



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

**ALERT Doctoral School 2018**  
*Energetical Methods in Geomechanics*

Editors:

Itai Einav

Eleni Gerolymatou



---

## Editorial

---

The ALERT Doctoral School 2018 on “Energetical Methods in Geomechanics” will take place as usual in Aussois, from October 4th to 6th, 2018. The School has been organized by Prof. Itai Einav (University of Sydney) and Prof. Eleni Gerolymatou (Chalmers University of Technology). I sincerely thank the organizers and all the authors of the contributions to this book for their effort!

If it is clear that energy is requested to deform soils and rocks, it exists under many forms and can transform from one to another. All these phenomena may respect the laws of thermodynamics, that constrain the development of constitutive models. Often seen as a limitation to the imagination of engineers, thermodynamic constraints can actually help avoiding flaws in the methodology. I am therefore convinced that this school will be beneficial to the ALERT community.

Lectures will include topics ranging from basic concepts of energetical methods (Thermodynamics, Principle of virtual power), to specific examples and applications that illustrate how energetical methods lead development of constitutive models for soil and rock mechanics. Practical sessions will be organized on the last day of the school, in order to evidence the effect of constitutive choices on the energetics.

As usual, the pdf file of the book can be downloaded for free from the website of ALERT Geomaterials – <http://alertgeomaterials.eu/>.

On behalf of the ALERT Board of Directors I wish all participants a successful ALERT Doctoral School 2018!

Frédéric Collin  
Director of ALERT Geomaterials  
University of Liege



## Contents

Foreword	
I. Einav, E. Gerolymatou .....	1
Thermodynamics and Constitutive Modeling	
M. Liu .....	3
Definition and Uses of the Principle of Virtual Power	
I. Stefanou .....	43
Energetics in Discrete Element Modelling	
F. Guillard .....	71
Hierarchical guide for constructing thermodynamically admissible constitutive models	
I. Einav .....	95
Energetical background of common approaches in geomechanics	
E. Gerolymatou .....	115
An Energy Based Constitutive framework for Multiphysics Geomechanics	
M. Sari, J. Lin, S. Alevizos, T. Poulet, M. Veveakis .....	137
Energetics of Crushable Granular Materials: from Particle Fracture to Breakage Mechanics	
Y. Zhang and G. Buscarnera .....	167



---

## **Energetical Methods in Geomechanics: Foreword**

---

*What is energy? Some say coffee, others say the sun. People connect many different ideas to energy, so maybe there is something unifying behind that word, energy. Why do we get sleepless from a cup of coffee? Coffee gets us going, activating both our mind and body, a motion which turns into a finer motion, of jiggling atoms that warm up our faces and fingertips. Energy flows and transforms, so to speak. These seemingly simple observations have been ingeniously encapsulated by the laws of thermodynamics, which frame mathematically the conservation and transformation of energy from one form to another. This may not be exciting for the layman, but engineers and scientists have benefitted from satisfying such universal laws tremendously, as these limit their otherwise unbounded freedom to make mistakes.*

*How do we understand energy and what are the benefits of energetical methods in geomechanics? The purpose of this book is to address these questions, by assembling contributions from eminent researchers who have been working actively in this field. The various chapters deal with different energetical aspects in geomechanics, from the very theoretical background to specific examples and applications.*

*The first chapter "Thermodynamics and constitutive modeling" layouts the hydrodynamic procedure of thermodynamics with which it is possible to mathematically formulate the constitutive relationships of any continua, as shown for Newtonian fluids, elasticity, and granular media. The derivation initiates from local conservation laws, with global conservations automatically being satisfied through integration.*

*The second chapter a "Hierarchical guide for constructing thermodynamically admissible constitutive models" continues with the hydrodynamic procedure of thermodynamics. Specifically, it lists the generic constraints thermodynamics put on the structure of famous constitutive frameworks in geomechanics, from elasticity to hyperplasticity, from hypoplasticity to  $h^2$ plasticity, and rate dependent models.*

*The third chapter explores the "Energetics in Discrete Element Modelling" (DEM). Taking the well-known DEM as a physically simulated granular medium, coarse grained averaging is then developed to investigate the particle-scale origins of continuum energetic properties in such media. In this way, the relative abstractness of hydrody-*

## 2 Foreword

*dynamic concepts is grounded down.*

*The next four chapters focus on establishing hands on examples and applications.*

*The fourth chapter "Definition and uses of the principle of virtual power" describes a principle consistent with hydrodynamics that is typically more familiar for engineers. In this approach the general derivation begins from global balance of virtual powers, from which local relationships could be retrieved through localization of information, as demonstrated for constructing micromorphic models and finite elements.*

*The fifth chapter gives an "Energetical background of common approaches in geomechanics", looking further into well-established methods and applications of thermodynamics in the field. Further topics include the meaning of properties such as stress, the notion of energy based upscaling, minimum potential energy, maximum plastic work and second order work.*

*The next chapter provides "An energy based constitutive framework for multiphysics geomechanics". Whereas the previous chapters focused on processes where the material dependence on the temperature could be mostly ignored, this chapter explores the thermodynamics of geomaterials under extreme conditions, involving both large shear deformation and high pressures, as applicable for fault mechanics.*

*The final chapter looks into the "Energetics of crushable granular materials: from particle fracture to breakage mechanics". By connecting those processes through the eyes of thermodynamics, this chapter demonstrates the role of energetics on the development of surface-creations at very different length scales, as they govern the development of various coupled mechanical and environmental phenomena.*

*Finally, we would like to thank the outstanding contribution from all the authors of this volume. We thoroughly enjoyed reading their chapters, and hope the readers will also find them engaging, inspiring and useful for their own work in the field.*

I. Einav  
E. Gerolymatou

---

# Thermodynamics and Constitutive Modeling

**Mario Liu**

*Institute of Theoretical Physics  
Universität Tübingen  
72076 Tübingen, Germany*

---

**Abstract:** *Setting up any continuum mechanical theory, especially a constitutive model for systems as complex as granular media, it is useful to ensure explicit compliance with all constraints provided by general principles of physics, such as energy conservation and the second law of thermodynamics. A formalism that systematizes this endeavor – developed by Landau and Khalatnikov in the context of superfluid helium and referred to as the “hydrodynamic procedure” – is presented, explained and applied to polymers and granular media in this lecture.*

*Starting with the familiar example of Newtonian fluids, questions such as what entropy and energy are, and how their consideration limit the “freedom” of constitutive modeling, are answered. Remarkably, these insights suffice for a cogent derivation of the hydrodynamic theory for Newtonian fluids, ie. the Navier-Stokes equation. Next, by including the displacement as an additional state variable, this hydrodynamic theory is generalized to account for elastic media. Then, following the same procedure and introducing the concept of “transient elasticity,” two hydrodynamic theories are derived, for polymers and granular media, respectively accounting for a wide range of typical experiments. This agreement, to a large part, is the consequence of compliance with conservation laws and thermodynamics, as detailed experimental data are not part of the theory’s input.*

*The lecture is divided into two parts. Part I contains thermodynamics (section 1) and the derivation of the equations for Newtonian fluids (section 2). In Part II, the elasticity theory is first derived (section 3), then modified to suit the cases of polymers (section 4) and granular media (section 5). For polymers, one simply make the elasticity transient. For granular media, which are also transiently elastic, one needs in addition a state variable that quantifies granular jiggling. As we shall see, it is to be treated in close analogy to the temperature, by suitably extending the second law of thermodynamics.*

# 1 Thermodynamics

## 1.1 Conventional constitutive models

It is useful to have mathematical formulas for a concise account of material behavior. These are referred to as *constitutive models*. To set up such a model, one typically starts from the mass and momentum conservation,

$$\partial_t \rho + \nabla_j g_j = 0, \quad g_i = \rho v_i, \quad (1)$$

$$\partial_t g_i + \nabla_j \sigma_{ij} = -\rho G \hat{z}_i, \quad (2)$$

where  $\partial_t a \equiv \partial a / \partial t$  for any quantity,  $\rho$  is the density,  $v_i$  the velocity,  $g_i$  the momentum density,  $G$  the gravitational constant, and  $\hat{z}$  a unit vector pointing upwards. One then proposes an expression for the stress tensor  $\sigma_{ij}$  – usually taken as a function of the density  $\rho$ , the temperature  $T$ , and shear rate  $D_{ij}$  – such that the observed data are rendered as realistically as possible. If this fails to work, one looks for a differential equation for  $\sigma_{ij}$  in time. The best-known constitutive model is the Navier-Stokes stress,

$$\sigma_{ij} = P_T \delta_{ij} + \rho v_i v_j - \eta A_{ij}^* - \zeta D_{kk} \delta_{ij}, \quad (3)$$

$$\text{where } D_{ij} \equiv \frac{1}{2}(\nabla_i v_j + \nabla_j v_i), \quad (4)$$

$P_T$  is the pressure, and  $*$  denoting the traceless part of any tensor. Clearly, realism is the main advantage of this empirically driven approach. Fundamental understanding is somewhat lacking: The Navier-Stokes stress holds, to great accuracy, for any Newtonian fluid – which is a material circularly defined as one for which this expression holds. It would certainly be more satisfying to understand when and why this expression holds, and to derive equation (3) from some general considerations.

This is where thermodynamics comes in. It includes the explicit consideration of energy conservation and provides many useful constraints, strongly reducing the arbitrariness of constitutive modeling. And it always yields an explicit expression for the stress, there is never the need to resort to a differential equation. Because a constitutive model necessarily contains unphysical structure if it violates thermodynamics, and because that structure will wreck havoc somewhere in its predictions, even if it seems innocuous in the context under focus, a model derived from thermodynamics stands a much better chance of providing formulas that remain valid for the whole range of observation – rather than, as usual, being confined to the types of experiments that the empirical data are drawn from to set up the model.

## 1.2 Energy and entropy

The basic ideas of thermodynamics is easily stated. First is the fact that the total energy  $W$  of a closed system is a conserved quantity. Second, that  $W$  may be divided into a

macroscopic part  $W^M$  and a microscopic one,  $H$ , called heat. Neither is conserved, but one can only convert  $W^M$  into  $H$ , never backwards,

$$\frac{d}{dt}W = \frac{d}{dt}(W^M + H) = 0, \quad (5)$$

$$\frac{d}{dt}W^M \leq 0, \quad \frac{d}{dt}H \geq 0. \quad (6)$$

This is the essence of the second law of thermodynamics. Clearly, with  $W^M$  diminishing and  $H$  increasing continually in a closed system, we will eventually arrive at the minimum of  $W^M$  and the maximum of  $H$ , where the system will stop changing, and equilibrium reigns. This is the case of a pendulum hanging down, motionless, with all its kinetic and potential energy ( $W^M$ ) converted into heat.

Now we refine these ideas, making them more precise and useful. First, we divide all degrees of freedom into two categories, macroscopic and microscopic. Typical macroscopic degrees are the conserved densities of mass  $\rho(\vec{r}, t)$  and momentum  $g_i(\vec{r}, t)$ , coarse-grained over many particles. An example for microscopic ones is the fluctuating momentum of an atom around the average value  $g_i(\vec{r}, t)$  of many atoms.

In a macroscopic theory (such as the Navier-Stokes equation), there is no need to track every microscopic degree of freedom, hence we lump all microscopic ones into a single state variable, the entropy density  $s$ , by considering only their coarse-grained energy contributions. But we attribute a state variable to every macroscopic degree, each with its own characteristic energy contribution. (We shall return to the concept of entropy later. For now, it suffices to know that it quantifies the energy of all microscopic degrees.) As the state variables may all have a distribution in space  $\vec{r}$ , and an evolution with time  $t$ , we write the energy density as a function of all state variables,

$$w(\vec{r}, t) = w[\rho(\vec{r}, t), g_i(\vec{r}, t), \dots, s(\vec{r}, t)], \quad (7)$$

and introduce the chemical potential  $\mu$ , velocity  $v_i$  and temperature  $T$  as

$$dw = \mu d\rho + v_i dg_i + \dots + T ds, \quad \text{where} \quad (8)$$

$$\mu \equiv \frac{\partial w}{\partial \rho}, \quad v_i \equiv \frac{\partial w}{\partial g_i}, \quad \dots, \quad T \equiv \frac{\partial w}{\partial s}. \quad (9)$$

Summarily referred to as *conjugate variables*, these are useful quantities for making general considerations independent of  $w$ 's explicit form (which is a constitutive input).

If there are no more state variables, and we may drop the dots in equation (7), the equations for energy conservation and entropy balance,

$$\partial_t w + \nabla_i Q_i = 0, \quad \partial_t s + \nabla_i f_i = R/T \geq 0, \quad (10)$$

in addition to mass and momentum conservation equations (1, 2), are a complete theory. They form a closed set of equations independent of the function  $w$  if the fluxes  $\sigma_{ij}, Q_i, f_i$  and the source  $R$  are given in terms of the variables and conjugate variables. As we shall see, this will turn out to be the hydrodynamic theory of Newtonian fluids.

## 6 Thermodynamics and Constitutive Modeling

To obtain the expressions for  $\sigma_{ij}$ ,  $Q_i$ ,  $f_i$ ,  $R$ , we shall note that since the energy is a dependent quantity,  $\partial_t w = \mu \partial_t \rho + v_i \partial_t g_i + \dots + T \partial_t s$ , see equation (8), the fact that it is conserved puts a strong constraint on the possible form for the unknown fluxes.

The second law of thermodynamics, precisely and locally formulated, is:

$$\mathbf{R} = 0 \text{ in equilibrium, and } \mathbf{R} > 0 \text{ off equilibrium.}$$

The reason is the following: Consider a volume with no entropy flux  $f_i \equiv 0$ . We then have  $\partial_t H = \int d^3r (T \partial_t s) = \int d^3r R$ . Since the heat always increases,  $\partial_t H \geq 0$ , and the volume being integrated over is arbitrary, we also have  $R \geq 0$ , and may identify  $R$  as the local quantity doing the conversion from macro- to microscopic energy, usually referred to as the *entropy production*. Turning on the entropy flux does not change this identification, because the flux  $f_i$  only transfers the entropy from point to point, to equalize the temperature. In equilibrium,  $S = \int d^3r s$  is maximal, same as the heat, and  $R$  necessarily vanishes. Even though dissipation may sometimes be negligibly small, the reverse is strictly speaking also true,  $R = 0$  implies equilibrium.

### 1.3 Global and local equilibrium

The following three chapters: 1.3, 1.4, 1.5, deal with some basic thermodynamic concepts of continuum mechanical modeling. The approach may appear unfamiliar, even outlandish, and difficult at times, but the results are very useful, and easy to remember. Aim to grasp the gist of the arguments, possibly skipping the footnotes, at first read.

Elementary thermodynamics considers equilibrium states, in which the entropy,  $S = \int s d^3r$  is maximal for given energy  $\int w d^3r$ , mass  $\int \rho d^3r$  and momentum  $\int g_i d^3r$ . Varying  $S$  by changing the distribution of  $w, \rho, g_i$ , we shall soon find (in section 2) that  $S$  is maximal for  $w, s, \rho, g_i = \text{uniform}$ . This is *global equilibrium*. If they do vary,  $S$  is not maximal, and the distribution not “optimal.” The system may still be in equilibrium, but only point for point, with neighboring points in slightly different equilibria. This is referred to as *local equilibrium*. All thermodynamic relations remain unchanged – though the state variables vary in space and time, as depicted in equation (7). They dissipate, to maximize  $S$  and to approach uniformity.<sup>1</sup>

It is crucial to realize that local equilibrium is quickly reached: After a perturbation, all degrees are off equilibrium – a chaos the account of which needs a theory for all  $10^{23}$  degrees of freedom  $f_i$  in a macroscopic body. Yet most relax quickly to their equilibrium values given by the locally conserved variables,

$$f_i(\vec{r}, t) = f_i^{eq}[w(\vec{r}, t), \rho(\vec{r}, t), g_i(\vec{r}, t) \dots]. \quad (11)$$

<sup>1</sup>The equilibrium distribution of  $w, s, \rho, g_i$  is uniform only if gravitation is neglected and the system does not possess an angular momentum, executing a solid-body rotation in equilibrium. Including either, we shall find nonuniform equilibrium distributions. Clearly, local equilibrium is then given by the state variables deviating from this distribution, see section 2.

The evolution of  $w(\vec{r}, t), \rho(\vec{r}, t), g_i(\vec{r}, t)$  are much slower. To arrive at a uniform distribution, because  $w, \rho, g_i$  are conserved and cannot be locally produced, transport over macroscopic distances, of the order of the system size, is necessary.

One of the numerous  $f_i$  is the entropy density, the energy contribution of the microscopic degrees of freedom. And the relation  $s(\vec{r}, t) = s^{eq}[w(\vec{r}, t), \rho(\vec{r}, t), g_i(\vec{r}, t) \cdots]$  holds only after local equilibrium is reached. It is the inverse of equation (7), which has the same range of validity. Therefore, the knowledge of  $w(\vec{r}, t)$  and  $s^{eq}(\vec{r}, t)$  are equivalent. We may use either  $w, \rho, g_i$  or  $s^{eq}, \rho, g_i$  as the independent set of variables for thermodynamics (if there are no more state variables).

Neglecting the brief time span between total chaos and local equilibrium, we may take equation (11), or (7), to hold instantaneously. Then we only need a theory accounting for the evolution of the state variables  $\rho(\vec{r}, t), g_i(\vec{r}, t), \cdots, s(\vec{r}, t)$ , instead of all  $10^{23}$  degrees. The price is an upper bound in the theory for the frequency  $\omega$ , and wave vector  $q$ :

$$\omega\tau \ll 1, \quad q\xi \ll 1. \quad (12)$$

The first equation, with  $\tau$  being the largest of all relaxation times for  $f_i$ , states that the frequency has to be small enough for local equilibrium to always hold. The second equation states that the smallest spatial resolution, the pixel of the theory, needs to be macroscopic enough for thermodynamics to be valid.<sup>2</sup>

Occasionally, the locally conserved quantities are by themselves insufficient to characterize local equilibrium. One also need eg. the displacement vector for solids, polymers and granular media, and the director for nematic liquid crystals.<sup>3</sup>

Then there is the frequent case of *slowly relaxing state variables*. If a few relaxation times are much larger than all the others,

$$\tau_1 > \tau_2 \gg \tau_k, \quad k = 3, 4, \cdots \quad (13)$$

the range of validity for the hydrodynamic theory,  $\omega\tau_1 \ll 1$ , is unnecessarily small. To amend this, one includes the associated variables  $f_1, f_2$  as additional state variables, in effect expanding the concept of local equilibrium, with

$$w(\vec{r}, t) = w[\rho(\vec{r}, t), g_i(\vec{r}, t), s(\vec{r}, t), f_1(\vec{r}, t), f_2(\vec{r}, t)]. \quad (14)$$

The range of validity is then restored to equation (12), with  $\tau$  being the largest of  $\tau_k$ . These *slowly relaxing state variables* are important for both polymers and granular media.

Local equilibrium is the appropriate scenario for constitutive modeling. All models assume local equilibrium (though some do it implicitly) and account for how global equilibrium is established. If a constitutive model is derived employing the hydrodynamic procedure, it is referred to as a *hydrodynamic theory*.

<sup>2</sup>In fact, the pixel is typically larger, because in the time span  $\tau$  to establish local equilibrium, some transport of the conserved quantities, over the distance  $\xi(\tau)$ , has occurred. If that transport is diffusive,  $|\omega| = Dq^2$ , the inequality  $\omega\tau \ll 1$  implies  $q\xi(\tau) \ll 1$  with  $\xi = \sqrt{D\tau}$ .

<sup>3</sup>These are symmetry variables, and they are also slow if the associated continuous symmetries are spontaneously broken.

## 1.4 Statistical mechanics

Let us turn to the question: “what exactly is entropy?” The concise answer, with  $\rho_i$  the probability of the system being in the state  $i$ , and  $\langle \rangle$  being the average over all  $g$  states, is

$$S = -\langle \ln \rho_i \rangle \equiv -\sum_{i=1}^g \rho_i \ln \rho_i, \quad \sum_{i=1}^g \rho_i = 1. \quad (15)$$

As defined,  $S$  is in essence some “objective ignorance” about the system: First, note that the information is maximal and ignorance minimal if the system is with certainty in state  $i = 1$ . Or  $\rho_1 = 1$ , and  $\rho_i = 0$  for  $i \neq 1$ . The entropy  $S = -(1 \ln 1) = 0$  is also minimal then. Next, the information is minimal and ignorance maximal if the system has equal probability in any of the  $g$  state, or  $\rho_i = 1/g$ . Then the entropy

$$S = \ln g$$

is also maximal.<sup>4</sup> The last formula is usually referred to as the entropy of a *micro-canonical ensemble*. Between those two limits, ignorance grows with the entropy.

