layer deviates significantly from that predicted by the linear stability theory. Rather than increasing exponentially with time, the velocity of

the diapirs becomes approximately constant after a finite amount of time. A simple theory combines the Stokes velocity and the exponential theory, and fits the numerical data with an acceptable accuracy. Numerical simulations also show that purely 2D (linear) modes amplify slower than 3D modes. Only 3D modes with aspect ratios $\lambda_x/\lambda_y > 100$ will behave as a 2D mode. This result indicates that 2D simulations may underestimate the rising velocity of plumes, or the rate of lithospheric delamination.

An obvious limitation of the present work is its restriction to isoviscous flows. Previous workers, concentrating on thin, highly viscous, 2D layers did not observe a slowing-down of velocity and rather predicted an over-exponential growth which ultimately resulted in an amplitude 'blow-up' [e.g. Houseman and Molnar, 1997]. Bercovici and Kelly [1997] studied the initiation of a diapir from a thin, low-viscosity layer. They demonstrated, using simplified analytical theory and laboratory experiments, that a 'stalling'-period exist between the exponential and the Stokes-rising stages, which has not been observed here for isoviscous cases. None of the mentioned studies compared the transition from 2D to 3D modes.

More work is obviously necessary to better understand the effect of viscosity contrasts on the breakdown of exponential growth.

Bibliography

- D. Bercovici and A. Kelly. The non-linear initiation of diapirs and plume heads. *Physics of the Earth and Planetary Interiors*, 101:119–130, 1997.
- M.A. Biot. Three-dimensional gravity instability derived from two-dimensional solutions. *Geophysics*, 31(153-166), 1966.
- D. Canright and S. Morris. Buoyant instability of a viscous film over a passive fluid. *Journal of Fluid Mechanics*, 255:349–372, 1993.
- S. Chandrasekhar. *Hydrodynamic and hydromagnetic instabilities*. Oxford University Press, 1961.
- Danes. Mathematical formulation of salt-dome dynamics. *Geophysics*, 29(3):414–424, 1964.
- J. Happel and H. Brenner. *Low Reynolds number hydrodynamics*. Noordhoff international publishing, Leyden, 1975.
- G. Houseman and P. Molnar. Gravitational (Rayleigh-Taylor) instability of a layer with non-linear viscosity and convective thinning of continental lithosphere. *Geophysical Journal International*, 128:125–150, 1997.
- B. J. P. Kaus and Y. Y. Podladchikov. Forward and reverse modeling of the three-dimensional viscous Rayleigh-Taylor instability. *Geophysical Research Letters*, 28(6): 1095–1098, 2001.
- B.J.P. Kaus. *Forward and reverse modeling of the three-dimensional viscous Rayleigh-Taylor instability*. Master thesis, ETH Zürich, 2000.

- N.M. Ribe. Spouting and planform selection in the Rayleigh-Taylor instability of miscible viscous fluids. *Journal of Fluid Mechanics*, 377:27–45, 1998.
- Joerg Schmalzl and Alexander Loddoch. Using subdivision surfaces and adaptive surface simplification algorithms for modeling chemical heterogeneities in geophysical flows. *Geochemistry Geophysics Geosystems*, 4(9):DOI 10.1029/2003GC000578, 2003.
- G.G. Stokes. On the effect of internal friction of fluids on the motion of pendulums. *Cambridge Philosophical Transactions*, 9:8–106, 1851.
- P. E. van Keken. Evolution of starting mantle plumes: a comparison between laboratory and numerical models. *Earth and Planetary Science Letters*, 148:1–14, 1997.
- R. F. Weinberg and Y. Podladchikov. Diapiric ascent of magmas through power-law crust and mantle. *Journal of Geophysical Research-Solid Earth*, 99(B5):9543–9559, 1994.
- J.A. Whitehead and D.S. Luther. Dynamics of laboratory diapir and plume models. *Journal of Geophysical Research*, 91:705–717, 1975.

Chapter 7

2-D and 3-D pattern formation of the viscous Rayleigh-Taylor instability: a numerical study

Abstract The Rayleigh-Taylor (RT) instability arises when a high density fluid overlies a fluid of lower density. This density-driven instability may be significant for a set of geological processes ranging from lithospheric delamination to pluton emplacement and the formation of load-cast structures. This work focuses on 2-fluid models. We study the effects of boundary conditions (no-slip, free-slip, no-stress, fast-erosion) on the geometries that develop in two dimensions. Whereas no-slip, free-slip and no-stress boundary conditions result in similarly looking, mushroom-shaped structures, a fast-erosion boundary condition results in chimney-like structures. We also study the patterns and structures that form in 3D settings. In this case, a linear stability analysis predicts an infinite number of dominant modes that amplify with nearly identical rates. Linear superposition of these dominant modes results in complex-looking patterns. A fast-erosion upper boundary condition triggers the formation of finger-like structures, which in map-view, range from circular, to curved and elongated. If an initial 2D perturbation is imposed at the interface between the two boundaries, 2D wall-like structures may form. We explore the transition from wall-like to finger-like structures, and show that an intermediate mode exist of aligned 3D diapirs. All 3D modes grow with a rate predicted by linear stability analysis.

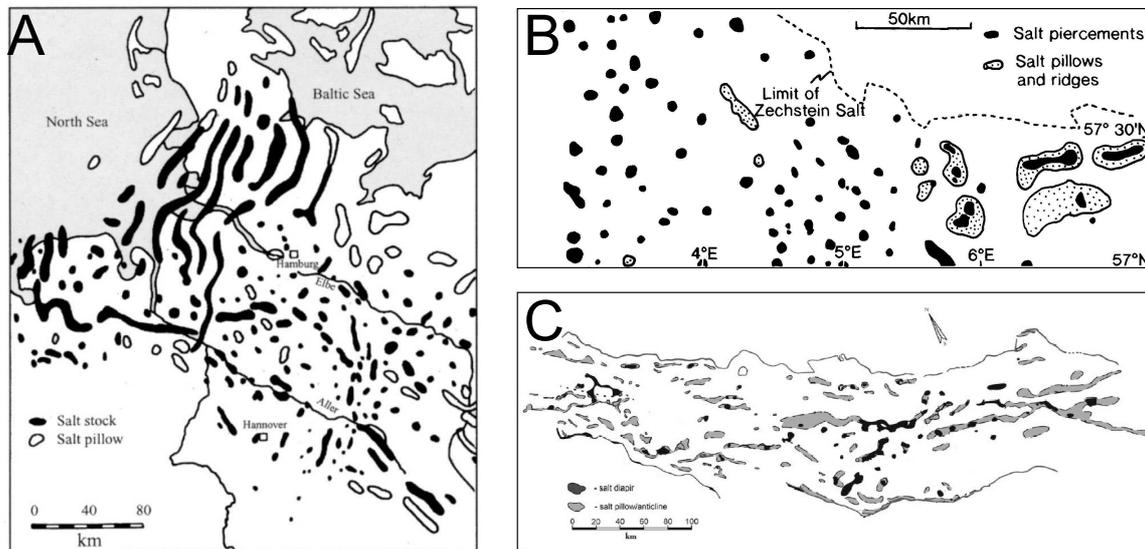


Figure 7.2: Salt domes patterns in A) The North Sea Zechstein basin [after Trusheim, 1960], B) Norwegian-Danish basin [from Hospers et al., 1988]. C) Dniepr-Donets basin [after Stovba and Stephenson, 2003].

little theoretical work has been carried out on diapirism with a brittle visco-elasto-plastic overburden. We will however discuss the applicability of our results in the light of previous numerical work that employed a brittle overburden.

Previous numerical models of the RT-instability have mainly been two-dimensional [e.g. Woitdt, 1978, Schmeling, 1987, Romer and Neugebauer, 1991, Weinberg and Schmeling, 1992, Zaleski and Julien, 1992, Poliakov and Podladchikov, 1992, Podladchikov et al., 1993] with various complexities like erosion and sedimentation [Poliakov et al., 1993b, Podladchikov et al., 1993], nonlinear rheologies [van Keken et al., 1993], brittle overburden [Poliakov et al., 1993a, Daudré and Cloething, 1994, Poliakov et al., 1996], differential loading [Gemmer et al., 2004], and dynamic restoration [Ismail-Zadeh et al., 2001]. Only few workers [Kaus and Podladchikov, 2001, Ismail-Zadeh et al., 2004] have studied the three-dimensional, viscous RT instability. Kaus and Podladchikov [2001] have shown that relatively complex three-dimensional patterns may develop even if an initially smooth two-dimensional perturbation, such as one caused by folding of the interface between salt and overburden, is present. Ismail-Zadeh et al. [2004] showed examples of how wall-like and mushroom-shaped diapiric structures may develop from

various initial conditions. Their results, however, give little quantitative insight in why these specific patterns develop. Three-dimensional pattern formation of the RT instability has been studied by means of analogue experiments [e.g. Talbot et al., 1991]. They observed spoke-like polygonal patterns and also pointed out the difficulty in obtaining wall-like structures without strong 2D perturbations.

The results of this work are generally applicable to the Rayleigh-Taylor instability, which has also been applied to mantle plumes [Whitehead and Luther, 1975], layered intrusions [Gerya et al., 2003], gneiss domes [Fletcher, 1972], plutons [Miller and Paterson, 1999, Dietl and Koyi, 2002], loading structures [Anketell et al., 1970], lithospheric delamination [e.g. Houseman and Molnar, 1997] and lower crustal convection [Gerya et al., 2000, Jull and Kelemen, 2001].

7.2 Model and numerical methods

Several numerical codes are used in this work. All assume the materials to be incompressible:

$$\frac{\partial v_i}{\partial x_i} = 0 \quad (7.1)$$

where v_i denotes velocity and x_i the spatial coordinates. Inertial terms are negligible on geological timescales. Therefore, force balance is given by:

$$\frac{\partial \sigma_{ij}}{\partial x_j} = \rho g \vec{e}_z \quad (7.2)$$

where σ_{ij} is stress, ρ density, g gravitational acceleration and \vec{e}_z the unit vector pointing in the vertical direction. In this work, we further assume a linear viscous rheology:

$$\sigma_{ij} = -P\delta_{ij} + 2\mu\dot{\epsilon}_{ij} \quad (7.3)$$

where $P = -\bar{\sigma}$, $\dot{\epsilon}_{ij} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right)$, μ is viscosity and the Einstein notation is used throughout. The model domain consists of a 2-D or 3-D box of initial height H which is filled with two fluids having a density difference $\Delta\rho$. With μ_1 being the viscosity of the lower, low-density fluid, non-dimensionalization was done taking H , $H^2g\Delta\rho/\mu_1$, $Hg\Delta\rho$ and $\mu_1/(Hg\Delta\rho)$ as characteristic scales for length, velocity, stress and time, respectively.

Two-dimensional computations are performed with a newly-developed hybrid finite-difference/spectral code (Gango), which solves the governing equations on an Eulerian computational grid [Kaus et al., 2004]. At each timestep, the related density and viscosity structures are computed from the markerlines. Timestepping is done with an implicit method. The code has been extensively tested from 0D rheology tests, to stress distributions around inclusions [e.g. Schmid and Podladchikov, 2003] and linear and nonlinear stages of folding in viscous and viscoelastic materials. Relevant to the present work are comparison with the linear stability theory of the RT-instability (see fig. 7.3) and comparisons with results obtained with a recently developed two-dimensional finite element code (SloMo).

The three-dimensional code (Diapir3D) also employs a spectral/finite-difference primitive variables technique, with the difference that the finite difference scheme is solved using only v_z , rather than v_x and σ_{zz} as unknowns. Horizontal directions are approximated by a 2D Fourier expansion. Since viscosity varies only in the vertical direction, the solution of the system of equations can be performed wavenumber-by-wavenumber and the numerical method has a high performance. The interface between the two fluids is represented by a triangular surface. At each timestep-iteration, the 3D density field is recomposed from this surface and the known location of the two fluids [see Schmalzl and Loddoch, 2003, for a more thorough description of the method]. Velocity is computed for the given density distribution, and this velocity is used to advect the surface. An implicit timestep algorithm is employed to accurately track the initiation of the instability. The advantage of this method is that the interface between the fluids always remains sharp. A disadvantage is that, due to numerical errors, the amount of mass does not remain exactly constant throughout a simulation. The maximum change in mass is less than 0.5% during a typical model run, and therefore we assume the method to be sufficiently accurate. Other widely employed methods in geodynamics, like tracking of tracers [e.g. Tackley and King, 2003], or characteristic-based methods [e.g. Ismail-Zadeh et al., 2004] also suffer from this problem. Furthermore, the markerpoint method is expensive in 3D and the characteristic-based method allows a less accurate determination of the interface deformation. Resolution tests have been performed, that pointed out an additional advantage of the spectral methods used in this work. Even at surprisingly low resolutions (e.g. $51 \times 32 \times 32$), the main features of the RT instability are caught. Increasing resolution has an almost negligible effect on

the spacing of the diapirs. Accordingly, most runs presented here have been performed for a resolution of $101 \times 128 \times 128$.

7.3 2D simulations

In order to get a better understanding of the three-dimensional behavior of the RT-instability, we will first give an overview of 2D numerical simulations and linear stability theory. The results partly reproduce previous work [e.g. Woidt, 1978, Schmeling, 1987, Poliakov et al., 1993a] but the comparison between fast-erosion and free-surface boundary conditions is new.

7.3.1 Linear stability analysis

A perturbation analysis is the classical way to study the Rayleigh-Taylor instability [e.g. Chandrasekhar, 1961, Danes, 1964, Ribe, 1998]. The interface between the low-density and the high-density fluids is assumed to be initially horizontal but with sinusoidal perturbations of infinitesimally small amplitude and wavelength λ . The analysis shows that the amplification rate (called growthrate) of such a sinusoidal perturbation depends on its wavelength. For a two-dimensional setup, the maximum growthrate corresponds to the dominant wavelength. This dominant wavelength and the corresponding growthrate are a function of parameters such as the viscosity contrast, the relative thicknesses of upper and lower layers, and the surface and bottom boundary conditions (e.g. no-slip or stress-free). Expressions for the 2D dominant growthrate as a function of wavelength are given in the Appendix for various boundary conditions, arbitrary viscosity contrasts and layer thicknesses.

Plots of the growthrate as a function of non-dimensional wavelength for different surface boundary conditions, different viscosity contrasts and different relative layer thicknesses are shown in figure 7.3. In general, a no-stress (or fast-erosion) upper boundary condition results in a faster growth of the instability, than in cases with free-slip or no-slip boundary conditions. Moreover, the no-stress case has a wider range of dominant wavelengths. Good agreement exists between numerically computed growthrates and the rates predicted by the linear stability theory (fig. 7.3).

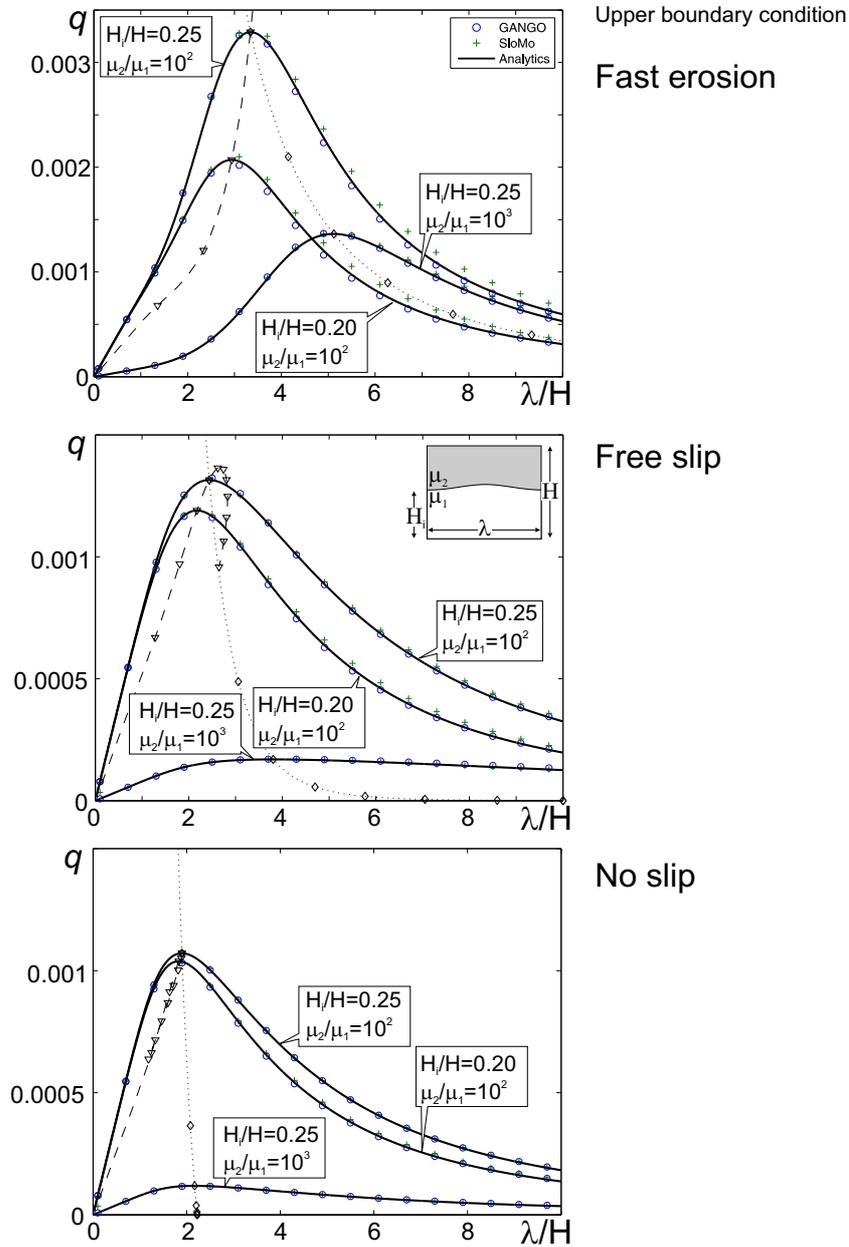


Figure 7.3: Growthrate q versus wavelength λ for a no-slip lower boundary condition and fast-erosion, free-slip and no-slip upper boundary condition. Analytical results are shown as well as numerical verification with a finite-difference/spectral method (GANGO) and with a finite element code (SloMo). Dashed lines indicate dominant wavelengths and growthrates for various thickness ratios at constant viscosity ratio 100. Dotted line indicates the dominant wavelengths and growthrates for various viscosities for constant thickness ratio=0.25. Results with a no-stress upper boundary condition are identical to the ones with fast erosion. Numerical computations are performed with an initial amplitude of the perturbations of $10^{-6}H$.

7.3.2 Initial random perturbations

The linear stability theory is only valid for initially infinitesimally small perturbations. Nonlinear effects may play a role during evolved stages, which significantly complicates the mathematical analysis. Therefore, numerical methods are employed to study the mature stages of the RT-instability. A series of numerical experiments have been performed starting from an initially horizontal interface perturbed with random noise. The lower boundary has a no-slip condition whereas the upper boundary has been varied with no-slip, free-slip, fast-erosion and free surface conditions. In each case, the same initial random noise has been imposed. The linear stability theory predicts diapirs to form with a characteristic spacing, which is in agreement with numerical results (fig. 7.4). The free-slip and the no-slip simulations produce similar spacing although the instabilities grow slightly faster in the free-slip case. The fast erosion-case develops a larger spacing than the free-slip and no-slip cases. Moreover, erosion considerably speeds up the formation of the RT-instability (by almost a factor two). The free-surface case initially behaves like the fast-erosion case, i.e. with similar spacing, but ultimately slows down when the upper boundary influences the raising instabilities. Specially noteworthy is the late-stage geometry in the case of a fast-erosion upper boundary, which is chimney-like rather than mushroom-like.

Further insight in the development of the diapirs, as well as on the applicability of linear stability analysis during evolved stages, can be obtained by analyzing the interfacial perturbations with a Fourier transformation [see also Schmeling, 1987]. Such an analysis reveals that wavelengths close to the dominant wavelength do amplify faster than other wavelengths (fig. 7.5), in agreement with linear stability theory. Ultimately, however, all simulations slow down and grow sub-exponentially. Note that the absolute dominant wavelength ($= 2.95$) in the case of a fast-erosion upper boundary does not develop. This is due to the box width of $10H$, which allows only an integer number of diapirs to develop (either three or four for the given setup). Non-dominant initial perturbations exhibit periods of over-exponential growth probably caused by interference [Schmeling, 1987], but owing to their small amplitude, they do not significantly influence the late-stage geometry and the spacing.

The 2-D simulations reiterate the importance of erosion during the formation and growth of diapirs [see also e.g. Poliakov et al., 1993a, Podladchikov et al., 1993]. It is

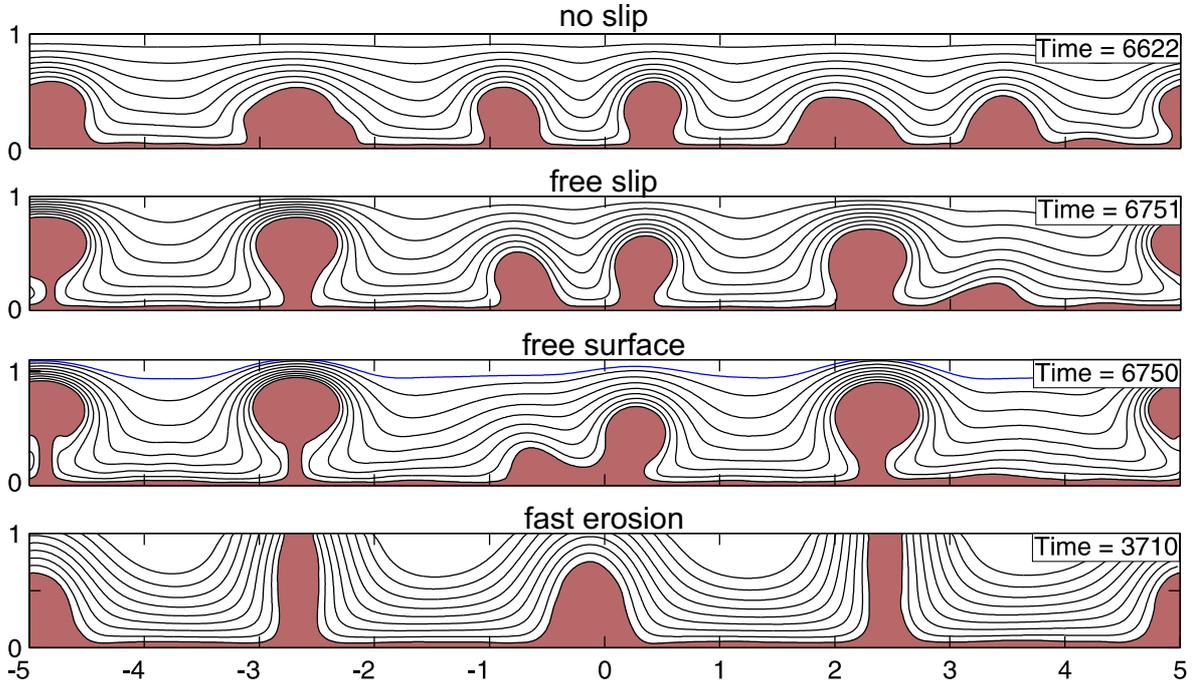


Figure 7.4: Geometries that develop from an initially horizontal interface at height 0.2, which has been perturbed with random noise of maximum amplitude 10^{-2} . A no-slip lower boundary condition is employed; the upper boundary condition has been varied. The viscosity contrast between upper and lower layer is 100 and the density difference 1. The same initial random noise distribution has been used in all simulations. Black lines represent initially horizontal passive markerlines.

important to know which of the above-mentioned boundary conditions is most applicable to natural salt systems. Since the Earth's surface is free and simulations with a free-slip or no-slip have significantly different dynamics, we argue that both the free-surface and the fast-erosion conditions are applicable. Seismic records and borehole information [Davison et al., 2000b] give evidence that buoyant doming of salt creates sufficient topographic relief to allow downslope sliding of unlithified sediments and lithified chalk slabs. The topographic relief above these domes through most of their history is tens to hundreds of meters [Davison et al., 2000b], which creates a negative force of $F_{topo} = (\rho_{wat} - \rho_{sed})gA_{topo}$, where ρ_{wat} denotes water density, ρ_{sed} sediment density and A_{topo} the topographic relief. Simultaneously, the salt diapir has a positive buoyancy force: $F_{salt} = (\rho_{sed} - \rho_{salt})gA_{salt}$, where A_{salt} is the amplitude of the salt dome. The approximation of infinitely fast erosion is justified if $|F_{topo}| \ll |F_{salt}|$.