If the system is initially in one state, with  $n$  particles evolving according to the Schrödinger or Newtonian equation, we have  $S = 0$ . (This is exactly what being in a “state” means.) Any small perturbation will have the system transit to a different state. A sufficient rate of perturbations – never avoidable, however the insulation – will render the transitions so frequent, that the system is practically in a number of states simultaneously, implying an increase of the entropy.  $S$  is maximal and the system in equilibrium, if it has equal probability being in any state – all characterized by the same energy, mass and other conserved quantities. The Schrödinger or Newtonian equation (with potential energy and reversible forces) do not account for these transitions and the increase of the entropy. Constitutive modeling does, and hydrodynamic theories do it explicitly.

In local equilibrium,  $S = \ln g$  holds, with the number of accessible states  $g$  a function of the local values of  $w, \rho, g_i$ . Since all states, being in equilibrium with one another, have the same energy,  $g$  and  $\ln g$  are measures of the heat in that pixel. Finally, note that the entropy is, same as the energy, additive: If we have two subsystems, 1 and 2, in different equilibria and with  $g_1, g_2$  accessible states, the total number is  $g = g_1 \times g_2$ , and the total entropy is  $S = \ln g = \ln g_1 + \ln g_2 = S_1 + S_2$ . Or more generally,

$$S = \ln g = \ln \prod_i g_i = \sum_i \ln g_i = \sum_i S_i \longrightarrow \int s(\vec{r}, t) d^3r.$$

<sup>4</sup>To see that, realize that varying the entropy  $S$  for constant  $\sum \rho_i = 1$ , or

$$\delta(S - \lambda_1 \sum \rho_i) = 0,$$

the Euler-Lagrange condition is  $\ln \rho_i + 1 + \lambda_1 = 0$  for  $\forall i$ , or  $\rho_i = \text{const}$ . Since there are  $g$  states the system may be in, we have  $\rho_i = 1/g$  and  $S = \ln g$ .

## 1.5 Energy and angular momentum

Next, we address the question: “what exactly is energy?” noting that, same as time, energy is a basic concept hard to define concisely. There are a few steps involved. First, we define a *locally conserved quantity*  $a$  as one that satisfies the continuity equation,  $\partial_t a + \nabla_i A_i = 0$ , implying  $a$  cannot be created or destroyed locally. To change its local value, it must be transported from, or to, the neighboring volume element,  $\partial_t \int a d^3r = - \oint A_i ds_i$ . (Note that only the value of the flux  $A_i$  at the volume’s surface is relevant.) The equation for the entropy,  $\partial_t s + \nabla_i f_i = R/T \geq 0$ , on the other hand, possesses the source term  $R/T$  that accounts for the local creation of heat.

Then there is the fact that, in a closed system of particles with a known interaction (ie. with a given Hamiltonian  $\hat{H}$ ), a quantity calculable from  $\hat{H}$  that we call energy is locally conserved – *if the interaction does not depend on the absolute time*, given (say) by an externally ticking clock. But this we believe to be generally true for any physically meaningful interaction: Time is uniform, and no interaction, however complex, may depend on the absolute time. Hence there is always a locally conserved quantity called energy  $w$  – irrespective whether we have any idea about the interaction and the Hamiltonian  $\hat{H}$ .

Generally speaking,  $w$  depends on all degrees of freedom, though we subsume all microscopic ones into the entropy  $s$ . The functional form of  $w$ , since unknown, needs to be appropriately postulated to fit experiments. These are the thoughts behind equation (7).

However, the dependence of the energy  $w$  on the momentum density  $g_i$  is universal, and must not be postulated. It is simply the kinetic energy that is added to any rest-frame energy  $w_0$ , when it is being moved en bloc, whatever its interaction is. (In fact, the system may stay at rest, with the observer walking past it. Then he will also register an increase of the system’s energy by  $g_i^2/2\rho$ .) We therefore write

$$w(s, \rho, g_i) = w_0(s, \rho) + g_i^2/(2\rho), \quad \text{with} \quad (16)$$

$$dw_0(s, \rho) = T_0(s, \rho)ds + \mu_0(s, \rho)d\rho. \quad (17)$$

It is  $w_0$ , depending on one less variable, which may, and needs to, be postulated.

As a result, the conjugate variables satisfy the following relations that will prove useful later on:

$$v_i \equiv \frac{\partial w}{\partial g_i} = \frac{g_i}{\rho}, \quad T \equiv \left. \frac{\partial w}{\partial s} \right|_{\rho, g_i} = \left. \frac{\partial w_0}{\partial s} \right|_{\rho} \equiv T_0, \quad (18)$$

$$\mu \equiv \frac{\partial w}{\partial \rho} = \left. \frac{\partial}{\partial \rho} \left( w_0 - \frac{g_i^2}{2\rho} \right) \right|_{\rho, g_i} = \mu_0 - \frac{v^2}{2}. \quad (19)$$

That time’s uniformity leads to energy conservation in a close system is a special case of a general principle: The invariance of the interaction (ie. of the Hamiltonian

$\hat{H}$ ) under any continuous symmetry of time and space always gives rise to a locally conserved quantity.

Space is uniform and isotropic. Therefore, interaction must not depend on an absolute point in space (translation) or an absolute direction (rotation). It may and does depend on particle coordinates in relation to one another.) The first gives rise to the conservation of the momentum density  $\vec{g} = \rho\vec{v}$ , the second to the conservation of the angular momentum density  $\vec{\ell} = \vec{r} \times \vec{g}$ . Note that in contrast to the energy, these two do have a definite form. (The source term  $-\rho G \hat{z}_i$  in equation (18) is present, because no system of particles on earth forms a close system, even when it is isolated from other systems, as all particles interact gravitationally with the earth's mass.)

Finally, there is Galilean invariance, or the fact that empty space does not possess a preferred absolute velocity: There is no ether, and no way to determine the absolute velocity of any physical object. This gives rise to the conservation of the booster,  $\vec{b} = \rho\vec{r} - \vec{g}t$ . The discussion of the booster is usually confined to relativistic physics, but as we see below, it is more generally relevant.<sup>5</sup>

Starting from  $\partial_t \rho + \nabla_j j_j = 0$  and  $\partial_t g_i + \nabla_j \sigma_{ij} = 0$ , the local conservation of mass and momentum, and requiring the local conservation of  $\vec{\ell}$  and  $\vec{b}$ , we find

$$\begin{aligned} \partial_t \ell_m &= (\vec{r} \times \partial_t \vec{g})_m = \epsilon_{mki} r_k \partial_t g_i = -\epsilon_{mki} r_k \nabla_j \sigma_{ij} = -\nabla_j [\epsilon_{mki} r_k \sigma_{ij}] + \epsilon_{mki} \sigma_{ik}, \\ \partial_t b_i &= \partial_t \rho r_i - \partial_t g_i t - g_i = -r_i \nabla_j j_j + t \nabla_j \sigma_{ij} - g_i = \nabla_j (t \sigma_{ij} - r_i j_j) + j_i - g_i, \end{aligned}$$

and conclude that  $\epsilon_{mki} \pi_{ik}$  and  $j_i - g_i$  must vanish. This means that – independent of the system's interaction – the stress tensor is always symmetric, and the mass current is always equal to the momentum flux  $g_i = \rho v_i$ , cf. equations (1,2),

$$\sigma_{ij} = \sigma_{ji}, \quad j_i = g_i = \rho v_i. \quad (20)$$

As  $\vec{\ell}$  and  $\vec{b}$  are not independent from  $\vec{g}$  and  $\rho$ , we do not need their values for fixing the local equilibrium. And as long as equations (20) are satisfied, they are also conserved.<sup>6</sup>

## 1.6 Hydrodynamic theories

In this lecture on how to capture thermodynamic insights in constitutive modeling, we shall derive the hydrodynamic theory for four systems: Newtonian fluids, elastic media, polymers, and granular media, using the first two cases to illustrate the approach,

<sup>5</sup>Note  $\vec{\ell}$  and  $\vec{b}$  are closely related:  $\vec{b}$  is the zeroth component of the 4-angular momentum, the conservation of which is a result of the Lorentz invariance:  $\ell^{\alpha,\beta} = x^\alpha g^\beta - x^\beta g^\alpha$ ,  $x^\alpha = (ct, \vec{r})$ ,  $g^\alpha = (\varepsilon/c \approx \rho c, \vec{g})$ . Since angular momentum conservation holds independent of the inertial system, the zeroth component (that mixes with the other three under a Lorentz transformation) has to be conserved as well. In other words, taking the booster to be non-conserved is similarly illogical as saying that one component of the angular momentum (that mixes with the other two under a rotation) were non-conserved.

<sup>6</sup>The above proof depends on the explicit form of the angular momentum  $\vec{\ell} = \vec{r} \times \vec{g}$ , and the booster,  $\vec{b} = \rho\vec{r} - \vec{g}t$ . Deviations are conceivable, see [Ko98] for the associated considerations.

and showing that only a few easy steps are involved to generalize the results to the next two cases. Starting from a set of state variables, we shall always first consider the consequences of  $S = \max$  in equilibrium, and what happens when the system deviates from it. Then these results are combined with the energy conservation to setup the complete theory. We shall not apply the derived hydrodynamic theories to any experiments here, because of the lack of time and space, and because this is well rendered in the cited references.

An important advantage of the hydrodynamic approach is its clear separation between general principles and constitutive assumptions. The basic starting point is a set of state variables that defines the class of systems being described. General principles are then used to derive the evolution equations for these state variables (which are referred to as the structure of the theory). Constitutive assumptions are then needed to specify the functional dependence for the energy and the transport coefficients.

For Newtonian fluids, we have  $\rho, g_i, s$  as the state variables. The resulting Navier-Stokes equations are completely cogent and do not involve any constitutive assumptions. Adding the strain as an additional state variable, we arrive at both the elasticity theory for solids and the polymeric hydrodynamics – depending on whether the strain may remain finite in equilibrium, or vanishes (transient elasticity). As we shall see, this is equivalent to whether the system allows a plastic strain rate. Granular media is also transiently elastic, but requires the additional state variable of  $s_g$ , the granular entropy density.

In constitutive modeling, some researchers tend to postulate additional variables, for the purpose of improving the agreement with some data. Frequently, specific microscopic mechanisms are inferred to support this introduction (though the respective energy contribution is rarely specified). This is unwise. Introducing a new variable, changing the descriptive class, is a most consequential step. Without overseeing the many implications such a step entails, it is frivolous to go down this path, for the sole purpose of fitting a few experiments. The costs are untrue, possibly even grotesque, predictions elsewhere. Instead, one should rather alter the expressions for the energy and transport coefficients.

Finally, some words on history and literature. The *hydrodynamic procedure* was pioneered by Landau [LL87] and Khalatnikov [Kha65] in the context of superfluid helium, and introduced to complex fluids, specifically liquid crystals, by de Gennes [DePG93]. (Most physicists take *hydrodynamics* to mean the long-wave-length continuum theory of any condensed system, while engineers typically use it as a synonym for the Navier-Stokes equations.) Hydrodynamic theories [DeM84] have been derived for, and successfully applied to, many condensed systems, including

- liquid crystals [MPP72, Lub72, Liu79, Liu94, liu94b, PB96],
- superfluid  $^3\text{He}$  [Gra74, GP75, Liu75, LC78, LC79, Liu79],
- superconductors [Liu98, JL01, Liu02],
- macroscopic electromagnetism [HL93, Liu93, Liu95, JL96, LS09, SL15],

## 12 Thermodynamics and Constitutive Modeling

- ferrofluids [Liu95, Liu98, Liu99, ML01, ML02, MHL06, MIL08],
- polymers [TPLB00, TPLB01, PLB04, Mul06, MLPH16, MLPH16b],
- granular media [JL07, JL09, KV09, JL14, GJL11, JL03, JL04, JL07, KPBJL06, BPKMJL06, JL08, JL09, ML12, KML12, JZPFSSML12, ZLHJL12, JL13, JL13, JL15, JL16b, JL17].

## 2 Newtonian Fluids

In this section, we derive the full set of evolution equations for the Newtonian fluid, including energy conservation and the balance equation for the entropy. The purpose is to demonstrate the cogency of the hydrodynamic procedure: Given the set of state variables,  $s, \rho, g_i$ , the hydrodynamic theory, including especially the form for the stress  $\sigma_{ij}$ , is a derived result. Clearly, the definition of a Newtonian fluid is that  $s, \rho, g_i$  are a complete set of state variables, which fixes the local equilibrium state unambiguously.

### 2.1 Equilibrium conditions

We start with Eqs(8,9), but without the dots,

$$dw = Tds + \mu d\rho + v_i dg_i, \quad (21)$$

because these are the complete set of state variables for Newtonian fluids. First, we derive the equilibrium conditions by maximizing the entropy  $\int s d^3r$ , for constant energy  $\int w d^3r$ , mass  $\int \rho d^3r$  and momentum  $\int g_i d^3r$ . This is equivalent to minimizing the energy for constant entropy  $S = \int s d^3r$  (and mass, momentum), similar to the fact that a circle is either the figure of largest area for given circumference, or one of the smallest circumference for given area. Varying the energy in a closed system at rest, of given volume  $V = \int d^3r$ , entropy  $S = \int s d^3r$ , and mass  $\int \rho d^3r$ , employing  $T_L, \mu_L$  as constant Lagrange parameters, we have

$$\begin{aligned} \delta \int (w_0 - T_L s - \mu_L \rho) d^3r &= 0, \quad (22) \\ \text{or } \int [T \delta s + \mu_0 \delta \rho - T_L \delta s - \mu_L \delta \rho] d^3r \\ &= \int [(T - T_L) \delta s + (\mu_0 - \mu_L) \delta \rho] d^3r = 0. \end{aligned}$$

Because  $\delta s$  and  $\delta \rho$  vary independently, both brackets must vanish. And because  $T_L, \mu_L$  are constant,  $T, \mu_0$  also need to be. So the equilibrium conditions are

$$\nabla_i T = 0, \quad \nabla_i \mu_0 = 0. \quad (23)$$

Since  $T, \mu_0$  are functions of  $s, \rho$ , see equation (17), the latter are also constant, and with them the energy  $w$ .

Rewriting equation (17) as  $d(W_0/V) = \mu_0 d(M/V) + T d(S/V)$ , and keeping the volume  $V$  constant, we arrive at  $dW_0 = \mu_0 dM + T dS$ . Keeping instead the mass  $M$  constant, we obtain the more familiar form, with  $P_T$  denoting the thermodynamic pressure,

$$dW_0 = T dS - P_T dV, \quad P_T \equiv \mu_0 \rho + T s - w_0. \quad (24)$$

Because  $\nabla_i w_0 = T\nabla_i s + \mu_0 \nabla_i \rho$ , we also have the relation,

$$\nabla_i P_T = s\nabla_i T + \rho\nabla_i \mu_0. \quad (25)$$

Further, we note that  $P_T$  is a Galilean-invariant quantity: With  $\mu_0 \rho + Ts - w_0 = (\mu_0 - v^2/2)\rho + v_i g_i - (w_0 + \rho v^2/2) = \mu\rho + v_i g_i - w$ , see equations (16,19), we may also write

$$P_T = \mu\rho + Ts + v_i g_i - w, \quad (26)$$

$$\nabla_i P_T = s\nabla_i T + \rho\nabla_i \mu + g_j \nabla_i v_j. \quad (27)$$

### 2.1.1 gravitation

Including the gravitational energy  $\rho\phi$ , the conserved energy is  $\bar{w}_0 = w_0 + \rho\phi$ . Varying  $\int \bar{w}_0 d^3r$  as in the last section, keeping in addition  $\phi = \text{const}$ , we obtain

$$\nabla_i T = 0, \quad \nabla_i \bar{\mu}_0 = 0, \quad \text{with } \bar{\mu}_0(\rho) \equiv \partial \bar{w}_0 / \partial \rho = \mu_0 + \phi. \quad (28)$$

With  $T, \mu_0$  functions of  $s, \rho$ , these conditions imply nonuniform  $s, \rho$  being the optimal distributions that minimize the energy, or maximize the entropy. Taking  $\phi = Gh$  ( $h$  being the height) for the earth surface, we have, with equation (25) and  $\hat{z}_i$  pointing up,

$$\nabla_i \mu_0(\rho) = -\nabla_i \phi = -G\hat{z}_i, \quad (29)$$

$$\nabla_i P_T = \left. \frac{\partial P_T}{\partial \rho} \right|_T \nabla_i \rho = \rho \nabla_i \mu_0 = -\rho \nabla_i \phi. \quad (30)$$

The equation  $\nabla_i P_T = -\rho \nabla_i \phi$  is usually taken as an expression of force equilibrium, though we now realize that it expresses equilibrium and maximal entropy.

If  $\partial P_T / \partial \rho|_T$  is a constant, we have

$$\frac{P_T}{P_0} = \frac{\rho}{\rho_0} = \exp \frac{\phi}{\partial P_T / \partial \rho}. \quad (31)$$

For  $\phi = \rho Gh$  and in an ideal ideal gas,  $\partial P_T / \partial \rho = kT/m$ , this is the barometric formula.

### 2.1.2 macroscopic motion

Now we include macroscopic motion, but for simplicity neglect gravitation. We shall deduce the fact that the only motion permitted in equilibrium is a solid body rotation, or equivalently, that  $D_{ij} \equiv \frac{1}{2}(\nabla_i v_j + \nabla_j v_i) = 0$  holds. In addition,  $\nabla_i \mu = 0$  is altered to  $\partial_t v_i + \nabla_i \mu = 0$ , which is very close to the Navier-stokes equation in equilibrium. The consideration involves three additional conserved quantities: momentum, angular momentum, and booster. More details may be found in reference [Ko98], see also [LL80].

Now, we minimize the energy  $\int w d^3r$  for given entropy  $S = \int s d^3r$ , mass  $M = \int \rho d^3r$ , momentum  $\vec{G} = \int \vec{g} d^3r$ , angular momentum  $\vec{L} = \int (\vec{r} \times \vec{g}) d^3r$ , and booster  $\vec{B} = \int (\rho \vec{r} - \vec{g}t) d^3r$ , employing 11 constant Lagrange parameters (all with the subscript  $L$ ),

$$\delta[\int w d^3r - T_L \int s d^3r - \mu_L \int \rho d^3r - \vec{U}_L \cdot \int \vec{g} d^3r \quad (32)$$

$$- \vec{\Omega}_L \cdot \int (\vec{r} \times \vec{g}) d^3r + \vec{\Lambda}_L \cdot \int (\rho \vec{r} - \vec{g}t) d^3r] = 0,$$

$$\text{or } \int [(T - T_L) \delta s + (\mu + \vec{\Lambda}_L \cdot \vec{r} - \mu_L) \delta \rho \quad (33)$$

$$+ (\vec{v} - \vec{U}_L - \vec{\Omega}_L \times \vec{r} - \vec{\Lambda}_L t) \cdot \delta \vec{g} \\ + (\vec{\Lambda}_L \rho - \vec{g} \times \vec{\Omega}_L) \cdot \delta \vec{r} - \vec{\Lambda}_L \cdot \vec{g} \delta t] d^3r = 0.$$

As  $\delta s, \delta \rho, \delta \vec{g}, \delta \vec{r} = \text{const}, \delta t = \text{const}$  all vary independently, we conclude

$$M \vec{\Lambda}_L = \vec{G} \times \vec{\Omega}_L, \quad \vec{\Lambda}_L \cdot \vec{G} = 0, \quad T = T_L, \quad (34) \\ \mu = \mu_L - \vec{\Lambda}_L \cdot \vec{r}, \quad \vec{v} = \vec{U}_L + \vec{\Omega}_L \times \vec{r} + \vec{\Lambda}_L t,$$

where  $T(\vec{r}, t), \mu(\vec{r}, t), \vec{v}(\vec{r}, t)$  are the fields,  $M, \vec{G}$  the conserved quantities, and  $T_L, \mu_L, \vec{U}_L, \vec{\Omega}_L, \vec{\Lambda}_L$  the constant Lagrange parameters. (We note that the condition  $\vec{\Lambda}_L \cdot \vec{G} = 0$  may be obtained by multiplying the first with  $\vec{G}$ .) A temporal derivative of the last condition yields

$$\vec{\Lambda}_L = \partial_t \vec{v}. \quad (35)$$

Spatial derivatives then yield the equilibrium conditions,

$$\vec{\nabla} T = 0, \quad \vec{\nabla} \mu + \partial_t \vec{v} = 0, \quad D_{ij} \equiv \frac{1}{2} (\nabla_i v_j + \nabla_j v_i) = 0. \quad (36)$$

To understand equilibrium motion better, we consider a steady motion of the center of mass,  $\vec{R} = \vec{R}_0 + t \partial_t \vec{R}$ , in addition to a solid-body rotation

$$\vec{v} = \partial_t \vec{R} + \vec{\Omega} \times (\vec{r} - \vec{R}) = \quad (37) \\ (\partial_t \vec{R} - \vec{\Omega} \times \vec{R}_0) + \vec{\Omega} \times \vec{r} + (\partial_t \vec{R} \times \vec{\Omega}) t,$$

$$\text{finding } \vec{U}_L = \partial_t \vec{R} - \vec{\Omega} \times \vec{R}_0, \quad (38)$$

$$\vec{\Omega} = \vec{\Omega}_L, \quad \vec{\Lambda}_L = \partial_t \vec{R} \times \vec{\Omega}. \quad (39)$$

For  $\partial_t \vec{R} \parallel \vec{\Omega}$ , or  $\partial_t \vec{R} = 0$ , or  $\vec{\Omega} = 0$ , we have  $\partial_t \vec{v} = \vec{\Lambda}_L = 0$ . Otherwise, the coordinate  $\vec{r}$  moves with respect to the center of mass  $\vec{R}$ , implying  $\partial_t \vec{v} \neq 0$ , and rendering the chemical potential nonuniform. With  $M \vec{R} = \int \rho \vec{r} d^3r$ , we have

$$\vec{G} = \int \rho \vec{v} d^3r = \int \rho [\partial_t \vec{R} + \vec{\Omega} \times (\vec{r} - \vec{R})] d^3r = M \partial_t \vec{R}, \quad (40)$$

$$\vec{B} = \int (\rho \vec{r} - \vec{g}t) d^3r = M \vec{R} - \vec{G}t = M(\vec{R} - \partial_t \vec{R}t) = M \vec{R}_0. \quad (41)$$

That the booster is given by  $\vec{B} = M\vec{R}_0$ , an initial condition, is not odd. The momentum is also given by its initial value,  $\vec{G} = \vec{G}_0$ . There are many initial conditions, but only a few locally conserved densities.

Choosing an inertial frame for which  $\vec{R}, \partial_t \vec{R} = 0$ , employing equation (25) and  $\nabla_i \mu = \nabla_i(\mu_0 - v^2/2) = 0$ , we obtain

$$\nabla_i P_T = \rho \nabla_i \mu_0 = \rho \nabla_i v^2/2 = \rho \vec{v} \cdot \nabla_i(\vec{\Omega} \times \vec{r}) = \rho(\vec{v} \times \vec{\Omega}). \quad (42)$$

Clearly, this time the expression of maximal  $S$  yields the centrifugal force. And the associated density distribution, same as in equation (44), is given by

$$\frac{P_T}{P_0} = \frac{\rho}{\rho_0} = \exp \frac{v^2/2}{\partial P_T / \partial \rho}, \quad (43)$$

if  $\partial P_T / \partial \rho$  may be approximated as constant.

Combining macroscopic motion with gravitation, we need to have  $\vec{G} \perp \hat{z}$  and  $\vec{\Omega} \parallel \hat{z}$  to preserve translational and rotational symmetry, and the conservation of momentum and angular momentum. Then  $\vec{\nabla} \mu + \partial_t \vec{v} = 0$  holds in equilibrium (with  $\bar{\mu}$ !), implying

$$\vec{\nabla} \mu + \partial_t \vec{v} = -\rho G \hat{z}, \quad (44)$$

see equations (29, 30).

### 2.1.3 thermodynamics in astronomy

It is worth noting that the reason for us earthlings being able to see only one side of the moon is because the moon's spin  $\omega$  and its orbital rotational velocity around the earth  $\Omega$  are equal,  $\omega = \Omega$ , an expression of a solid-body rotation. This means that the spinning of moon is already in equilibrium with its orbital rotation, while the earth's spinning, with a much larger kinetic energy, is not yet in equilibrium. To account for this dynamics, a dissipative force in the Newtonian equations  $\sim \omega - \Omega$  would be necessary.

## 2.2 The entropy production $R \geq 0$

Since the vanishing of  $\nabla_i T, D_{ij}$  and  $\partial_t v_i + \nabla_i \mu$  is necessary and sufficient for equilibrium to hold, and since the same is true for  $R$ , we take  $R$  as a function of these three fields. (They are usually referred to as *thermodynamic forces*.) Expanding  $R$  in them, the lowest order terms are of second order: A constant term would imply  $R \neq 0$  in equilibrium, while linear terms are not positive definite. However, as we shall see in the next section, given  $\nabla_i T, D_{ij} = 0$ , the vanishing of  $\partial_t v_i + \nabla_i \mu$  follows from the structure of the Navier-Stokes equations, or more generally, from that of momentum conservation. It is not independent. Therefore, we take  $R$  given as

$$R = \kappa(\nabla_i T)^2 + \eta A_{ij}^* A_{ij}^* + \zeta D_{\ell\ell}^2, \quad (45)$$

where  $D_{\ell\ell}$  is the trace of  $D_{ij}$ , and  $A_{ij}^* = D_{ij} - \frac{1}{3}D_{\ell\ell}\delta_{ij}$  its traceless part. Given the isotropy of Newtonian fluids in equilibrium, these are all possible quadratic terms. There are no others. Equation 45 is therefore the most general form.

The expansion coefficients  $\kappa, \eta, \zeta$  will be identified as transport (or Onsager) coefficients in the next section. They are functions of the state variables, or equivalently, the conjugate variables. Because of the expansion, they must not depend on the thermodynamic forces, especially not on  $D_{ij}$  (a frequent error). Moreover, because the entropy  $S = \ln g$  is Galilean-invariant and cannot depend on the inertial frame, so must its source term  $R$  be. Since the thermodynamic forces:  $\nabla_i T, D_{ij}$  are also Galilean-invariant, so must  $\kappa, \eta, \zeta$  be. Therefore, they only depend on  $s, \rho$ , or  $T, P_T$ , not  $v_i, g_i$ . These three functions, in addition to that for the energy, are the only constitutive choices one can make for any Newtonian Fluid – though its structure (as derived in the next section) holds independent of these choices.

### 2.3 The evolution equations

Neglecting gravitation, the evolution equations of the state variables have the form of continuity equations, except  $s$  that also possesses a source term,

$$\begin{aligned} \partial_t \rho + \nabla_i j_i &= 0, & \partial_t s + \nabla_i f_i &= R/T \geq 0, \\ \partial_t w + \nabla_i Q_i &= 0, & \partial_t g_i + \nabla_k \sigma_{ik} &= 0. \end{aligned} \quad (46)$$

We already know that  $j_i = g_i = \rho v_i$  and  $\sigma_{ik} = \sigma_{ki}$ , cf equation (20).

Next, we determine the the fluxes:  $f_i, Q_i$  and  $\sigma_{ik}$ . We do this in two steps, first what these fluxes are in equilibrium, then how they get modified off equilibrium.

In a macroscopic motion permitted in equilibrium, the fields  $\rho, s$ , of each mass point moving with  $v_i$ , remains unchanged in time,  $\partial_t \rho + v_i \nabla_i \rho = 0$  and  $\partial_t s + v_i \nabla_i s = 0$ . Since the last of equation (36) implies  $\nabla_k v_k = 0$ , we may also write

$$\partial_t \rho + \nabla_i (\rho v_i) = 0, \quad \partial_t s + \nabla_i (s v_i) = 0.$$

Next, starting from  $\rho(\nabla_i \mu + \partial_t v_i) = 0$ , and adding  $v_i[\partial_t \rho + \nabla_j (\rho v_j)] = 0$ , we find  $\partial_t g_i + \rho \nabla_i (\mu_0 - v^2/2) + v_i \nabla_j (\rho v_j) = 0$ , or equivalently,

$$\partial_t g_i + \nabla_j (P_T + \rho v_i v_j) = 0,$$

because  $\nabla_i P_T = \rho \nabla_i \mu_0$ , equation (25), and with  $\nabla_i v_j + \nabla_j v_i = 0$ , also  $-\rho \nabla_i (v_j^2/2) = -\rho v_j \nabla_i v_j = \rho v_j \nabla_j v_i$ . As a result, we conclude that the fluxes in equilibrium are

$$j_i^{\text{eq}} = \rho v_i, \quad f_i^{\text{eq}} = s v_i, \quad \sigma_{ij}^{\text{eq}} = P_T \delta_{ij} + \rho v_i v_j. \quad (47)$$

Including gravitation, all fluxes remain the same, though because of equation (44), momentum conservation is modified to

$$\partial_t g_i + \nabla_j (P_T + \rho v_i v_j) = -\rho G \hat{z}_i.$$

Off equilibrium, we denote the (dissipative) modifications of the fluxes with a superscript  $D$  (for dissipative),

$$f_i = f_i^{\text{eq}} - f_i^D, \quad \sigma_{ij} = \sigma_{ij}^{\text{eq}} - \sigma_{ij}^D, \quad (48)$$

noting that  $j_i^D = j_i - j_i^{\text{eq}} = g_i - \rho v_i = 0$ . To obtain  $f_i^D$ ,  $\sigma_{ij}^D$ , and  $Q_i$ , we differentiate equation (21),

$$\partial_t w = T \partial_t s + \mu \partial_t \rho + v_i \partial_t g_i = -\nabla_i Q_i, \quad (49)$$

and require this equality to hold generally, independent of how  $w$  depends on  $s, \rho$ . Next, inserting the expressions from Eqs(46,47,48), employing equation (27), and noting  $\sigma_{ik} \nabla_i v_k = \frac{1}{2}(\sigma_{ik} \nabla_i v_k + \sigma_{ki} \nabla_i v_k) = \sigma_{ik} D_{ik}$ , equation (20), we rewrite  $\nabla_i Q_i$  as

$$\begin{aligned} \nabla_i Q_i &= \nabla_i(\mu j_i + T f_i + v_k \sigma_{ik}) \\ &\quad - R + f_i^D \nabla_i T + \sigma_{ij}^D D_{ij} - (j_i - \rho v_i) \nabla_i \mu. \end{aligned} \quad (50)$$

This is a unique expression, because one needs to rewrite all terms such that they are either part of a divergence, or vanish in equilibrium. (For instance, we write  $T \partial_t s = -T \nabla_i f_i + \dots$  as  $f_i \nabla_i T - \nabla_i(T f_i)$ , and deduce that first term belongs to  $R$ , and the second to  $Q_i$ .) Therefore, we conclude

$$Q_i = \mu j_i + T f_i + v_k \sigma_{ik}, \quad (51)$$

$$R = f_i^D \nabla_i T + \sigma_{ij}^D D_{ij}. \quad (52)$$

Finally, comparing equation (52) to equation (45), we obtain

$$f_i^D = \kappa \nabla_i T, \quad \sigma_{ij}^D = \eta D_{ij}^* + \zeta \delta_{ij} D_{\ell\ell}. \quad (53)$$

Note that, for  $\nabla_i T, D_{ij} = 0$ , the Navier-Stokes equations,  $\partial_t g_i + \nabla_k \sigma_{ik} = -\rho G \hat{z}_i$  automatically reduces to  $\partial_t v_i + \nabla_i \mu = -\rho G \hat{z}_i$ : Given the former, we are left with the equilibrium fluxes, and we know that  $\partial_t g_i + \nabla_k \sigma_{ik} = -\rho G \hat{z}_i$  and  $\partial_t v_i + \nabla_i \mu = -\rho G \hat{z}_i$  are then equivalent.

This concludes the derivation of the full hydrodynamic theory for Newtonian fluids.

### 3 Elastic Media

#### 3.1 Small deformations

Elasticity of small deformations is simpler and facilitates an easy grasp of its essence. We therefore consider it first, by introducing the displacement vector  $u_i$ , its evolution equation,

$$\partial_t u_i = v_i - y_i^D, \quad (54)$$

and the elastic strain, on which the energy depends,

$$\varepsilon_{ij}^e = -\frac{1}{2}(\nabla_i u_j + \nabla_j u_i). \quad (55)$$

Note there are different sign notations in elasticity, for both the stress and the strain. We use the convention of Soil Mechanics, given by equation (55) above, implying

$$\dot{\varepsilon}_{ij} \equiv -D_{ij} = -\frac{1}{2}(\nabla_i v_j + \nabla_j v_i), \quad (56)$$

$$\partial_t \varepsilon_{ij}^e = -D_{ij} + \frac{1}{2}(\nabla_i y_j^D + \nabla_j y_i^D), \quad (57)$$

$$\equiv \dot{\varepsilon}_{ij} - \dot{\varepsilon}_{ij}^{pl} \quad (58)$$

We write  $\dot{\varepsilon}_{ij}$ ,  $\dot{\varepsilon}_{ij}^{pl}$  instead of  $\partial_t \varepsilon_{ij}$ ,  $\partial_t \varepsilon_{ij}^{pl}$  because  $\varepsilon_{ij}^{pl}$  does not usually exist, and  $\varepsilon_{ij} = \varepsilon_{ij}^e$  exists only if  $\dot{\varepsilon}_{ij}^{pl} \equiv 0$ . It is preferable to call  $\varepsilon_{ij}^e$ , as defined, the *elastic strain*, because the elastic energy depends on it. Besides, the plastic rate is clearly non-zero.

The stress sign remains the same as given above, see equation (2). This is the same sign notation as in [EL18], but differ from all other references on polymeric dynamics and GSH by a minus sign with respect to the elastic strain fields, referred to as  $u_{ij} = -\varepsilon_{ij}^e$  there.

Terms of second order in  $\nabla_i u_j$ , and convective terms such as  $v_k \nabla_k u_i$  are neglected in this section, because we focus on small deformations here.

In the evolution equation (54), the equilibrium flux is given by  $v_i$ : Moving a solid with a constant velocity  $v_i$ , the displacement changes as  $v_i t$ . The dissipative contribution  $y_i^D$ , to be determined below, will turn out to be one that aims to render the stress uniform (in one dimension). The quantity  $y_i^D$  is frequently small, and typically neglected. But it is nonzero generally, and useful for understanding plastic strain rates.

Since the energy  $w$  now depends on  $\varepsilon_{ij}^e$ , we add a term to equation (21),

$$dw = T ds + \mu d\rho + v_i dg_i + \pi_{ij} d\varepsilon_{ij}^e. \quad (59)$$

Because  $\pi_{ij} \equiv \partial w / \partial \varepsilon_{ij}^e$  is symmetric,  $\pi_{ij} = \pi_{ji}$ , we may also write

$$dw = T ds + \mu d\rho + v_i dg_i - \pi_{ij} d\nabla_j u_i. \quad (60)$$

Minimizing this energy, same as we did in equation (22),  $\delta \int (w - T_L s - \mu_L \rho) dV = 0$ , with the additional constraint that  $\delta u_i$  vanish at the surface,  $\oint \pi_{ij} \delta u_i dA_i = 0$ , we find

$$\begin{aligned} & \int [T \delta s + \mu \delta \rho - \pi_{ij} \delta \nabla_j u_i - T_L \delta s - \mu_L \delta \rho] dV = \\ & \int [(T - T_L) \delta s + (\mu - \mu_L) \delta \rho + (\nabla_j \pi_{ij}) \delta u_i] dV = 0. \end{aligned}$$

Since  $\delta s$ ,  $\delta \rho$  and  $\delta u_j$  vary independently, we have the same equilibrium conditions as in equation (23) – or in equations (28, 36, 44), and in addition

$$\nabla_j \pi_{ij} = 0. \quad (61)$$

The entropy production  $R$  is now a quadratic function of three fields,  $\nabla_i T$ ,  $D_{ij}$  and  $\nabla_j \pi_{ij}$ . Neglecting possible mix terms for simplicity (but see below), the quadratic form is

$$R = \kappa_{ij} \nabla_i T \nabla_j T + \eta_{ijkl} D_{ij}^* D_{kl}^* + \zeta D_{\ell\ell}^2 + \xi_{ij} \nabla_k \pi_{ik} \nabla_m \pi_{jm}. \quad (62)$$

Next, we look for the equilibrium form of the stress tensor, adding tentatively the term  $\pi_{ij}$  to the expression in equations (47), with  $P_T$  still defined as in equation (26),

$$\sigma_{ij}^{\text{eq}} = \pi_{ij} + P_T \delta_{ij} + \rho v_i v_j. \quad (63)$$

Because  $\nabla_j \pi_{ij} = 0$  in equilibrium, momentum conservation  $\partial_t g_i + \nabla_j \sigma_{ij} = 0$  reduces, as before, to  $\partial_t v_i + \nabla_i \mu = 0$ . (Note that the pressure gradient now has an extra term,  $\nabla_i P_T = \dots + \pi_{kj} \nabla_i \varepsilon_{kj}$ , because the energy also does, cf. equation (59). This term may be neglected for small deformations.)

Next, starting from equations (46,54), differentiating the energy, equation (60),

$$\partial_t w = T \partial_t s + \mu \partial_t \rho + v_i \partial_t g_i - \pi_{ij} \nabla_j \partial_t u_i = -\nabla_i Q_i. \quad (64)$$

concentrating on the equilibrium fluxes, and inserting the expressions from equations (46,47) as before, and now also equation (54), with  $y_i^D = 0$ , we find that, indeed, only with  $\pi_{ij}$  in  $\sigma_{ij}^{\text{eq}}$  is energy conserved,

$$\partial_t w = v_i \partial_t g_i - \pi_{ij} \nabla_j \partial_t u_i + \dots = -v_i \nabla_j \pi_{ij} - \pi_{ij} \nabla_j v_i + \dots = -\nabla_j (v_i \pi_{ij}) + \dots$$

Also, we see that there is an additional term in the energy flux,  $Q_i = v_i \pi_{ij} + \dots$ .

To obtain the dissipative, off-equilibrium terms, we also insert the expressions from equations (48), including  $y_i^D \neq 0$ , and employ equation (27), note  $\sigma_{ik} \nabla_i v_k = \sigma_{ik} D_{ik}$  [cf. equation (20)], to arrive at

$$\begin{aligned} \nabla_i Q_i &= \nabla_i [\mu j_i + T f_i + v_k \sigma_{ik} - \pi_{ij} y_i^D], \\ &- R + f_i^D \nabla_i T + \sigma_{ij}^D D_{ij} + y_i^D \nabla_j \pi_{ij} \end{aligned} \quad (65)$$

and conclude

$$Q_i = \mu j_i + T f_i + v_k \sigma_{ik} - \pi_{ij} y_i^D, \quad (66)$$

$$R = f_i^D \nabla_i T + \sigma_{ij}^D D_{ij} + y_i^D \nabla_j \pi_{ij}. \quad (67)$$

Because of equation (62), the fluxes:  $f_i^D, \sigma_{ij}^D, y_i^D$  are linear functions of forces:  $\nabla_i T, D_{ij}, \nabla_j \pi_{ij}$ , and we conclude

$$f_i = \kappa_{ij} \nabla_j T, \quad \sigma_{ij}^D = \eta_{ijkl} D_{kl}, \quad y_i^D = e_{ij} \nabla_k \pi_{jk}, \quad (68)$$

with  $\kappa_{ij}, \eta_{ijkl}, e_{ij}$  being transport coefficients, functions of the state variables. The thermodynamic force  $\nabla_j \pi_{ij}$  redistributes strain and stress, toward the equilibrium condition of  $\nabla_j \pi_{ij} = 0$ . As long as this condition is not satisfied, we have  $y_i^D \neq 0$ , and neither does the plastic strain rate vanish.

More generally, cross coupling are allowed and given by the Onsager matrix,

$$\begin{pmatrix} f_i^D \\ y_i^D \\ \sigma_{ij}^D \end{pmatrix} = \begin{pmatrix} \kappa_{ik} & b_{ik} & c_{ikl} \\ b_{ki} & e_{ik} & h_{ikl} \\ -c_{kli} & -h_{kli} & \eta_{ijkl} \end{pmatrix} \cdot \begin{pmatrix} \nabla_k T \\ \nabla_l \pi_{kl} \\ D_{kl} \end{pmatrix}. \quad (69)$$

Note the time-inversion properties of the forces:  $\nabla_i T, \nabla_j \pi_{ij}$  are positive,  $D_{ij}$  is negative. Hence  $b_{ki}$  is symmetric, while  $c_{ikl}, h_{ikl}$  are antisymmetric. The non-negativeness of  $R$  is ensured by the following procedure: First, taking both  $\sigma_{ij}^D$  and  $D_{kl}$  as a six-tuple vector, and writing  $\eta_{ijkl}$  as a 6x6 matrix,  $\eta_{\alpha,\beta}$  ( $\alpha, \beta$  going from 1 to 6), we require it to have only positive Eigenvalues. Next, taking  $(f_i, y_i^D)$  and  $(\nabla_k T, \nabla_l \pi_{kl})$  as two six-tuple vectors, we again require the 6x6 matrix of coefficients connecting them to have positive Eigenvalues. Note that the antisymmetric coefficients,  $c_{kli}$  and  $h_{kli}$ , do not contribute to  $R$ , there are no constraints for them. For details see [DeM84].