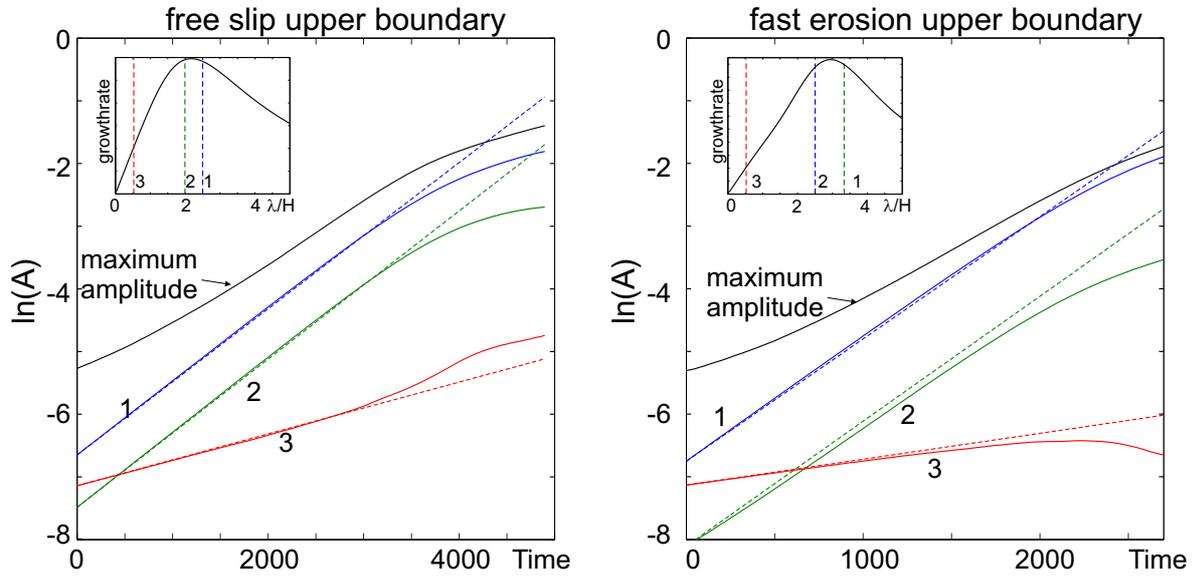


Figure 7.5: Time evolution of three selected Fourier coefficients of the interface for the free-slip and the fast-erosion simulations of Fig. 7.4. Numbers 1 and 2 refer to the amplitude of the two most pronounced wavelengths at $t \approx 5000$. Number 3 refers to the second most-pronounced wavelength at $t = 0$ (by chance the most pronounced wavelength at $t = 0$ is number 1). Dashed (straight) lines are the time-amplitude evolution as predicted by the linear stability theory. The inset shows the growthrate versus wavelength predicted by the linear stability theory, together with wavelengths 1-3. Wavelengths with the largest growthrate during the linear stage dominate the flow.

Taking $\rho_{sed} = 2650 \text{ kgm}^{-3}$, $\rho_{salt} = 2250 \text{ kgm}^{-3}$, $\rho_{wat} = 1000 \text{ kgm}^{-3}$, $A_{topo} = 10 - 200$ m, $A_{salt} = 3000 - 8000$ m, allows estimating $|F_{topo}/F_{salt}| = 0.004 - 0.2$. Thus the fast erosion boundary condition, with $|F_{topo}/F_{salt}| = 0$, is a reasonable end-member model for many natural cases.

7.3.3 Initial step-like perturbation

If the interface has a strong initial perturbation, this perturbation may dominate during the later stages. Schmeling [1987] showed in 2D and Kaus and Podladchikov [2001] in 3D that 2D *sinusoidal* perturbations (for example caused by buckling) may survive during later stages if their growthrate and amplitude is sufficiently large compared to

the dominant growthrate. This is expressed by [Kaus and Podladchikov, 2001]:

$$\frac{q_{initial}}{q_{dom}} > \frac{\ln(A_{init}^{2D}/(1-H_i))}{\ln(A_{dom}^0/(1-H_i))} \quad (7.4)$$

where it has been assumed that linear stability theory is valid throughout, A_{dom}^0 is the initial amplitude of the dominant wavelength ('hidden' in background noise), q_{dom} the dominant growthrate for the given setup and $q_{initial}$ the growthrate of the initial non-dominant perturbation.

Eq. (7.4) is no longer valid if the initial perturbation is step-like. This is because a step-like perturbation (for example caused by brittle faulting of the overburden) behaves like 'red noise'; in other words: the whole Fourier-spectra is present rather than only one wavelength. Wavelengths close to the dominant wavelength will thus be selected and dominate the late-stage geometry. Numerical simulations (fig. 7.6) demonstrate that geometries with a dominant wavelength form preferentially close to the initial step, even if the amplitude of this step is comparable to the amplitude of the background noise. In addition, there is an over-exponential growth stronger than for simulations without a step-like initial perturbation.

7.4 3D simulations

The 2D case-studies give insight in the nature of the Rayleigh-Taylor instability. This instability is however inherently 3D, and a detailed understanding of the similarities and differences between 2D and 3D cases is essential, especially when applying the model to case studies. Results of 3D numerical simulations and stability analysis are now presented.

7.4.1 Linear and nonlinear stability

Linear stability analysis in 3D is derivable from the results of the 2D analysis [Biot, 1966, see also appendix]. Perturbations on the interface of the form $\Gamma(x, y) = H_i + A_0 \cos(2\pi x/\lambda_x) \cos(2\pi y/\lambda_y)$ have a growthrate depending on both λ_x and λ_y . Different than in the 2D case, where only a single dominant wavelength exist, linear stability analysis shows that the 3D case exhibits an infinite number of dominant modes. These modes range from 2D ($\lambda_x = \infty$ or $\lambda_y = \infty$) to purely 3D ($\lambda_x = \lambda_y$).

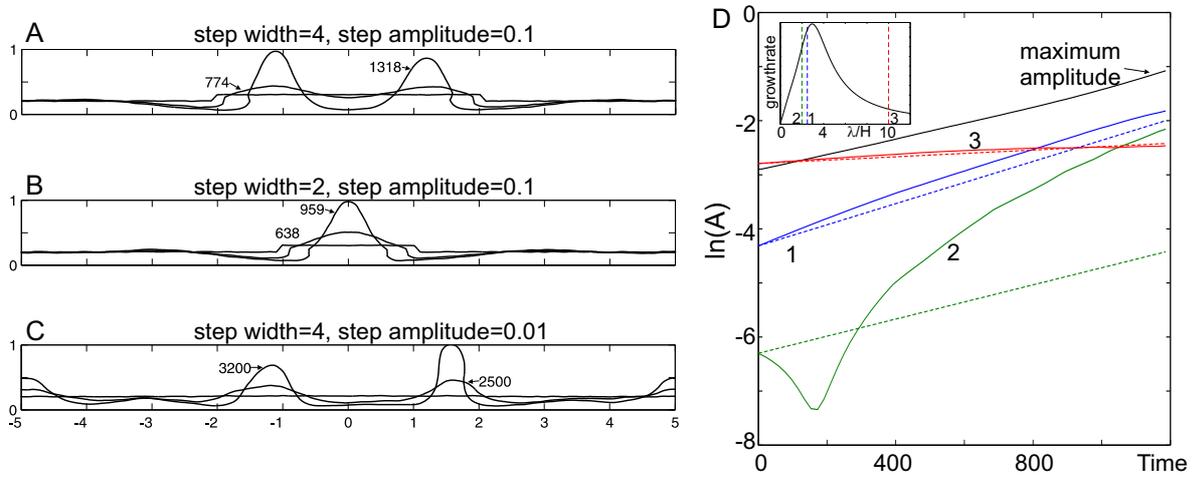


Figure 7.6: Effect of an initial step-like perturbation on late-stage geometry. The initial interface is perturbed with the same noise distribution as on Fig. 7.4, but additionally a step-like perturbation of variable width and height is added. D) Time-evolution of three Fourier components and maximum amplitude of the simulation shown in A. The fact that the amplitude of the initial step is larger than the maximum amplitude (red curve at $t = 0$), is an artefact of Fourier-analysis of step-like perturbations.

A nonlinear stability analysis by Ribe [1998], however, shows that purely 3D modes (in the form of squares or hexagons) grow slightly faster than 2D modes, the difference increasing with increasing amplitude. This analysis has been confirmed numerically for equal layer thicknesses and no-slip boundary conditions [Kaus and Podladchikov, 2001]. Results for different boundary conditions and unequal layer thicknesses are shown on figure 7.7. 3D modes grow slightly faster for finite amplitude perturbations. Also noteworthy is a shift in the absolute dominant growthrate for initial perturbation amplitudes which are large compared to the initial layer thickness.

A more detailed comparison of the time-evolution of 2D rolls and 3D perturbations is given in fig. 7.9. In all cases, 3D modes amplify faster than 2D modes. Good agreement exists between the linear stability theory and numerical results for initial layer thicknesses larger than 0.3. For thinner layers, however, the linear stability theory overestimates the amplitude growth versus time. This is related to the fact that a transition from exponential- to buoyancy-controlled growth occurs, which is studied in more details in chapter 6. The difference between 2D and 3D modes is now very pronounced.

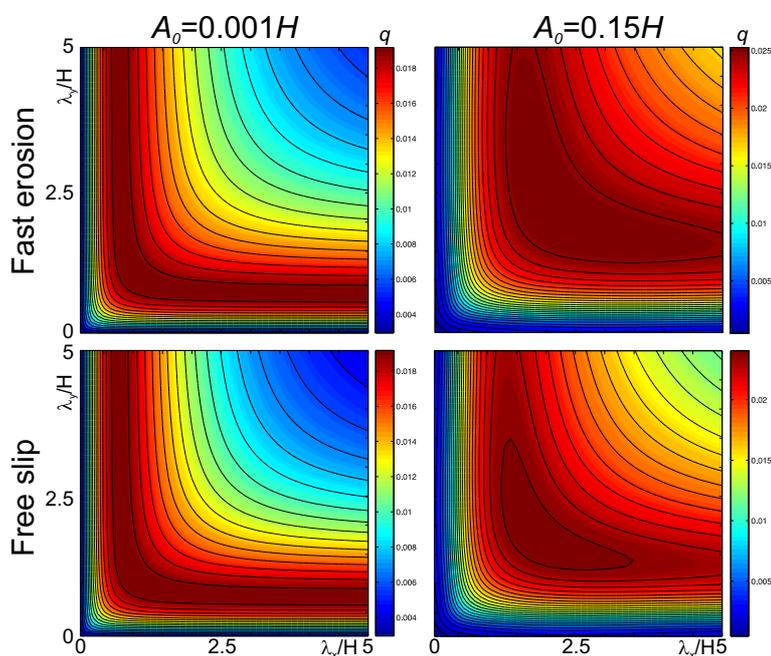


Figure 7.7: Numerically computed growthrates for different sinusoidal perturbations of the form $z/H = H_i + A_0 \cos(2\pi/\lambda_x x) \cos(2\pi/\lambda_y y)$, where A_0 and the upper boundary condition have been varied (other parameters are: $\mu_2/\mu_1 = 1$, $H_i = 0.2$, $\Delta\rho = 1$, no-slip lower boundary condition). For small initial amplitudes, a large number of dominant perturbations (i.e. with maximum growthrate) exist. At larger initial amplitudes, purely 3D modes ($\lambda_x = \lambda_y$) amplify slightly faster than 2D modes. The preference for 3D modes is more pronounced in cases with a free-slip upper boundary condition. Also note the absolute shift in dominant wavelength at larger initial amplitudes, which is due to the fact that the two layers are initially of unequal thickness.

The fact that in 3D a large number of dominant modes amplify with almost identical rates may result in complex patterns, due to the superposition of different modes (see fig. 7.9). A regular spacing of 3D domes will only occur if purely 3D modes amplify significantly faster than 2D modes, which is generally not the case for finite layer thicknesses.

7.4.2 Initial random noise

The 3D geometries that evolve from an initial horizontal interface with random noise, for different initial layer thicknesses and upper boundary conditions (fast erosion and

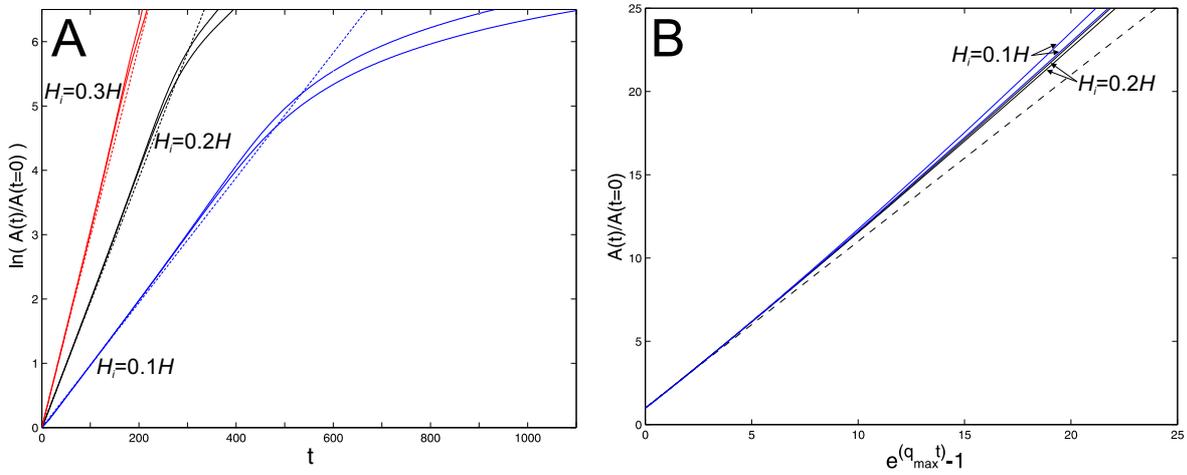


Figure 7.8: A) Normalized maximum amplitude versus time for 2D and 3D structures with different initial layer-thicknesses and a fast-erosion upper boundary condition. The upper line is for 3D, the lower for 2D structures (which grow slightly slower). Dashed line represents the amplitude behavior predicted by linear theory. B) Normalized amplitude versus normalized time during initial stages. All instabilities grow super-exponentially (dashed line represents the exponential solution). The thinner the layer, the more pronounced the preference for 3D structures. $\lambda_{3D} = \lambda_x = \lambda_y = 1.73$, $\lambda_{2D} = 1.22$, $q_{max} = 0.03$ if $H_i = 0.3H$; $\lambda_{3D} = 1.05$, $\lambda_{2D} = 0.74$, $q_{max} = 0.019$ if $H_i = 0.2H$; $\lambda_{3D} = 0.52$, $\lambda_{2D} = 0.37$, $q_{max} = 0.01$ if $H_i = 0.1H$.

free slip), are illustrated in figure 7.10. Simulations with a free-slip upper boundary ultimately form mushroom-shaped structures, whereas a fast-erosion boundary condition leads to chimney-like structures. The patterns that evolve from a thin initial layer are identical for both boundary conditions, which is related to the fact that the dominant growthrate and wavelength are nearly identical for the two cases (since they are dominated by the no-slip lower boundary condition). If the initial layer is thicker, the dominant wavelength and spacing are different for the two cases.

Simulations with a fast erosion boundary condition and random initial perturbations for different initial layer thicknesses are shown in figure 7.11. The resulting structures are not perfectly circular in mapview. They vary in size and may even form elongated and curved shapes, especially for thicker initial layer thicknesses. Spokes occur between the different domes. A Fourier analysis of the interfacial perturbations (fig. 7.12) reveals that the largest 10 Fourier components are all close to the dominant

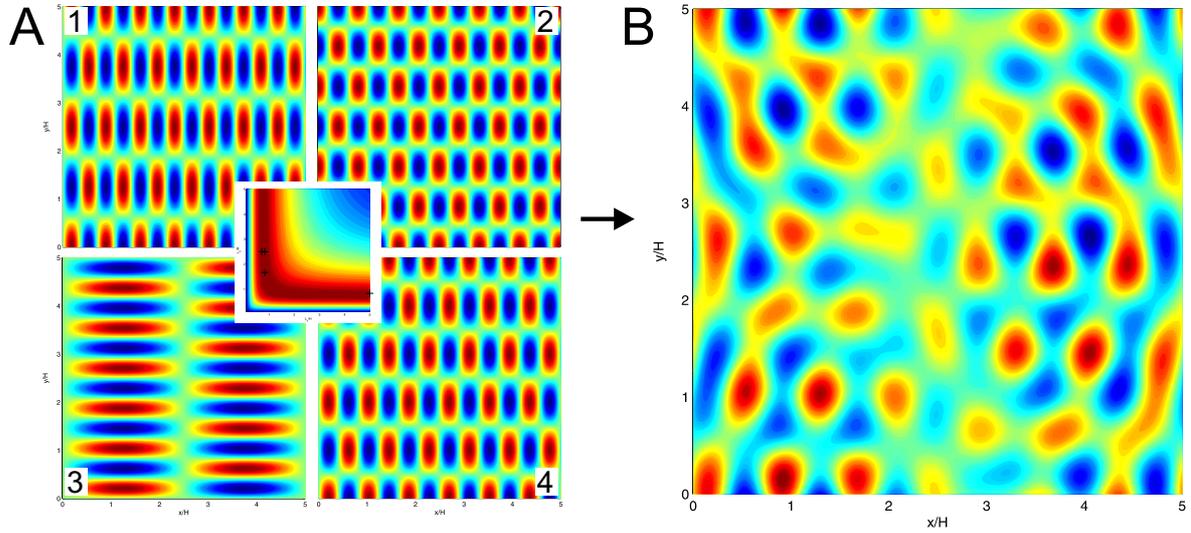


Figure 7.9: Illustration of the interference of different 3-D dominant wavelengths (A), which results in an -at first sight- fairly complex pattern (B). The following equations have been used to compute the four structures in A): 1: $z = \sin\left(\frac{2\pi}{0.71}x\right) \cos\left(\frac{2\pi}{2.5}y\right)$ 2: $z = \cos\left(\frac{2\pi}{0.83}x\right) \cos\left(\frac{2\pi}{1.67}y\right)$ 3: $z = \sin\left(\frac{2\pi}{5}x\right) \sin\left(\frac{2\pi}{0.83}y\right)$ 4: $z = \sin\left(\frac{2\pi}{0.83}x\right) \cos\left(\frac{2\pi}{2.5}y\right)$, with x, y, z in nondimensional form. The structure in B) is obtained by adding A1-4. The inset in A) shows the growthrate for a fast-erosion upper boundary condition and an interface height of $0.2H$. All structures in A) are close to the dominant wavelength.

modes. Thus, the range of patterns that can be observed is due to the superposition of different dominant modes. The analysis also reveals that all modes grow with a nearly constant rate, as predicted by the linear stability rate. It should be noted that a Fourier analysis is only possible until the overhangs begin to form. The amplitude evolution during later stages may be controlled by a buoyancy-controlled velocity and may be slower than that predicted by linear stability analysis.

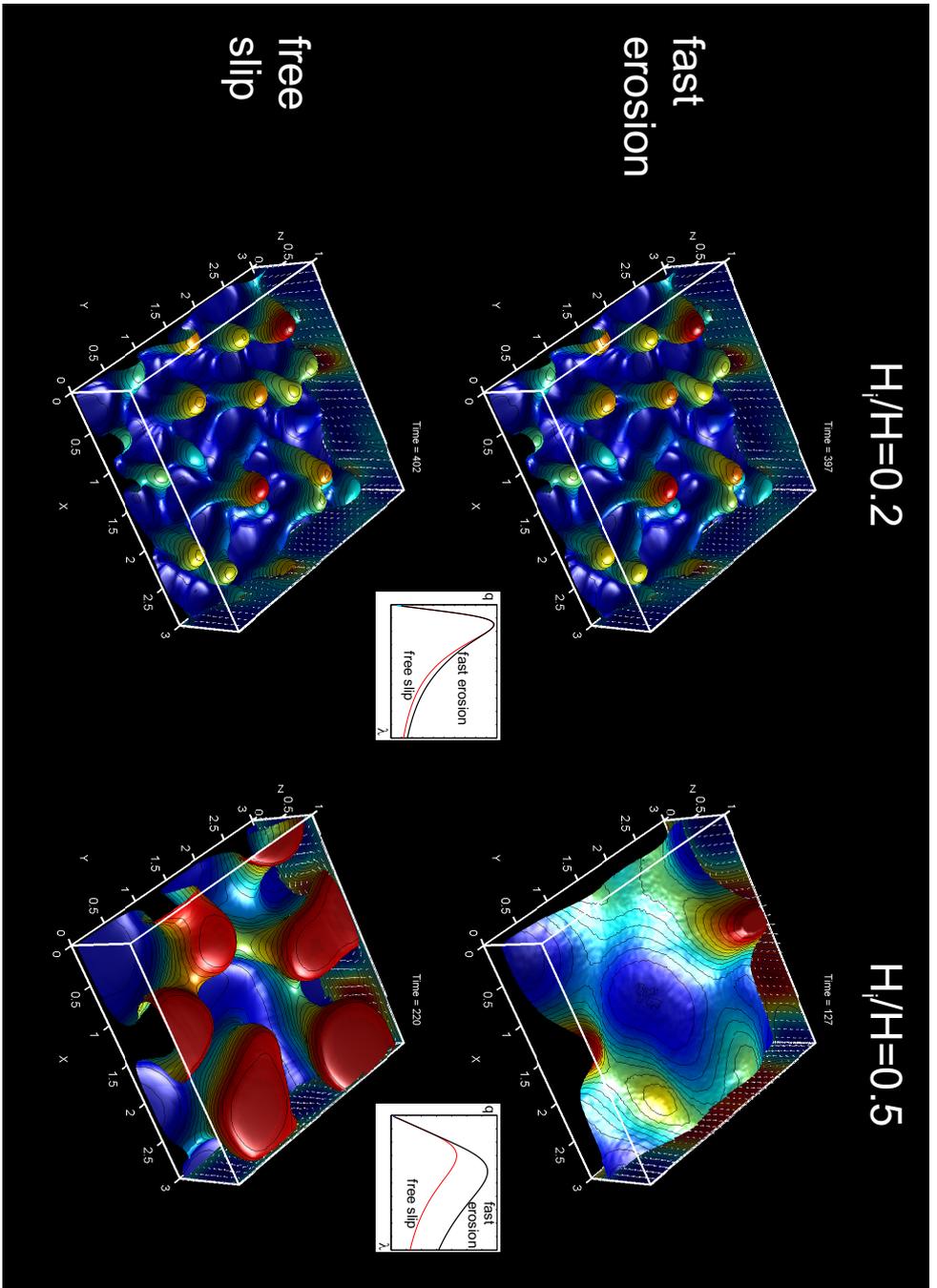


Figure 7.10: 3-D numerical simulations of the RT instability, starting from an initially horizontal interface of height H_i which is perturbed with random noise (the in all simulations). A no-slip lower boundary condition has been employed; the upper boundary condition was varied. Other parameters are $\mu_2/\mu_1 = 1$, $\Delta\rho g = 1$, noise-amplitude $= 10^{-3}H$. Inset shows the growthrate as a function of λ for the different cases. Colors on the interface indicate height (red=heightest), contour-spacing is 0.05. Colors at the back indicate magnitude of velocity. Simulations with a free-slip upper boundary condition always develop overhangs, whereas simulations with a fast erosion upper boundary condition ultimately result in chimney-like structures.

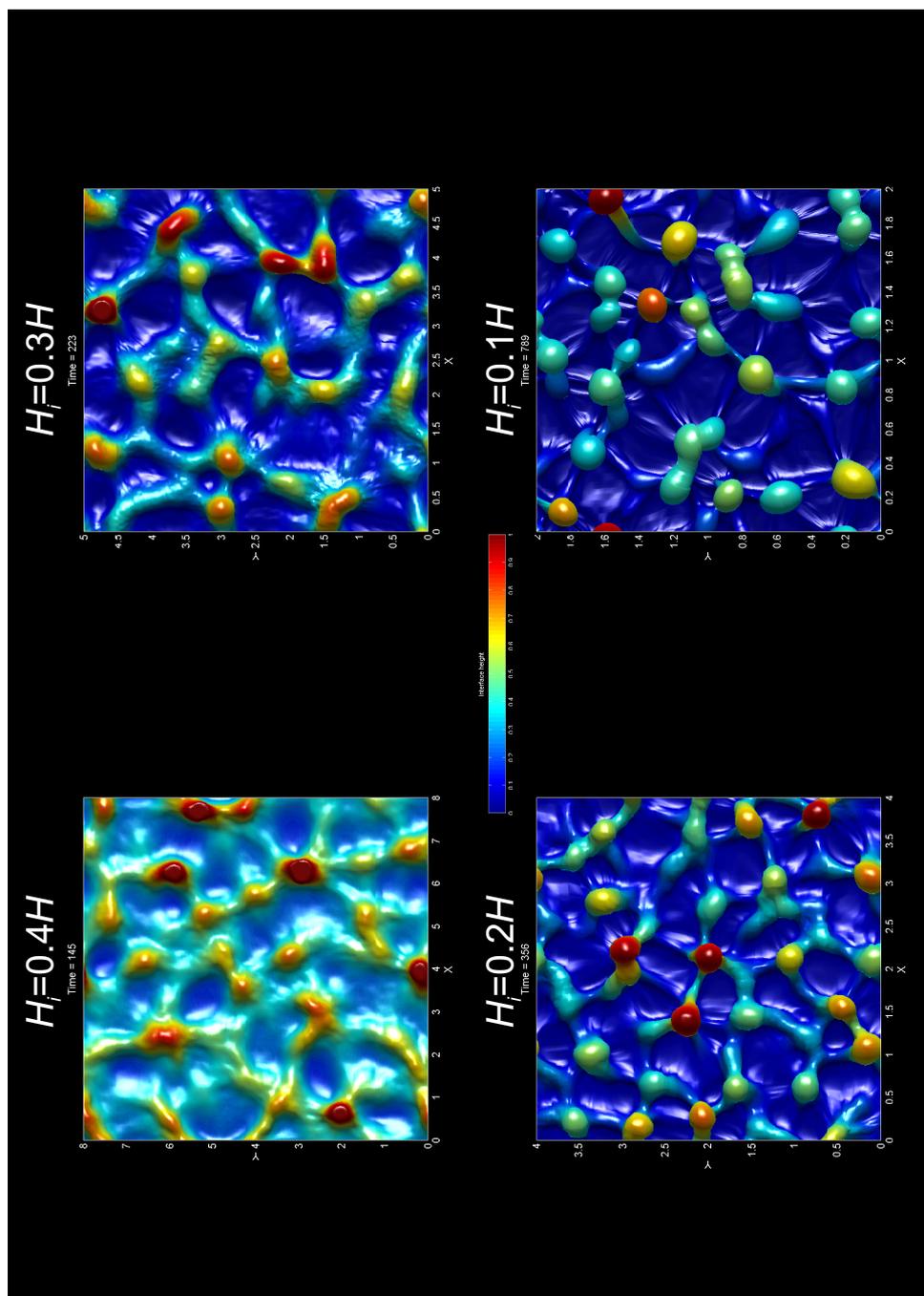


Figure 7.11: Top view of the interface between a low and a high-density fluid for different initial thicknesses of the low-density fluid. The initial interface was horizontal and but perturbed by random noise of maximum amplitude $2 \times 10^{-2}H$ (varied in all cases). A fast erosion upper and a no-slip lower boundary condition have been imposed. Other parameters are $\mu_2/\mu_1 = 1$, $\Delta\rho g = 1$. Computations are terminated when the tallest structure reaches the top of the box. Some of the initial random noise is still visible in the upper two figures. The preference for purely 3D (finger-like) structures is slightly larger for thin source-layers.