In the above treatment, we took the density  $\rho$  as a variable independent from the elastic strain  $\varepsilon_{ij}^e$ . This is the general case. But if one neglects the effect of mass defects [MPP72], smallish in solids, the density is a dependent quantity, the variation of which is given by

$$d\rho/\rho = d\varepsilon_{\ell\ell}^e, \quad (70)$$

Eliminating  $\rho$  as a variable implies  $P_T \equiv 0$ , especially in equation (63). To see that, take the rest-frame expression of equation (59),  $dw_0 = T ds + \mu_0 d\rho + \pi_{ij} d\varepsilon_{ij}^e$ , and rewrite it as

$$d(w_0/\rho) = T d(s/\rho) + (P_T/\rho^2) d\rho + (\pi_{ij}/\rho) d\varepsilon_{ij}^e. \quad (71)$$

Since  $P_T/\rho^2$  is a density derivative taken at constant  $\varepsilon_{ij}^e$ , and constant  $s/\rho$ , it vanishes. (The chemical potential  $\mu$  vanishes at constant  $\varepsilon_{ij}^e$  and constant  $s$ . Eq.(70), if taken as without any specification of what is being held constant, is not well defined.)

### 3.2 Large deformations

Generally speaking, the description of an elastic body relies on two coordinates: the actual spatial coordinate  $r_i$ , specifying a point in an elastic body, and the initial coordinate  $a_i$ , see [TPLB00], also [TPLB01, PLB04]. Starting from a stress-free elastic body, we consider a point with the coordinate  $a_i$ . As the body is translated, rotated, compressed and sheared,  $a_i$  is displaced to  $r_i$ . In general,  $r_i$  may be quite remote from  $a_i$  – especially in soft matter or a metal sheet. The function  $r_i(a_j)$  is unique and invertible to  $a_j(r_i)$ . The first is the Lagrangian description, the second the Eulerian one. Their difference is relevant only for large deformations. We shall refer to  $r_i$  as the real space, and  $a_i$  as the initial space, and both  $\partial r_j/\partial a_i$  and  $\nabla_i a_j \equiv \partial a_j/\partial r_i$  as the deformation tensor (since only one of either is employed in any description).

As discussed in most books on elasticity theory, see eg [LL86], the elastic energy depends on the change in distance between two neighboring points, from  $da_i^2$  to  $dr_i^2$ . Defining the displacement vector as  $u_i(a_j) = r_i(a_j) - a_i$ , and the strain tensor as

$$\varepsilon_{ik}^L = -\frac{1}{2}[\partial u_i/\partial a_k + \partial u_k/\partial a_i + (\partial u_j/\partial a_i) \cdot (\partial u_j/\partial a_k)], \quad (72)$$

we have  $dr_i^2(a_m) - da_i^2 = -2\varepsilon_{ik}^L da_i da_k$ . The elastic energy is a function of  $\varepsilon_{ik}^L$ . It is, to lowest order, simply  $w = \frac{1}{2} K_{ikjlm} \varepsilon_{ik}^L \varepsilon_{jm}^L$ .

The special point here is that both the strain tensor and the energy density are functions of the initial coordinate  $a_j$  – hence the superscript in  $\varepsilon_{ik}^L$ , for Lagrangian. Contrast this with the energy density of an isotropic liquid in its rest frame, a function of the mass and entropy density,  $w(\rho, s)$  – or equivalently,  $dw = Tds + \mu d\rho$ . All variables, including the conjugate ones, temperature  $T$  and chemical potential  $\mu$ , are here functions of the real coordinate  $r_m$ , see equation (7). This is the Euler notation, the basic advantage of which is that physics, which we insist must be local, is also expressed in local terms, accounted for by quantities at the real coordinates  $r_m$ . Consider for instance the diffusive heat current, which is given by the local gradient of the temperature,  $\sim \partial T(r_m)/\partial r_k$ , only in the Eulerian description.

Returning to elastic media, we have two choices: First, take all variables including especially the temperature and chemical potential as functions of  $a_m$ , and employ them with the strain tensor  $\varepsilon_{ik}^L$ . This would be consistent, but highly inconvenient. For instance, the heat current  $\sim \partial T(r_m)/\partial r_i$  at the real space point  $r_m$  now presumes the knowledge (not usually available) of the global transformation,  $r_m \leftrightarrow a_m$ , as  $\partial T(r_m)/\partial r_i = [\partial T(a_m)/\partial a_k](\partial a_k/\partial r_i)$ . Similarly, with  $\mathbf{g}$  the momentum density, the angular momentum density is  $\mathbf{r}(a_m) \times \mathbf{g}(a_m)$ , not  $\mathbf{a} \times \mathbf{g}(a_m)$ . (If the system is only weakly deformed, with  $u_i = r_i - a_i$  small, the above differences between  $r_i$  and  $a_i$  may be neglected to linear order, as we did in the last section. )

The second, and the only actually viable, choice is to take all variables including the strain tensor in the local, Eulerian notation, as functions of  $r_m$ . We therefore employ the Eulerian strain tensor [CL95], introduced via  $dr_i^2 - da_i^2(r_m) = 2\varepsilon_{ik}^e(r_m) dr_i dr_k$ ,

where the superscript also stands for elastic,  $u_i(r_m) = r_i - a_i(r_m)$ , and

$$\varepsilon_{ik}^e(r_m) = -\frac{1}{2}[\partial u_i/\partial r_k + \partial u_k/\partial r_i - (\partial u_j/\partial r_i)(\partial u_j/\partial r_k)]. \quad (73)$$

There is a second point, essential yet somewhat subtle: We need to eliminate the displacement field  $u_i$ , to deal exclusively with the initial coordinate  $a_i(r)$  and the elastic strain  $\varepsilon_{ik}^e$ . This is possible because starting again from  $dr_i^2 - da_i^2(r_m) = 2\varepsilon_{ik}^e(r_m)dr_idr_k$ , we find

$$\varepsilon_{ik}^e = -\frac{1}{2}[\delta_{ik} - (\partial a_\alpha/\partial r_k)(\partial a_\alpha/\partial r_i)], \quad (74)$$

with no need whatever for a detour via  $u_i$ . This is necessary because the introduction of  $u_i$  destroys an inbuilt symmetry and represents an *ad hoc* choice. As discussed,  $a_i$  and  $r_i$  are vectors of different spaces, and they transform as vectors, under rotations in initial and real space, respectively. The introduction of the displacement fixes both spaces with respect to each other, and prohibits the rotation of either space alone. No physical field depends on the orientation of the initial space, the fictitious unstressed body. Given any relation  $a_m \leftrightarrow r_m$ , we are still free to take a global but arbitrary rotation of all  $a_i$ , to rotate the initial space with respect to the real space. Therefore, we must treat  $a_\alpha$  and  $r_i$  as vectors of two different spaces, and a quantity such as  $\nabla_k a_\alpha \equiv \partial a_\alpha/\partial r_i$  is a bi-vector, not a second rank tensor. We use Latin indices to denote the components (x,y,z) in real space, and Greek indices for (1,2,3) in initial space. Clearly, this renders the fact that the displacement  $r_i - a_\alpha$  is an oxymoron rather obvious. The equation of motion for  $a_\alpha$  is

$$\frac{d}{dt}a_\alpha \equiv \frac{\partial}{\partial t}a_\alpha + v_k \nabla_k a_\alpha = \hat{y}_\alpha^D, \quad (75)$$

because in equilibrium, with the dissipative contribution  $\hat{y}_\alpha^D$  vanishing, this equation states the simple fact that the initial coordinate  $a_\alpha$  of a mass point does not change when one moves with it. The energy density is now

$$dw = Tds + \mu d\rho + v_i dg_i + \psi_{\alpha i} d\nabla_i a_\alpha, \quad (76)$$

with the equilibrium condition [obtained with essentially identical algebra as for equation (61)]

$$\nabla_i \psi_{\alpha i} = 0. \quad (77)$$

Entropy production  $R$  is now a quadratic function of  $\nabla_i T$ ,  $D_{ij}$  and  $\nabla_i \psi_{\alpha i}$ . The pressure, still  $P_T = \mu\rho + Ts + v_i g_i - w$ , equation (27), has the gradient

$$\nabla_i P_T = s\nabla_i T + \rho\nabla_i \mu + g_j \nabla_i v_j - \psi_{\alpha k} \nabla_i \nabla_k a_\alpha. \quad (78)$$

The equilibrium part of the stress tensor is given as

$$\sigma_{ij}^{eq} = \psi_{\alpha j} \nabla_i a_\alpha + P_T \delta_{ij} + \rho v_i v_j, \quad (79)$$

because  $\nabla_j(\psi_{\alpha j} \nabla_i a_\alpha) + \dots = \psi_{\alpha j} \nabla_j \nabla_i a_\alpha + \dots$  cancels the last term in equation (78), again reducing  $\partial_t g_i + \nabla_j \sigma_{ij}^{eq} = 0$  to  $\nabla_i \mu + \partial_t v_i = 0$ . We also note that  $\psi_{\alpha j} \nabla_i a_\alpha$  reduces to  $\pi_{ij}$  for small deformations, see equation (85).

The new term  $(\psi_{\alpha i} \nabla_j a_\alpha)$  is symmetric: The energy  $w$ , a scalar, is invariant under a real space rotation of the angle  $d\theta_i$ , but  $\nabla_i a_\alpha$ , a vector, is not,  $d\nabla_i a_\alpha = \epsilon_{ijk} \nabla_j a_\alpha d\theta_k$ , where  $\epsilon_{ijk}$  is the total antisymmetric, Levi-Civita tensor. Since  $0 = dw = \psi_{\alpha i} d\nabla_i a_\alpha = \psi_{\alpha i} \epsilon_{ijk} \nabla_j a_\alpha d\theta_k$ , we have the rotation identity,

$$(\psi_{\alpha i} \nabla_j a_\alpha) = (i \leftrightarrow j). \quad (80)$$

The hydrodynamic procedure, just as in the last section, delivers the energy flux and the dissipative contributions [as given in equations (48, 75)],

$$Q_i = \mu j_i + T f_i + v_k \sigma_{ik} + \psi_{i\alpha} \hat{y}_\alpha^D, \quad (81)$$

$$R = f_i^D \nabla_i T + \sigma_{ij}^D D_{ij} + \hat{y}_\alpha^D \nabla_i \psi_{\alpha i}. \quad (82)$$

For an isotropic medium, without any cross couplings, this implies [cf. equation (68)]

$$f_i^D = \kappa \nabla_i T, \quad \sigma_{ij}^D = \eta D_{ij}^* - \zeta \delta_{ij} D_{\ell\ell}, \quad \hat{y}_\alpha^D = \hat{e} \nabla_i \psi_{\alpha i}. \quad (83)$$

Again, the thermodynamic force  $\nabla_i \psi_{\alpha i}$  redistributes the strain and stress, toward the equilibrium condition of  $\nabla_i \psi_{\alpha i} = 0$ . More generally, one needs to reconsider the Onsager matrix, as in equation (69).

Confining to isotropic elastic media (such as glass or polymers) that do not depend on the orientation, and are equally compliant being compressed along  $\alpha = 1, 2$  or  $3$ , allows us to rewrite the large-deformation elasticity theory in terms of the more familiar  $\pi_{ij}$  and  $\epsilon_{ij}^e$ . This is what we shall do next. [We note that the bi-vector  $\nabla_i a_\alpha$ , with nine components, not only contains the information about the strain  $\epsilon_{ij}^e$  (six components), but also that about the local orientation, ie. the rotation matrix  $R_{\alpha i}$  that rotates real space to initial space (three components). Neglecting the dependence of the energy on the latter, we may, by comparing equation (59) to (76), write  $-\pi_{ij} d\epsilon_{ij}^e = \psi_{\alpha i} d\nabla_i a_\alpha$ , or

$$\psi_{\alpha i} = -\pi_{km} (\partial \epsilon_{km}^e / \partial \nabla_i a_\alpha).$$

Employing equation (74), we deduce

$$\psi_{\alpha i} = \pi_{ij} \nabla_j a_\alpha, \quad \psi_{\alpha i} \nabla_j a_\alpha = \pi_{ij} - 2\pi_{ik} \epsilon_{kj}^e, \quad (84)$$

$$\sigma_{ij}^{eq} = \pi_{ij} + 2\pi_{ik} \epsilon_{kj}^e + P_T \delta_{ij} + \rho v_i v_j. \quad (85)$$

The rotation identity,  $0 = dw = \pi_{ij} d\epsilon_{ij}^e = \pi_{ij} (\epsilon_{imk} \epsilon_{mj}^e d\theta_k + \epsilon_{jmk} \epsilon_{im}^e d\theta_k)$ , again shows that the stress is symmetric,

$$\pi_{ik} \epsilon_{kj}^e = (i \leftrightarrow j). \quad (86)$$

Because of  $\pi_{ij} d\epsilon_{ij}^e = \psi_{\alpha i} d\nabla_i a_\alpha$ , the pressure gradient is, see equation (78),

$$\nabla_i P_T = s \nabla_i T + \rho \nabla_i \mu + g_j \nabla_i v_j - \pi_{kj} \nabla_i \epsilon_{kj}^e. \quad (87)$$

In equilibrium and in the rest frame, we have  $\nabla_j \sigma_{ij}^{eq} = 0$ . With equation (78), it delivers the correction terms to  $\nabla_j \pi_{ij} = 0$ , the equilibrium condition for small deformations. The condition for large deformations in isotropic elastic media is

$$\nabla_j \pi_{kj} (\delta_{ik} + 2\varepsilon_{ik}^e) - \pi_{jk} (\nabla_i \varepsilon_{jk}^e - 2\nabla_j \varepsilon_{ik}^e) = 0. \quad (88)$$

Differentiating equation (74),

$$\begin{aligned} 2\partial_t \varepsilon_{ik}^e &= \nabla_k a_\alpha \nabla_i \partial_t a_\alpha + (i \leftrightarrow k), \\ 2\nabla_m \varepsilon_{ik}^e &= \nabla_k a_\alpha \nabla_i \nabla_m a_\alpha + (i \leftrightarrow k), \end{aligned}$$

and inserting these into equation (75), we find,

$$\frac{D}{Dt} \varepsilon_{ik}^e + D_{ik} - \left(\frac{1}{2} \nabla_i a_\alpha \nabla_k \hat{y}_\alpha^D\right) - (i \leftrightarrow k) = 0, \quad (89)$$

where

$$\frac{D}{Dt} \varepsilon_{ik}^e \equiv (\partial_t + v_m \nabla_m) \varepsilon_{ik}^e + \varepsilon_{im}^e \nabla_k v_m + \varepsilon_{km}^e \nabla_i v_m \quad (90)$$

is the – derived and hence only permissible – objective time derivative, including all convective derivatives. This result is especially useful for polymeric hydrodynamics.

In transiently elastic media, ie. polymeric fluids and granular media, however, there is a subtle, constructive choice that one needs to be aware of. Taking

$$\Omega_{ij} \equiv \frac{1}{2} (\nabla_i v_j - \nabla_j v_i), \quad D_{ij} \equiv \frac{1}{2} (\nabla_i v_j + \nabla_j v_i) \quad (91)$$

we may rewrite the last two terms of equation (90) as

$$\varepsilon_{im}^e \nabla_k v_m + \varepsilon_{km}^e \nabla_i v_m = \varepsilon_{im}^e \Omega_{km} + \varepsilon_{km}^e \Omega_{im} + \alpha (\varepsilon_{im}^e D_{km} + \varepsilon_{km}^e D_{im}),$$

with  $\alpha = 1$ . Yet the value of  $\alpha$  is in fact a constructive choice. By changing  $h_{ijkl}$  in equation (115), we may choose any value for  $\alpha$ , especially  $\alpha = 0$ , implying

$$\begin{aligned} \frac{D}{Dt} \varepsilon_{ik}^e &\equiv (\partial_t + v_m \nabla_m) \varepsilon_{ik}^e + \varepsilon_{im}^e \Omega_{km} + \varepsilon_{km}^e \Omega_{mk} \\ &= (\partial_t + v_m \nabla_m) \varepsilon_{ik}^e + \varepsilon_{im}^e \Omega_{km} - \Omega_{km} \varepsilon_{mk}^e. \end{aligned} \quad (92)$$

Due to the counter term  $-h_{ijkl}$  in equation (115), this also simplifies the Cauchy stress as,

$$\sigma_{ij}^{eq} = \pi_{ij} + P_T \delta_{ij} + \rho v_i v_j. \quad (93)$$

More specifically, equations (85, 90) were successfully used for polymeric fluids. Though in exploring large rotational velocities in granular media (not yet done), we believe one should first try the simpler expressions of equations (92, 93).

Note that the same freedom does not exist for elastic media. A change of  $h_{ijkl}$  in equation (69) cannot possibly compensate  $D_{im}$  in equation (90), because of the gradient in front of  $y_k^D$ , see equation (57).

The equilibrium condition (88) is fairly complicated to solve. Also, the dissipative contribution  $(\frac{1}{2} \nabla_i a_\alpha \nabla_k \hat{y}_\alpha^D) + (i \leftrightarrow m)$  in equation (89) cannot easily be given in terms of  $\pi_{ij}$  and  $\varepsilon_{ij}^e$  alone. Fortunately, neither will be needed for polymers or granular media, because there is a stronger, dominating dissipative mechanism, given by the equilibrium condition  $\pi_{ij} = 0$ , or equivalently,  $\psi_{\alpha i} = 0$ .

## 4 Transient Elasticity: Polymeric Fluids

The hydrodynamic method is, as discussed above, a powerful top-down approach to obtain constitutive relations. In this section, we use it to derive the hydrodynamic theory for polymeric fluids, a set of equations that we shall refer to as *polymeric hydrodynamics*.

Our understanding of polymers' basic physics has two points: First that there is elasticity in polymeric fluids, because the polymeric strands may be elastically extended. This leads to reversible energy storage, an elastic strain and an elastic stress. But they vanish over time, irreversibly and dissipatively, as the strands untangle. This is the second point. We therefore call these systems transiently elastic (TE), and we make use of the elasticity theory as derived in the last section, with the modification that the elastic stress  $\pi_{ij}$  is allowed to relax to zero. In solids, the equilibrium condition  $\nabla_j \pi_{ij} = 0$  is satisfied by a uniform stress, which may persist for ever. In polymers, this is replaced by  $\pi_{ij} = 0$ .

The hydrodynamic theory derived with this prejudice reproduces a large number of polymeric phenomena, in shear and elongational flows, including stationary, oscillatory and transient ones, and the Weissenberg effect [MLPH16,MLPH16b]. Therefore, taking TE to be the basic physics of polymeric solutions is clearly justified. In fact, polymeric hydrodynamics is simply viscoelasticity consistently implemented to comply with thermodynamics – including especially the objective time derivative obtained in equation (90). (We call it TE, because people tend to have rather definite ideas of what viscoelasticity is, leading to mostly fruitless controversies over vocabulary.)

The rheological behavior of polymeric fluids is typically characterized as *non-Newtonian*, which stresses the difference to *Newtonian* behavior, but does not provide an explanation for the differences. Non-Newtonian effects include linear viscoelasticity, shear thinning, elongational hardening, rod-climbing (Weissenberg effect), yield stress, viscoplasticity, and thixotropy. There are many textbook models to account for them, including Maxwell, Jeffrey, Oldroyd, Giesekus, Leonov and the intricate KBKZ [BAH77, Lar88, Str97, Tan00]. All are designed bottom-up, with the non-Newtonian effects in mind, by skillfully combining fluid dynamic insights, elasticity theory, and the frame-indifference principle. Starting from momentum conservation, these models provide *constitutive relations* that specify  $\partial_t \sigma_{ij}$  as functions of  $D_{kl}$  and  $\sigma_{kl}$ . No model is encompassing, or generally accepted as authoritative. In recent years, the approach of constitutive relations has been supplemented by mesoscopic models of local and transient substructures, such as the tube model for polymer melts [GLMc03]. Consequently, rather more complicated equations have been employed that, although better at reproducing even the slightest details of an experiment, frequently require separate descriptions for every new experiment. For instance, the description of shear flow [DE86] is different from that of elongational ones [WYT03], and branched polymers require a treatment distinct from linear ones [McL98].

## 4.1 Elastic strain field

We introduce a state variable that quantifies the (coarse-grained) elastic deformation of the medium, which we again call the *elastic strain*  $\varepsilon_{ij}^e$ . We also note that the plastic strain rate is rather more important in polymers. With  $\dot{\varepsilon}_{ij}^{pl} = -\frac{1}{2}(\nabla_i y_j^D + \nabla_j y_i^D)$ , equation (57), and  $y_i^D \sim \nabla_k \pi_{jk}$ , equation (68), the plastic strain rate vanishes for uniform stress,  $\nabla_k \pi_{jk} = 0$ . It is therefore only a small correction for non-uniform situations such as in wave propagation. As we see below, this is not the case in polymeric hydrodynamics, in which the plastic rate may remain non-zero in uniform situations and frequently dominates.

The elastic energy is, by definition, a function of the elastic strain,  $w = w(\varepsilon_{ij}^e)$ , and the elastic stress remains

$$\pi_{ij}(\varepsilon_{kl}^e) \equiv \partial w / \partial \varepsilon_{ij}^e. \quad (94)$$

This is a result of the derivation in the last section, but it is also readily understandable: Differentiating an energy with respect to the geometric quantity that leads to its storage (distance, strain, elastic strain), one always obtains the restoring force or the stress. One example is a car driving up a slippery slope. The rotation of the wheels has a gripping portion  $\theta_g$ , and a slipping portion  $\theta_s$ . The force on the car is  $\partial w_g / \partial \ell$ , with  $w_g$  the gravitational energy and  $\ell$  the distance traveled, and the torque on the wheels is  $\partial w_g / \partial \theta_g$ . Clearly,  $\theta_g$  is, in this case, the ‘‘elastic portion.’’

The equilibrium condition is altered, it now reads

$$\pi_{ij}^{eq}(\varepsilon_{kl}^e) = 0, \quad \text{equivalently} \quad \varepsilon_{ij}^e = 0, \quad (95)$$

implying that dissipative processes forces stress and strain to relax, not only to equalize. Since a stable energy is convex,  $\pi_{ij}$  is a monotonically increasing function of  $\varepsilon_{ij}^e$ , hence a monotonic relaxation of one implies that of the other. Therefore, we rewrite equation (57), for small deformations, as

$$\partial_t \varepsilon_{ij}^e - \dot{\varepsilon}_{ij} = -\dot{\varepsilon}_{ij}^{pl} = -\varepsilon_{ij}^e / \tau.$$

Note three points. First, the plastic strain rate, the part of the total strain rate that neither changes the energy nor the stress, is proportional to the elastic strain. Second, it is always present and influences the dynamics strongly, but it is a dependent quantity, and must not be taken as an independent state state variable. Third, writing it as  $\dot{\varepsilon}_{ij}^{pl}$ , we do not imply that a plastic strain  $\varepsilon_{ij}^{pl}$  exists. It does not. There is only an elastic strain field  $\varepsilon_{ij}^e$ .

For large deformations, and allowing for two different relaxation times,  $\tau_1$  and  $\tau$ , for the trace and the traceless part of the strain, we rewrite the evolution equation as

$$\frac{D}{Dt} \varepsilon_{ij}^e + D_{ij} = -(\varepsilon_{ij}^{*e} / \tau + \varepsilon_{\ell\ell}^e \delta_{ij} / 3\tau_1) = -\dot{\varepsilon}_{ij}^{pl}. \quad (96)$$

see equation (90).

A TE-system accounted for by this relaxation equation, is elastic for time spans  $t \ll \tau$ , much smaller than the relaxation time  $\tau$ , and behaves as a Newtonian fluid for time spans much larger  $t \gg \tau$ . (One may also write  $\omega\tau \gg 1$  or  $\omega\tau \ll 1$ , respectively, for the frequency  $\omega$  of the external perturbation.) Because there is a regime of elastic behavior, we have to copy the objective time derivative  $D/Dt$  from equation (90). Polymeric fluids are soft elastic media, with typically large deformations. The objective derivative is therefore almost always relevant. For  $t \gg \tau$ ,  $\varepsilon_{ij}^e = 0$ , the elastic stress vanishes, and the system naturally reverts to fluid behavior. Under a steady shear or an elongational flow, for  $t \rightarrow \infty$  (ie.  $\omega \rightarrow 0$ ), equation (96) is in its steady state, with a steady state value for  $\varepsilon_{ij}^e$  and  $\pi_{ij}$ . These are the quantities that control the complex rheological behavior of polymeric fluids.

The dissipative contribution  $\sim \hat{y}_i^D$  may typically be neglected, because it is of order  $\nabla_i^2$  smaller than  $\varepsilon_{ij}^e/\tau$ , a small correction for non-uniform stresses.

## 4.2 Polymeric hydrodynamics

Given equation (96), and setting  $\hat{y}_i^D = 0$ , the rest of the equations from section 3 remain unchanged, aside from one contribution in  $R$  that results naturally from the hydrodynamic procedure. We summarize the polymeric hydrodynamic equations.

First, all equilibrium fluxes remain unchanged, especially the equilibrium stress  $\sigma_{ij}^{eq}$ , which is still given by equations (85,86,87),

$$\begin{aligned} \sigma_{ij}^{eq} &= \pi_{ij} + 2\pi_{ik}\varepsilon_{kj}^e + P_T\delta_{ij} + \rho v_i v_j, & \pi_{ik}\varepsilon_{kj}^e &= \pi_{jk}\varepsilon_{ki}^e, \\ \nabla_i P_T &= s\nabla_i T + \rho\nabla_i \mu + g_j\nabla_i v_j - \pi_{kj}\nabla_i \varepsilon_{kj}^e \end{aligned} \quad (97)$$

Note since true equilibrium requires  $\pi_{ij} = 0$ , this expression is not, strictly speaking, an equilibrium one. Still, in the sense discussed around equation (13), we have a generalized equilibrium, in which  $\varepsilon_{ij}^e$  is included as a slowly relaxing state variable. For this generalized equilibrium,  $\varepsilon_{ij}^e$  and  $\pi_{ij}$  can have any values.

Employing the hydrodynamic procedure as above, we see that only  $R$  has one term changed,  $y_i^D \nabla_j \pi_{ij} \rightarrow \pi_{ij} \dot{\varepsilon}_{ij}^{pl}$ . This is because from  $\partial_t w = T\partial_t s \cdots + \pi_{ij}\partial_t \varepsilon_{ij}^e = R \cdots + \pi_{ij}(D_{ij} - \dot{\varepsilon}_{ij}^{pl})$ , we quickly deduce  $R = \cdots + \pi_{ij}\dot{\varepsilon}_{ij}^{pl}$ . The equations are therefore:

$$\begin{aligned} \frac{D}{Dt}\varepsilon_{ij}^e &= -D_{ij} - \dot{\varepsilon}_{ij}^{pl} = \dot{\varepsilon}_{ij} - \dot{\varepsilon}_{ij}^{pl}, \\ \partial_t \rho + \nabla_i(\rho v_i) &= 0, & \partial_t s + \nabla_i(sv_i - f_i^D) &= R/T \geq 0, \\ \partial_t w + \nabla_i Q_i &= 0, & \partial_t g_i + \nabla_k(\sigma_{ik}^{eq} - \sigma_{ik}^D) &= 0, & g_i &= \rho v_i \\ Q_i &= \mu\rho v_i + T(sv_i - f_i^D) + v_k(\sigma_{ik}^{eq} - \sigma_{ik}^D), & f_i^D &= \kappa\nabla_i T, \\ R &= f_i^D \nabla_i T + \sigma_{ij}^D D_{ij} + \pi_{ij}\dot{\varepsilon}_{ij}^{pl}. & \sigma_{ij}^D &= \eta D_{ij}^* + \zeta \delta_{ij} D_{\ell\ell}, \end{aligned} \quad (98)$$

with the only new terms (again assuming isotropy and absence of off-diagonal Onsager coefficients) given as

$$\dot{\varepsilon}_{\ell\ell}^{pl} = \beta_1 \pi_{\ell\ell}, \quad \dot{\varepsilon}_{ij}^{pl*} = \beta \pi_{ij}^*. \quad (100)$$

Given any appropriate energy expression  $w$  – along with its conjugate variables,  $\mu \equiv \partial w / \partial \rho$ ,  $T \equiv \partial w / \partial s$ , and  $\pi_{ij} \equiv \partial w / \partial \varepsilon_{ij}^e$  – and the Onsager coefficients  $\kappa, \eta, \zeta, \beta, \beta_1$  as functions of the state variables, these equations are closed and may be solved.

However, for calculational convenience, one would prefer the form of equation (96): For stationary cases,  $\partial_t \varepsilon_{ij}^e = 0$ , equation (96) yields a general and typically analytic relation between  $\varepsilon_{ij}^e$  and  $\nabla_k v_m$ , independent of the energy, which is an input that one would have to find by trial and error. Therefore, we assume the energy is such that  $\pi_{ij}^* / |\pi_{ij}^*| = \varepsilon_{ij}^{*e} / |\varepsilon_{ij}^{*e}|$  holds (where  $|B_{ij}^*| \equiv \sqrt{B_{ij}^* B_{ij}^*}$  for any matrix  $B_{ij}^*$ ), such that we may write

$$\begin{aligned} \dot{\varepsilon}_{\ell\ell}^{pl} &= \beta_1 \pi_{\ell\ell} = \beta_1 (\pi_{\ell\ell} / \varepsilon_{\ell\ell}^e) \varepsilon_{\ell\ell}^e \stackrel{!}{=} \tau_1 \varepsilon_{\ell\ell}^e, \\ \dot{\varepsilon}_{ij}^{pl*} &= \beta \pi_{ij}^* = \beta (|\pi_{ij}^*| / |\varepsilon_{ij}^{*e}|) \varepsilon_{ij}^{*e} \stackrel{!}{=} \tau \varepsilon_{ij}^{*e}, \end{aligned} \quad (101)$$

assuming the dependence of  $\beta, \beta_1$  on  $\varepsilon_{ij}^e$  is such that  $\tau, \tau_1 > 0$  are constant.

### 4.3 Applications

Application of these equations to experimentally relevant circumstances is well rendered in [MLPH16, MLPH16b], we shall not repeat it here. The calculation is executed assuming the polymeric fluid is incompressible and isothermal:  $T, \rho = \text{const}$ , with  $D_{\ell\ell} = 0$  and  $(1 - 2\varepsilon_1)(1 - 2\varepsilon_2)(1 - 2\varepsilon_3) = 1$ , where  $\varepsilon_i$  is the eigenvalue of  $\varepsilon_{ij}^e$ . (The last relation reduces to  $\varepsilon_{\ell\ell}^e = 0$  for linear elasticity.) We found these equations well capable of accounting for characteristic polymeric effects, including shear thinning, normal stress differences, and the Weissenberg effect. Hereby, the convective derivatives contained in  $\frac{D}{Dt} \varepsilon_{ij}^e$  are especially important, even for the first paper, treating small to moderate deformations.

Starting from a general fourth-order expansion of the elastic energy, the first paper [MLPH16] considers shear and elongational flow, both stationary and oscillating, also their relaxation and onset, including the elongation rate dependence of the Trouton viscosity. And it considers surface effects including the Weissenberg effect and the surface curvature for the flow down a slightly tilted channel. The second paper [MLPH16b] starts from a postulated energy expression that we take to be valid to any order of the elastic strain. We employ it to consider viscosity overshoot near the onset of shear flow, the onset of elongational flows in situations for which there is no stationary solution, as well as shear thinning and normal stress differences for a large range of shear rates. In addition, we find that the presented equations satisfy empirical relations including the Cox-Merz rule, the Yamamoto relation and Gleissle's mirror relations – in several cases quantitatively.

## 5 Granular Media, GSH

### 5.1 Two-stage irreversibility

Two notions govern the basic physics of granular media: *two-stage irreversibility* and *variable transient elasticity*. The first is related to the system's three spatial scales: (a) macroscopic, (b) granular or mesoscopic, and (c) microscopic. Dividing all degrees of freedom into these three categories, we account for the energy of (b) and (c) with two temperatures, the true temperature  $T$  and the granular temperature  $T_g$ , each with an associated entropy. The first is the temperature one measures with a thermometer, and  $T_g$  the one quantifying the average energy of a jiggling grain. Both (a) and (b) consist of kinetic and elastic contributions, with the difference that the former is coarse-grained over many particles, the latter the fluctuating leftover. The conserved energy  $W$  is divided into  $W^M$  and  $H$ , as in equation (5), and in addition, into granular heat  $H_g$ ,

$$\frac{d}{dt}W = \frac{d}{dt}(W^M + H_g + H) = 0. \quad (102)$$

The second law of thermodynamics is modified to: One can convert  $W^M$  into  $H_g$  or  $H$ , and convert  $H_g$  into  $H$ , but never backwards, implying

$$\frac{d}{dt}W^M \leq 0, \quad \frac{d}{dt}H = T \frac{d}{dt}S \geq 0, \quad \frac{d}{dt}H_g = T_g \frac{d}{dt}S_g, \quad (103)$$

where the third quantity does not have a definite sign. However, since  $W^M$  diminishes continually in a closed system, we will eventually arrive at the minimum of  $W^M$ . Because then  $\frac{d}{dt}H_g = T_g \frac{d}{dt}S_g < 0$  is negative, we will soon arrive at the minimum of  $H_g$ , and simultaneously the maximum of  $H$ , or  $S = \max$ , implying equilibrium, see figure 1.

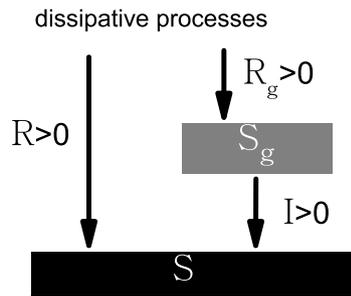


Figure 1: Two-stage irreversibility: Macroscopic energy either decays directly into heat, or in a two-step process, first decays into granular heat (jiggling of the grains), then into true heat (random atomic motion).

## 5.2 Variable transient elasticity

*Variable transient elasticity* is the second notion. First of all, grains are transiently elastic – they may be elastically deformed, but their elastic stress are transient when the grains jiggle. Everything we learned in the context of polymers apply here as well: The elastic strain  $\varepsilon_{ij}^e$  is a state variable, with the evolution equation (96), and with the equilibrium stress as given in equation (97).

The qualifier “variable” addresses the following fact: The stable free surface of a granular system at rest may be tilted, and this will, unperturbed, stay forever. When perturbed, when the vessel is being tapped and the grains jiggle, ie. when  $T_g \neq 0$ , the tilted surface will move toward the horizontal, same as polymeric fluids. The stronger the grains jiggle, the faster this process is. This is indicative of a system that is elastic for  $T_g = 0$ , transiently elastic for  $T_g \neq 0$ , with a stress relaxation rate that grows with  $T_g$ . We take it to be  $\propto T_g$ , because granular media are then, as observed, rate-independent in certain regimes. Note that this aspect only changes a Onsager coefficient, the stress relaxation rate. It does not at all change the structure of the hydrodynamics. Therefore, except for the additional state variable of the granular temperature  $T_g$ , the hydrodynamic theory for polymeric fluids and granular media are very similar.

We call the hydrodynamic theory derived starting from these two notions: GSH, for “granular solid hydrodynamics.” We take a detailed look at GSH below.

## 5.3 The two temperatures

We separate granular, mesoscopic from microscopic degrees, instead of lumping them into one group and one temperature, not only because of the difference in length scales. Equally important is the fact that  $T_g$  is a quantity, on which the dynamics critically depends: The elastic stress relaxes when the grains jiggle, when  $T_g \neq 0$ , while it is (within limits) independent of  $T$ . Lumping both into one temperature would obscure this crucial difference.

On the other hand, introducing a third temperature would be frivolous. Extending the set of state variables quickly complicates the hydrodynamic theory, it must never be done without compelling reasons. For instance, it is not useful to introduce a temperature for the degrees characterizing the surface roughness of grains, though the length scale is distinctly smaller, because no qualitative aspect of macroscopic granular dynamics depends on an associated temperature. Besides, surface roughness is given, and does not change with the state. It may influence the energy and the transport coefficients, but it is not a state variable. Similarly, it is futile to separate  $T_g$  into two temperatures, one for the kinetic and the other the elastic part. In every collision, the two energy contributions are quickly converted into each other. If it were not for dissipation into true heat, their sum would be constant. Separating both does not serve any conceivable purpose.

A division into three scales with two temperatures (as was done above for GSH) works, when the scales are well separated – though this is a problem of accuracy, not viability. Scale separation [between (a) and (b)] is well satisfied in large-scaled, engineering-type experiments, less so in small-scaled ones. Using glass or steel beads (typically larger) aggravates the problem. Nevertheless, when the system is too small for spatial averaging, one may still average over the time and over runs, to rid the data of fluctuations.

Both  $T_g$  and  $T$  are genuine temperatures, as each characterizes the energy of a group of degrees of freedom. Taking the energy density as a function of the two entropy densities,  $w = w(s, s_g)$ , we write

$$dw = Tds + T_g ds_g = T ds_{tot} + (T_g - T) ds_g, \quad (104)$$

$$T \equiv \partial w / \partial s, \quad T_g \equiv \partial w / \partial s_g, \quad s_{tot} = s + s_g. \quad (105)$$

While  $Tds$  and  $T_g ds_g$  in the first expression of equation (104) are equivalent, the two terms in the second one are not: One may identify  $T ds_{tot}$ , the equilibrium energy change for changes of the total entropy, as equivalent to  $Tds$  in Newtonian fluids. It possesses a contribution to the energy that is minimal for  $\nabla_i T = 0$ .  $T$  does not vanish in equilibrium. The second one,  $(T_g - T) ds_g$ , is the non-equilibrium energy for  $T_g \neq T$ , which reflects the non-optimal energy distribution between the granular and microscopic degrees of freedom. Its energy contribution is  $w \propto (T_g - T)^2$  if expanded, and has a minimum at  $T_g = T$ .  $T_g - T$  relaxes until it vanishes.

Yet since  $s \gg s_g$ , and any granular jiggling at all occurs at  $T_g \gg T$ , we have  $T_g - T \approx T_g$ ,  $s_{tot} \approx s$ , and the two expressions of equation (104) are very similar. So there is in fact a practical difference between both temperatures. In equilibrium,  $T$  is uniform, but the granular temperature  $T_g \approx T_g - T$  vanishes.

The pair of entropy densities obey similar equations,

$$\partial_t s + \nabla_i (s v_i - \kappa \nabla_i T) = R/T \geq 0, \quad (106)$$

$$\partial_t s_g + \nabla_i (s_g v_i - \kappa_g \nabla_i T_g) = (R_g - \gamma T_g^2)/T_g, \quad (107)$$

each with a convective, diffusive, and source term, where

$$\begin{aligned} R &= f_i^D \nabla_i T + \sigma_{ij}^D D_{ij} + \pi_{ij} \varepsilon_{ij}^{pl} + \gamma T_g^2 \\ &= \kappa (\nabla_i T)^2 + \eta D_{ij}^* D_{ij}^* + \zeta D_{\ell\ell}^2 \\ &\quad + \beta (\pi_{ij}^*)^2 + \beta_1 \pi_{\ell\ell}^2 + \gamma T_g^2. \end{aligned} \quad (108)$$

$$\begin{aligned} R_g &= f_i^{gD} \nabla_i T + \sigma_{ij}^{gD} D_{ij} \\ &= \kappa_g (\nabla_i T_g)^2 + \eta_g D_{ij}^* D_{ij}^* + \zeta_g D_{\ell\ell}^2. \end{aligned} \quad (109)$$

Compare  $R$  to the polymeric  $R$ , equations (99), the only difference is the term  $\gamma T_g^2$ . It is equal to the counter term in equation (107), which accounts for the  $T_g$ -relaxation, converting granular heat to true heat. They are the same because energy has to be conserved:  $\partial_t w = T \partial_t s + T_g \partial_t s_g + \dots = \gamma T_g^2 - \gamma T_g^2 = 0$ . Granular entropy

production  $R_g$  – with  $f_i^{gD}, \sigma_{ij}^{gD}$  denoting the granular dissipative contributions – is again similar. There is no granular plastic contribution, no term  $\sim \dot{\varepsilon}_{ij}^{g,pl}$  in  $R_g$ , because otherwise the elastic stress relaxes even for  $T_g = 0$ . Also, because  $\eta \ll \eta_g, \zeta \ll \zeta_g$  in dry sand, they are usually neglected in applications. (These are all constitutive choices, see [JL09] for more detailed reasoning.)