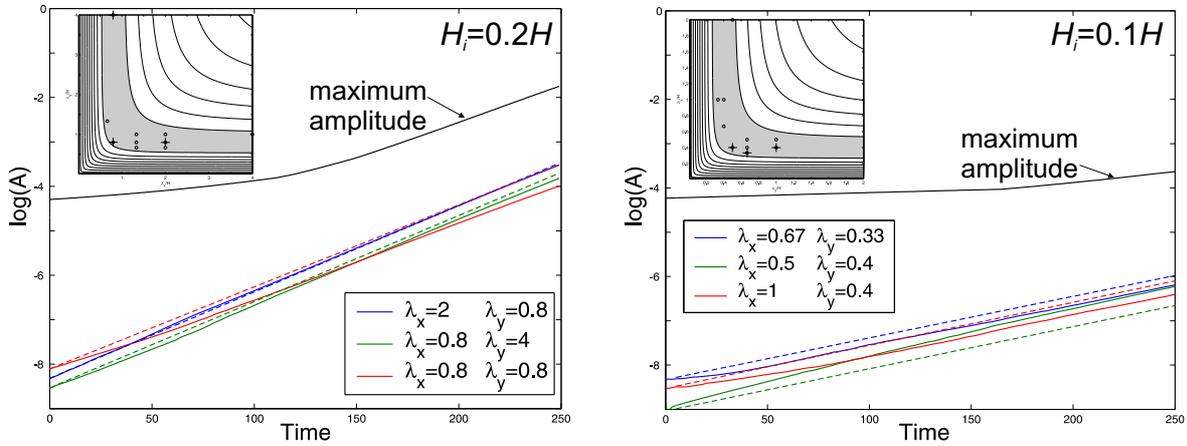


Figure 7.12: Amplitude versus time of selected Fourier components for the simulations of Fig. 7.11. with $H_i = 0.2H$ and $H_i = 0.1H$. The plotted components have maximum amplitude at the time when overhangs form ($t \approx 250$). Inset shows the growthrate versus wavelength together with 10 Fourier components of maximum amplitude at $t \approx 250$. All Fourier components are close to the dominant wavelength; the simulation with $H_i = 0.1H$ has slightly more preference for 3D structures.

7.4.3 Initial step-like perturbations

If a step-like perturbation with amplitude larger than the amplitude of background noise affects the initial interface, linear wall-like structures develop during later stages (fig. 7.13). If the step amplitude has the same order of magnitude than the amplitude of background noise, aligned domes result; if the step-offset is smaller than the background noise, 3D chimney-like diapirs are produced. Fourier analysis of the amplitude versus time reveals that 2D modes grow sub-exponentially in case the 2D mode does not survive during later stages, whereas they grow slightly super-exponential otherwise (fig. 7.14).

7.5 Discussion

Before discussing the relevance of the results to salt diapirism, it is useful to review the different opinions on salt tectonics.

Jackson [1995], divided the research on salt tectonics in three eras: (1) the pioneering era (1856-1933), (2) the fluid era (1933– \sim 1989) and (3) the brittle era

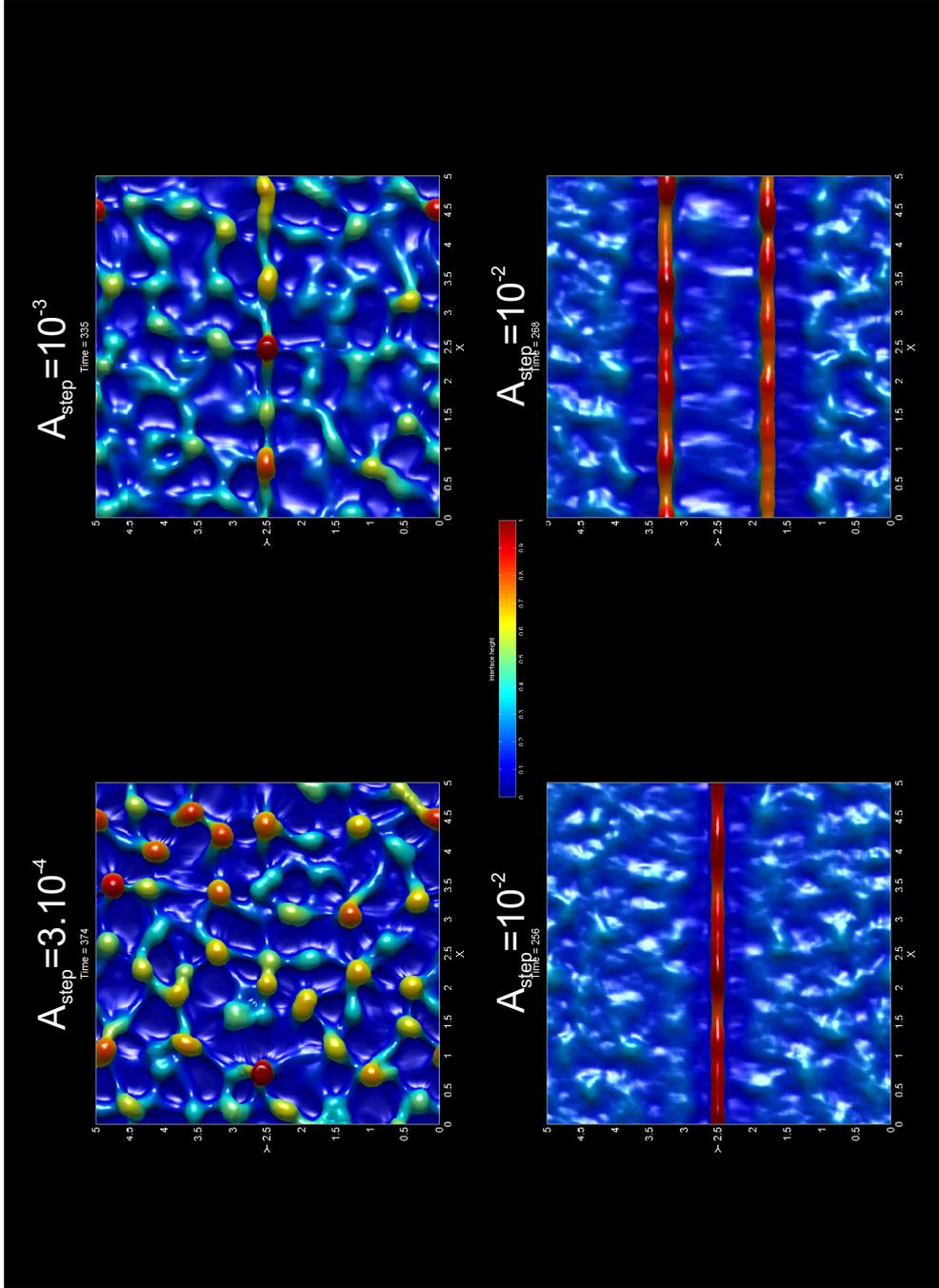


Figure 7.13: Pattern formation of the 3D RT-instability, when a step-like perturbation of amplitude A_{step} and random noise of maximum amplitude $10^{-3}H$ are imposed on the initial interface. The width of the step-like perturbation was fixed to 0.2 in the first three simulations and to 3 in the last simulation. Other parameters are $\mu_2/\mu_1 = 1$, $\Delta\rho g = 1$, noise-amplitude $= 10^{-3}H$ and $H_i = 0.2H$.

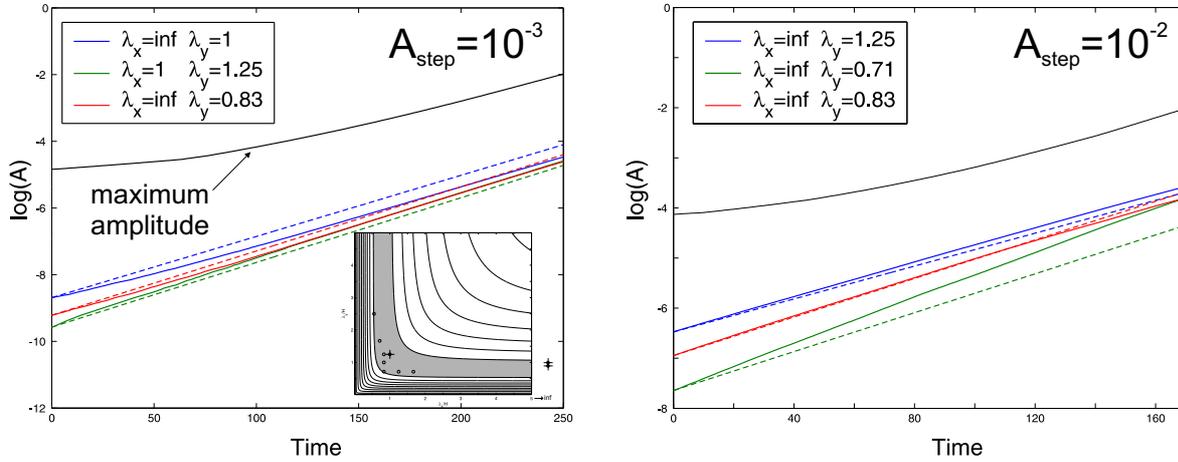


Figure 7.14: Amplitude versus time of selected Fourier components for the simulations of Fig. 7.13. with $A_{step} = 10^{-3}$ and $A_{step} = 10^{-2}$ and a width of 3. If $A_{step} = 10^{-2}$, 2-D perturbations dominate the late-stage flow.

(1989–now). During the *pioneering era*, salt structures were recognized and different formation mechanisms were discussed. The *fluid era* started with the laboratory experiments of Nettleton [1934] explaining the formation of salt diapirs with a Rayleigh-Taylor (RT) instability. One argument in favor of this mechanism is the fact that the salt density is nearly pressure-independent whereas the density of overlying sediments increases with depth. Thus, *once the overburden reaches a certain thickness*, its mean density becomes larger than the density of salt, and if both salt and overburden behave effectively like viscous fluids over geological times, a RT-instability is an appropriate model. Many salt diapirs display regular spacing in map view over large areas (e.g. fig. 7.2B), which is in agreement with the RT model and has been used to estimate the effective viscosity of the overburden ($\sim 10^{20} - 10^{23}$ Pas, [e.g. Rönnlund, 1989]). But do sediments actually behave like viscous fluids? The frequent occurrence of brittle faults in sedimentary basins suggests that sediments behave brittle rather than ductile in the upper crust. At the same time, laboratory experiments indicate that pressure solution is an important, creep-like deformation mechanism in sedimentary rocks. Analogue experiments of a ductile layer (usually silicone), overlain by a brittle layer (typically represented by cohesionless sand) demonstrated that one may simulate salt structures if extension is applied [e.g. Vendeville and Jackson, 1992]. This marked the onset of

the *brittle era* of salt tectonics. One of the main arguments in favor of brittle salt tectonics is the fact that salt seems to be mainly mobilized during periods of regional compression or extension [Jackson, 1995].

However, many other factors have been shown to contribute to salt diapirism, including sedimentation, erosion, and differential loading. The idea that salt-tectonics is due to brittle deformation of the overburden has recently been challenged by a number of field observations, borehole information, seismics and numerical studies. A detailed study of salt diapirs in the Central North-Sea Basin showed that most of the diapirs have been formed by down-building (sedimentation), and little brittle deformation was observed above the domes. Salt diapirs in the Dniepr-Donets Basin were shown to be initiated during periods of rifting, but continued to rise long after rifting had ceased [Stovba and Stephenson, 2003]. Overburden deformation is predominantly ductile above salt layers. There is a strong correlation between basement faults and salt structures, but the regular spacing of the structures cannot be explained by the spacing of these basement faults.

Thus, the evidence for both brittle and ductile behaviors is not unambiguous. Perhaps the best way to model the sedimentary overburden is allowing the sediments to behave ductile *and* brittle, depending on the ambient conditions (e.g. strainrate, temperature). Such a rheological model (visco-elasto-plastic) accounts for the finite strength of rocks and is applicable to models of lithospheric-scale [Burg and Podladchikov, 1999, Gerbault, 2000, Burov et al., 2003], and crustal-scale [Poliakov et al., 1996, Lavier et al., 2000, Lavier and Buck, 2002] deformation. Modelling results of fault patterns in a brittle layer above a ductile layer in extension are in excellent agreement with laboratory experiments that use clay as analogue brittle material [Bellahsen et al., 2003].

A similar rheology has been employed in numerical models designed to study the interaction between a brittle overburden and a viscous salt layer [Poliakov et al., 1996]. Results showed that the location of salt-diapirs is governed by brittle faults in compression or in extension [Poliakov et al., 1993a]. If large-scale tectonic deformation is absent, the effects of buoyancy are sufficient to deform the overburden and to create salt diapirs [Poliakov et al., 1996], especially if sedimentation and erosion are taken into account. Physically, these observations can be explained if the dominating deformation mode depends on the stress state of the overburden. Brittle faulting will only occur if

the differential stress exceeds the yield criteria given by Byerlee's law. If the overburden has a visco-elastic rheology, the maximum differential stress is proportional to strain-rate. During periods of extension, this differential stress will exceed the strength of rocks and lead to faulting. However, the differential stress due to buoyancy is proportional to $A\Delta\rho g$ (with A the amplitude of the dome, $\Delta\rho$ the density difference between salt and overburden and g the gravitational acceleration), which is on the order of a few MPa for natural parameters [Poliakov et al., 1996]. Such a stress will only exceed the yield criteria close to the surface, which means that the governing mode during the initial stages of salt-diapir formation (when A is small) is dominated by ductile (or pressure-solution) creep rather than by brittle faulting [Poliakov et al., 1996].

Our 3D results employing a viscous overburden show that the linear, wall-like salt structures observed in many sedimentary basins (e.g. fig. 7.2A) are unlikely to have formed spontaneously unless a strong initial 2D perturbation was present. Such a perturbation could be caused by brittle faulting of the overburden, that offsets the salt-overburden interface. The frequently observed three-dimensional patterns, however, are in agreement with a RT model [see e.g. Rönnlund, 1989, Hughes and Davison, 1993]. The fact that not all domes are perfectly circular or of equal size and shape cannot be used as an argument to reject the RT model as a cause for the pattern formation. Yet, more work is required to understand the mechanics of the interaction of a brittle (visco-elasto-plastic) overburden with a viscous substratum, especially in 3D.

7.6 Conclusions

The viscous RT instability has been studied for a no-slip lower boundary condition and for no-slip, free-slip, free surface and fast redistribution upper boundary conditions both in 2D and 3D cases. Emphasis has been put on the differences between previous 2D modelling and the 3D results. The following conclusions are reached:

1. Erosion and deposition have a profound effect on the final geometry of diapiric structures. Rather than mushroom-shaped they become chimney-like.
2. Whereas only a single dominant wavelength exist in the 2D Newtonian viscous case, an infinite number of dominant modes exist in the 3D case.
3. The spacing of diapirs both in 2D and 3D cases closely follows the predictions of linear stability theory if they grow from an initial horizontal interface perturbed with

random noise.

4. The superposition of different 3D dominant modes results in complex patterns with diapirs of different thicknesses, having a variety of shapes, from finger-like (circular or elliptical in mapview) to curved elongated.

5. Wall-like diapirs may be initiated by a strong sinusoidal or step-like initial perturbation of the interface between low and high density fluids. Thus, salt-walls can only be explained by a RT-type model if they are initiated by a strong initial linear perturbation, for example caused by brittle faulting or folding. Simultaneously, the 3D salt-dome patterns observed in many basins seem to be in good agreement with a RT model.

This work has some limitation. Firstly the effects of sedimentation and strongly varying overburden thicknesses are not taken into account. Secondly the overburden is assumed to be viscous. Although this may be relevant for several cases, more analytical and numerical work is required to better understand the mechanics of a brittle overburden superposed on a ductile substratum. This is a challenging problem, especially in 3D.

7.7 Acknowledgements

The Swiss National Science Foundation (project 21-61912.00) supported this work.

7.8 Appendix

The linear stability analysis for the 2D Rayleigh-Taylor instability can be derived by means of a perturbation method. Both the technique and some results are relatively well documented and will therefore not be repeated here [see for example Chandrasekhar, 1961, Biot and Odé, 1965, Turcotte and Schubert, 1982, Fletcher, 1972, Canright and Morris, 1993, Ribe, 1998, Conrad and Molnar, 1997, Burg et al., 2004, for more details]. The resulting expressions, however, are in the most general case (variable thickness, viscosity ratio etcetera) rather lengthy. In order to make the results available for a wider audience, we have chosen to present the solutions in the form of a MAPLE script (fig. 7.15). The expressions have been used to verify two independently developed numerical codes (fig. 7.3). Three different boundary conditions are

considered. A no-slip condition:

$$v_z = 0, v_x = 0 \quad (7.5)$$

A free-slip condition:

$$v_z = 0, \sigma_{xz} = 0 \quad (7.6)$$

and a fast-erosion (horizontal no-stress) condition:

$$\sigma_{zz} = 0, \sigma_{xz} = 0 \quad (7.7)$$

The growthrate q is non-dimensionalized by $Hg\Delta\rho/\mu_1$. Thus the amplitude behavior versus time for 'real parameters' is given by $A(t) = A_0 \exp(qHg\Delta\rho t/\mu_1)$, with t in seconds and A in meters. The 2D solution can be transformed into the 3D solution after noticing that [Biot, 1966]

$$\left(\frac{2\pi}{\lambda_{2D}}\right)^2 = \left(\frac{2\pi}{\lambda_x}\right)^2 + \left(\frac{2\pi}{\lambda_y}\right)^2 \quad (7.8)$$

```

> #RT instability without compression for a linearly viscous rheology
> restart;
> # assume velocity to be periodic:
> Vx := -1/omega*vz(z)*sin(omega*x):
> Vz :=          vz(z)*cos(omega*x):
> P  :=          p(z) *cos(omega*x):
> # Rheology:
> Sxx := -P + 2*mu*diff(Vx,x):
> Szz := -P + 2*mu*diff(Vz,z):
> Sxz :=          mu*(diff(Vx,z)+diff(Vz,x)):
> # Equilibrium equations:
> eq1 := diff(Vx,x)+diff(Vz,z)=0:
> eq2 := diff(Sxx,x)+diff(Sxz,z)=0:
> eq3 := diff(Sxz,x)+diff(Szz,z)=0:
> # Use incompressibility equation to find expression for vx(z):
> vx(z):=solve(eq1,vx(z)):
> # Eliminate pressure from force-balance equations
> eq := simplify(diff(eq2,z)-diff(eq3,x)):
> # Solve the equation
> sol:=dsolve(eq,vz(z)):
> vz(z) := _C1*exp(-omega*z)+_C2*exp(-omega*z)*z+_C3*exp(omega*z)+_C4*exp(omega*z)*z:
> # Compute velocities and stresses from given solution for vz(z):
> vx(z) := eval(vx(z)): p(z) := solve(eq2,p(z)):
> Sxx := eval(Sxx): Sxz := eval(Sxz):
> Szz := eval(Szz): P := eval(P):
> Vx := eval(Vx): Vz := eval(Vz):
> # Our system has two layers: upper and lower.
> # Set viscosity of lower layer to 1.
> # Spell out velocities and stresses in every layer
> sbs := {_C1=C[1],_C2=C[2],_C3=C[3],_C4=C[4],mu=R}:
> Vz_up := subs(sbs,Vz): Vx_up := subs(sbs,Vx): Sxx_up := subs(sbs,Sxx):
> Sxz_up := subs(sbs,Sxz): Szz_up := subs(sbs,Szz):
> sbs := {_C1=C[5],_C2=C[6],_C3=C[7],_C4=C[8],mu=1}:
> Vz_lo := subs(sbs,Vz): Vx_lo := subs(sbs,Vx): Sxx_lo := subs(sbs,Sxx):
> Sxz_lo := subs(sbs,Sxz): Szz_lo := subs(sbs,Szz):
> # Set boundary conditions at top of domain
> # In case of fast erosion:
> eq[1] := subs(z=1,Sxz_up) = 0:
> eq[2] := subs(z=1,Szz_up) = 0:
> # In case of no-slip
> #eq[1] := subs(z=1,Vz_up) = 0:
> #eq[2] := subs(z=1,Vx_up) = 0:
> # In case of free-slip
> #eq[1] := subs(z=1,Vz_up) = 0:
> #eq[2] := subs(z=1,Sxz_up) = 0:
> # Matching conditions at the interface
> eq[3] := subs(z=Hi,Vz_up - Vz_lo) = 0:
> eq[4] := subs(z=Hi,Vx_up - Vx_lo) = 0:
> eq[5] := subs(z=Hi,Sxz_up - Sxz_lo) = 0:
> eq[6] := subs(z=Hi,Szz_up - Szz_lo + 1*Vz_up/q) = 0:
> # Lower boundary (no slip)
> eq[7] := subs(z=0,Vx_lo) = 0:
> eq[8] := subs(z=0,Vz_lo) = 0:
> # Only keep coefficients
> eq[1]:=simplify(eq[1]/sin(omega*x)): eq[2]:=simplify(eq[2]/cos(omega*x)):
> eq[3]:=simplify(eq[3]/cos(omega*x)): eq[4]:=simplify(eq[4]/sin(omega*x)):
> eq[5]:=simplify(eq[5]/sin(omega*x)): eq[6]:=simplify(eq[6]/cos(omega*x)):
> eq[7]:=simplify(eq[7]/sin(omega*x)): eq[8]:=simplify(eq[8]/cos(omega*x)):
> # collect coefficients in matrix form. Note that we now have q in the equations (the growthrate)
> # The problem now is: A*C=RHS, and since RHS=0, the only solution occurs if det(A)=0.
> # Thus the growthrate can be calculated by assuming that the determinant of A=0
> N := 8:
> A := matrix(N,N,[]):
> RHS := vector(N,0):
> for i from 1 to N do
>   RHS[i] := eval(lhs(eq[i])):
>   for j from 1 to N do
>     A[i,j] := simplify(coeff(eval(lhs(eq[i])),C[j])):
>     RHS[i] := simplify(RHS[i] - A[i,j]*C[j]):
>   od:
> od:
> with(linalg):
> #Calculate determinant of the matrix A:
> det_A:=det(A):
> # Compute growthrate:
> q_num:=solve(det_A,q):
> #Plot the solution:
> plot([subs({omega=2*Pi/lambda,Hi=.25,H=1,R=1e2}, q_num),
>      subs({omega=2*Pi/lambda,Hi=.20,H=1,R=1e2}, q_num),
>      subs({omega=2*Pi/lambda,Hi=.25,H=1,R=1e3}, q_num)],lambda=.2..10);

```

Figure 7.15: MAPLE (version 8.0) script to compute the growthrate as a function of wavelength and material parameters for the 2D Rayleigh-Taylor instability for various boundary conditions.