## 5.4 Granular equilibrium conditions

The state variables of any granular media are the density  $\rho$ , the momentum density  $\rho v_i$ , the two entropy densities  $s, s_g$ , and the elastic strain  $\varepsilon_{ij}^e$ . Denoting the energy density in the rest frame ( $v_i = 0$ ) as  $w_0 = w_0(s, s_g, \rho, \varepsilon_{ij}^e)$ , we have

$$dw_0 = T ds + T_g ds_g + \mu d\rho + \pi_{ij} d\varepsilon_{ij}^e. \quad (110)$$

A further useful conjugate variable is the pressure  $P_T$ , cf. equation (24), where

$$P_T \equiv -w_0 + sT + s_g T_g + \mu_0 \rho = - \left. \frac{\partial(wV)}{\partial V} \right|_M = - \frac{\partial(w/\rho)}{\partial(1/\rho)}. \quad (111)$$

Equilibrium conditions are obtained by requiring  $\int w d^3r = \text{minimum}$ , for given entropy  $\int s d^3r$  and mass  $\int \rho d^3r$ , with  $T_g$  allowed to relax (cf. section 5.3). We first obtain

$$\nabla_i T = 0, \quad T_g = 0. \quad (112)$$

Usually  $T_g$  vanishes quickly. After this has happened,  $d\rho$  and  $du_{\ell\ell} = -d\rho/\rho$  no longer vary independently. They therefore share a solid-like equilibrium condition,

$$\nabla_i (\pi_{ij} + P_T \delta_{ij}) = -\rho G \hat{z}_i, \quad (113)$$

rather than separately,  $\nabla_i \pi_{ij} = 0$  and  $\nabla_i P_T = \rho \nabla_i \mu_0 = -\rho G \hat{z}_i$ , cf. Section 2.1.1. If  $T_g$  is kept finite, say by continual tapping, the system may further decrease its energy (or increase its entropy) by independently varying  $\rho$  and  $\varepsilon_{ij}^e$ . It then arrives at two fluid-like equilibrium conditions,

$$\pi_{ij} = 0, \quad \nabla_i P_T = \rho \nabla_i \mu_0 = -\rho G \hat{z}_i, \quad (114)$$

the first of which requires the shear stress to vanish, and a free surface to be horizontal.

## 5.5 The evolution equations

Aside from the addition of the evolution equation for  $s_g$ , equation (107), GSH has the same structure as polymeric hydrodynamics, see the discussion in section 5.2. However, some off-diagonal Onsager coefficients are needed, and the strain relaxation rate are  $\propto T_g$ . These are what we concentrate on in this section.

Denoting, for granular media,  $\frac{D}{Dt}\varepsilon_{ij}^e = \dot{\varepsilon}_{ij} - \dot{\varepsilon}_{ij}^{pl}$  [cf. equation (98)], we revisit the Onsager matrix, as given in equation (69). Starting from equation (108), the entropy production  $R$ , we write

$$\begin{pmatrix} f_i^D \\ \dot{\varepsilon}_{ij}^{pl} \\ \sigma_{ij}^D \end{pmatrix} = \begin{pmatrix} \kappa_{ik} & b_{ikl} & c_{ikl} \\ b_{kli} & \beta_{ijkl} & h_{ijkl} \\ -c_{kli} & -h_{kli} & \eta_{ijkl} \end{pmatrix} \cdot \begin{pmatrix} \nabla_k T \\ \pi_{kl} \\ D_{kl} \end{pmatrix}, \quad (115)$$

with the constitutive choices (that was referred to as the ‘‘minimalist GSH,’’ [JL17]):

$$\kappa_{ik} = \kappa\delta_{ik}, \quad \eta_{ijkl} = \eta\delta_{ik}\delta_{jl} + (\zeta - \eta/3)\delta_{ij}\delta_{kl}, \quad (116)$$

$$b_{ikl}, c_{ikl} = 0, \quad \beta_{ijkl} = \beta\delta_{ik}\delta_{jl} + (\beta_1 - \beta/3)\delta_{ij}\delta_{kl}, \quad (117)$$

$$h_{ijkl} = \alpha_1\delta_{ij}\varepsilon_{kl}^*/3. \quad (118)$$

(The two forces  $\pi_{kl}$  and  $D_{kl}$  have different time reversal property, hence  $h_{ijkl} = -h_{klij}$ , and there is no contribution to  $R$ , cf. [DeM84].) We assume, in addition, relaxation rates linear in  $T_g$ ,

$$\dot{\varepsilon}_{\ell\ell}^{pl} = \beta_1\pi_{\ell\ell} = \tau_1\varepsilon_{\ell\ell}^e = \lambda T_g\varepsilon_{\ell\ell}^e, \quad (119)$$

$$\dot{\varepsilon}_{ij}^{*pl} = \beta\pi_{ij}^* = \tau\varepsilon_{ij}^e = \lambda_1 T_g\varepsilon_{ij}^e, \quad (120)$$

see equations (101) and the discussion in section 5.2 on *variable transient elasticity*. Together, we have, denoting  $\Delta \equiv \varepsilon_{\ell\ell}^e$  and taking the expressions from equations (92, 93)

$$\frac{D}{Dt}\varepsilon_{ij}^{*e} + D_{ij}^* = -\lambda T_g\varepsilon_{ij}^{*e}, \quad (121)$$

$$\frac{D}{Dt}\Delta + D_{\ell\ell} + \alpha_1\varepsilon_{ij}^{*e}D_{ij}^* = -\lambda_1 T_g\Delta. \quad (122)$$

$$\begin{aligned} \sigma_{ij} &= \pi_{ij} + P_T\delta_{ij} + \rho v_i v_j + \alpha_1\varepsilon_{ij}^{*e}\pi_{\ell\ell}/3 \\ &\quad - (\eta + \eta_g)D_{ij}^* - (\zeta + \zeta_g)D_{\ell\ell}\delta_{ij}. \end{aligned} \quad (123)$$

For small deformations, these equations reduce to

$$\partial_t\varepsilon_{ij}^{*e} + D_{ij}^* = -\lambda T_g\varepsilon_{ij}^{*e}, \quad (124)$$

$$\partial_t\Delta + D_{\ell\ell} + \alpha_1\varepsilon_{ij}^{*e}D_{ij}^* = -\lambda_1 T_g\Delta. \quad (125)$$

$$\sigma_{ij} = \pi_{ij} + P_T\delta_{ij} + \rho v_i v_j - (\eta + \eta_g)D_{ij}^* - (\zeta + \zeta_g)D_{\ell\ell}\delta_{ij}. \quad (126)$$

We note that in contrast to polymeric fluids, most granular phenomena do not have large elastic strain fields. Exceptions are cases with a rotating sand pile, such as in a tumbler. The term preceded by  $\alpha_1$  in equation (125) is formerly of the same order as the eliminated ones. but it accounts for dilatancy and contractancy, is qualitatively important and must not be neglected. Without it, the steady state solution under shear flow  $D_{\ell\ell} = 0$  is  $\Delta = 0$ , and there cannot be any density or pressure change in the approach to the critical state. Besides,  $\alpha_1$  turns out to be fairly large, of the order of 30. The counter term in the stress, also preceded by  $\alpha_1$ , changes the zeroth order term  $\pi_{ij}$  slightly, it is not as qualitatively important.

Finally, we note that the elastic strain cannot be obtained by coarse-graining its mesoscopic counterpart,  $\varepsilon_{ij}^e \neq \langle \varepsilon_{ij}^{mes} \rangle$ . The reason is, both the energy and stress are given by averaging:  $w = \langle w^{mes} \rangle$ ,  $\pi_{ij} = \langle \pi_{ij}^{mes} \rangle$ . Since

$$dw = \langle dw^{mes} \rangle = \langle \pi_{ij}^{mes} d\varepsilon_{ij}^{mes} \rangle = \pi_{ij} d\varepsilon_{ij}^e = \langle \pi_{ij}^{mes} \rangle d\varepsilon_{ij}^e$$

and since  $w$  is always a nonlinear expression,

$$dw = \langle \pi_{ij}^{mes} d\varepsilon_{ij}^{mes} \rangle \neq \langle \pi_{ij}^{mes} \rangle d\langle \varepsilon_{ij}^{mes} \rangle,$$

we have  $d\varepsilon_{ij}^e \neq d\langle \varepsilon_{ij}^{mes} \rangle$ .

## 5.6 Applications

The minimalist GSH (cf. [JL17]) consists of:

- the continuity equation:  $\partial_t \rho + \nabla_i (\rho v_i) = 0$ ,
- momentum conservation:  $\partial_t (\rho v_i) + \nabla_j \sigma_{ij} = -\rho G \hat{z}_i$ , with  $\sigma_{ij}$  given in equation (126),
- balance equation for  $s_g$ , equation (107),
- the evolution equation for the elastic strain, equations (124,125).

The balance equation for the entropy  $s$  is not included, as effects such as thermal expansion or diffusion of true temperature are not at present of primary concern.

A more complete version with more off-diagonal Onsager coefficients and a more complicated energy  $w$  has also been employed to yield a more quantitative study of granular phenomena, see [JL15]. The basic structure of both versions are identical.

Given an expression for the energy  $w$  and the transport coefficients [not discussed here, see [JL15, JL17]], these equations are closed and enable calculation of many granular phenomena, grouped as followed:

- $T_g = 0$ : static stress distribution, velocity and damping of elastic waves, clogging;
- moderate  $T_g$ , elasto-plastic motion: loading and unloading, approach to the critical state, angle of stability and repose;
- elevated  $T_g$ , rapid dense flow: the  $\mu$ -rheology, Bagnold scaling and shear band;
- compaction (both reversible and irreversible branch, memory effect), wide and narrow shear band, shear jamming, visco-elastic behavior and nonlocal fluidization.

Only in the second of the four regimes are grains rate-independent (see next section).

It is useful to realize that two qualitatively different types of yield surface are encoded in GSH. The first is in the convexity transition of the energy  $w$ . Usually, the elastic

energy is convex, and the 6x6 matrix of its second derivatives (with respect to the 6-tuple vector of elastic strain) has only positive eigenvalues. Elastic solutions are then stable. When, however, one of the eigenvalues turns negative, there are no stable elastic solutions. In GSH, the elastic energy needs to have yield surfaces at which this convexity transition takes place, to account for the fact that elastic solutions may be maintained only below a critical density, or that a steep slope will crumble.

The second is given by the stationary solution of equations (124,125), or the so-called critical state, for which the elastic strain (hence also the stress) is finite at non-zero total strain rate,  $D_{ij} \neq 0$ . Although this is also referred to as yield surface, there is no instability involved.

With all these phenomena ordered, related, explained and accounted for, though as yet frequently qualitatively, we believe that GSH may be taken as a unifying framework, providing the appropriate macroscopic vocabulary and mindset that help one coming to terms with the breadth of granular physics.

## 5.7 Rate-independence

Granular phenomena are frequently rate-independent: The observed stress remains the same, however fast the experiment is executed. This has led to the wide-spread believe that rate-independence (RI) is a basic granular feature of granular media. Constructing constitutive relations, one therefore needs to start from it. Such approaches are efficient, but they prevent an understanding of why RI holds, and where it does (ie., to what range RI is limited). Moreover, such approaches preclude the construction of constitutive relations that remain valid outside the RI regime, in phenomena such as chute flows.

To get a better understanding of RI, it is useful to first recognize the difference in RI between that of elasticity and hypoplasticity. They are different because the first is a static property, and the second a dynamic one. In elasticity, the strain  $\varepsilon_{ij}$  is a state variable,<sup>7</sup> and both the elastic energy  $w$  and the stress  $\pi_{ij} \equiv -\partial w / \partial \varepsilon_{ij}$  are unique functions of it (if, as discussed above, mass defects are neglected). This is the reason the stress remains the same for a given strain,  $\pi_{ij} = \pi_{ij}(\varepsilon_{kl})$ , however fast  $\varepsilon_{ij}$  is built up, ie., whatever the shear rate  $D_{ij}$  is. Dynamics does not enter this consideration. (Of course, since the Cauchy stress always has a viscous contribution,  $\sigma_{ij} = \pi_{ij} - \eta D_{ij}$ , it is not rate-independent if this term becomes significant at larger shear rates.)

Hypoplasticity is a constitutive model, defined as

$$\partial_t \sigma_{ij} = H_{ijkl} \dot{\varepsilon}_{kl} + \Lambda_{ij} \sqrt{\dot{\varepsilon}_{kl} \dot{\varepsilon}_{kl}}, \quad (127)$$

with the two tensors  $H_{ijkl}, \Lambda_{ij}$  functions of  $\sigma_{ij}$  and  $\rho$ . (For the simplicity of display, here and below, we always set  $\dot{\varepsilon}_{kk} = -D_{kk} \equiv 0$ .) As a result, the stress  $\sigma_{ij}$  remains

<sup>7</sup>Neglecting  $y_i^D$  in equation (54), we have  $\varepsilon_{ij}^e = \varepsilon_{ij}$ .

the same for any given path  $\int \dot{\epsilon}_{kl}(t)dt$ , whatever the rate  $\dot{\epsilon}_{kl}(t)$  is.<sup>8</sup> So RI is a result of the postulated dynamics, of the structure of the evolution equation for  $\sigma_{ij}$ .

Mathematically speaking, equation (127) is invariant under a transformation rescaling the time,

$$t \rightarrow t/A, \quad (128)$$

because all three terms of equation (127) get multiplied by the same constant factor  $A$ . Therefore, the solution for the stress also needs to be scale invariant and cannot depend on  $\dot{\epsilon}_{kl}$  alone, but may depend on  $\int \dot{\epsilon}_{kl}(t)dt$ .

More generally, the dynamics of neither elasticity nor hypoplasticity is scale invariant. Consider for instance momentum conservation,  $\partial_t(\rho v_i) + \nabla_j(\pi_{ij} + P_T + \rho v_i v_j) = 0$ , in which  $\partial_t(\rho v_i)$  and  $\nabla_j(\rho v_i v_j)$  scale as  $A^2$ , and  $\nabla_j \pi_{ij}$  as  $A^0$ . Therefore, wave phenomena are never rate-independent. And when one speaks of RI in the context of resonance column experiments, one needs to carefully spell out what is implied.

Next, we discuss when and why GSH is rate-independent, or more precisely, scale invariant. GSH reduces to hypoplasticity, equation (127), for constant (or slowly varying) slow shear rates [JL16b]. GSH is trivially rate-independent then, because hypoplasticity is. If the shear rate oscillates, such as in cyclic loading, GSH is more complicated than hypoplasticity, as more terms appear. But as we shall see, it is then less plastic but remains rate-independent. Moreover, the gradient term accounting for shear bands is also scale-invariant and maintains RI.

A minimalist version of the GSH equations is collected here, to facilitate the discussion. (The complete version renders the description of granular phenomena more realistic. Yet it suffices to consider the former, since all conclusions on RI drawn below remain valid for the latter.) The state variables of any granular system are the density  $\rho$ , the momentum density  $\rho v_i$ , the granular entropy density  $s_g$  and the elastic strain  $\epsilon_{ij}^e$ . (The entropy  $s$  does not play a role and is neglected here.) Their close set of equations is

$$\partial_t \rho + \nabla_i(\rho v_i) = 0, \quad (129)$$

$$\partial_t(\rho v_i) + \nabla_i \sigma_{ij} = -\rho G \hat{z}_i, \quad (130)$$

$$\sigma_{ij} = \pi_{ij} + P_T \delta_{ij} - \eta_1 T_g D_{ij} + \rho v_i v_j, \quad (131)$$

$$\pi_{ij}(\epsilon_{ij}^e) \equiv \partial w / \partial \epsilon_{ij}^e, \quad P_T \equiv -\partial(w/\rho) / \partial(1/\rho), \quad (132)$$

$$\partial_t T_g = -R_T [T_g (1 - \xi_T^2 \nabla_i^2) T_g - f^2 D_{ij} D_{ij}], \quad (133)$$

$$\partial_t \epsilon_{ij}^{*e} + D_{ij} = -\lambda T_g \epsilon_{ij}^{*e}, \quad \partial_t \epsilon_{kk}^e + \alpha_1 \epsilon_{ij}^{*e} D_{ij} = -\lambda_1 T_g \epsilon_{kk}^e, \quad (134)$$

where  $\epsilon_{ij}^{*e}$  is the traceless part of the elastic strain, and  $\dot{\epsilon}_{kk} = 0$  is again assumed. Equation (129) is the continuity equation, equation (130) the momentum balance, with the Cauchy stress  $\sigma_{ij}$  explicitly given if the free energy  $w = w(\rho, T_g, \epsilon_{ij}^e)$  is known. (As the specific form of the free energy we employ is irrelevant to the discussion about RI, it is not reproduced here.)

<sup>8</sup>We note equation (127) is not of the most general form, because terms such as  $\sqrt{\dot{\epsilon}_{ik} \dot{\epsilon}_{kj}}$  or  $\sqrt[3]{\dot{\epsilon}_{ik} \dot{\epsilon}_{kl} \dot{\epsilon}_{lj}}$  would retain RI.

Equation (133) is the balance of the granular entropy, and equations (134) are the evolution equations for the elastic strain. The seven scalar coefficients:  $\lambda, \lambda_1, \alpha_1, R_T, \xi_T, f, \eta_1$  are functions of the density alone, providing the only leeway GSH has for fitting experimental data.

For  $\dot{\epsilon}_{kk} = 0$ , equation (129) is redundant. Excluding acceleration,  $\partial_t(\rho v_i) = 0$ , and wave propagation leaves the dynamics given by equations (133,134) alone, which are, as we shall see, invariant under the scale transformation of equation (128).

Note first that in the stationary limit,  $\partial_t T_g = 0$ , and for uniform samples,  $\xi_T^2 \nabla_i^2 T_g = 0$ , we have

$$T_g = f \sqrt{\dot{\epsilon}_{ij} \dot{\epsilon}_{ij}}. \quad (135)$$

Inserted into equations (134), this expression yields two equations of hypoplastic structure, implying RI of the elastic strain.

Next, we go an important step further. Equation (135) implies that  $T_g$  scales with  $A$ , same as  $\dot{\epsilon}_{ij}$ . Then, clearly, under the scale transformation of equation (128), every single term of equation (133) scales with  $A^2$ , and every one of equation (134) with  $A$ . This leaves both equations scale invariant. As above, we again conclude that no solution for the elastic strain  $\epsilon_{ij}^e$  may depend on  $\dot{\epsilon}_{ij}$  or  $T_g$  alone, though  $\int \dot{\epsilon}_{kl}(t) dt$  or  $\int T_g dt$  are fine.

In the Cauchy stress of equation (131), both the viscous contribution and the pressure  $P_T \sim T_g^2$  are not scale invariant. However, they are important only at higher shear rates, such as in chute flows or the  $\mu(I)$ -rheology. At lower shear rates, they may be neglected, rendering the Cauchy stress a function of the elastic strain alone,  $\sigma_{ij} = \pi_{ij}(\epsilon_{kl}^e)$ , scale invariant and rate-independent.

In cyclic loading, both  $T_g$  and  $\dot{\epsilon}_{ij} = -D_{ij}$  vary with time. If the amplitude is too small for  $T_g$  to reach the stationary limit, the term  $\partial_t T_g = 0$  becomes important, typically rendering  $T_g$  smaller than given by equation (135). This reduces the plasticity that is accounted for by the terms  $\sim T_g$  in equations (134), and is how GSH accounts for cyclic loading rate-independently. Similarly, the term  $T_g \xi_T^2 \nabla_i^2 T_g$  in equation (133) accounts for shear bands with no rate-dependence.

## References

- [Ko98] P. Kostädt and M. Liu, *Three ignored Densities, Frame-independent Thermodynamics, and Broken Galilean Symmetry*, Phys. Rev. **E 58**, 5535, (1998).
- [LL87] L. D. Landau and E. M. Lifshitz. *Fluid Mechanics*. Butterworth-Heinemann, 1987.
- [Kha65] I. M. Khalatnikov. *Introduction to the Theory of Superfluidity*. Benjamin, New York, 1965.