Bibliography

- G. I. Alsop, D.J. Blundell, and I. Davison, editors. *Salt tectonics*, volume 100 of *Geological Society Special Publication*. The Geological Society, London, 1996.
- J.M. Anketell, J. Cegla, and S. Dzulynski. On the deformational structures in systems with reversed density gradients. *Annales of the Geological Society of Poland*, XL: 3–30, 1970.
- N. Bellahsen, J. M. Daniel, L. Bollinger, and E. Burov. Influence of viscous layers on the growth of normal faults: insights from experimental and numerical models. *Journal of Structural Geology*, 25:1471–1485, 2003.
- M.A. Biot. Three-dimensional gravity instability derived from two-dimensional solutions. *Geophysics*, 31(153-166), 1966.
- M.A. Biot and H. Odé. Theory of gravity instability with variable overburden and compaction. *Geophysics*, 30:153–166, 1965.
- J. P. Burg, B.J.P. Kaus, and Y. Podladchikov. Dome structures in collision orogens. Mechanical investigation of the gravity/compression interplay. In D.L. Whitney, C. Teyssier, and C.S. Siddoway, editors, *Gneiss domes in orogeny*, volume 380, pages 47–66. Geological Society of America, Boulder, 2004.
- J. P. Burg and Y. Podladchikov. Lithospheric scale folding: numerical modeling and application to the Himalayan syntaxes. *International Journal of Earth Sciences*, 88 (2):190–200, 1999.
- E. Burov, C. Jaupart, and L. Guillou-Frottier. Ascent and emplacement of buoyant magma bodies in brittle- ductile upper crust. *Journal of Geophysical Research-Solid Earth*, 108(B4):art. no.–2177, 2003.

- D. Canright and S. Morris. Buoyant instability of a viscous film over a passive fluid. *Journal of Fluid Mechanics*, 255:349–372, 1993.
- S. Chandrasekhar. *Hydrodynamic and hydromagnetic instabilities*. Oxford University Press, 1961.
- C.P. Conrad and P. Molnar. The growth of Rayleigh-Taylor-type instabilities in the lithosphere for various rheological and density structures. *Geophysical Journal International*, 129:95–112, 1997.
- Danes. Mathematical formulation of salt-dome dynamics. *Geophysics*, 29(3):414–424, 1964.
- B. Daudré and S. Cloetingh. Numerical modelling of salt diapirism: influence of the tectonic regime. *Tectonophysics*, 240(1-4):59–79, 1994.
- I. Davison, G. I. Alsop, P. Birch, C. Elders, N. Evans, H. Nicholson, P. Rorison, D. Wade, J. Woodward, and M. Young. Geometry and late-stage structural evolution of Central Graben salt diapirs, North Sea. *Marine and Petroleum Geology*, 17:499–522, 2000a.
- I. Davison, G. I. Alsop, A. G. Evans, and M. Safaricz. Overburden deformation patterns and mechanisms of salt diapir penetration in the Central Graben, North Sea. *Marine and Petroleum Geology*, 17:601–618, 2000b.
- C. Dietl and H.A. Koyi. Emplacement of nested diapirs: results of centrifuge modelling. *Journal of Virtual Explorer*, 7:81–88, 2002.
- R. C. Fletcher. Application of a mathematical-model to emplacement of mantled gneiss domes. *American Journal of Science*, 272(3):197–216, 1972.
- L. Gemmer, S.J. Ings, S. Medvedev, and C. Beaumont. Salt tectonics driven by differential sediment loading: Stability analysis and finite element experiments. *Basin Research*, 16(2):199–218, 2004.
- M. Gerbault. At what stress level is the central Indian Ocean lithosphere buckling? *Earth and Planetary Science Letters*, 178(1-2):165–181, 2000.

- T. V. Gerya, L. L. Perchuk, D.D. van Reenen, and C.A. Smit. Two-dimensional numerical modeling of pressure-temperature-time paths for the exhumation of some granulite facies terrains in the precambrian. *Journal of Geodynamics*, 30(1-2):17–35, 2000.
- T. V. Gerya, R. Uken, J. Reinhardt, M.K. Watkeys, W.V. Maresch, and C. Brendan. Cold fingers in hot magma: numerical modeling of country-rock diapirs in the Bushveld Complex, South Africa. *Geology*, 31:753–756, 2003.
- J. Hospers, J.S. Rathore, F. Jianhua, E.G. Finnstrom, and J. Holthe. Salt tectonics in the Norwegian-Danish Basin. *Tectonophysics*, 149:35–60, 1988.
- G. Houseman and P. Molnar. Gravitational (rayleigh-taylor) instability of a layer with non-linear viscosity and convective thinning of continental lithosphere. *Geophysical Journal International*, 128:125–150, 1997.
- M. Hughes and I. Davison. Geometry and growth kinematics of salt pillows in the southern North Sea. *Tectonophysics*, 228:239–254, 1993.
- A. Ismail-Zadeh, I. Tsepelov, C. Talbot, and A. Korotkii. Three-dimensional forward and backward modelling of diapirism: numerical approach and its applicability to the evolution of salt structures in the pricaspian basin. *Tectonophysics*, 387(1-4): 81–103, 2004.
- A. T. Ismail-Zadeh, C. J. Talbot, and Y. A. Volozh. Dynamic restoration of profiles across diapiric salt structures: numerical approach and its applications. *Tectonophysics*, 337(1-2):23–38, 2001.
- M.P.A. Jackson. Retrospective salt tectonics. In M.P.A. Jackson, D.G. Roberts, and D. Snelson, editors, *Salt tectonics: a global perspective*, pages 1–28. AAPG memoir, 1995.
- M.P.A. Jackson, D.G. Roberts, and S. Snelson, editors. *Salt tectonics: a global perspective.*, volume 65 of *AAPG Memoir*. Tulsa, Oklahoma, U.S.A., 1995.
- M. Jull and P. B. Kelemen. On the conditions for lower crustal convective instability. *Journal of Geophysical Research*, 106:6423–6446, 2001.

- B. J. P. Kaus and Y. Y. Podladchikov. Forward and reverse modeling of the three-dimensional viscous Rayleigh-Taylor instability. *Geophysical Research Letters*, 28(6): 1095–1098, 2001.
- B.J.P. Kaus, Y.Y. Podladchikov, and D. W. Schmid. Eulerian spectral/finite difference method for large deformation modelling of visco-elasto-plastic geomaterials. *Bolletino Geofisico*, 45:346–349, 2004.
- L Lavier and W.R. Buck. Half graben versus large-offset low-angle normal fault: importance of keeping cool during normal faulting. *Journal of Geophysical Research*, 107(B6):10.1029/2001JB000513, 2002.
- L. Lavier, W.R. Buck, and A.N.B. Poliakov. Factors controlling normal fault offset in an ideal brittle layer. *Journal of Geophysical Research*, 105(B10):23431–23442, 2000.
- R.B. Miller and S.R. Paterson. In defence of magmatic diapirs. *Journal of Structural Geology*, 21:1161–1173, 1999.
- L.L. Nettleton. Fluid mechanics of salt domes. *Bulletin of the American Association of Petroleum Geologists*, 18:1175–1204, 1934.
- Y. Podladchikov, C. Talbot, and A. Poliakov. Numerical models of complex diapirs. *Tectonophysics*, 228:349–362, 1993.
- A. Poliakov and Y. Podladchikov. Diapirism and topography. *Geophysical Journal International*, 109:553–564, 1992.
- A. Poliakov, Y. Podladchikov, J. B. Dawson, and C. Talbot. Salt diapirism with simultaneous brittle faulting and viscous flow. *Geological Society Special Publication*, 100:291–302, 1996.
- A. Poliakov, Y. Podladchikov, and C. Talbot. Initiation of salt diapirs with frictional overburdens: numerical experiments. *Tectonophysics*, 228:199–210, 1993a.
- A. Poliakov, R. T. van Balen, Y. Podladchikov, B. Daudré, S. Cloething, and C. Talbot. Numerical analysis of how sedimentation and redistribution of surficial sediments affects salt diapirism. *Tectonophysics*, 226(1-4):199–216, 1993b.

- N.M. Ribe. Spouting and planform selection in the Rayleigh-Taylor instability of miscible viscous fluids. *Journal of Fluid Mechanics*, 377:27–45, 1998.
- M. Romer and H.J. Neugebauer. The salt dome problem: A multilayered approach. *Journal of Geophysical Research*, 96(B2):2389–2396, 1991.
- P. Rönnlund. Viscosity ratio estimates from natural Rayleigh-Taylor instabilities. *Terra Nova*, 1(344-348), 1989.
- Joerg Schmalzl and Alexander Loddock. Using subdivision surfaces and adaptive surface simplification algorithms for modeling chemical heterogeneities in geophysical flows. *Geochemistry Geophysics Geosystems*, 4(9):DOI 10.1029/2003GC000578, 2003.
- H. Schmeling. On the relation between initial conditions and late stages of the Rayleigh-Taylor instability. *Tectonophysics*, 133:65–80, 1987.
- D. W. Schmid and Y. Podladchikov. Analytical solutions for deformable elliptical inclusions in general shear. *Geophysical Journal International*, 155:269–288, 2003.
- S. Stovba and R. Stephenson. Style and timing of salt tectonics in the Dniepr-Donets Basin (Ukraine): implications for triggering and driving mechanisms of salt movement in sedimentary basins. *Marine and Petroleum Geology*, 19:1169–1189, 2003.
- P.J. Tackley and S. D. King. Testing the tracer ratio method for modeling active compositional fields in mantle convection simulations. *Geochemistry Geophysics Geosystems*, 4:doi:10.1029/2001GC000214, 2003.
- C. Talbot, P. Ronnlund, H. Schmeling, H.A. Koyi, and M.P.A. Jackson. Diapiric spoke patterns. *Tectonophysics*, 188:187–201, 1991.
- F. Trusheim. Mechanism of salt migration in northern Germany. *Bulletin of the American Association of Petroleum Geologists*, 44:1519–1540, 1960.
- D. L. Turcotte and G. Schubert. *Geodynamics*. Wiley, 1982.
- R. J. Twiss and E.M. Moores. *Structural Geology*. Freeman and Company, New York, 1992.

-
- P. E. van Keken, C. J. Spiers, A. P. van den Berg, and E.J. Muzyert. The effective viscosity of rocksalt: Implementation of steady state creep laws in numerical models of salt diapirism. *Tectonophysics*, 225:457–476, 1993.
- B.C. Vendeville and M.P.A. Jackson. The rise of diapirs during thin-skinned extension. *Marine and Petroleum Geology*, 9:331–353, 1992.
- R. F. Weinberg and H. Schmeling. Polydiapirs: multiwavelength gravity structures. *Journal of Structural Geology*, 14(4):425–436, 1992.
- J.A. Whitehead and D.S. Luther. Dynamics of laboratory diapir and plume models. *Journal of Geophysical Research*, 91:705–717, 1975.
- W. D. Woidt. Finite element calculations applied to salt dome analysis. *Tectonophysics*, 50:369–386, 1978.
- S. Zaleski and P. Julien. Numerical simulation of Rayleigh-Taylor instability for single and multiple salt diapirs. *Tectonophysics*, 206(1-2):55–69, 1992.

Chapter 8

A finite difference/spectral method for large deformation modelling of visco-elasto-plastic geomaterials¹

Abstract

Many problems that occur in geodynamics can be reduced to the solving of visco-elasto-plastic rheological equations. Here we describe a spectral/finite-difference method that can deal with a Maxwell viscoelastic rheology combined with Mohr-Coulomb plasticity in an Eulerian framework. The method approximates derivatives in vertical direction by finite-differences and in horizontal direction(s) by a pseudospectral approach. Material boundaries are tracked by marker chains that are moved through a fixed grid and allow large deformations. Lateral viscosity variations are treated by either a direct solver or by an iterative method. Advection of smoothly varying properties is done with a semi-Lagrangian advection scheme. Performance issues of the iterative versus the direct scheme are discussed and benchmark results are presented for a range of different problems.

¹Part of this work has been published in: Kaus, B.J.P, Podladchikov, Y.Y., Schmid, D.W. (2004) Eulerian spectral/finite difference method for large deformation modelling of visco-elasto-plastic geomaterials. *Extended abstracts of the GeoMod2004 conference. Bolletino di Geofisica*. Vol. 45. p346-349.)

8.1 Introduction

One of the goals of geoscientists is to understand the physical processes that formed and reworked the earth. This is a difficult problem since the current state of earth is a result of billions of years of accumulated deformation and insight in the main driving mechanisms is at least indirect or approximate. In recent years numerical modelling has been proven useful on the matter. Much of our understanding about mantle convection, for example, comes from numerical simulations that treat the mantle as a creeping, highly viscous fluid. The lithospheric plates, on the other hand, are colder than the mantle and behave more like a viscoelastic, brittle solid. Numerical modelling of the deformation of these plates and their coupling to the mantle thus requires the solving of visco-elasto-plastic rheological equations. In addition, it is important that a numerical method developed for this purpose, is efficient, can handle large deformations, strong variations of material properties, free surface effects and non-linear rheologies. Previous workers utilized finite element methods [Huisman et al., 2001], dynamic Lagrangian remeshing methods [Braun and Sambridge, 1994], particle-in-cell finite element techniques [Moresi et al., 2003] and control-volume methods [Poliakov et al., 1993, 1996] to deal with the problems described above.

We follow a different approach and solve the governing equations by an Eulerian finite-difference/spectral method [Schmalholz et al., 2001]. Our method differs from the method of Schmalholz et al. [2001] since we (1) employ a different discretization technique, (2) have a different advection scheme, (3) solve for a visco-elasto-plastic rheology rather than for a viscoelastic rheology, (4) modified the iterative scheme, and (5) couple the energy equation to the mechanical equations. In the following parts of this chapter, details are given about the mathematical and numerical aspects of the code and several benchmark tests are presented.

8.2 Mathematical model

For many geodynamic applications involving slow deformation inertial effects are not important and the balance equations can be written as:

$$\frac{\partial \rho v_i}{\partial x_i} + \frac{\partial \rho}{\partial t} = 0$$

$$\frac{\partial \sigma_{ij}}{\partial x_j} = \rho g_i \quad (8.1)$$

here v_i is velocity, σ_{ij} stress, ρ density, g gravitational acceleration and t is time.

A Maxwell viscoelastic rheology for stress- and strain rate deviators is assumed [e.g. Schmalholz et al., 2001]:

$$\tilde{\varepsilon}_{ij}^{ve} = \tilde{\varepsilon}_{ij}^v + \tilde{\varepsilon}_{ij}^e = \frac{1}{2\mu_{vis}} \tau_{ij} + \frac{1}{2G} \frac{D\tau_{ij}}{Dt} \quad (8.2)$$

where $\tau_{ij} = \sigma_{ij} + \delta_{ij}P$, $\tilde{\varepsilon}_{ij} = \dot{\varepsilon}_{ij} - \bar{\varepsilon}$, $\dot{\varepsilon}_{ij} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right)$, $P = -\frac{1}{3} \text{tr}(\sigma_{ij})$, $\bar{\varepsilon} = \frac{1}{3} \text{tr}(\dot{\varepsilon}_{ij})$, G is the elastic shear modulus, μ_{vis} the shear viscosity (which may have a nonlinear stress- and temperature-dependence) and $\frac{D}{Dt}$ denotes the objective derivative of the stress tensor.

Rocks cannot sustain high stresses; instead they will fail plastically. Since in this case elastic strains are small, it is adequate to make the additive strain rate decomposition, which states that:

$$\dot{\varepsilon}_{ij} = \dot{\varepsilon}_{ij}^{ve} + \dot{\varepsilon}_{ij}^p \quad (8.3)$$

where $\dot{\varepsilon}_{ij}^{ve}$ and $\dot{\varepsilon}_{ij}^p$ denote viscoelastic and plastic strain rates respectively. The plastic strain rate can be calculated according to:

$$\dot{\varepsilon}_{ij}^p = \dot{\lambda} \frac{\partial Q}{\partial \sigma_{ij}} \quad (8.4)$$

where $\frac{\partial Q}{\partial \sigma_{ij}}$ is the direction of plastic flow and $\dot{\lambda}$ the plastic multiplier. The yield criterion can be expressed in Kuhn-Tucker form as [Simo and Hughes, 2000]

$$\dot{\lambda} \geq 0, F \leq 0, \dot{\lambda} F = 0 \quad (8.5)$$

For rocks under upper-crustal conditions, a Mohr-Coulomb yield function with non-associated flow rule is the minimum model. Spelled out for a 2D case, it can be written as [Vermeer and de Borst, 1984]:

$$\begin{aligned} F &= \tau^* - \sigma^* \sin(\phi) - c \cos(\phi) \\ Q &= \tau^* - \sigma^* \sin(\psi) \end{aligned} \quad (8.6)$$

where ϕ is the friction angle, ψ the dilation angle (in general smaller than ϕ), c the cohesion of the rocks, τ^* is the radius and σ^* the center of the Mohr-circle, given by

$$\tau^* = \sqrt{\left(\frac{\sigma_{xx} - \sigma_{zz}}{2} \right)^2 + \sigma_{xz}^2}$$

$$\sigma^* = -\frac{\sigma_{xx} + \sigma_{zz}}{2} \quad (8.7)$$

Substituting eqs. (8.3) and (8.4) into eq. (8.2), we obtain the rheological equation for a visco-elasto-plastic material:

$$\dot{\varepsilon}_{ij} - \dot{\varepsilon}_{ij}^p - \bar{\varepsilon}^{ve} = \tilde{\varepsilon}_{ij}^{ve} = \frac{1}{2\mu_{vis}} (\sigma_{ij} + P) + \frac{1}{2G} \frac{D(\sigma_{ij} + P)}{Dt} \quad (8.8)$$

In addition, we also solve the energy equation, given by

$$\rho c_p \left(\frac{\partial T}{\partial t} + v_i \frac{\partial T}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(k \frac{\partial T}{\partial x_j} \right) + H + \tau_{ij} (\dot{\varepsilon}_{ij} - \dot{\varepsilon}_{ij}^{el}) \quad (8.9)$$

where c_p is the heat capacity, k the thermal conductivity, H radioactive heat production and the last term denotes shear-heating due to dissipative, non-recoverable processes. We have ignored adiabatic heating, which should however be taken into account in models of mantle convection with the extended Bousinesq approximation. Currently, the numerical code is written for an incompressible rheology, thus $\partial\rho/\partial t = 0$ in eq. (8.1).

8.3 Numerical method

8.3.1 Discretization

The system of equations, eqs. (8.1), (8.8), and eq. (8.9) are discretized using a finite-difference/spectral approach. Differently to conventional approach leading to a 4th-order ordinary differential equation [e.g. Schmalholz et al., 2001], our formulation results in two 2^{nd} -order equations for two unknown functions. The balance equations are always satisfied analytically by choosing σ_{zz} and v_z as the primitive unknown functions. Moreover, the formulation makes the implementation of stress boundary-conditions straightforward. Variables are approximated by a Fourier series in the horizontal direction and by a finite-difference scheme in the vertical direction. Examples of Fourier expansions are [see e.g. Trefethen, 2001] :

$$v_z(x, z) = \sum_{k=-nk/2}^{nk/2} \hat{v}_z^k(z) e^{i\omega x} \quad (8.10)$$

$$v_x(x, z) = \sum_{k=-nk/2}^{nk/2} \hat{v}_x^k(z) e^{i\omega x} \quad (8.11)$$

$$\sigma_{zz}(x, z) = \sum_{k=-nk/2}^{nk/2} \hat{\sigma}_{zz}^k(z) e^{(i\omega x)} \quad (8.12)$$

$$\mu(x, z) = \sum_{l=-nk}^{nk} \hat{\mu}^l(z) e^{(i\omega_l x)} \quad (8.13)$$

$$T(x, z) = \sum_{m=-nm/2}^{nm/2} \hat{T}^m(z) e^{(i\omega_m x)} \quad (8.14)$$

where $i = \sqrt{-1}$, $\omega = 2\pi k/L$, $\omega_l = 2\pi l/L$, $\omega_m = 2\pi m/L$ are wavenumbers and L is the length of the domain under consideration. Note that the length of the viscosity expansion is twice the length of the stress and velocity expansions. Substituting the Fourier expansions into the balance equations (eqns. (8.1)) and collecting factors in front of the exponential term gives:

$$\begin{aligned} i\omega \hat{v}_x^k + \frac{\partial \hat{v}_z^k}{\partial z} &= 0 \\ i\omega \hat{\sigma}_{xx}^k + \frac{\partial \hat{\sigma}_{xz}^k}{\partial z} &= 0 \\ i\omega \hat{\sigma}_{xz}^k + \frac{\partial \hat{\sigma}_{zz}^k}{\partial z} &= \hat{\rho}^k g \end{aligned} \quad (8.15)$$

From these equations, analytical expressions for $\hat{\sigma}_{xx}$ and $\hat{\sigma}_{xz}$ as a function of $\hat{\sigma}_{zz}$ can be written (for $\omega \equiv \omega_m \equiv \omega_l \neq 0$):

$$\begin{aligned} \hat{\sigma}_{xz}^k &= -\frac{1}{i\omega} \left(\frac{\partial \hat{\sigma}_{zz}^k}{\partial z} - \hat{\rho}^k g \right) \\ \hat{\sigma}_{xx}^k &= -\frac{1}{i\omega} \left(\frac{\partial \hat{\sigma}_{xz}^k}{\partial z} \right) = \frac{1}{i\omega} \left(\frac{\partial}{\partial z} \left(\frac{1}{i\omega} \left(\frac{\partial \hat{\sigma}_{zz}^k}{\partial z} - \hat{\rho}^k g \right) \right) \right) \end{aligned} \quad (8.16)$$

The same applies for \hat{v}_x and \hat{v}_z :

$$\hat{v}_x^k = -\frac{1}{i\omega} \left(\frac{\partial \hat{v}_z^k}{\partial z} \right) \quad (8.17)$$

The rheological equation (8.8) involves the time derivative of the stress tensor. This objective derivative is discretized, e.g.

$$\frac{D\tau_{xx}}{Dt} = \frac{\tau_{xx} - \tau_{xx}^{old}}{dt} + \overset{\nabla^{old}}{\tau}_{xx} \quad (8.18)$$

$$\overset{\nabla^{old}}{\tau}_{xx} = \underbrace{\left(v_x \frac{\partial \tau_{xx}}{\partial x} + v_z \frac{\partial \tau_{xx}}{\partial z} \right)}_{\text{advection}} + \underbrace{\left(\tau_{xx}^{old} \cos^2(\theta) + \tau_{zz}^{old} \sin^2(\theta) - \tau_{xz}^{old} \sin(2\theta) \right)}_{\text{rotation}} \quad (8.19)$$

Numerically, the advective terms are solved by a semi-Lagrangian advection scheme, with a second order Runge-Kutta algorithm for backwards tracing [Malevsky and Yuen, 1991]. The rotational terms denote the rotation by an angle θ , computed from the vorticity:

$$\theta = -\frac{1}{2} \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) dt \quad (8.20)$$

We employ the full rotational formula [e.g. Turcotte and Schubert, 1982], which can be shown to simplify to the Jaumann rotation formula for $\theta \rightarrow 0$ (Appendix). Using the definition of eq. 8.18, the rheological equation (8.8) can be spelled out as:

$$\sigma_{xx} + P = 2\mu_{eff} (\dot{\epsilon}_{xx} - \dot{\epsilon}_{xx}^{pl}) + \eta_{eff} \left(\sigma_{xx}^{old} + P^{old} + \nabla_{xx}^{old} \right) \quad (8.21)$$

$$\sigma_{zz} + P = 2\mu_{eff} (\dot{\epsilon}_{zz} - \dot{\epsilon}_{zz}^{pl}) + \eta_{eff} \left(\sigma_{zz}^{old} + P^{old} + \nabla_{zz}^{old} \right) \quad (8.22)$$

$$\sigma_{xz} = 2\mu_{eff} (\dot{\epsilon}_{xz} - \dot{\epsilon}_{xz}^{pl}) + \eta_{eff} \left(\sigma_{xz}^{old} + \nabla_{xz}^{old} \right) \quad (8.23)$$

where

$$\mu_{eff} = \frac{1}{\frac{1}{\mu_{vis}} + \frac{1}{Gdt}} \quad (8.24)$$

$$\eta_{eff} = \frac{1}{1 + \frac{Gdt}{\mu_{vis}}} \quad (8.25)$$

Thus the effective viscosities become dependent on the timestep dt . The rheological equations involve multiplication of Fourier series (convolution), e.g. of the form $\mu_{eff} \frac{\partial v_x}{\partial x}$:

$$\begin{aligned} \mu_{eff}(x, z) \frac{\partial v_x(x, z)}{\partial x} &= \sum_{l=-nk}^{nk} \hat{\mu}_{eff}^l(z) e^{(i\omega_l x)} \sum_{k=-nk/2}^{nk/2} i\omega \hat{v}_x^k(z) e^{(i\omega x)} \\ &= \sum_{l=-nk}^{nk} \sum_{k=-nk/2}^{nk/2} i\omega \hat{\mu}_{eff}^l(z) \hat{v}_x^k(z) e^{(i\frac{2\pi}{L}(k+l)x)} \end{aligned} \quad (8.26)$$

if both velocity harmonics \hat{v}_x^k and viscosity harmonics $\hat{\mu}_{eff}^l$ are known, such a convolution can be performed either by evaluating equation (8.26), or by employing a FFT-based transformation [see e.g. Trefethen, 2001]. In practice, however, only the viscosity harmonics are known, and the velocity harmonics are to be computed. In this

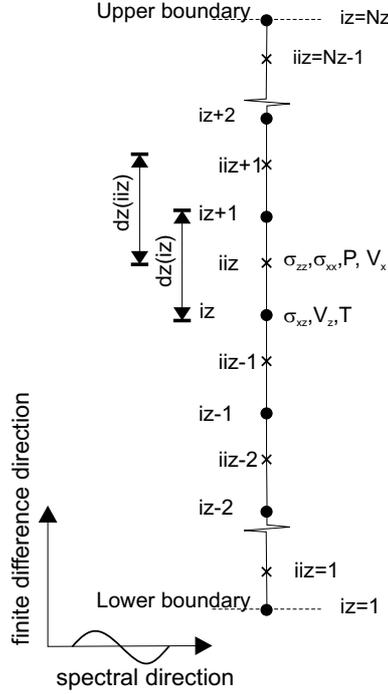


Figure 8.1: Discretization scheme employed in this work.

case, it is advantageous to write the convolution (8.26) in matrix-form, e.g.

$$\begin{pmatrix}
 \hat{\mu}_{eff}^0 & \hat{\mu}_{eff}^{-1} & \hat{\mu}_{eff}^{-2} & \hat{\mu}_{eff}^{-3} & \cdots & \hat{\mu}_{eff}^{-nk+2} \\
 \hat{\mu}_{eff}^1 & \hat{\mu}_{eff}^0 & \hat{\mu}_{eff}^{-1} & \hat{\mu}_{eff}^{-2} & \cdots & \vdots \\
 \hat{\mu}_{eff}^2 & \hat{\mu}_{eff}^1 & \hat{\mu}_{eff}^0 & \hat{\mu}_{eff}^{-1} & \cdots & \vdots \\
 \vdots & \cdots & \hat{\mu}_{eff}^1 & \hat{\mu}_{eff}^0 & \hat{\mu}_{eff}^{-1} & \hat{\mu}_{eff}^{-2} \\
 \vdots & \cdots & \hat{\mu}_{eff}^2 & \hat{\mu}_{eff}^1 & \hat{\mu}_{eff}^0 & \hat{\mu}_{eff}^{-1} \\
 \hat{\mu}_{eff}^{nk-2} & \cdots & \hat{\mu}_{eff}^3 & \hat{\mu}_{eff}^2 & \hat{\mu}_{eff}^1 & \hat{\mu}_{eff}^0
 \end{pmatrix}
 \begin{pmatrix}
 i\omega \hat{v}_x^{-nk/2+1} \\
 \vdots \\
 i\omega \hat{v}_x^{-1} \\
 i\omega \hat{v}_x^1 \\
 \vdots \\
 i\omega \hat{v}_x^{nk/2-1}
 \end{pmatrix}
 =
 \begin{pmatrix}
 \hat{R}^{-nk/2+1} \\
 \vdots \\
 \hat{R}^{-1} \\
 \hat{R}^1 \\
 \vdots \\
 \hat{R}^{nk/2-1}
 \end{pmatrix}
 \quad (8.27)$$

Note that convolved harmonics with $R^k \leq -nk/2$ and $R^k \geq nk/2$ are dropped in order to reduce the aliasing effect [Trefethen, 2001]. Moreover, this spectral approach only solves for perturbations, so that the 0^{th} harmonics are not computed.

The rheological equations can be split into a part for normal and a part for shear stresses:

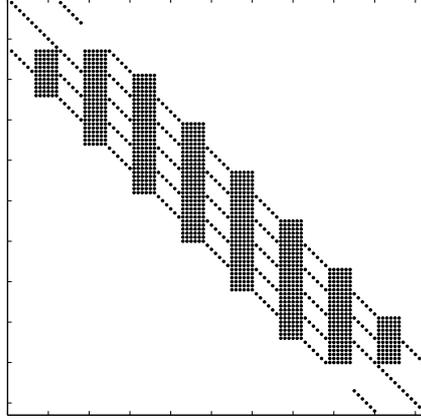


Figure 8.2: Sparseness of solution matrix \mathbf{A} for a problem with variable viscosity.

$$\begin{aligned}
 \hat{\sigma}_{xx}^k - \hat{\sigma}_{zz}^k &= 2\hat{\mu}_{eff}^l \left(\left(i\omega\hat{v}_x^k - \hat{\varepsilon}_{xx}^{p,k} \right) - \left(\frac{\partial\hat{v}_z^k}{\partial z} - \hat{\varepsilon}_{zz}^{p,k} \right) \right) + \hat{\eta}_{eff}^l \left(\begin{matrix} \nabla^{old,k} \\ \hat{\sigma}_{xx} - \hat{\sigma}_{zz} \end{matrix} \right) \\
 \hat{\sigma}_{xz}^k &= \hat{\mu}_{eff}^l \left(\frac{\partial\hat{v}_x^k}{\partial z} + i\omega\hat{v}_z^k - \hat{\varepsilon}_{xz}^{p,k} \right) + \hat{\eta}_{eff}^l \left(\begin{matrix} \nabla^{old,k} \\ \hat{\sigma}_{xz} \end{matrix} \right)
 \end{aligned} \tag{8.28}$$

Substituting eq. (8.16) and (8.17) into eqns (8.28) results in two equations for the two unknowns \hat{v}_z^k and $\hat{\sigma}_{zz}^k$. These equations can be discretized with a conservative finite difference scheme on a staggered grid (see fig. 8.1). If the convolution matrix (8.27) is used throughout, the solution of the system of equations, together with the appropriate boundary conditions, can be written as

$$\mathbf{A}\mathbf{f} = \mathbf{RHS} \tag{8.29}$$

where \mathbf{A} is a sparse matrix (illustrated in fig. 8.2), \mathbf{f} is a vector composed of unknown coefficients $\hat{\sigma}_{zz}^k$ and \hat{v}_z^k , and \mathbf{RHS} is a vector that can be computed from eqns. 8.28 taking as an initial guess, e.g. $\sigma_{zz}(x, z) = v_z(x, z) = 0$.

The energy equation is also solved with a spectral/finite difference scheme. The spectral resolution of temperature, however, may be larger than the resolution of the mechanical equations, which is advantageous when performing computations with a

temperature-dependent viscosity. Advective terms in the energy equation are solved with a semi-Lagrangian method [Malevsky and Yuen, 1991]. The remaining (diffusive) part of the energy equation can thus be written as:

$$\frac{\hat{T}^m - \hat{T}_{old}^m}{dt} = \kappa \left(\frac{\partial^2 \hat{T}^m}{\partial z^2} - \omega^2 \hat{T}^m \right) + \hat{H}_{source}^m \quad (8.30)$$

where \hat{H}_{source}^m are heat sources due to mechanical dissipation and radioactive heating, and κ is the thermal diffusivity. This equation can be solved on a wavenumber-by-wavenumber basis.

8.3.2 Iterative approach

If the direct method is employed, the system of equations that has to be solved scales with the number of gridpoints in vertical direction n_z and the number of Fourier harmonics in horizontal direction n_k as $n_k(2n_z - 1)$. Such a system can quickly become large, and it may be advantageous to switch to an iterative solution scheme. The iterative scheme does not use the full bandwidth matrix (eq. 8.29), but employs a reduced matrix instead. There are several ways to compose this matrix. Christensen and Harder [1991] and Marquart [2001] employ as initial guess μ_{eff}^0 (mean viscosity). Here we employ a more general method and allow the bandwidth to be varied (for example 2,4 etc.). The algorithm is best illustrated by an example. Suppose we want to solve eq. (8.27) by an iterative approach and assume that $nk = 8$. If only the mean viscosity is employed on the left-hand-side a pseudocode gives:

- 1: **while** max(|**RHS**|) > tolerance **do**
- 2: **RHS** = **RHS**^{*i*} - **Af**
- 3: Compute for $k=-3..3$ (but $k \neq 0$)

$$(\mu_{eff}^0) (\mathbf{df}^{-3}) = (\mathbf{RHS}^{-3})$$

$$(\mu_{eff}^0) (\mathbf{df}^{-2}) = (\mathbf{RHS}^{-2})$$

etc.

- 4: **f** = **f** + χ **df**
- 5: **end while**

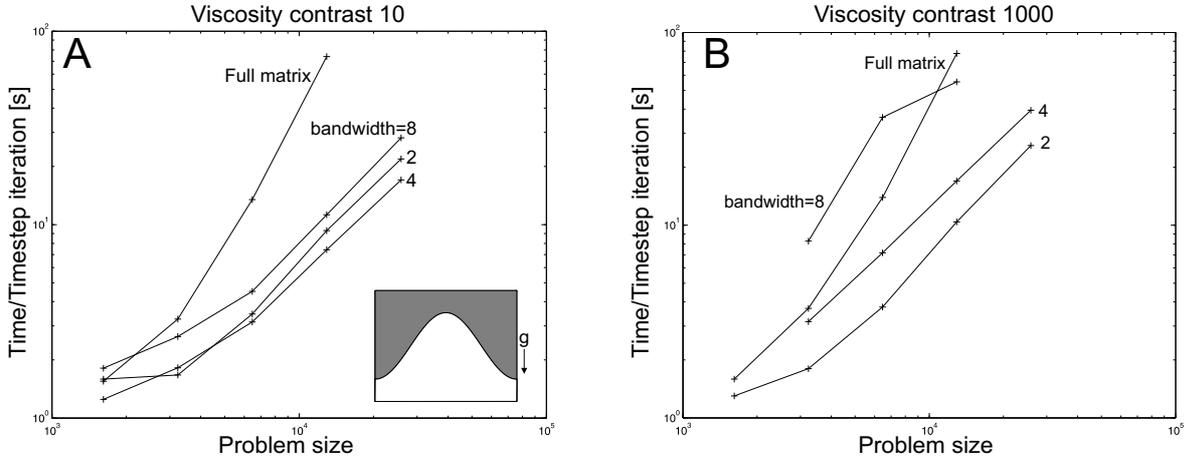


Figure 8.3: Performance of the code for a Rayleigh-Taylor type problem with a highly viscous overburden for two different viscosity contrasts as a function of problem size. Note that this example represents a 'worse-case' scenario, since in typical model runs the instability is initiated with a smaller initial amplitude and that thus a better initial guess exists for the iterative solver.

where \mathbf{RHS}^i is the initial right-hand-side, and χ is a relaxation parameter ($\chi < 1$), required for the iterations to converge.

If a bandwidth of 2 is employed, a pseudocode is given by:

- 1: **while** $\max(|\mathbf{RHS}|) > \text{tolerance}$ **do**
- 2: $\mathbf{RHS} = \mathbf{RHS}^i - \mathbf{A}\mathbf{f}$
- 3: Compute

$$\begin{pmatrix} \mu_{eff}^0 & \mu_{eff}^{-1} \\ \mu_{eff}^1 & \mu_{eff}^0 \end{pmatrix} \begin{pmatrix} \mathbf{f}^{-1} \\ \mathbf{f}^1 \end{pmatrix} = \begin{pmatrix} \mathbf{RHS}^{-1} \\ \mathbf{RHS}^1 \end{pmatrix}$$

$$\begin{pmatrix} \mu_{eff}^0 & \mu_{eff}^{-2} \\ \mu_{eff}^2 & \mu_{eff}^0 \end{pmatrix} \begin{pmatrix} \mathbf{f}^{-2} \\ \mathbf{f}^2 \end{pmatrix} = \begin{pmatrix} \mathbf{RHS}^{-2} \\ \mathbf{RHS}^2 \end{pmatrix}$$

etc.

- 4: $\mathbf{f} = \mathbf{f} + \chi \mathbf{df}$
- 5: **end while**

With such an approach, we can arbitrarily vary the bandwidth and switch from a direct solver to an iterative solver. Whether the direct solver or the iterative solver has a better performance, is dependent on the type of problem (periodic or non-periodic),

the effective lateral viscosity contrast (better performance for small viscosity contrasts), the problem size, the timestep, the 'initial guess', the smoothness of the lateral viscosity contrast, etc. Examples of the performance of the code are given in figure 8.3 for a RT-type problem. Interestingly, the use of two harmonics is most efficient for large viscosity contrasts and problem sizes.

8.3.3 Plasticity formulation

Under large differential stresses rocks may fail plastically. For upper-crustal conditions, the maximum deviatoric stress is given by the Byerlee law, which can be described by a Mohr-Coulomb failure criteria. Under high-stress, upper mantle conditions, a Peierls-creep mechanism is more applicable. This mechanism can be described, as a first approximation, by a pressure-independent Von Mises yield criteria (see Chapter 4), which can in turn be shown to be identical to a Mohr-Coulomb model with zero friction angle (in 2D plane strain). Therefore, we have chosen to incorporate the Mohr-Coulomb plasticity model in the current code.

Plastic yielding occurs if stresses are higher than the yield envelop (i.e. $F > 0$ in equation (6)). If this is detected, stresses are pointwise returned to the yield envelop in the direction of the plastic flow (equation 4). Plastic strain-rates for the Mohr-Coulomb model in the two-dimensional case can be computed from (8.4) and (8.6) together with the requirement that they are objective (independent on reference-frame):

$$\begin{aligned}\dot{\varepsilon}_{xx}^{pl} &= \dot{\lambda} \left(\frac{\sigma_{xx} - \sigma_{zz}}{4\tau^*} + \frac{1}{2} \sin(\psi) \right) \\ \dot{\varepsilon}_{zz}^{pl} &= \dot{\lambda} \left(-\frac{\sigma_{xx} - \sigma_{zz}}{4\tau^*} + \frac{1}{2} \sin(\psi) \right) \\ \dot{\varepsilon}_{xz}^{pl} &= \dot{\lambda} \left(\frac{\sigma_{xz}}{2\tau^*} \right)\end{aligned}\tag{8.31}$$

where $\dot{\lambda}$ is a to-be-determined scalar. From these expressions one can verify that the incompressible case is retained only if $\psi = 0$ (since only then $\dot{\varepsilon}_{xx}^{pl} + \dot{\varepsilon}_{zz}^{pl} = 0$). The change in deviatoric stress due to plasticity is then proportional to $\mu_{eff} \dot{\varepsilon}^{pl}$. Since in this case pressure does not change during return mapping to the yield surface, the change in stress is of equal magnitude as the change in deviatoric stress. A given stress state $(\sigma_{xx}^{tr}, \sigma_{zz}^{tr}, \sigma_{xz}^{tr})$ which is outside the yield surface $F(\sigma^{tr}) > 0$, can be returned to the yield surface $(\sigma_{xx}^y, \sigma_{zz}^y, \sigma_{xz}^y)$ with a plastic stress increment $(\Delta\sigma_{xx}^{pl}, \Delta\sigma_{zz}^{pl}, \Delta\sigma_{xz}^{pl})$. These

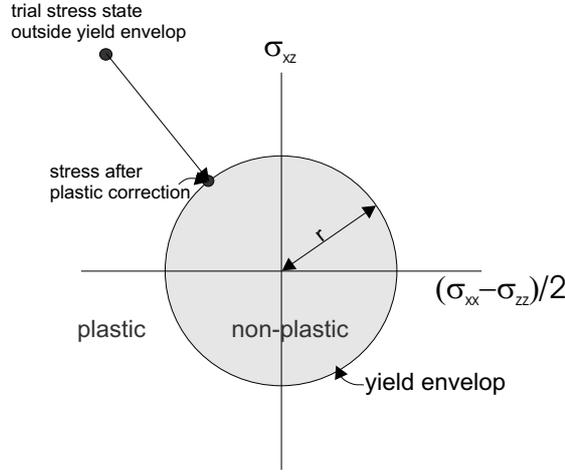


Figure 8.4: Mohr-Coulomb return mapping. The yield envelop for a Mohr-Coulomb material plotted in $\sigma_{xz} - (\sigma_{xx} - \sigma_{zz})/2$ space is a circle with radius $r = -\sin(\phi)(\sigma_{xx} + \sigma_{zz})/2 + C\cos(\phi)$. A given stress $(\sigma_{xx}^{tr}, \sigma_{zz}^{tr}, \sigma_{xz}^{tr})$ that is outside the yield envelop, should be returned to the yield envelop $(\sigma_{xx}^y, \sigma_{zz}^y, \sigma_{xz}^y)$. The plastic flow direction points towards the center of the circle, and the correct stress state is thus the crossing point between the circle and the radial line. If the dilation angle is zero, the circle radius stays constant during return mapping. In the other case the radius increases in size.

plastic stress increments can be found by computing the intersection of a line with a circle (fig. 8.4) and are given by:

$$\begin{aligned}\Delta\sigma_{xz}^{pl} &= (1-f)\sigma_{xz}^{tr} \\ \Delta\sigma_{xx}^{pl} &= -(1-f)\left(\frac{\sigma_{xx}^{tr} - \sigma_{zz}^{tr}}{2}\right) \\ \Delta\sigma_{zz}^{pl} &= (1-f)\left(\frac{\sigma_{xx}^{tr} - \sigma_{zz}^{tr}}{2}\right)\end{aligned}\quad (8.32)$$

where

$$f = \frac{-\frac{(\sigma_{xx}^{tr} + \sigma_{zz}^{tr})}{2}\sin(\phi) + C\cos(\phi)}{\sqrt{\left(\frac{\sigma_{xx}^{tr} - \sigma_{zz}^{tr}}{2}\right)^2 + (\sigma_{xz}^{tr})^2}}\quad (8.33)$$

Once plastic stress increments are known, plastic strainrates can be computed. Our current plasticity implementation is similar to the explicit FLAC way of implementing plasticity [e.g. Poliakov et al., 1994, Huismans et al., 2001], i.e. all plastic terms go

to the right-hand-side. In future, this approach should be made implicit, which would allow to take larger timesteps during plastic yielding. Moreover, more work is required to understand the effect of visco-elasticity on stabilizing the width of shearzones. Additional effects like fluid flow, or shear-heating (see chapter 4), are probably required to obtain a numerically resolvable stable width of shearzones.

8.3.4 Further numerical issues

The computational grid is moved in a Lagrangian fashion if pure-shear boundary conditions are employed. Boundaries between materials with a sharp variation of properties are described by marker-chains that are moved through the computational grid by an implicit time step algorithm. Markers are added if the distance between two markers exceeds a given threshold. Fourier coefficients of effective viscosity and density fields are calculated directly from this marker-chain. An advantage of such a method over particle-based methods is that less points are needed to describe materials, and hence the algorithm is faster. A possible disadvantage is that it is more difficult to handle cases where one continuous layer becomes divided into several domains (although it is possible to check for self-intersections of the markerchain, and construct these domains). In the problems we solved sofar, there has been no need for such a method.

The current formulation makes it straightforward to implement no-slip or free-slip boundary-conditions. In addition, the implementation of stress-free boundary-conditions or any condition involving normal stresses does not require 3^{rd} derivatives. A free surface boundary condition is difficult to incorporate in Eulerian fixed grid numerical codes such as the one described here. Therefore we approximate the free surface by putting a layer of low viscosity in the upper part of the numerical box, which is overlain by a horizontal stress-free (also called fast-erosion or fast-redistribution) upper boundary condition. This boundary condition has an advantage over free-slip or no-slip conditions, since it makes the upper boundary 'permeable', and hence the 'air' layer can be thinner than in the other cases.

The energy equation and the mechanical equations are solved sequentially, with iterations to account for the nonlinear effects of shear-heating. Time-stepping is done with an implicit algorithm. Since we employ a semi-Lagrangian advection scheme, there are no numerical timestep-restrictions (e.g. by the CFL-criteria). If the effects

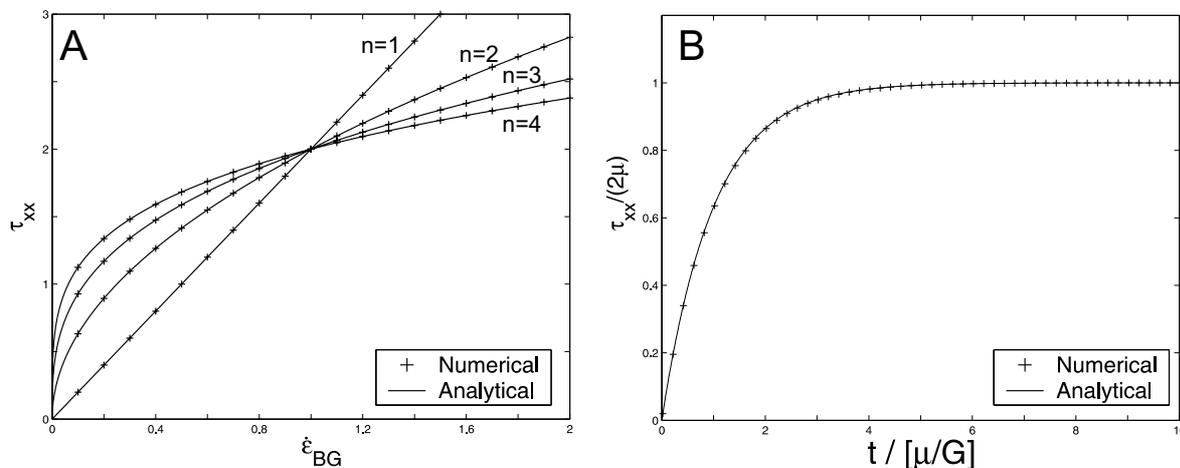


Figure 8.5: A) Deviatoric stress as a function of strainrate for a viscous rheology and different values of the stress exponent n . B) Deviatoric stress versus time for a Maxwell linear viscoelastic rheology.

of visco-elasticity are studied, however, the timestep should be less than the Maxwell relaxation time ($= \mu/G$).

8.4 Verification of the numerical code

This section presents several benchmark tests that have been made to verify various aspects of the numerical code.

8.4.1 0D rheology tests

The first test considers the effect of viscous or powerlaw rheology on deviatoric stress. If pure-shear is applied to a homogeneous, viscous, material with a strainrate $\dot{\epsilon}_{BG}$, the deviatoric stress should depend on the strainrate as:

$$\tau_{xx} = 2B\dot{\epsilon}_{BG}^{\frac{1}{n}} \quad (8.34)$$

where n is the stress exponent, B a material constant and compressive stresses are taken to be positive (fig. 8.5A).

If the rheology is linear Maxwell viscoelastic, rather than viscous, and pure-shear

deformation is applied, the governing equations are given by:

$$\frac{1}{2\mu}\tau_{xx}(t) + \frac{1}{2G}\frac{D\tau_{xx}(t)}{Dt} = \dot{\epsilon}_{BG} \quad (8.35)$$

under pure-shear, rotational terms of the Jaumann derivative disappear, and assuming that $\tau_{xx}(0) = 0$ and $\dot{\epsilon}_{BG} = \text{constant}$, a solution to (8.36) is given by:

$$\tau_{xx}(t) = 2\mu\dot{\epsilon}_{BG} \left(1 - e^{-\frac{G}{\mu}t}\right) \quad (8.36)$$

A comparison between numerics and analytics is shown on figure 8.5B.

A Maxwell rheology under simple shear experiences the effect of the rotational components. The governing rheological equations for a 2D plane strain case are (with a Jaumann derivative of the stress tensor):

$$\begin{aligned} \frac{1}{2\mu}\tau_{xx} + \frac{1}{2G}\left(\frac{\partial\tau_{xx}}{\partial t} - \tau_{xz}\dot{\gamma}\right) &= 0 \\ \frac{1}{2\mu}\tau_{zz} + \frac{1}{2G}\left(\frac{\partial\tau_{zz}}{\partial t} + \tau_{xz}\dot{\gamma}\right) &= 0 \\ \frac{1}{2\mu}\tau_{xz} + \frac{1}{2G}\left(\frac{\partial\tau_{xz}}{\partial t} + \left(\frac{\tau_{zz} - \tau_{xx}}{2}\right)\dot{\gamma}\right) &= \frac{1}{2}\dot{\gamma} \end{aligned} \quad (8.37)$$

where $\dot{\gamma} = \frac{\partial v_x}{\partial z}$ is assumed to be constant. Under the assumption that $\tau_{xx}(0) = \tau_{zz}(0) = \tau_{xz}(0) = 0$, one can derive an analytical solution for the stress evolution versus time:

$$\begin{aligned} \tau_{xx}(t) &= -\frac{\dot{\gamma}\mu G \left((\mu\dot{\gamma}\cos(\dot{\gamma}t) + G\sin(\dot{\gamma}t)) \exp\left(-\frac{Gt}{\mu}\right) - \mu\dot{\gamma} \right)}{\mu^2\dot{\gamma}^2 + G^2} \\ \tau_{xz}(t) &= -\frac{\dot{\gamma}\mu G \left((G\cos(\dot{\gamma}t) - \mu\dot{\gamma}\sin(\dot{\gamma}t)) \exp\left(-\frac{Gt}{\mu}\right) - G \right)}{\mu^2\dot{\gamma}^2 + G^2} \\ \tau_{zz}(t) &= -\tau_{xx}(t) \end{aligned} \quad (8.38)$$

These expressions exhibit a periodic oscillation which is damped on the Maxwell time-scale. A comparison of numerically computed values and the analytical solution is given on figure 8.6A. Stress rotation is significant for the given parameters. But does this mean that the rotational terms are important for geodynamic settings? In order to answer this question, it is useful to study the stress evolution during initial stages. A Taylor expansion of eqns. (8.38) around $t = 0$, reveals that $\tau_{xz} \propto \dot{\gamma}Gt$ whereas $\tau_{xx} \propto 0.5\dot{\gamma}^2Gt^2$. Rocks (or basically any material) cannot sustain differential stresses

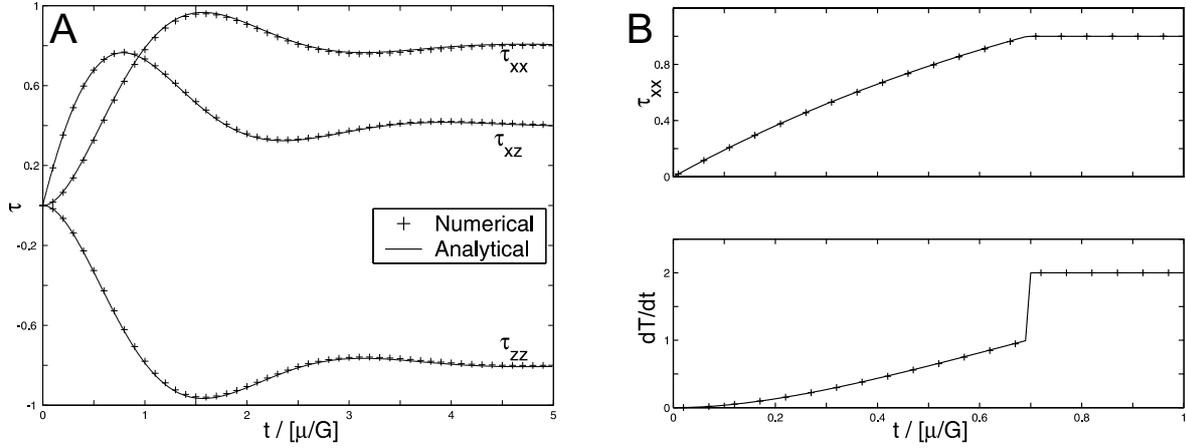


Figure 8.6: A) Simple shear deformation of a Maxwell viscoelastic material with $\mu = 2$, $G = 1$, $\dot{\gamma} = 1$. B) Stress evolution (upper panel) and increase in temperature (lower panel) during pure-shear deformation of a visco-elasto-plastic material with $G = 1$, $\mu = 2$, $\sigma^y = 1$, $\dot{\epsilon}_{BG} = 1$, $\rho c_p = 1$.

larger than roughly one-tenth of the elastic shear modulus ($0.1/G$), from where the bonds between atoms start to break [see e.g. Scholz, 2002]. In the given experiment such a stress is reached after $t \approx 0.1/\dot{\gamma}$. The other stress components are then on the order of $0.005G$ (which is 5% of the shear stress). Thus for the given setup and for geodynamically reasonable stresses, rotational terms make a rather small contribution to the overall stress state.