- [DePG93] P.G. de Gennes and J. Prost. *The Physics of Liquid Crystals*. Clarendon Press, Oxford, 1993.
- [DeM84] S. R. de Groot and P. Masur, *Non-Equilibrium Thermodynamics*, (Dover, New York 1984).
- [MPP72] P.C. Martin, O. Parodi, and P.S. Pershan, *Unified Hydrodynamic Theory for Crystals, Liquid Crystals, and Normal Fluids*, Phys. Rev. A **6**, 2401 (1972).
- [Lub72] T.C. Lubensky, *Hydrodynamics of Cholesteric Liquid Crystals*, Phys. Rev. A **6**, 452 (1972).
- [Liu79] M. Liu, *Hydrodynamic Theory near the Nematic Smectic-A Transition*, Phys. Rev. A **19**, 2090 (1979);
- [Liu94] M. Liu, *Hydrodynamic theory of biaxial nematics*, Phys. Rev. A **24**, 2720 (1981).
- [liu94b] M. Liu, *Maxwell equations in nematic liquid crystals*, Phys. Rev. E **50**, 2925, (1994).
- [PB96] H. Pleiner and H.R. Brand, in *Pattern Formation in Liquid Crystals*, edited by A. Buka and L. Kramer (Springer, New York, 1996).
- [Gra74] R. Graham, *Hydrodynamics of  $^3\text{He}$  in Anisotropic A Phase*, Phys. Rev. Lett. **33**, 1431 (1974).
- [GP75] R. Graham and H. Pleiner, *Spin Hydrodynamics of  $^3\text{He}$  in the Anisotropic A Phase*, Phys. Rev. Lett. **34**, 792 (1975).
- [Liu75] M. Liu, *Hydrodynamics of  $^3\text{He}$  near the A-Transition*, Phys. Rev. Lett. **35**, 1577 (1975).
- [LC78] M. Liu and M.C. Cross, *Broken Spin-Orbit Symmetry in Superfluid  $^3\text{He}$  and the B-Phase Dynamics*, Phys. Rev. Lett. **41**, 250 (1978).
- [LC79] M. Liu and M.C. Cross, *Gauge Wheel of Superfluid  $^3\text{He}$* , Phys. Rev. Lett. **43**, 296 (1979).
- [Liu79] M. Liu, *Relative Broken Symmetry and the Dynamics of the  $A_1$ -Phase*, Phys. Rev. Lett. **43**, 1740 (1979).
- [Liu98] M. Liu, *Rotating Superconductors and the Frame-independent London Equations*, Phys. Rev. Lett. **81**, 3223, (1998).
- [JL01] Jiang Y.M. and M. Liu, *Rotating Superconductors and the London Moment: Thermodynamics versus Microscopics*, Phys. Rev. B **6**, 184506, (2001).
- [Liu02] M. Liu, *Superconducting Hydrodynamics and the Higgs Analogy*, J. Low Temp. Phys. **126**, 911, (2002)
- [HL93] K. Henjes and M. Liu, *Hydrodynamics of Polarizable Liquids*, Ann. Phys. **223**, 243 (1993).

- [Liu93] M. Liu, *Hydrodynamic Theory of Electromagnetic Fields in Continuous Media*, Phys. Rev. Lett. **70**, 3580 (1993).
- [Liu95] *Mario Liu replies*, Phys. Rev. Lett. **74**, 1884, (1995).
- [JL96] Y.M. Jiang and M. Liu, *Dynamics of Dispersive and Nonlinear Media*, Phys. Rev. Lett. **77**, 1043, (1996).
- [LS09] M. Liu and K. Stierstadt, Thermodynamics, Electrodynamics and Ferrofluid Dynamics, in: Colloidal Magnetic Fluids, edited by S. Odenbach, Lect. Notes Phys. Vol. 763 (Berlin, Heidelberg, 2009), pp. 83–156.
- [SL15] Klaus Stierstadt and Mario Liu, Maxwell's stress tensor and the forces in magnetic liquids. *Z. Angew. Math. Mech.* 95, No. 1, 4 – 37 (2015)
- [Liu95] M. Liu, *Fluidynamics of Colloidal Magnetic and Electric Liquid*, Phys. Rev. Lett. **74**, 4535 (1995).
- [Liu98] M. Liu, *Off-Equilibrium, Static Fields in Dielectric Ferrofluids*, Phys. Rev. Lett. **80**, 2937, (1998).
- [Liu99] M. Liu, *Electromagnetic Fields in Ferrofluids*, Phys. Rev. **E 59**, 3669, (1999).
- [ML01] H.W. Müller and M. Liu, *Structure of Ferro-Fluidynamics*, Phys. Rev. **E 64**, 061405 (2001).
- [ML02] H.W. Müller and M. Liu, *Shear Excited Sound in Magnetic Fluid*, Phys. Rev. Lett. **89**, 67201, (2002).
- [MHL06] O. Müller, D. Hahn and M. Liu, *Non-Newtonian behaviour in ferrofluids and magnetization relaxation*, J. Phys.: Condens. Matter **18**, 2623, (2006).
- [MIL08] S. Mahle, P. Ilg and M. Liu, *Hydrodynamic theory of polydisperse chain-forming ferrofluids*, Phys. Rev. **E 77**, 016305 (2008).
- [TPLB00] H. Temmen, H. Pleiner, M. Liu and H.R. Brand, *Convective Nonlinearity in Non-Newtonian Fluids*, Phys. Rev. Lett. **84**, 3228 (2000).
- [TPLB01] H. Temmen, H. Pleiner, M. Liu and H.R. Brand, *Temmen et al. reply*, Phys. Rev. Lett. **86**, 745 (2001).
- [PLB04] H. Pleiner, M. Liu and H.R. Brand, *Nonlinear Fluid Dynamics Description of non-Newtonian Fluids*, Rheologica Acta **43**, 502 (2004).
- [Mul06] O. Müller, *Die Hydrodynamische Theorie Polymerer Fluide*, PhD Thesis University Tübingen (2006).
- [MLPH16] Oliver Müller, Mario Liu, Harald Pleiner, Helmut R. Brand, Phys. Rev. **E 93**, 023113 (2016).
- [MLPH16b] Oliver Müller, Mario Liu, Harald Pleiner, Helmut R. Brand, Phys. Rev. **E 93**, 023114 (2016).

- [JL07] Y.M. Jiang and M. Liu. A brief review of granular elasticity. *Eur. Phys. J. E* 22, 255 (2007).
- [JL09] Y.M. Jiang and M. Liu. Granular solid hydrodynamics. *Granular Matter*, 11:139 (2009). Free download: [www.springerlink.com/content/a8016874j8868u8r/fulltext](http://www.springerlink.com/content/a8016874j8868u8r/fulltext)
- [KV09] Y.M. Jiang and M. Liu. The physics of granular mechanics. In D. Kolymbas and G. Viggiani, editors, *Mechanics of Natural Solids*, pages 27–46. Springer, 2009.
- [JL14] Y.M. Jiang and M. Liu. Granular Solid Hydrodynamics (GSH): a broad-ranged macroscopic theory of granular media. *Acta Mech.* 225, 2363–2384 (2014) DOI 10.1007/s00707-014-1131-3.
- [GJL11] G. Gudehus, Y.M. Jiang, and M. Liu. Seismo- and thermodynamics of granular solids. *Granular Matter*, 1304:319–340, 2011.
- [JL03] Y.M. Jiang, M. Liu. Granular Elasticity without the Coulomb Condition *Phys. Rev Lett*, **91**, 144301(2003).
- [JL04] Y.M. Jiang, M. Liu. Energetic Instability Unjams Sand and Suspension. *Phys. Rev Lett*, **93**, 148001(2004).
- [JL07] Y.M. Jiang, M. Liu. From Elasticity to Hypoplasticity: Dynamics of Granular Solids. *Phys. Rev Lett*, **99**, 105501 (2007).
- [KPB JL06] Krimer, D.O., Pfitzner, M., Bräuer, K., Jiang, Y., Liu, M.: Granular elasticity: general considerations and the stress dip in sand piles. *Phys. Rev. E* 74, 061310 (2006)
- [BPKMJL06] Bräuer, K., Pfitzner, M., Krimer, D.O., Mayer, M., Jiang, Y., Liu, M.: Granular elasticity: stress distributions in silos and under point loads. *Phys. Rev. E* 74, 061311 (2006)
- [JL08] Jiang, Y.M., Liu, M.: Incremental stress-strain relation from granular elasticity: Comparison to experiments. *Phys. Rev. E* 77, 021306 (2008)
- [JL09] Y.M. Jiang and M. Liu. GSH, or Granular Solid Hydrodynamics: on the Analogy between Sand and Polymers. *AIP Conf. Proc.* 7/1/2009, Vol. 1145 Issue 1, p1096.
- [ML12] Mayer, M., Liu, M.: Propagation of elastic waves in granular solid hydrodynamics. *Phys. Rev. E* 82, 042301 (2010)
- [KML12] D. Krimer, S. Mahle and M. Liu, Dip of the Granular Shear Stress, *Phys. Rev. E* 86, 061312 (2012)
- [JZPFSSML12] Jiang, Y.M., Zheng, H.P., Peng, Z., Fu, L.P., Song, S.X., Sun, Q.C., Mayer, M., Liu, M.: Expression for the granular elastic energy. *Phys. Rev. E* 85, 051304 (2012)

- [ZLHJL12] Zhang, Q., Li, Y.C., Hou, M.Y., Jiang, Y.M., Liu, M.: Elastic waves in the presence of a granular shear band formed by direct shear. *Phys. Rev. E* 85, 031306 (2012).
- [JL13] Jiang, Y., Liu, M.: Proportional path, barodesy, and granular solid hydrodynamics. *Granul. Matter* 15, 237–249 (2013)
- [JL13] Yimin Jiang and Mario Liu, *AIP Conf. Proc.* 1542, 52 (2013); *Stress- and rate-controlled granular rheology*, <http://dx.doi.org/10.1063/1.4811867>
- [JL15] Y.M. Jiang and M. Liu. *Applying GSH to a wide range of experiments in granular media*. *Eur. Phys. J. E* (2015) 38:15.
- [JL16b] Yimin Jiang and Mario Liu. *Similarities between GSH, hypoplasticity and KCR*. *Acta Geotech.* (2016) 11:519-537, DOI 10.1007/s11440-016-0461-9
- [JL17] Y.M. Jiang and M. Liu. *Why granular media are thermal, and quite normal, after all* *Eur. Phys. J. E* (2017) 40: 10
- [LL80] L.D. Landau, and E.M. Lifshitz, *Statistical Physics*, Butterworth-Heinemann, 1980
- [EL18] Itai Einav and Mario Liu, *Hydrodynamic derivation of the work input to fully and partially saturated soils*, *J of Mechanics and Physics of Solids* 110 (2018) 205–217
- [LL86] L.D. Landau and E.M. Lifshitz, *Theory of Elasticity* (Pergamon Press, Oxford, 1986).
- [CL95] P.M. Chaikin and T.C. Lubensky, *Principles of condensed matter physics*, Cambridge University Press, Cambridge, 1995, Chap. 6.5;
- [BAH77] R.B. Bird, R.C. Armstrong, O. Hassager, *Dynamics of Polymeric Liquids, Vol.1* (John Wiley, New York, 1977).
- [Str97] G. Strobl, *The Physics of Polymers* (Springer-Verlag Berlin Heidelberg, 1997).
- [Tan00] R.I. Tanner, *Engineering Rheology* (Oxford University Press, New York, 2000).
- [Lar88] R.G. Larson, *Constitutive Equations for Polymer Melts and Solutions*, (Butterworth) 1988.
- [GLMc03] R.S. Graham, A.E. Likhtman, and T.C.B. McLeish, *J. Rheol.* 47, 1171 (2003).
- [DE86] M. Doi, S.F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, New York, 1986).
- [WYT03] M.H. Wagner, M. Yamaguchi, and M. Takahashi, *J. Rheol.* 47, 779 (2003).
- [McL98] T.C.B. McLeish and R.G. Larson, *J. Rheol.* 42, 81 (1998).

---

# Definition and Uses of the Principle of Virtual Power

**Ioannis Stefanou**

*Laboratoire Navier, UMR 8205, Ecole des Ponts ParisTech, IFSTTAR, CNRS, Champs-sur-Marne, France*

---

*The Principle of Virtual Power (PVP) offers a systematic approach for studying the equilibrium of complex systems. This chapter aims at showing the importance of the principle and its use in discrete and continuum systems. Simple examples are given throughout the chapter for helping understanding. The general theory of micromorphic continua is also derived using the principle of virtual power. Finally, PVP being global rather than local, is directly amenable to numerical schemes such as the Finite Element Method (FEM). An example is given using the open-source FEM library FEniCS.*

## 1 Introduction

The Principle of Virtual Power (PVP) offers a systematic approach for studying the equilibrium of complex systems following an energy- and velocity-based approach. PVP shows an important advantage compared to the conventional approach of force equilibrium equations as it is simpler to apply in complex systems. Moreover, forces play a secondary role in PVP as they appear only as conjugate in energy quantities to velocities, which are in principle measurable.

After studying this chapter, the reader will be able to understand the importance of the principle of virtual power, its use in problems of discrete systems and continuum mechanics and its application to the Finite Element Method (FEM). Simple examples are given throughout the chapter for helping understanding.

The current chapter has the following structure. In the beginning of section 2 the statement of the principle is given, accompanied by a brief historical note, in order to clarify the fundamental ideas of the principle and its connection with the equilibrium equations. Examples are then given for showing the application of PVP to simple problems involving discrete systems of one and several degrees of freedom. In section 3, the principle is extended to continuum mechanics, where the equivalence between

the strong form of the differential equilibrium equations and the weak/variational form of the problem is shown. Further, in section 4, the principle is used for deriving more advanced continuum theories, i.e. the micromorphic continuum theory. The equations of micromorphic continua are presented in a general form and the hierarchical structure of the theory is illustrated. Cosserat and strain gradient theories are presented as special cases. Applications are also presented for showing the advantages of micromorphic continua and the use of PVP in upscaling. Finally, the use of the principle in the Finite Element Method is shown and the example of shearing of an infinite layer modeled as a Cauchy and a Cosserat continuum is presented, using the FEniCS open source FEM library.

## 2 The Principle of Virtual Power

### 2.1 Statement

There is no doubt that the intuition for the Principle of Virtual Power (PVP) existed from ancient times. However, the principle is used in a more systematic time in the 18<sup>th</sup> century. It is difficult to say who is the father of the principle, maybe Galileo Galilei, René Descartes, Jean-Baptiste le Rond d'Alembert or Johann Bernoulli. The first, clear, general and direct statement of the principle seems to be by Joseph-Louis Lagrange in 1788 in his seminal work. The principle reads [Lag88; p.10-11]:

*“Si un système quelconque de tant de corps ou points que l'on veut tirés, chacun des puissances quelconques, est en équilibre et qu'on donne à ce système un petit mouvement quelconque, en vertu duquel point parcourt un espace infiniment petit qui exprimera sa vitesse virtuelle ; la somme des puissances, multipliées chacune par l'espace que le point où elle est appliquée, parcourt suivant la direction de cette même puissance, sera toujours égale à zéro, en regardant comme positifs les petits espaces parcourus dans le sens des puissances, et comme négatifs les espaces parcourus dans le sens opposé.”<sup>1</sup>*

It is worth noting that Lagrange uses the term “*puissance*” (power in English) instead of the term “*force*”, which is commonly used today. The word “*puissance*” comes from the old French verb “*pouvoir*”, which means “be able to”. In this sense a force is a quantity that enables motion, the cause.

Another central notion in the above principle is the notion of virtual velocities. According to Lagrange [Lag88]:

*“On doit entendre par vitesse virtuelle, celle qu'un corps en équilibre est disposé à recevoir; en cas que l'équilibre vient d'être rompu ; c'est-à-dire la*

<sup>1</sup>Translation in English by G.A. Maugin [Mau14]: *If any system of bodies or points as we want, is acted upon by any system of powers, is in equilibrium, and we give to this system any small motion, then by virtue of the fact that each point travels an infinitesimally small space that expresses its virtual velocity, the sum over powers each multiplied by the space that the point where it is applied travels along the direction of the same power, will always be equal to zero, regarding as positive the small distances followed in the direction of the powers and as negative those travelled in the opposite direction.*

*vitesse que le corps prendroit réellement dans le premier instant de son mouvement*<sup>2</sup>

Virtual velocities are “virtual” and not “real”, in the sense that they are possible velocities that could be developed, if the equilibrium is not anymore satisfied. Virtual velocities should be thought as variations from the reference/equilibrium state.

It is worth mentioning that the principles of virtual power (velocities) and virtual work (displacements) are practically the same. However, there is a small difference. The principle of virtual power has the advantage to be applied without considering infinitesimal displacements and rotations.

## 2.2 PVP applied on a single body

In the case of only one body, the principle says that the body will be in equilibrium if, and only if, the power generated by the forces acting on it is null under any possible (virtual) velocity of the body. The power is said to be “virtual”, because there is no need for the particle to actually move to apply the principle. It only needs to be in equilibrium (steady state), either moving, or at rest.

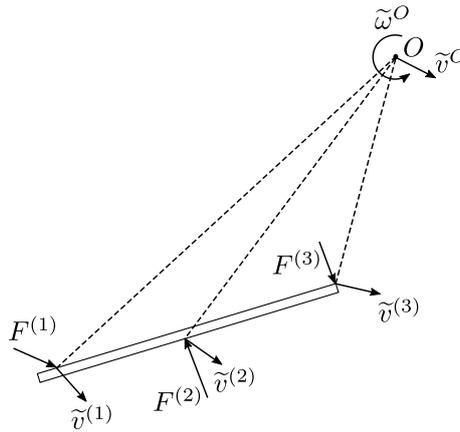


Figure 1: A rigid body under the action of three forces and virtual velocities.

Take for example the undeformable solid of Figure 1, where three forces are applied. The body will be in equilibrium if, and only if, the total power of the applied forces  $F^{(i)}$  is zero for any kinematically admissible virtual velocity  $\tilde{v}^{(i)}$ , i.e.:

$$\sum_i \tilde{\mathcal{P}}^{(i)} = \sum_i \mathbf{F}^{(i)} \cdot \tilde{\mathbf{v}}^{(i)} = 0, \quad \forall \tilde{\mathbf{v}}^{(i)}. \quad (1)$$

<sup>2</sup>Translation in English: *We mean with virtual velocity, the velocity that a body in equilibrium will develop if its equilibrium is broken, i.e. the velocity that the body will really have in the first instance of its movement*

The fact that the solid is not deformable imposes the following constraint to the virtual velocities:

$$\tilde{\mathbf{v}}^{(i)} = \tilde{\mathbf{v}}^O + \tilde{\boldsymbol{\omega}}^O \times \mathbf{r}^{(O,i)}, \quad (2)$$

where  $\tilde{\mathbf{v}}^O$  and  $\tilde{\boldsymbol{\omega}}^O$  represent, respectively, virtual translational and rotational velocities with respect to an arbitrary point  $O$  and  $\mathbf{r}^{(O,i)}$  is the position vector of the application point of force  $i$  with respect to  $O$ . The virtual velocities,  $\tilde{\mathbf{v}}^{(i)}$ , have to respect equation (2) or any other restrictions (e.g. boundary conditions), i.e. to be kinematically admissible. Using equation (2), equation (1) becomes:

$$\left( \sum_i \mathbf{F}^{(i)} \right) \cdot \tilde{\mathbf{v}}^O + \left( \sum_i \mathbf{r}^{(O,i)} \times \mathbf{F}^{(i)} \right) \cdot \tilde{\boldsymbol{\omega}}^O = 0, \quad \forall \tilde{\mathbf{v}}^O, \tilde{\boldsymbol{\omega}}^O. \quad (3)$$

As this equation is valid for any virtual displacement and rotation, we can derive (deduce) the standard force and moment equilibrium equations:

$$\begin{aligned} \sum_i \mathbf{F}^{(i)} &= 0 \\ \sum_i \mathbf{M}^{(O,i)} &= 0 \end{aligned} \quad (4)$$

where  $\mathbf{M}^{(O,i)} = \mathbf{r}^{(O,i)} \times \mathbf{F}^{(i)}$  is the moment of force  $i$  w.r.t. point  $O$ .

Notice that the reverse procedure is also possible. Starting from the equilibrium equations, equations (4), multiplying each one by arbitrary quantities  $\tilde{\mathbf{v}}^O$  and  $\tilde{\boldsymbol{\omega}}^O$ , adding them and using equation (2) we retrieve equation (1). This shows the equivalence between the principle of virtual power and the equilibrium equations.