The last 0-D test discussed here considers the interaction between shear-heating, plasticity and viscoelasticity in pure-shear settings. The governing equations for an incompressible material are:

$$\begin{aligned}
 \rho c_p \frac{\partial T}{\partial t} &= 2\tau_{xx}(t) \left(\dot{\epsilon}_{BG} - \frac{1}{2G} \frac{\partial \tau_{xx}(t)}{\partial t} \right) \\
 \dot{\epsilon}_{BG} &= \frac{1}{2\mu} \tau_{xx}(t) + \frac{1}{2G} \frac{\partial \tau_{xx}(t)}{\partial t} \\
 \tau_{xx}(t) &= \min(\tau_{xx}(t), \sigma^y)
 \end{aligned} \tag{8.39}$$

where a Von Mises plasticity model with yield stress σ^y was employed. Temperature increases due to dissipative (viscous and plastic) processes. This set of ODE's is most readily solved numerically (MATLAB scripts are given in the appendix of Chapter 4). A good agreement exists between the (semi-) analytical solution and results obtained

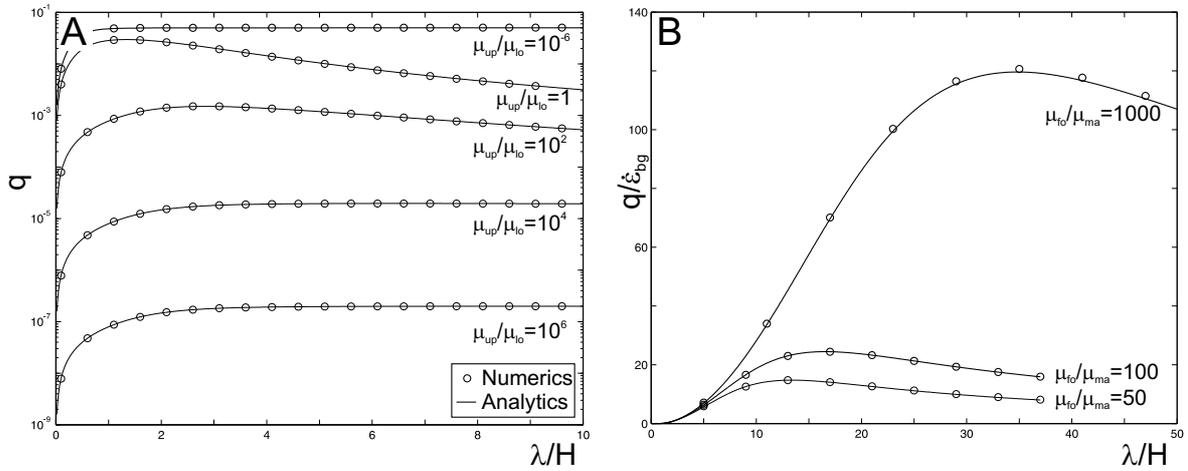


Figure 8.7: A) Growthrate spectra as a function of dimensionless wavelength for the RT instability with a free-slip upper and lower boundary conditions, and a mean thickness of the low-density layer of $0.2H$. Resolution employed is 201×8 . B) Growthrate spectra for single-layer viscous folding under pure shear compression (with a background strainrate $\dot{\epsilon}_{BG}$). Resolution employed is 1001×8 .

with the finite-difference/spectral method (fig. 8.6B and C). Note that shear-heating increases significantly after plastic yielding, which is caused by vanishing elastic strain rates.

8.4.2 Linear stability analysis of folding and RT-instability

A classical (quasi one-dimensional) test for viscous flow problems is the linear stage of the Rayleigh-Taylor instability, for which analytical solutions are available for a range of boundary conditions [e.g. Biot and Odé, 1965, Turcotte and Schubert, 1982, appendix chapter 7]. Results are shown for a free-slip upper and lower boundary condition and for viscosity contrasts between the upper and the lower fluid, which vary from 10^6 to 10^{-6} (fig. 8.7A). The numerical results have a maximum deviation of 1.5% with the analytical results. Additional benchmarks can be found in Chapters 3, 6 and 7.

Similarly, analytical solutions exist for single layer folding of a highly viscous layer embedded in an infinite viscous matrix [e.g. Fletcher, 1977] under overall pure-shear conditions. Good agreement exists between numerical and analytical results (fig. 8.7B).

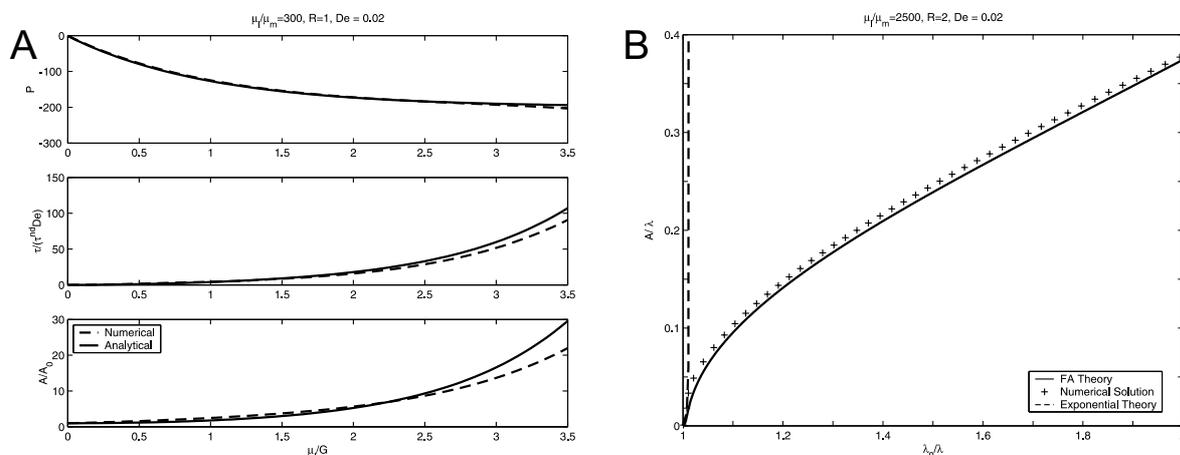


Figure 8.8: A) Initial stages of buckling of a viscoelastic layer embedded in a viscoelastic matrix, with $R = 1$. B) Finite amplitude behavior of buckling of a layer with $R = 2$. $\dot{\epsilon}_{BG} = 1$, $\mu_l = 100$ in both cases.

8.4.3 Viscoelastic buckling

The importance of elasticity in a Maxwell viscoelastic material is classically indicated by the Deborah number, $De = 2\mu_l\dot{\epsilon}_{BG}/G$ (where μ_l is the viscosity, G the elastic shear module and $\dot{\epsilon}_{BG}$ the applied background strainrate). If $De > 1$, elasticity is assumed to be important, whereas for $De < 1$, viscous deformation dominates. Schmalholz and Podladchikov [1999] studied the effect of elasticity on the folding instability and demonstrated that the De -number gives a weak indication, since strainrates during highly unstable initial phases of folding are typically much larger than the background strainrate. They demonstrated that a different parameter, $R = [\mu_l/6\mu_m]^{1/3}(4\mu_l\dot{\epsilon}_{BG}/G)$, should be employed to indicate the importance of elasticity for folding of low- De materials (where $R > 1$ denotes elastic response). Figure 8.8A shows a comparison for mean stress, coefficient of fibre stress and amplitude versus time for the initial stages of folding of a viscoelastic layer with $R = 1$ [see also Schmalholz et al., 2001].

In Schmalholz [2000] a finite amplitude theory of folding of viscous and viscoelastic layers was derived. Good agreement exists between numerical results and the FA theory (fig. 8.8B).

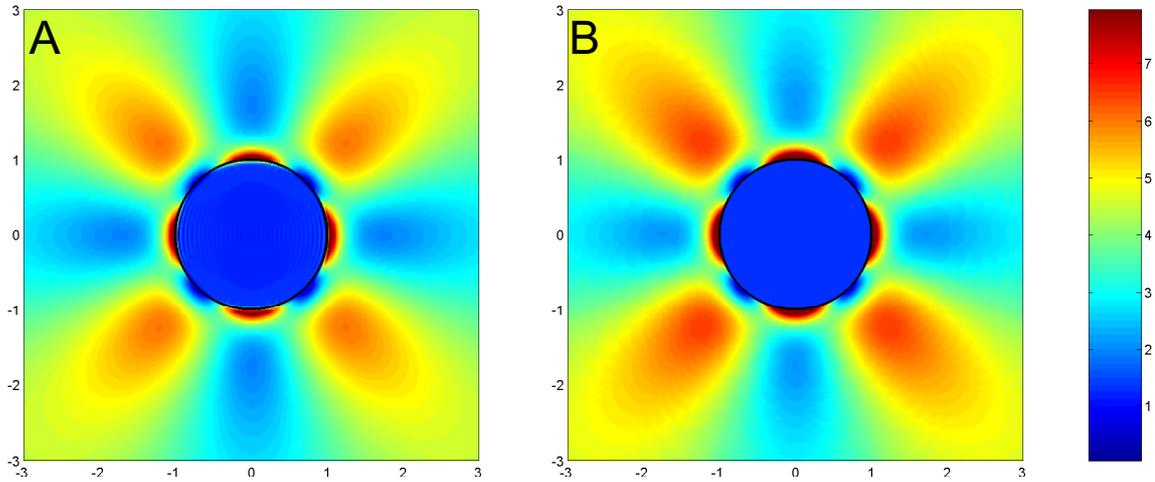


Figure 8.9: Shear heating ($\tau_{ij}\dot{\epsilon}_{ij}$) around a weak circular inclusion under pure-shear compression. Rheology is incompressible viscous and the viscosity contrast between the matrix and the inclusion is 10. A) Numerical result. B) Analytical result. Resolution 501×512 .

8.4.4 Stress around a weak inclusion

Most of the previous tests are 0D or 1D. Only few analytical solutions are two-dimensional. A solution for the stress and strainrate distribution around 2D elliptical inclusions in an incompressible viscous or elastic material is described in Schmid [2002] and Schmid and Podladchikov [2003]. A comparison of numerical results of the amount of shear heating with the analytical solution is shown on fig. 8.9.

8.4.5 Diapirism

A benchmark study for the finite amplitude stages of the RT-instability was presented by van Keken et al. [1997]. A thin viscous low-density layer underlies a high-density material. The interface between the two layers is given by $\Gamma(x) = 0.2 + 0.02\cos(\pi x/0.9142)$. The results are shown on figure 8.10 and compare well with results of [van Keken et al., 1997].

It is interesting to study the evolution of the RT instability with a free-surface upper boundary. The Eulerian-based numerical code described here cannot treat such a surface in a self-consistent way. Instead, a low viscosity and density layer is added on

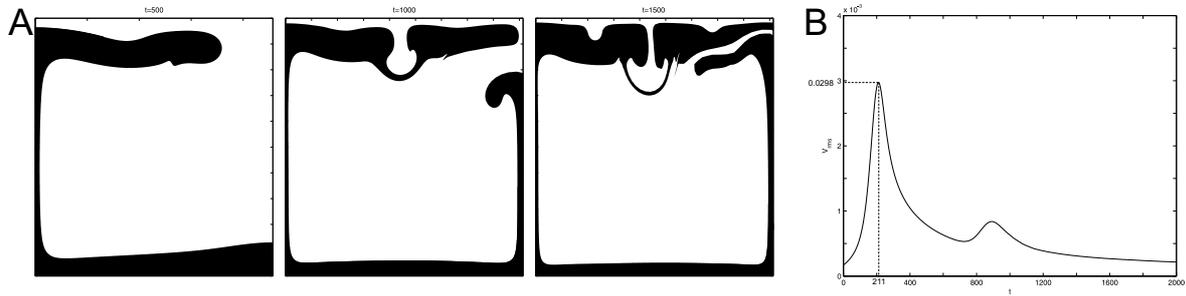


Figure 8.10: RT-benchmark for an isoviscous case. A) geometry at various times. B) Evolution of mean velocity versus time. Results compare well with the results of van Keken et al. [1997]. Resolution 101×128 . Only half of the computed domain is shown.

top of the numerical domain to mimic a layer of 'air'. It is however unclear what the thickness and viscosity of this air layer should be in order to mimic the free surface successfully. For this purpose, we compare solutions with different air viscosity and thicknesses with a solution obtained with a standard FEM code (SloMo, see Appendix B), which is able to compute the free-surface evolution naturally. The initial setup consists of a thin layer of viscosity 0.1 and density 9, overlain by a thicker layer of viscosity 1 and density 10 [see also Poliakov and Podladchikov, 1992]. The initial interface between the two layers is given by $\Gamma(x) = 0.3 + 0.04 \cos(2\pi x/1.68)$. Both the finite element and the finite difference/spectral code are run with an implicit timestepping algorithm and convergence tests have been performed w.r.t. timestep. The time-evolution of the interface reveals that the Eulerian code always slightly 'lags behind' the Lagrangian code (fig. 8.11A). This can also be observed in the time-evolution of the topography and diapir-head (fig. 8.11B). It can be explained by the fact that the Eulerian approach involves interpolation of velocity from nodal points to the markerchain at every timestep. This results in interpolation errors which typically reduce the maximum velocity, and thus reduce the speed of the interface movement. Other explanations are interpolation errors due to the reconstruction of density and viscosity fields from the markerchain or the iterative solver. These last two explanations may be ruled out since (1) no significant differences have been observed between runs that were converged to various accuracy-levels and (2) the same effect has been observed in constant viscosity cases, with a free slip upper surface (where no iterations have to be performed).

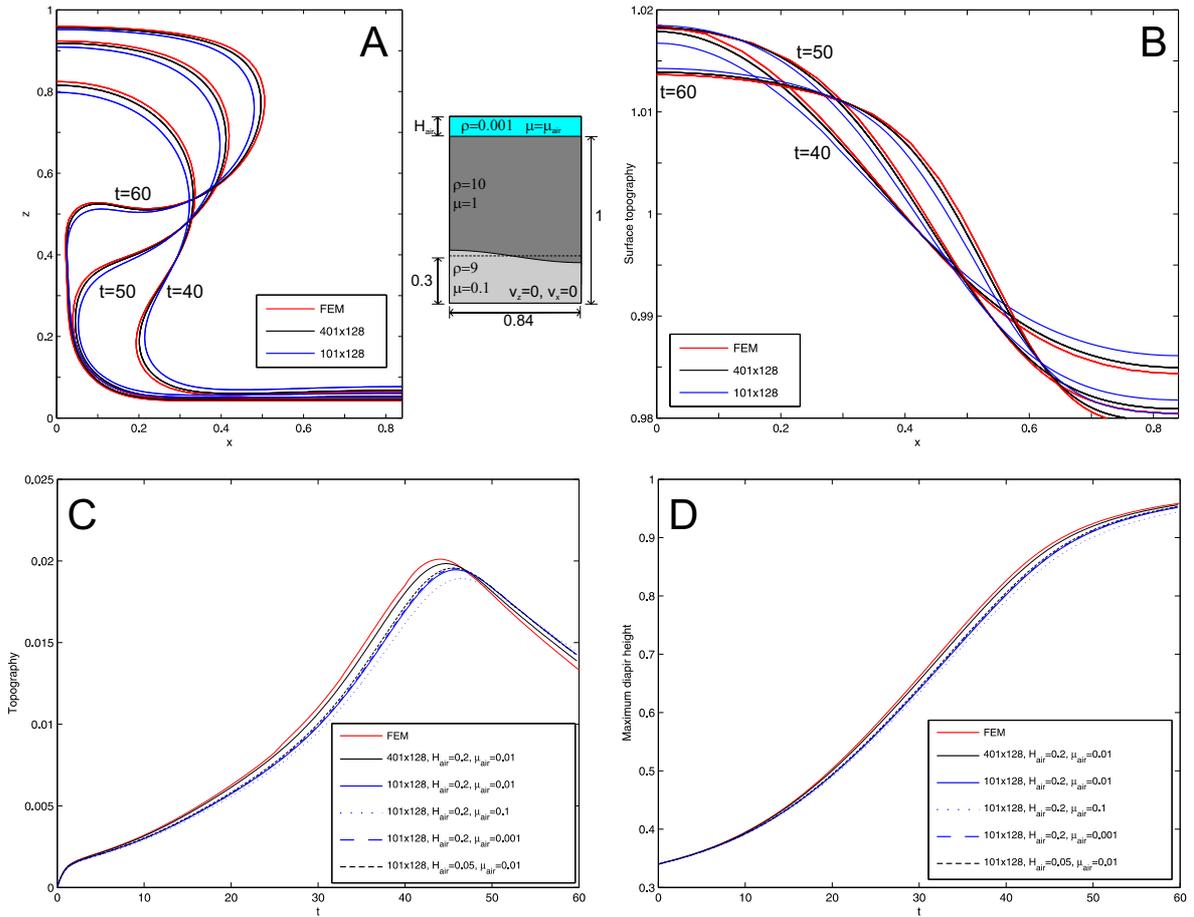


Figure 8.11: A) Evolution of the interface from a FEM simulation and from the spectral/finite-difference code at various resolutions. Inset illustrates the initial setup. Different upper boundary conditions are employed: the FEM code employs a free-surface condition. The eulerian code adds a air layer with thickness H_{air} and viscosity μ_{air} . On top of this air layer a fast-erosion ($\sigma_{xz} = 0, \sigma_{zz} = 0$) boundary condition is employed. B) Topography (both A) and B) with $\mu_{air} = 0.01$ and $H_{air} = 0.2$). C) Maximum amplitude of the diapir versus time. D) Maximum topography versus time.

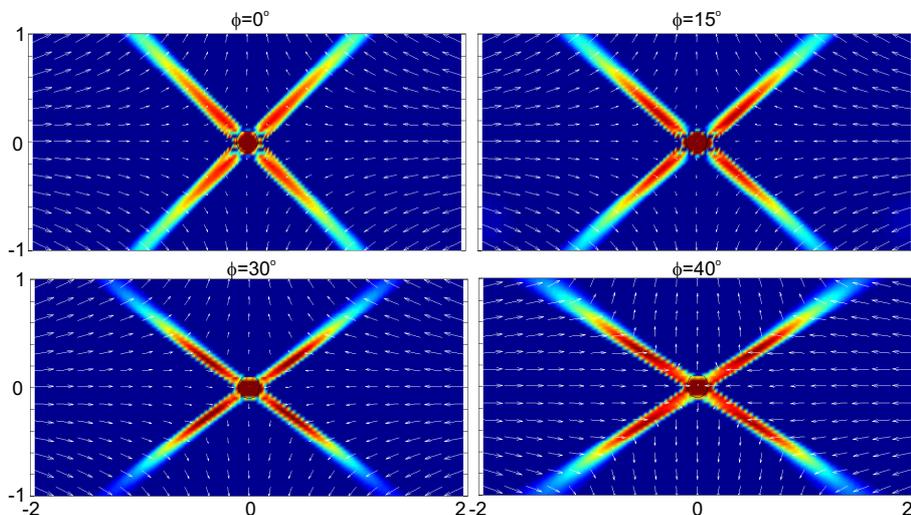


Figure 8.12: Second invariant of strainrate tensor (red=large) for shearzones that initiate from a weak inclusion with $\mu = 0.01$, and for different friction angles ϕ . Other parameters are $\dot{\epsilon}_{BG} = 1$, $G = 10$, $\mu = 100$ and cohesion is 0.1. Snapshots are shown after 2.2% compression. Resolution is 101×256 .

An interesting point is the required viscosity and thickness of the upper layer in order to successfully mimic a free surface. The results (fig. 8.11C and D) indicate that the viscosity should be at least 100-1000 times smaller than the viscosity of the upper layer. The thickness of the upper layer is not crucial if a no-stress (or 'fast erosion') upper boundary condition is employed.

8.4.6 Shearband inclination

Non-associated Mohr-Coulomb plasticity causes shearbands to nucleate at angles different than 45° from the maximum compression direction [e.g. Vermeer and de Borst, 1984, Vermeer, 1990, Poliakov et al., 1994]. Vermeer [1990] demonstrated analytically that the optimal orientation of a shearband is between the Roscoe angle ($45^\circ - \psi/2$) and the Coulomb angle ($45^\circ - \phi/2$). Here, we perform tests in which the friction angle is varied and the dilation angle $\psi = 0$ (consistent with the incompressible formulation). The shearbands (fig. 8.12) initiate at an angle θ with $45^\circ - \phi/2 < \theta < 45^\circ - \psi/2$, which is consistent with bifurcation theory [e.g. Vermeer, 1990] and previous numerical results [e.g. Poliakov et al., 1994].

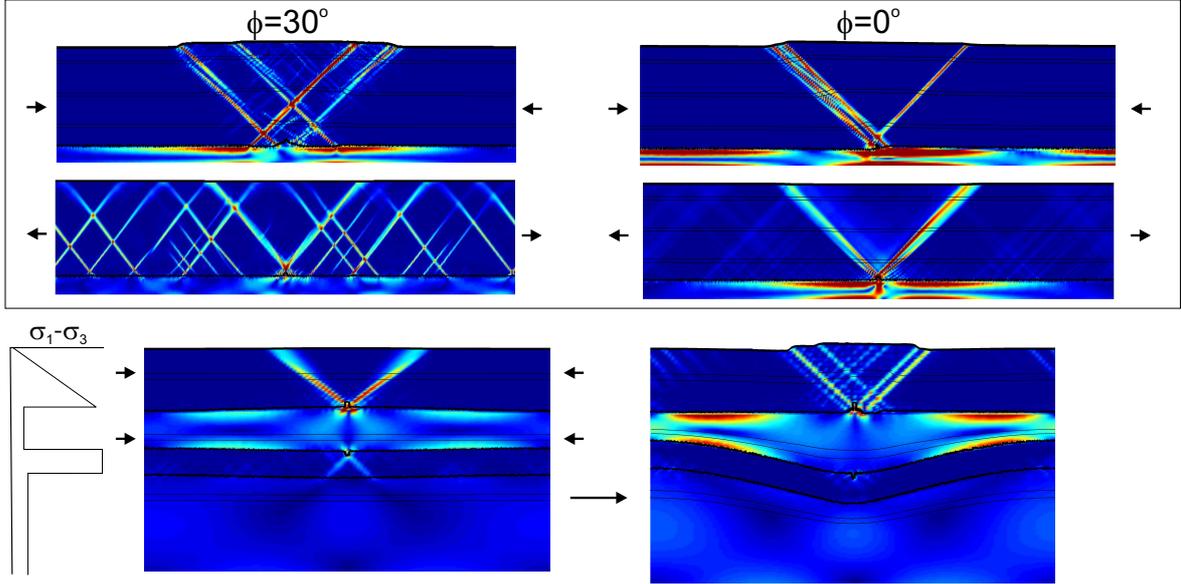


Figure 8.13: Second invariant of strainrate tensor for both compression and extension of a brittle upper crust, deforming on top of a viscous lower crust for different values of the friction angle ϕ (upper panels). Lower panels shows the same, but on a lithospheric scale where the upper mantle has a friction angle of 0° , and the upper crust a friction angle of 30° .

An example of combined visco-elasto-plastic deformation on a lithospheric scale is given on figure 8.13. Plastic yielding is initiated by a small notch, and by random noise at the interface between brittle and ductile regimes. Regions with a friction angle > 0 tend to deform by plastic yielding; the upper mantle (with zero friction angle) initially deforms by plastic yielding but later by buckling. Incorporation of additional strain weakening mechanisms (e.g. by shear heating) may change this picture.

8.4.7 Thermal advection and diffusion

The energy equation has been verified versus analytical solutions for steady-state and non-steady state temperature evolution for both 1D and 2D cases, with and without radioactive heating. Figure 8.14 shows an example of a 2D test for diffusion. A 2D Gaussian thermal perturbation of initial amplitude A_0 and of half-width R evolves in time t as

$$T(x, z, t) = \frac{A_0}{\left(1 + \frac{4t}{R^2}\right)} \exp\left(-\frac{(x - x_0)^2 + (z - z_0)^2}{4t + R^2}\right) \quad (8.40)$$

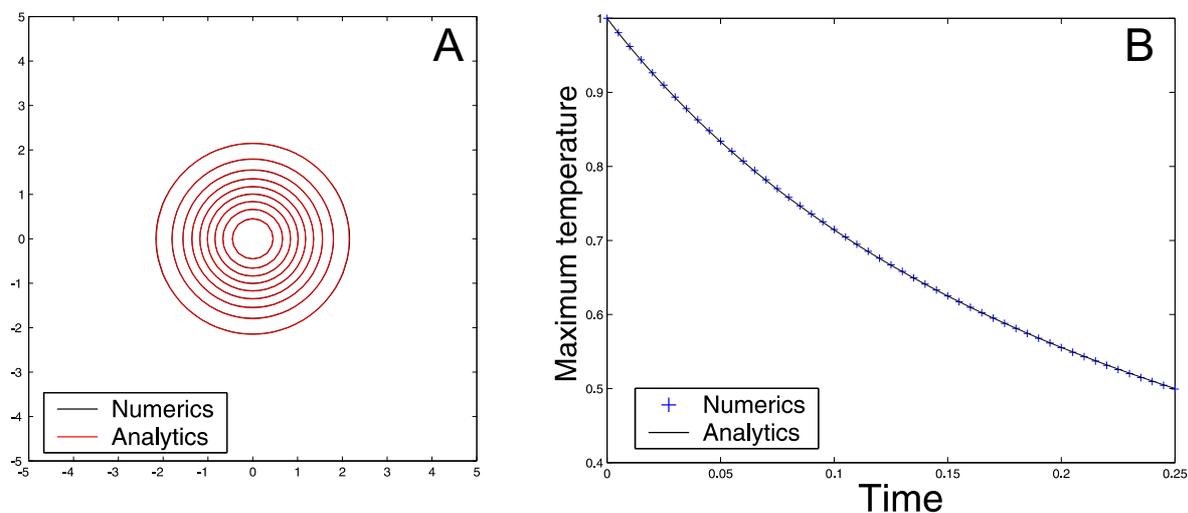


Figure 8.14: Example of 2D diffusion test of a Gaussian wave. A) Comparison of numerical solution with analytical solution at $t = 0.25$, for a resolution of 65×64 . Spacing of contours is 0.05. B) Maximum temperature versus time.

where diffusivity $\kappa = 1$ and the center of the perturbation is located at (x_0, z_0) . The semi-Lagrangian advection scheme is tested by two different methods. The first one is a rigid-body rotation test, during which a Gaussian perturbation is moved with the following velocity field:

$$\begin{aligned} V_x(x, z) &= z \\ V_z(x, z) &= -x \end{aligned} \quad (8.41)$$

The second one is a shear-test, in which the velocity field is given by:

$$\begin{aligned} V_x(x, z) &= -\sin(\pi(x - 0.5)) \cos(\pi(z - 0.5)) \\ V_z(x, z) &= \cos(\pi(x - 0.5)) \sin(\pi(z - 0.5)) \end{aligned} \quad (8.42)$$

Results (fig. 8.15) after one rotation ($t = 2\pi$) show that both methods are free of artificial oscillations. The bicubic spline interpolation method has a better accuracy, with the drawback of an increased computational time (by a factor 11 for a resolution of $201 \times 64, 201 \times 128$ on the mechanical respectively the thermal grid). In addition it should be realized that bicubic spline interpolation may result in artificial oscillations if the interpolation is performed over step-like features, like they occur in flows with

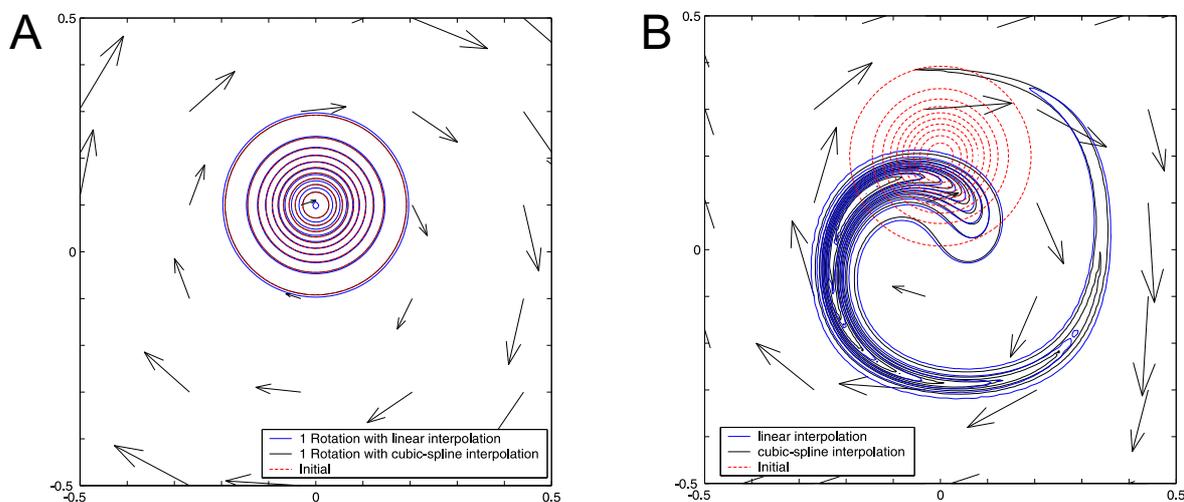


Figure 8.15: A) Solid body rotation test of a gaussian wave with maximum amplitude 2. 50 timesteps are employed, and the resolution is 201×128 . B) Shear test at the same resolution.

strongly variable viscosity (e.g. folding, or weak inclusions). Linear interpolation will diffuse these fields slightly, but will not produce artificial oscillations (which may lead to blowup of the code during computation of viscoelastic flows). A comparison of the semi-Lagrangian method with other advection schemes (see e.g. the webpage of M. Spiegelman, <http://www.ldeo.columbia.edu/~mspieg/mmm/>) learns that the semi-Lagrangian method is at least as accurate as other methods for the rigid body rotation test, and is more accurate for shear-flows. We typically employ linear interpolation in our simulations.

8.5 Conclusions

The Eulerian spectral/finite difference method is a promising tool for modeling geodynamic problems, since it can deal with large deformations. Here we describe an algorithm which solves the governing equations for a slowly-moving incompressible fluid with a nonlinear Maxwell viscoelastic rheology coupled with non-associated Mohr-Coulomb plasticity. Benchmark tests are described for various aspects of the code, ranging from 0D rheology-tests through RT and folding-instabilities, and the initiation of frictional shearbands.

8.6 Appendix A: stress rotation

Rotation of a stress tensor about an angle θ can be expressed in 2D as [Turcotte and Schubert, 1982]:

$$\begin{aligned}\tau_{xx}^{new} &= \tau_{xx} \cos^2(\theta) + \tau_{zz} \sin^2(\theta) - \tau_{xz} \sin(2\theta) \\ \tau_{zz}^{new} &= \tau_{xx} \sin^2(\theta) + \tau_{zz} \cos^2(\theta) + \tau_{xz} \sin(2\theta) \\ \tau_{xz}^{new} &= \frac{1}{2} (\tau_{xx} - \tau_{zz}) \sin(2\theta) + \tau_{xz} (2 \cos(\theta)^2 - 1)\end{aligned}\quad (8.43)$$

A Taylor expansion for $\theta \rightarrow 0$ gives:

$$\begin{aligned}\tau_{xx}^{new} &= \tau_{xx} - 2\theta\tau_{xz} \\ \tau_{zz}^{new} &= \tau_{zz} + 2\theta\tau_{xz} \\ \tau_{xz}^{new} &= \tau_{xz} + \theta(\tau_{xx} - \tau_{zz})\end{aligned}\quad (8.44)$$

Vorticity is given by equation 8.20. We can thus write:

$$\begin{aligned}\tau_{xx}^{new} &= \tau_{xx} + \tau_{xz} \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) dt \\ \tau_{zz}^{new} &= \tau_{zz} - \tau_{xz} \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) dt \\ \tau_{xz}^{new} &= \tau_{xz} + \left(\frac{\tau_{xx} - \tau_{zz}}{2} \right) \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) dt\end{aligned}\quad (8.45)$$

The Jaumann objective derivative of the stress tensor is given by [e.g. Altenbach and Altenbach, 1994]:

$$\tau_{ij}^{\nabla} = \frac{\partial \tau_{ij}}{\partial t} + \tau_{ik} W_{kj} - W_{ik} \tau_{kj} \quad (8.46)$$

where $W_{ij} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} \right)$. Spelled out for the 2D case, (8.46) reads:

$$\begin{aligned}\tau_{xx}^{\nabla} &= \frac{\partial \tau_{xx}}{\partial t} - 2W_{xz} \tau_{xz} \\ \tau_{zz}^{\nabla} &= \frac{\partial \tau_{zz}}{\partial t} + 2W_{xz} \tau_{xz} \\ \tau_{xz}^{\nabla} &= \frac{\partial \tau_{xz}}{\partial t} - W_{xz} (\tau_{xx} - \tau_{zz})\end{aligned}\quad (8.47)$$

After making an explicit timestep discretization, (8.45) is obtained from (8.47).

8.7 Appendix B: finite element formulation

The 'standard' finite element code employed to produce figure 8.11 is a personally developed two-dimensional code, which employs either the seven-node Crouzeix-Raviart triangle [Crouzeix and Raviart, 1973] or a nine-node quadrilateral element [Q2-P1 element, Boffi and Gastaldi, 2002] to solve the Stokes equations for incompressible materials under plane strain. The rheology is nonlinear temperature-dependent visco-elasto-plastic with Mohr-Coulomb plasticity. A mixed formulation is employed, with linear, discontinuous, shape-functions for pressure and quadratic shape functions for velocity to avoid spurious pressures [Pelletier et al., 1989]. Uzawa-type iterations are employed to satisfy the incompressibility constraint [Cuvelier et al., 1986]. In this work, the code has been used in a Lagrangian fashion, with regular remeshing to deal with large deformations [see also e.g. Poliakov and Podladchikov, 1992]. The code has been extensively tested, for example with most of the tests described in this work.

Bibliography

- J. Altenbach and H. Altenbach. *Einführung in die Kontinuumsmechanik*. Teubner, Stuttgart, 1994.
- M.A. Biot and H. Odé. Theory of gravity instability with variable overburden and compaction. *Geophysics*, 30:153–166, 1965.
- D. Boffi and L. Gastaldi. On the quadrilateral Q2-P1 element for the Stokes problem. *International Journal of Numerical Methods in Fluids*, 39:1001–1011, 2002.
- J. Braun and M. Sambridge. Dynamic Lagrangian Remeshing (DLR): A new algorithm for solving large strain deformation problems and its application to fault-propagation folding. *Earth and Planetary Science Letters*, 124:211–220, 1994.
- U.R. Christensen and H. Harder. Three dimensional convection with variable viscosity. *Geophysical Journal International*, 104:213–216, 1991.
- M. Crouzeix and P.A. Raviart. Conforming and nonconforming finite-element methods for solving stationary stokes equations. *Revue Francaise D'Automatique Informatique Recherche Operationnelle*, 7 DEC:33–75, 1973.
- C. Cuvelier, A. Segal, and A.A. van Steenhoven. *Finite Element Methods and Navier-Stokes Equations*. Mathematics and its applications. Reidel Publishing Company, Dordrecht, 1986.
- R. C. Fletcher. Folding of a single viscous layer - exact infinitesimal amplitude solution. *Tectonophysics*, 39(4):593–606, 1977.
- R. S. Huismans, Y. Y. Podladchikov, and S. Cloetingh. Transition from passive to active rifting: Relative importance of asthenospheric doming and passive extension

- of the lithosphere. *Journal of Geophysical Research-Solid Earth*, 106(B6):11271–11291, 2001.
- A. V. Malevsky and D.A. Yuen. Characteristics-based methods applied to infinite prandtl number thermal convection in the hard turbulent regime. *Physics of Fluids*, 3(9):2105–2115, 1991.
- G. Marquart. On the geometry of mantle flow beneath drifting plates. *Geophysical Journal International*, 144:356–372, 2001.
- L. Moresi, F. Dufour, and H.-B. Mühlhaus. A lagrangian integration point finite element method for large deformation modeling of viscoelastic geomaterials. *Journal of Computational Physics*, 184:476–497, 2003.
- D. Pelletier, A. Fortin, and R. Camarero. Are FEM solutions of incompressible flows really incompressible? (Or how simple flows can cause headaches!). *International Journal of Numerical Methods in Fluids*, 9:99–112, 1989.
- A. Poliakov, P. Cundall, Y.Y. Podladchikov, and V. Laykhovsky. An explicit inertial method for the simulation of visco-elastic flow: an evaluation of elastic effects on diapiric flow in two- and three- layers models. In D.B. Stone and S.K. Runcorn, editors, *Flow and Creep in the Solar System: observations, modeling and Theory*, Proceeding of NATO Advanced Study. Institute, pages 175–195. 1993.
- A. Poliakov and Y. Podladchikov. Diapirism and topography. *Geophysical Journal International*, 109:553–564, 1992.
- A. Poliakov, Y. Podladchikov, J. B. Dawson, and C. Talbot. Salt diapirism with simultaneous brittle faulting and viscous flow. *Geological Society Special Publication*, 100:291–302, 1996.
- A. N. B. Poliakov, H. J. Herrmann, and Y. Y. Podladchikov. Fractal plastic shear bands. *Fractals*, 2:567–581, 1994.
- S. M. Schmalholz. *Finite amplitude folding of viscoelastic rocks: importance of elasticity for slow tectonic motion*. Phd-thesis, ETH Zurich, 2000.

- S. M. Schmalholz and Y. Podladchikov. Buckling versus folding: Importance of viscoelasticity. *Geophysical Research Letters*, 26(17):2641–2644, 1999.
- S. M. Schmalholz, Y. Y. Podladchikov, and D. W. Schmid. A spectral/finite difference method for simulating large deformations of heterogeneous, viscoelastic materials. *Geophysical Journal International*, 145(1):199–208, 2001.
- D. W. Schmid. *Finite and infinite heterogeneities under pure and simple shear*. Phd, Swiss Federal Institute of Technology, 2002.
- D. W. Schmid and Y. Podladchikov. Analytical solutions for deformable elliptical inclusions in general shear. *Geophysical Journal International*, 155:269–288, 2003.
- C. H. Scholz. *The mechanics of earthquakes and faulting*. Cambridge University Press, 2nd edition, 2002.
- J.C. Simo and T. J. R. Hughes. *Computational Inelasticity*. Springer, New York, 2nd edition, 2000.
- L. Trefethen. *Spectral methods in MATLAB*. SIAM, 2001.
- D. L. Turcotte and G. Schubert. *Geodynamics*. Wiley, 1982.
- P. E. van Keken, S. D. King, H. Schmeling, U. Christensen, D. Neumeister, and M. P. Doin. A comparison of methods for modeling of thermochemical convection. *Journal of Geophysical Research*, 102:22477–22496, 1997.
- P. Vermeer and R. de Borst. Non-associated plasticity for soils, concrete and rock. 29: 1–65, 1984.
- P. A. Vermeer. The orientation of shear bands in bi-axial tests. *Geotechnique*, 40(2): 223–236, 1990.

Appendix A

Aftershocks driven by a high pressure CO₂ source at depth¹

Abstract Inspectors surveying the Assisi Cathedral following a magnitude 5.7 earthquake were trapped and killed when a magnitude 6 earthquake struck nine hours later and collapsed the weakened structure. These two earthquakes marked the beginning of a sequence that lasted more than 30 days, with thousands of aftershocks including four additional events with $5 < M_w < 6$. This normal-faulting sequence is not well-explained with models of elastic stress transfer [Stein, 1999, Cocco et al., 2000], particularly the persistence of hanging-wall seismicity [Chiaraluce et al., 2003] that included two events with $M_w > 5$. Here we show that this sequence was driven by a fluid pressure pulse generated from the co-seismic release of a known deep source [Chiodini and Cioni, 1989] of trapped high-pressure CO₂. Using precise hypocenter locations [Waldhauser and Ellsworth, 2000] and a simple model of non-linear diffusion, we show a strong correlation between the high pressure front and the aftershock hypocenters over a two week period. The 10-20 MPa triggering amplitude of the pressure pulse overwhelms the typical 0.1-0.2 MPa range from stress changes in the usual stress triggering arguments [Stein, 1999, Toda et al., 2002]. We propose that aftershocks of large earthquakes may in general be driven by the co-seismic release of trapped, high pressure fluids propagating through damaged zones created by the mainshock, thus providing a link between earthquakes, aftershocks, crust/mantle degassing and

¹This work has been published in: Miller, S.A, Collettini, C., Chiaraluce, L., Cocco, M., Barchi, M., Kaus, B.J.P. (2004) Aftershocks driven by a high pressure CO₂ source at depth., *Nature* 427. 724-727.

earthquake-triggered large scale fluid flow.

A.1 Introduction

The 1997 Umbria-Marche seismic sequence in the Northern Apennines, Italy (Fig. A.1a) occurred on shallow-dipping ($\sim 40^\circ$) normal faults, migrating from NNW to SSE along an area of about 40 km in length and 15 km in width. Structural geology studies show the earthquakes nucleate close to the deeper portion of a thrust, and that the faults were not optimally oriented relative to the regional stress field [Collettini, 2002]. Non-optimally oriented faults become seismically active either because of lower friction coefficients or the presence of fluid pressures in excess of hydrostatic [Sibson, 1992, Cox, 1995, Miller and Olgaard, 1996, Streit and Cox, 2001].

A geologic cross section integrating surface geology with seismic reflection profiles [Reuter et al., 1980, Mirabella and Pucci, 2002, Fig. A.1b] shows that the first two mainshocks nucleated in the Triassic evaporites (made up of alternating anhydrites and dolomites). All $M_w > 5$ earthquakes nucleated in the evaporites, the same lithologic unit where CO₂ at near-lithostatic pressure was encountered in the San Donato Borehole at a depth of 4.8 km about 50 km NW of Colfiorito [Chiodini and Cioni, 1989]. The tectonic environment of the Northern Apennines is suitable for trapping high pressure fluids derived from CO₂ mantle degassing [Chiodini et al., 2000, Quattrocchi, 1999], and in particular, the Rasiglia spring in the epicentral region shows an area-averaged deep CO₂ production rate [Reuter et al., 1980] of approximately $6 \times 10^5 \text{ mol m}^{-2} \text{ yr}^{-1}$.

We show that the driving mechanism for this earthquake and aftershock sequence is the co-seismic release and propagation of the trapped high pressure source into the overlying carbonates at hydrostatic pore pressure. Co-seismic fracturing of the seal separating these two distinct pressure states initiates the rapid propagation of a pressure pulse along the newly formed, highly permeable fault zone and into the adjacent damage zone. The newly fractured crust provides high permeability channels to propagate the pulse and trigger seismicity by significantly reducing the effective normal stress acting on incipient slip planes. In addition, recognizing that the earthquakes and aftershocks themselves create new fractures results in a large-scale permeability structure that increases significantly as the sequence evolves. This is shown by recast-

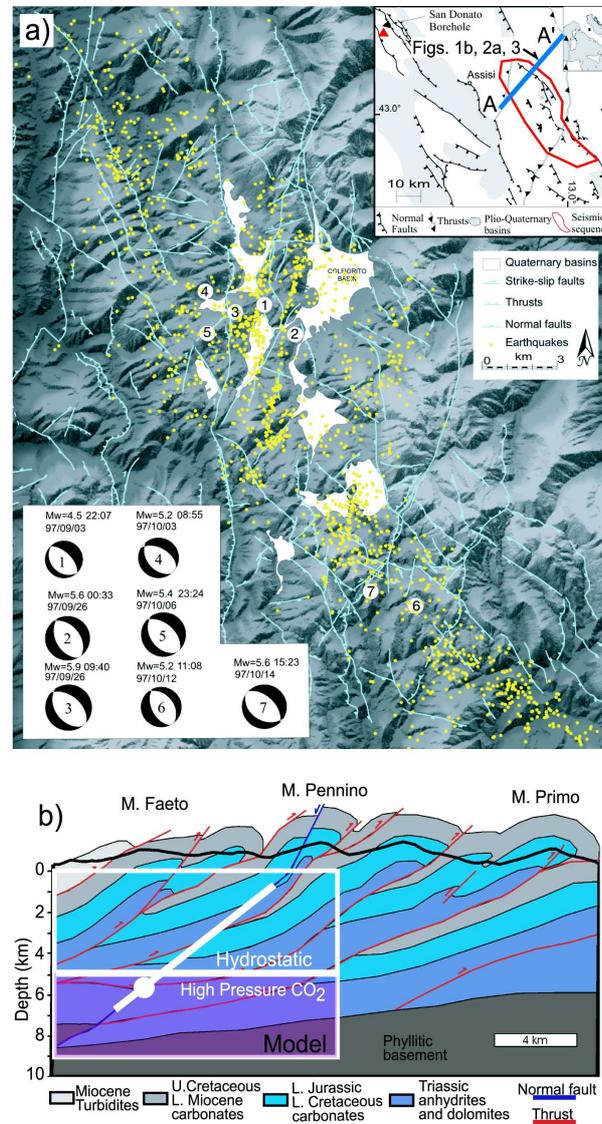


Figure A.1: Geologic setting of the study area. A) Seismicity and major structures of the Colforito region of the Northern Apennines, Italy. The region consists of a complex pattern of thrusts, folds and normal faults reflecting two main tectonic phases: A Miocene-Pliocene compressional phase forming E-NE verging thrusts and folds; and a superimposed upper Pliocene-Quaternary extensional phase forming basins bounded by NNW-SSE trending normal faults [Reuter et al., 1980, Mirabella and Pucci, 2002]. Near lithostatic pore pressure (CO₂) measured in the San Donato borehole (see inset) was encountered in the evaporates and just below the seal of a sub-horizontal thrust. All $M > 5$ earthquakes nucleated in the evaporates. B) Geologic cross-section calibrated from geology and surface profiles [Mirabella and Pucci, 2002], with the simplified model shown superposed (see also Fig. A.3)

ing the epicenters (Fig. A.1a) as the approximate slipped area of the sequence (Fig. A.2a), demonstrating how the entire region evolves to a complex system of fractures that provide conduits for propagating the pressure pulse.

A.2 Aftershock patterns

There are two regions of interest for this study. The first is the largest event of the entire sequence (Event 3), and its associated aftershocks (Fig. A.2a). We can approximate and model this sequence by projecting the aftershocks onto a 2-dimensional profile (A-A') that also corresponds to the geologic cross-section in Fig. A.1b. The second region of interest is the sequence propagating to the SSE, where the observed propagation velocity (Fig. A.2b) can be used to estimate the structural permeability of the system, and to show supporting evidence for the high permeability used in the model. Fig. A.2b shows that this sequence propagated at a relatively constant velocity about 1 km day⁻¹, similar to the velocity of CO₂-driven seismicity inferred from a swarm system [Bruer et al., 2003], and for induced seismicity in a deep borehole [Baisch and Harjes, 2003]. A fluid-pressure induced sequence with a constant propagation velocity is consistent with the wave-like solutions of a pressure pulse found for flow problems where permeability is a strong non-linear function of the effective normal stress acting on the fracture [Rice, 1992, Revil and Cathles, 2002]. Assuming that the observed propagation velocity reflects the velocity of a wave-like pulse, we can roughly estimate the large-scale permeability of the system using the relationship [Rice, 1992] $v = \frac{k\gamma}{\eta\phi}$, where v is the pulse velocity, k is the intrinsic permeability, γ is the weight density difference between the rock and fluid, η is the viscosity, and ϕ is the porosity. Taking $\gamma = 1.7$ MPa m⁻¹, $\eta = 10^{-3}$ Pa^{-s}, $\phi = .05$, and using the observed v , then $k = 4 \times 10^{-11}$ m². This is a very large permeability compared to the 10⁻¹⁶ m² inferred for crustal permeability in tectonically stable environments [Manning and Ingebritsen, 1999], but lower than the permeability inferred for the Dobi extensional earthquake sequence in Central Afar [Noir et al., 1997], and consistent with permeability measurements on rough fractures in granite and marble at low effective normal stress [Lee and Cho, 2002].

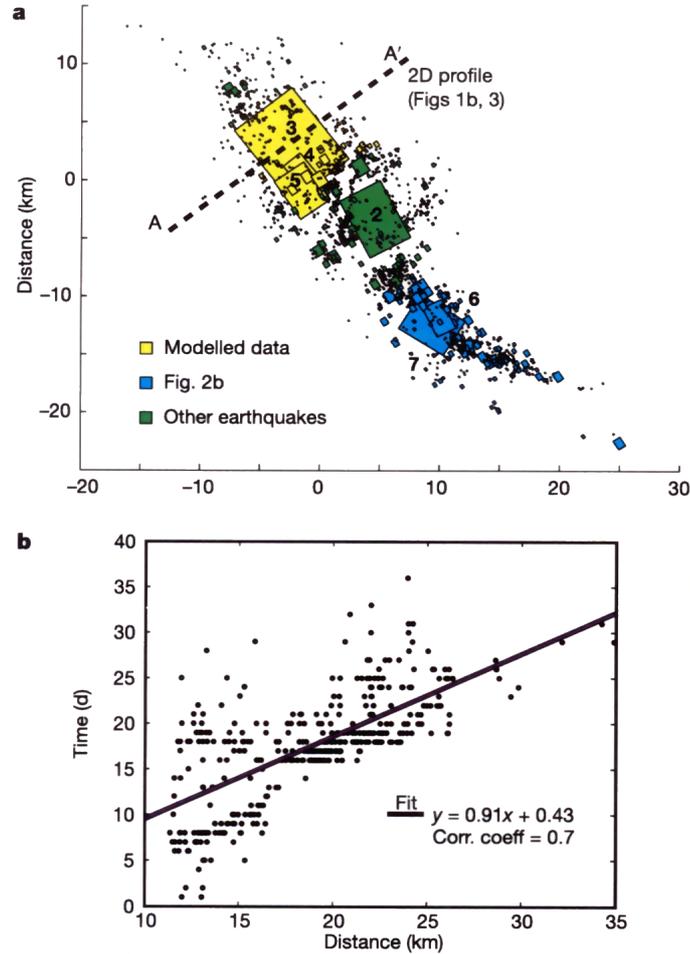


Figure A.2: Map view of the seismicity and the rate of propagation. (a) The seismicity in Fig. A.1a is recast as the approximate area fractured in the earthquake or aftershock to illustrate how the earthquake or aftershock themselves drastically alter the structural permeability of the system. Patches show the slipped area using the relation $M = GA\bar{u}$, where M is the scalar seismic moment, G is the shear modulus (30 GPa), $\Delta\sigma$ is the stress drop (assumed 1.5 MPa), and slip is calculated from $\Delta\sigma = \frac{2G}{\pi} \frac{\bar{u}}{W}$. The events are color-coded to show the events (yellow) compared to the model results, and include all events with $M > 2.4$ between the hypocenter of event 3 and about 7 km to the NW. Section A-A' corresponds to the cross section in Fig. A.1b and represents an approximate 2-D profile onto which the hypocenters of events shown in yellow are projected. Events shown in blue are plotted in (b) as the distance from the hypocenter of event 3 vs. time to estimate the structural permeability of the system. It is the length of the vector between hypocenters, and therefore includes both the horizontal and up-dip migration of the sequence. The linear correlation implies that the propagation velocity is faster than the $t^{1/2}$ diffusion time-scale, and a least squares fit shows that this sequence propagation at $\sim 1 \text{ km d}^{-1}$ (that is, $\sim 10^{-2} \text{ m s}^{-1}$).

A.3 Modeling results

Modelling the 3-dimensional flow field is beyond the scope of this letter, but we can simplify by considering the 2-dimensional profile perpendicular to the strike of Event 3. All aftershocks with $M_w > 2.4$ in volume surrounding Event 3 are projected onto the 2-dimensional profile (A-A') for comparing with model results. We model this sequence by numerically solving a non-linear diffusion equation with an effective-stress dependent permeability (see methods). In the initial conditions, an impermeable seal separates the upper region at hydrostatic fluid pressure from the high-pressure source region where the initial fluid pressure was taken at 85% of lithostatic [Chiodini and Cioni, 1989] (for example 70 MPa) at the upper boundary of the source region. A hydrostatic pressure gradient is imposed below the seal, corresponding to no-flow conditions prior to the earthquake. Flow initiates at $t = 0$, when a model fault and damage zone approximately 400m in width (field observations show damage zones from 200-600m) cuts the high pressure region and extends to about 1 km below the surface. This simulates the co-seismic fracturing of the pore pressure seal. The sudden communication between the high-pressure source and low-pressure surroundings initiates a pressure pulse that propagates along the fault and into the hanging and footwalls.

The evolution of the propagating pressure front and the fluid pressure field (Fig. A.3) are superposed with the hypocenters of aftershocks for period indicated in each figure. The steep front is a consequence of the wave-like solution for pressure-dependent permeability, in contrast to a more diffuse front resulting from linear diffusion. At early times (Fig. A.3a-b), the pulse propagates rapidly up the fault zone, leaving in its wake a slower moving pulse into the matrix material of the hanging and footwalls. The faster propagation into the hanging wall relative to the footwall is a result of the decreasing σ_n (thus increasing permeability) as the pulse propagates to shallower depths. As the pulse reaches the hydrostatic boundary condition imposed at the surface, the pressure is reduced but the pulse continues to propagate. The high-pressure front arrives and triggers Event 4 (Fig. A.3d), which we model by introducing a second fault (with the same properties as the original fault) extending into the source region (dotted line). Because this new high permeability fault extends into the source region, the system is recharged and generates additional aftershocks (Fig. A.3e). The subsequent pulse propagates to the location of event 5 (Fig. A.3f), triggering additional seismicity (Fig.

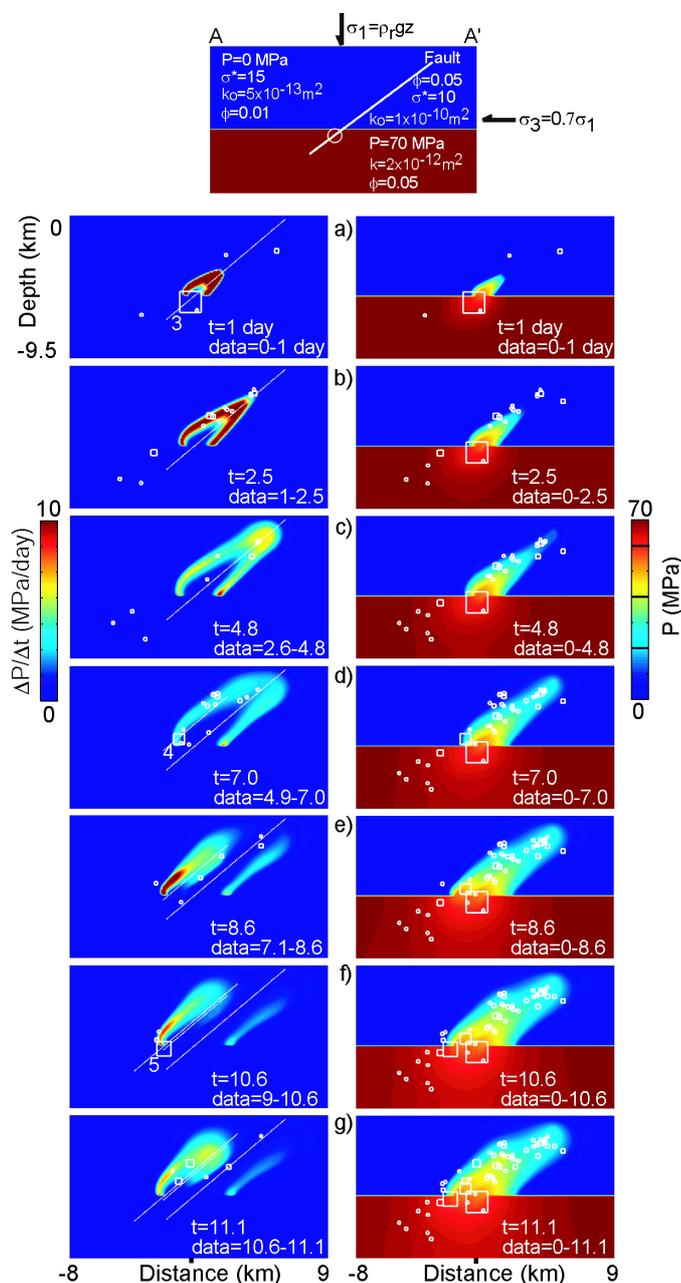


Figure A.3: Comparison of model results with initial conditions (top) to the hypocenters of aftershocks shown in yellow in Fig. A.2a. Model results are plotted as the rate of pore pressure increase to highlight propagation of the pressure front (left column), and the corresponding evolution of the entire fluid pressure field (right column). The left column compares the evolution of the pore pressure front to the aftershocks occurring during the times indicated. The overall fluid pressure field is superposed with the cumulative aftershock catalogue. The largest event in the sequence (Event 3) and subsequent large aftershocks in the hanging wall (Events 4 and 5) are indicated in a), c), and e).

A.3g). Aftershocks for events 4 and 5 correlate with the rapid propagation of a pulse into the already highly pressurized footwalls of these events.

The evolution of the total fluid pressure in the system (right column of Fig. A.3) shows that the data is matched both in space and through time and follows the structure of the evolving fluid pressure field. The aftershocks in the high-pressure source region appear to correlate with contours of reduced fluid pressure. In this case, the mechanism of triggering is more likely the transition from aseismic slip at high pore pressures to seismic slip as fluid pressure is reduced [Segall and Rice, 1995], or alternatively, a complex source region.

A.4 Aftershock triggering

Our model presents an alternative interpretation of the physical processes controlling earthquake triggering in the neighborhood of the causative fault. Several studies have shown correlations with stressing rate changes [Dieterich, 1994, Toda et al., 2002], static stress transfer [Stein, 1999], or poro-elastic effects [Nur and Booker, 1972, Bosl and Nur, 2002]. These models rely on extremely small stress changes (~ 0.1 MPa) and therefore have not unequivocally demonstrated that simple static stress changes or poro-elastic effects are the dominant mechanism of earthquake triggering or driving aftershocks. The change in Coulomb failure stress (ΔCFS) is defined as $\Delta CFS = \Delta\tau + \mu(\Delta\sigma_n + \Delta P_f)$, where $\Delta\tau$ and $\Delta\sigma_n$ are the shear and normal stress changes (positive in extension), and P_f is the change in pore pressure. Attempts to relate this earthquake sequence to ΔCFS from shear stress changes failed to explain this sequence [Cocco et al., 2000], particularly for the persistence of aftershocks in the hanging wall [Chiari et al., 2003]. Fig. A.4a shows the aftershock data with the ΔCFS (for stress changes only) for Event 3, and Fig. A.4b compares the same aftershock with the calculated pressure field.

In most ΔCFS formulations, the focus is primarily on changes in τ and σ_n , and it was found to be difficult to include poro-elastic effects [Cocco and Rice, 2003]. For Event 3, ΔCFS from shear and normal stress changes are on the order of a few tenths of MPa. Significantly, our results show that the effect from a 10-20 MPa direct fluid pressure loading (for example ΔP_f) overwhelms static stress transfer.

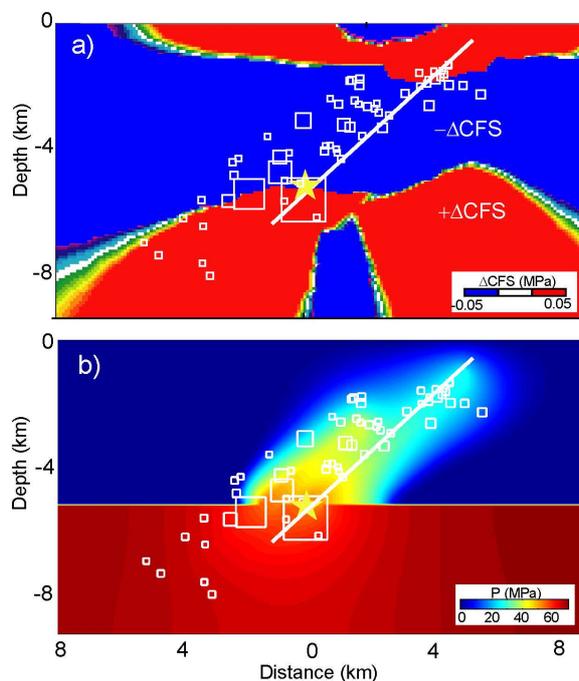


Figure A.4: Comparison of aftershock data to shear stress changes in the formulation and pore pressure changes. a) There is obviously no correlation between positive or negative regions and the aftershock locations. In contrast (b), the same aftershock data compared to the calculated fluid pressure state after 11 days shows a very strong correlation with the entire aftershock sequence (see also Fig. A.3).

A.5 Conclusions

The structural, seismic, and post-seismic response of this sequence support a scenario where high pressure CO_2 infiltrated the incipient seismic fault prior to the large earthquakes, followed by a large scale change in the hydraulic properties of the system. The co-seismic fracture generated a high amplitude pressure pulse initiating at the high-pressure/low pressure boundary, propagating into the damaged region caused by the mainshock. The increased fluid pressure triggered subsequent earthquakes and aftershocks by significantly reducing the effective normal stress. The results also suggest that the aftershocks in regions of increasing pore pressure occur along contours of constant $\bar{\sigma}_n$, implying that earthquakes occurred at the same shear stress (assuming a constant friction coefficient) independent of depth. Since the effect on ΔCFS due only

to pore pressure changes is orders of magnitude greater than the contribution of elastic stress transfer, we propose that this mechanism dominates some triggering phenomena and aftershock sequences. Although this sequence was driven by CO₂ out-gassing, the processes of fracture and co-seismic hydraulic property changes are general, suggesting this is an important general mechanism of aftershock generation. That is, earthquakes provide the trigger to hydraulically connect the upper crust at hydrostatic pore pressure with the lower crust at near-lithostatic pore pressure. The subsequent flow will be fast, high pressured, and will propagate readily into the new fractures created by the main event.

A.6 Methods

We adopt the model of Rice [1992] where permeability is a strongly decreasing function of effective normal stress, $k = f(\bar{\sigma}_n)$. Specifically, $k = k_0 \exp(-\bar{\sigma}_n/\sigma^*)$, where k_0 is the permeability at zero effective stress, and σ^* is a constant with lower values of corresponding to highly cracked rocks. Using this form for the permeability, we solve the diffusion equation with a spatially variable permeability [Wong et al., 1997]:

$$\frac{\partial P}{\partial t} = \frac{1}{\phi(\beta_f + \beta_\phi)} \left(\nabla \frac{k_0 \exp(-\frac{\bar{\sigma}_n}{\sigma^*})}{\eta} \nabla P + \dot{\Gamma}(P, T) \right) \quad (\text{A.1})$$

where P is the fluid pressure above hydrostatic, β_f and β_{phi} are the fluid and pore (crack) compressibility, and $\dot{\Gamma}$ is a source term. The source term is assumed zero here, but is included in Equation A.1 to show that the pressure-dependence of the dehydration (or de-carbonization) kinetics could provide an additional direct fluid source from co-seismic fluid pressure reductions [Miller et al., 2003]. The effective normal stress used in Equation (A.1) and acting on fault planes is calculated as [Jaeger and Cook, 1979]:

$$\bar{\sigma}_n = \frac{\sigma_1 + \sigma_3 - 2P_f}{2} + \frac{\sigma_1 - \sigma_3}{2} \cos 2\theta \quad (\text{A.2})$$

where σ_1 and σ_3 are the maximum and minimum principal stress, P_f is the total fluid pressure (for example, $P + \rho_w g z$), and θ is the dip angle. We take $\theta = 40^\circ$ (determined from the earthquake focal mechanisms), σ_1 as the weight of the overburden (for example $\rho_r g z$), where ρ_r is the rock density, and we assume $\sigma_3 = 0.7\sigma_1$ to reflect this extension tectonic environment.

We solve Equation A.1 with an implicit finite difference scheme, using the simplified model geometry and initial conditions shown at the top of Fig. A.3. A no-flow boundary condition is imposed on all boundaries except the upper surface, where a constant head (e.g. hydrostatic pore pressure) boundary condition is imposed. We use crack compressibility $\beta_\phi = 10^{-8} \text{ MPa}^{-1}$, fluid compressibility $\beta_f = 10^{-10} \text{ MPa}^{-1}$, and a temperature-dependent viscosity for water (assuming a temperature gradient of 25°C km^{-1}). We assume that the flow properties of supercritical CO_2 (the phase of CO_2 at the source depth) are the same as for water because CO_2 at this P-T condition is 10 times more compressible than water, but it is of the same order less viscous, resulting in similar flow properties. Note that the model can be made much more com-

plicated by considering two-phase flow, dual-porosity, anisotropic permeability, and other complexities. However, we find this simple model sufficient to show a very strong correlation between the calculated pressure field and the precise locations of aftershock hypocenters.

Bibliography

- S. Baisch and H.-P. Harjes. A model for fluid-injection induced seismicity at the KTB, Germany. *Geophys. J. Int.*, 152:160–170, 2003.
- W.J. Bosl and A. Nur. Aftershocks and pore fluid diffusion following the 1992 Landers earthquake. *J. Geophys. Res.*, 107:10.1029/2001JB000155, 2002.
- K. Bruer, H. Kmpf, G. Strauch, and S. Weise. Isotopic evidence ($^3\text{he}/^4\text{he}$, $^{13}\text{c}_{\text{CO}}$) of fluid-triggered intraplate seismicity. *J. Geophys. Res.*, 108:10.1029/2002JB002007, 2003.
- L. Chiaraluce, W.L. Ellsworth, C. Chiarabba, and M. Cocco. Imaging the complexity of an active normal fault system; the 1997 Colfiorito Central Italy case study. *J. Geophys. Res. B*, 6:10–1029/2002JB002166, 2003.
- G. Chiodini, F. Frondini, C. Cardellini, F. Parello, and L. Peruzzi. Rate of diffuse carbon dioxide Earth degassing estimated from carbon balance of regional aquifers: The case of central Apennine, Italy. *J. Geophys. Res.*, 105:8423–8434, 2000.
- R. Chiodini and G. Cioni. Gas geobarometry for hydrothermal systems and its application to various Italian geothermal areas. *Appl. Geochem.*, 4:564–572, 1989.
- M. Cocco, C. Nostro, and Ekstrom G. Static stress changes and fault interaction during the 1997 Umbria-Marche earthquake sequence. *J. Seismol.*, 4:501–516, 2000.
- M. Cocco and J.R. Rice. Pore pressure and poroelasticity in Coulomb analysis of earthquakes. *J. Geophys. Res.*, 107:10.1029/2000JB000138; correction 10.1029/2002JB002319, 2003.

- C. Collettini. Structural permeability control on the post-seismic fluid discharge and aftershocks triggering; hypotheses for the 1997 Colfiorito earthquakes. *Boll. Soc. Geol. It.*, 1:873–880, 2002.
- S.F. Cox. Faulting processes at high fluid pressures: an example of fault valve behavior from the Wattle Gully Fault, Victoria, Australia. *J. Geophys. Res.*, 100:12841–12859, 1995.
- J. Dieterich. A constitutive law for rate of earthquake production and its application to earthquake clustering. *J. Geophys. Res.*, 99:2601–2618, 1994.
- J.C. Jaeger and N.G.W. Cook. *Rock Mechanics*. Chapman and Hall, London, 1979.
- H.S. Lee and T.F. Cho. Hydraulic characteristics of rough fractures in linear flow under normal and shear load. *Rock Mech. Rock Eng.*, 35:299–318, 2002.
- C.E. Manning and S.E. Ingebritsen. Permeability of the continental crust: Implication of geothermal data and metamorphic systems. *Rev. Geophys.*, 37:127–150, 1999.
- S.A. Miller and D.L. Olgaard. Earthquakes as a coupled shear stress-high pore pressure dynamical system. *Geophys. Res. Lett.*, 23:197–200, 1996.
- S.A. Miller, W. van der Zee, D.L. Olgaard, and J.A.D. Connolly. A fluid-pressure controlled feedback model of dehydration reactions: Experiments modelling, and application to subduction zones. *Tectonophysics*, 370:241–251, 2003.
- F. Mirabella and S. Pucci. Integration of geological and geophysical data along two sections crossing the region of the 1997-98 Umbria-Marche earthquake sequence. *Boll. Soc. Geol. It.*, 1:891–900, 2002.
- J. Noir, F. Jaques, S. Bkri, P.M. Adler, and G.C.P. King. Fluid flow triggered migration of events in the 1989 Dobi earthquake sequence of central afar. *Geophys. Res. Lett.*, 24:2335–2338, 1997.
- A. Nur and J.R. Booker. Aftershocks caused by pore fluid flow. *Science*, 175:885–887, 1972.

- F. Quattrocchi. In search of evidence of deep fluid discharges and pore pressure evolution in the crust to explain the seismicity style of the Umbria-Marche , 1997-1998 seismic sequence (Central Italy). *Ann. Geophys.*, 42:609–636, 1999.
- K.J. Reuter, P. Giese, and H. Closs. Lithosphere split in the descending plate; observations from the Northern Apennines. *Tectonophysics*, 64:T1–T9, 1980.
- A. Revil and L.M. III Cathles. Fluid transport by solitary waves along growing faults. A field example from the South Eugene Island Basin, Gulf of Mexico. *Earth Planet. Sci. Lett.*, 202:321–335, 2002.
- J.R. Rice. *Fault Mechanics and Transport Properties of Rock*, pages 476–503. Academic, San Diego, 1992.
- P. Segall and J.R. Rice. Dilatancy, compaction, and slip instability of a fluid infiltrated fault. *J. Geophys. Res.*, 100:22155–22171, 1995.
- R.H. Sibson. Fault-valve behavior and the hydrostatic-lithostatic fluid pressure interface. *Earth Sci. Rev.*, 32:141–144, 1992.
- R.S. Stein. The role of stress transfer in earthquake triggering. *Nature*, 402:605–609, 1999.
- J.E. Streit and S.F. Cox. Fluid pressures at hypocenters of moderate to large earthquakes. *J. Geophys. Res.*, 106:2235–2243, 2001.
- S. Toda, R. Stein, and T. Saglya. Evidence from the AD 2000 Izu islands earthquake swarm that stressing rate governs seismicity. *Nature*, 419:58–61, 2002.
- F Waldhauser and W.L. Ellsworth. A double-difference earthquake location algorithm: Method and application to the Northern Hayward Fault, California. *Bull. Seismol. Soc. Am.*, 90:1353–1368, 2000.
- T.-f. Wong, S.-c. Ko, and D.L. Olgaard. Generation and maintenance of pore pressure excess in a dehydration system, 2, Theoretical analysis. *J. Geophys. Res.*, 102:841–852, 1997.

Acknowledgements

The first person to thank is the one who changed my (scientific) life by explaining me that mathematics is something that can be used by geologists. When I first met Yuri back in 1997 I was planning to stay for only a short time in Zurich. This short time became a bit longer when I discovered the fun in watching things moving on my computer-screen. Always being very motivating, and never tired of telling how easy mathematics is, it has been a pleasure having Yuri as an advisor. Jean-Pierre (who cares?), has provided help where possible, and was the one who showed me how to enjoy the wine-related cultural aspects of France. Jamie helped not only with the various petrological aspects of the work, but also made sure that, despite my swimming skills, the *Lame-Ducks* ended up with quite reasonable rankings on the various triathlons. With Dani and Stefan, I experienced the pleasure of working in a team, not only while they were in Zurich, but also after their movement to Far-North. Guy, Neil, Steve and Taras have always been present for lots of discussions and fun-times.

One of the nice aspects in Zurich was our weekly friday seminar in which I probably learned most. Thanks to Chris, Jeroen, Gabriele, Taras, and Thomas this has been very motivating, also in recent times. Thanks also to the PhD-students of the geological institute and my officemates, Katrin, Cristian, Jeannette and little Esteban. Mox helped me out with Latex issues. Many more people in Zurich and back in Holland made my PhD-time a pleasant time. Special thanks go to Florian, Beni, Edi, Lydia, Jos, Ivanka. My family in Holland and in Switzerland have always supported me during my studies and PhD.

Petra, without your love, the last years would never have been that beautiful.

Curriculum vitae

Name	Kaus
Vornamen	Boris Jozef Paul
Geburtsdatum	4. Dezember 1976
Geburtsort	Heerlen in den Niederlanden
Nationalität	Niederländische
Zivilstand	ledig
Schulbildung	
1983-1989	Grundschule, Gulpen Niederlande
1989-1995	Havo/VWO Sophianum, Gulpen Niederlande
Hochschulbildung	
1995-1997	Studium der Geologie an der Rheinisch-Westfälische Technische Hochschule Aachen (RWTH) Deutschland Vordiplom im Juli 1997
1997-2000	Studium der Geologie an der ETH Zürich Schweiz Diplom in September 2000
2001-2004	Doktorarbeit an der ETH Zürich