### 2.3 Generalized forces

The notion of power in the PVP allows the consideration of generalized forces such as moments (also called double forces, dipoles, couples) or even more general quantities (e.g. triple forces, tripoles, see section 4), which are conjugate in energy with generalized velocities. For instance, in the case of a moment,  $\mathbf{M}$ , the generalized velocity is an angular velocity  $\boldsymbol{\omega}$ . In the case of a system with  $j$  external forces and  $k$  external moments, equation (1) becomes ( $i = j + k$ ):

$$\sum_i \tilde{\mathcal{P}}^{(i)} = \sum_j \mathbf{F}^{(j)} \cdot \tilde{\mathbf{v}}^{(j)} + \sum_k \mathbf{M}^{(k)} \cdot \tilde{\boldsymbol{\omega}}^{(k)} = 0, \quad \forall \tilde{\mathbf{v}}^{(j)}, \tilde{\boldsymbol{\omega}}^{(k)}. \quad (5)$$

Following the same procedure as the previous paragraph, the standard equilibrium equations are retrieved:

$$\begin{aligned} \sum_j \mathbf{F}^{(j)} &= 0 \\ \sum_j \mathbf{M}^{(O,j)} + \sum_k \mathbf{M}^{(k)} &= 0 \end{aligned} \quad (6)$$

where  $\mathbf{M}^{(O,j)} = \mathbf{r}^{(O,j)} \times \mathbf{F}^{(j)}$  is the moment of the external force  $j$ ,  $\mathbf{F}^{(j)}$ , w.r.t. point  $O$ , while  $\mathbf{M}^{(k)}$  is a concentrated external moment.

## 2.4 PVP applied on a system of bodies

Consider now a system of several bodies that interact one with another through a system of internal forces  $F^{(int,i)}$ . These internal forces oppose the relative movement of the bodies (think of a spring for example that connects two bodies). Consequently, according to the PVP, their virtual power has to be considered with a negative sign. On the same system of bodies, we apply also external forces,  $F^{(ext,i)}$ . PVP says that the system will be in equilibrium if, and only if, the power of (all) the forces acting on it (both external and internal) is zero, under any virtual velocity:

$$\sum_i \tilde{\mathcal{P}}^{(i)} = \sum_j \tilde{\mathcal{P}}^{(ext,j)} - \sum_k \tilde{\mathcal{P}}^{(int,k)} = 0, \quad \forall \tilde{\mathbf{v}}^{(i)}. \quad (7)$$

*Example:* Consider the beam of length,  $L$ , presented in figure 2. Both the horizontal and vertical displacements are fixed at point A, while only the vertical displacement is fixed at point B. The system is statically determinate. A vertical external force is applied in the middle of the beam as shown in figure 2. For calculating the internal moment at point D, we remove the kinematical constraint that assures the continuity of the system at that point and we replace it with a system of internal moments  $M_D^{int, left} = M_D^{int, right} = M_D^{int}$ . The PVP (equation (7)) yields:

$$F\tilde{\delta}_C - \left( M_D^{int, left} \tilde{\alpha} + M_D^{int, right} \tilde{\beta} \right) = 0 \quad \forall \tilde{\delta}_C, \tilde{\alpha}, \tilde{\beta}. \quad (8)$$

From geometrical compatibility we obtain (see figure 2):

$$\tilde{\delta}_D = \tilde{\alpha} \frac{L}{4} = \tilde{\beta} \frac{3L}{4} \quad \text{and} \quad \tilde{\delta}_C = \tilde{\beta} \frac{L}{2}. \quad (9)$$

Replacing equation (9) to equation (8) we finally obtain that  $M_D^{int} = \frac{FL}{8}$ .

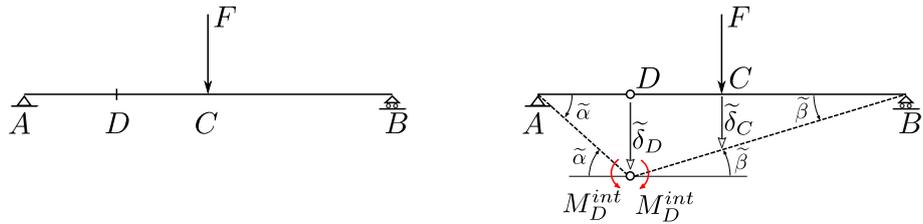


Figure 2: Calculation of internal moment at point D with the principle of virtual power.

## 3 The principle of virtual power in continuum mechanics

The above concepts are extended in continuum mechanics, provided that the various kinematic and stress fields show a certain mathematical regularity [Fré02, GFA10]. The first step is to define the internal and external virtual powers.

Consider a solid of volume  $V$ , with boundary  $S$ , as shown in figure 3. At  $S_v$  velocities are prescribed,  $v_i = v_i^{S_v}$ , while at  $S_p$  tractions are imposed,  $t_i = t_i^{S_p}$ .  $S = S_v \cup S_p$  and  $S_v \cap S_p = \emptyset$ . Body forces,  $f_i$ , are applied everywhere on the solid. In the rest of this manuscript, indicial notation is followed using Einstein's notation, i.e. repeated indices denote summation. In three dimensions  $i = 1, 2, 3$ .

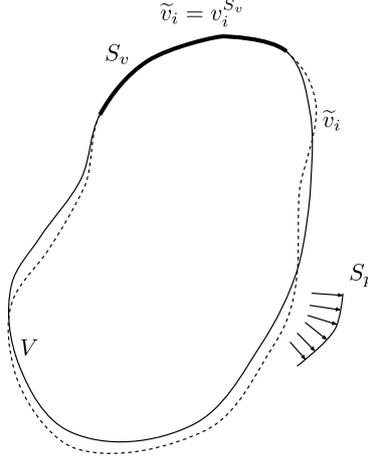


Figure 3: Calculation of internal moment at point D with the principle of virtual power.

Assuming the following forms for the internal and external virtual power densities:

$$\begin{aligned}\tilde{p}^{(int)} &= \sigma_{ij} \tilde{\epsilon}_{ij} \\ \tilde{p}^{(ext,t)} &= t_i \tilde{v}_i \\ \tilde{p}^{(ext,f)} &= f_i \tilde{v}_i,\end{aligned}\tag{10}$$

where  $\tilde{\epsilon}_{ij} = \frac{1}{2}(\tilde{v}_{i,j} + \tilde{v}_{j,i})$  and  $\sigma_{ij} = \sigma_{ji}$ , according to equation (7) and paragraph 2.1, the principle of virtual power says that any part of the system will be in equilibrium, if and only if for any sub-volume,  $D$ , of  $V$  with boundary,  $\partial D$ , the following equality is satisfied:

$$\int_D \sigma_{ij} \tilde{\epsilon}_{ij} dV - \int_D f_i \tilde{v}_i dV - \int_{\partial D} t_i \tilde{v}_i dS = 0,\tag{11}$$

for any kinematically admissible virtual velocity field  $v_i$ , i.e.  $\forall v_i$  with  $v_i = v_i^{S_v}$  on  $S_v$ .

Applying the divergence theorem and after some algebra, we obtain:

$$\int_D (\sigma_{ij,j} + f_i) \tilde{v}_i dV + \int_{\partial D} (t_i - \sigma_{ij} n_j) \tilde{v}_i dS = 0,\tag{12}$$

where  $n_i$  is the normal vector to  $S$ .

The above equation holds for any volume  $D \subset V$  and  $\forall v_i$  and therefore it has to hold:

$$\sigma_{ij,j} + f_i = 0 \quad \text{and} \quad t_i = \sigma_{ij}n_j. \quad (13)$$

The above equations form the standard equilibrium equations and stress boundary conditions in the classical, Cauchy continuum. As in the simpler case of discrete forces, the reverse procedure is possible, showing the full equivalence of the principle of virtual power (weak form) with the differential equilibrium equations (strong form). We say that equation (12) is the weak formulation of the strong form, equations (13), of the problem. The weak form of a Partial Differential Equation (PDE) is called also variational form and it is the starting point for obtaining numerical solutions with the Finite Element method. For an introduction to variational calculus and applications we refer to [Fun65]. Nowadays, several Finite Element codes exist that allow the user to automatically and efficiently solve PDE's in parallel by directly entering in symbolic, mathematical form the variational form of the problem (e.g. Fenics project [ABH<sup>+</sup>15], GetDP [GDL98], FreeFEM++ [Hec12], among others). An example is given in section 5.

Applying one more time the divergence theorem on the whole solid of volume  $V$ , we obtain:

$$\int_{\partial V} t_i dS + \int_V f_i dV = 0. \quad (14)$$

Multiplying equations (13) with  $\epsilon_{ijk}x_k$ , where  $\epsilon_{ijk}$  is the Levi-Civita symbol, and using once more the divergence theorem, we obtain the following equation:

$$\int_{\partial V} \epsilon_{ijk} t_j x_k dS + \int_V \epsilon_{ijk} f_j x_k dV = 0, \quad (15)$$

which represents a cross product (moment). Equations (14) and (15) are called, respectively, linear and angular momentum balance equations. They are the analogue of equations (4) in continuum mechanics.

## 4 Micromorphic continua and the method of virtual power in continuum mechanics

In the classical continuum theory, the material point is characterized by its position and velocity. This description is abandoned in non-Newtonian, quantum physics, where it is described by a single quantity, the wave function. Staying in the Newtonian context there are various situations where one might need to assign more than translational degrees of freedom to the material point, seen now as a particle.

A more general continuum theory that, by construction, can indeed account for an arbitrary number of degrees of freedom assigned at the material point is the Micromorphic theory. This theory is general enough to represent various heterogeneous systems with microstructure of non-negligible size and take into account various length

and time scales (internal lengths) that the classical Cauchy continuum fails to represent. The various features of the Micromorphic continuum theory were studied by many researchers in the past, showing several advantages compared to the classical continuum approach. Intrinsic wave dispersion, regularization in strain localization problems, non-singular fields in fracture mechanics, interesting properties related to the design of metamaterials, are some of the applications that emerge from the deep study of these continua.

According to Germain [Ger73], in the classical description, a continuum is a continuous distribution of particles, each of them being represented geometrically by a point and characterized kinematically by a velocity  $V_i$ . In a theory that takes microstructure into account, from the macroscopic point of view (which is the point of view of a continuum theory), each particle is still represented by a point  $M$ , but its kinematical properties are defined in a more detailed way.

At the microscopic level of observation, a particle appears itself as a continuum  $P(M)$  of small extent. Let  $M$  be the center of mass of the particle  $P(M)$ ,  $M'$  a point of  $P(M)$ ,  $V_i$  the displacement of  $M$ ,  $x'_i$  the coordinates of  $M'$  in a Cartesian frame parallel to the given, global frame and  $M$  its origin,  $V'_i$  the velocity of  $M'$  with respect to the given frame and  $x_i$  the coordinates of  $M$  in the given frame (see figure 4).  $D$  denotes the control volume. As  $P(M)$  is of small extent, it is natural to look at the asymptotic expansion of  $V'_i$  with respect to  $x'_i$ :

$$V'_i = V_i + \chi_{ij}x'_j + \chi_{ijk}x'_jx'_k + \chi_{ijkl}x'_jx'_kx'_l + \dots, \quad (16)$$

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

The virtual power density of the internal forces for a micromorphic continuum of order  $n$  is given as follows by [Ger73]:

$$\tilde{p}^{int} = \tau_{ij}\tilde{V}_{i,j} - (s_{ij}\tilde{\chi}_{ij} + s_{ijk}\tilde{\chi}_{ijk} + \dots) + (\nu_{ijk}\tilde{\kappa}_{ijk} + \nu_{ijkl}\tilde{\kappa}_{ijkl} \dots), \quad (17)$$

with  $\tau_{ij} \equiv \sigma_{ij} + s_{ij}$ , where  $\tau_{ij}$  is the stress tensor,  $\sigma_{ij}$  is the intrinsic stress tensor (symmetric due to objectivity requirement),  $s_{ij}$  is the intrinsic microstress tensor,  $\nu_{ijk}$  is the intrinsic second microstress tensor and  $s_{ij\dots m}$ ,  $\nu_{ij\dots ml}$  are higher order stress tensors that are conjugate in energy to  $\chi_{ij\dots m}$  and  $\kappa_{ij\dots ml} = \chi_{ij\dots m,l}$ , respectively.  $(\cdot)_{,i}$  denotes derivation in terms of  $x_i$  (macro-coordinate). The virtual power density of the external forces for a micromorphic continuum of order  $n$  is given as follows [Ger73]:

$$\begin{aligned} \tilde{p}^{(ext,t)} &= t_i\tilde{v}_i + \mu_{ij}\tilde{\chi}_{ij} + \mu_{ijk}\tilde{\chi}_{ijk} + \dots \\ \tilde{p}^{(ext,f)} &= f_i\tilde{v}_i + \psi_{ij}\tilde{\chi}_{ij} + \psi_{ijk}\tilde{\chi}_{ijk} + \dots, \end{aligned} \quad (18)$$

where  $f_i$ ,  $\psi_{ij\dots l}$  represent volumic (body) generalized forces and  $t_i$ ,  $\mu_{ij\dots l}$  generalized tractions. In particular,  $t_i$  is the surface traction,  $\mu_{ij}$  is the double surface traction

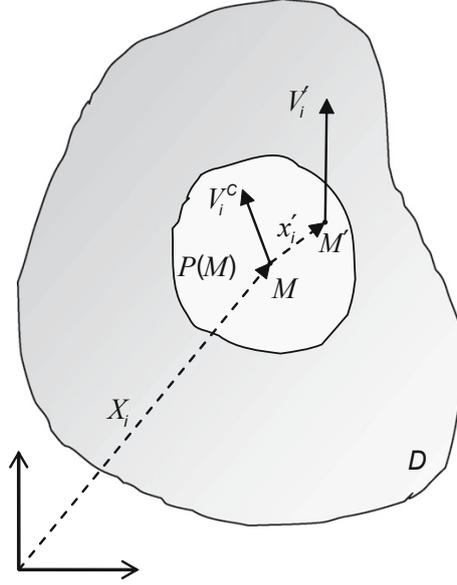


Figure 4: Continuum with microstructure.

(dipole, e.g. a concentrated moment density) and  $\mu_{ij\dots n}$  is the generalized surface traction of order  $n$  defined on the boundary  $\partial V$ . Similarly,  $f_i$  represent long-range volumic forces,  $\psi_{ij}$  double long-range volumic forces (e.g. due to an electromagnetic field) and  $\psi_{ij\dots n}$  higher order generalized volumic forces of order  $n$  defined on  $V$ .

Applying the principle of virtual power and using the divergence theorem, we obtain [Ger73]:

$$\begin{aligned}
 \tau_{ij,j} + f_i &= 0, & t_i &= \tau_{ij}n_j \\
 \nu_{ijk,k} + s_{ij} + \psi_{ij} &= 0, & \mu_{ij} &= \nu_{ijk}n_k \\
 \nu_{ijkl,l} + s_{ijk} + \psi_{ijk} &= 0, & \mu_{ijk} &= \nu_{ijkl}n_l \\
 & \dots, & &
 \end{aligned} \tag{19}$$

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

It is worth emphasizing that the above equations are derived systematically without any particular hypotheses besides the postulates of the internal and external power densities. No assumption was also made regarding the constitutive behavior of the system, i.e. the relation of the generalized stresses with the generalized deformations. The above approach can be easily generalized to take into account inertial effects [Ger73]. In this case, the additional degrees of freedom of micromorphic continua in-

roduce microinertia terms, whose presence leads to interesting wave dispersion properties, especially at short wavelengths (optic branch) [SSV10] and finite Lyapunov exponents in localization problems [SSV11].

Without doubt that micromorphic theory is rich enough to fit various physical situations. This is a strong point of the theory, but it is also its weak point. The discovery of the practical significance of some of the concepts which have been introduced (e.g. of the higher order terms), the design of a general method for exhibiting their physical validity (if any!) and their measurement in some specific physical situations is not obvious [Ger73] and remains an open research topic. The situation, though, becomes more tractable in some special cases of 1<sup>st</sup> order micromorphic continua, where applications exist in various disciplines.

#### 4.1 Special cases of micromorphic continua

In figure 6 we outline the various higher order (micromorphic) continuum theories and their special cases. Besides the classical continuum and the Cosserat continuum (called also micropolar continuum, see [Var09]), a special case of micromorphic continuum is also the second gradient and the indeterminate couple stress theory (called also restrained Cosserat medium).

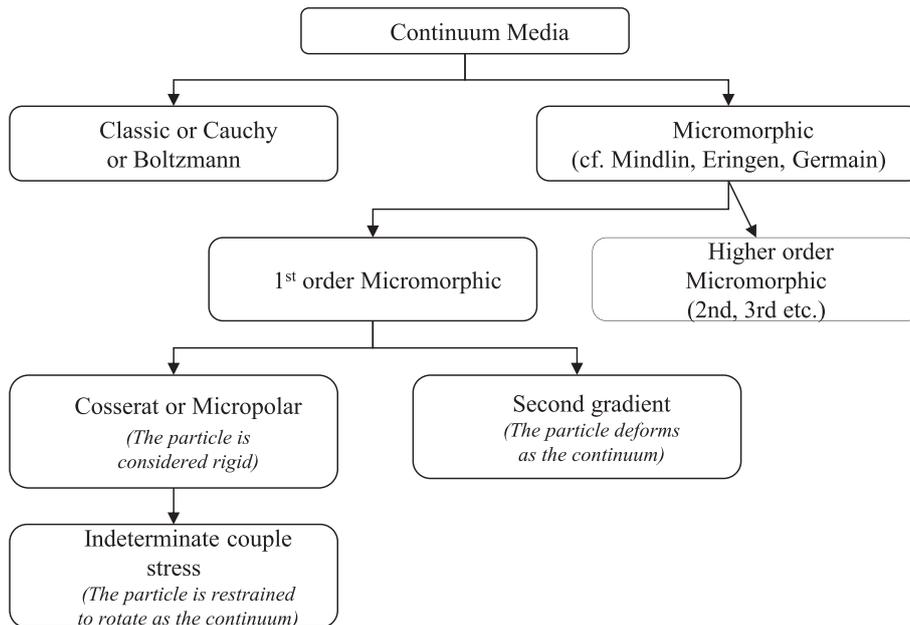


Figure 5: Higher order continuum theories according to Germain's terminology [Ger73]; see also [Min64, Eri99].

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

In the case that the particle  $P(M)$  is deformable and its microdeformation coincides with the deformation of the (macro-)continuum, i.e.  $\chi_{ij} = V_{i,j}$ , we obtain the so-called second gradient continuum theory. As in this case the microdeformation rate tensor is no more an independent generalized virtual velocity, one has to start from the very beginning and apply the principle of virtual power for deriving the strong form of the equilibrium equations and the appropriate boundary conditions. For more details we refer to [Ger73] and for some interesting applications of the theory to [DSMP93, CCE98, ZPV01, CCC06, KABC08, PZ16, DAD<sup>+</sup>17], among others.

## 4.2 The Cosserat continuum

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

The Cosserat continuum is easily retrieved by setting (hypothesis of rigid particle)  $\chi_{ij} = -\epsilon_{ijk}\omega_k^c$ ,  $k_{ij} = \omega_{i,j}^c$ ,  $s_{ij} = -\frac{1}{2}\epsilon_{ijk}s_k$ ,  $\mu_{ij} = -\frac{1}{2}\epsilon_{ijk}\mu_k$ ,  $\nu_{ijk} = -\frac{1}{2}\epsilon_{ijl}m_{lk}$ ,  $\psi_{ij} = -\frac{1}{2}\epsilon_{ijk}\psi_k$  and using  $\tau_{ij} \equiv \sigma_{ij} + s_{ij}$  (see equation (17)) and the property  $\epsilon_{ijp}\epsilon_{ijk} = 2\delta_{pk}$ , where  $\delta_{ij}$  is the Kronecker delta, equations (17) and (18) become:

$$\begin{aligned}\tilde{p}^{int} &= \tau_{ij}\tilde{\gamma}_{ij} + m_{ij}\tilde{k}_{ij} \\ \tilde{p}^{(ext,t)} &= t_i\tilde{v}_i + \mu_i\tilde{\omega}_i^c \\ \tilde{p}^{(ext,f)} &= f_i\tilde{v}_i + \psi_i\tilde{\omega}_i^c\end{aligned}\quad (20)$$

where  $\gamma_{ij} = u_{i,j} + \epsilon_{ijk}\omega_k^c$ .

The equilibrium equations (equations (19)) take the following form:

$$\begin{aligned}\tau_{ij,j} + f_i &= 0, & t_i &= \tau_{ij}n_j \\ m_{ij,j} - \epsilon_{ijk}\tau_{jk} + \psi_i &= 0, & \mu_i &= m_{ij}n_j\end{aligned}\quad (21)$$

This is the strong form of the Cosserat continuum equations. In figure 6 the stresses and couple-stresses (moments) of a Cosserat continuum are illustrated. It is worth emphasizing that the derivation is based on the principle of virtual power and not on the linear and angular momentum balance equations. These momentum balance equations can be deduced by integrating equation (21) over the volume  $V$  and by applying the divergence theorem, as in the case of the classical Cauchy continuum presented in section 3. Consequently, the method of PVP enables us to derive in a safe and systematic way complex, higher-order balance equations, that physical intuition can hardly bring us to.

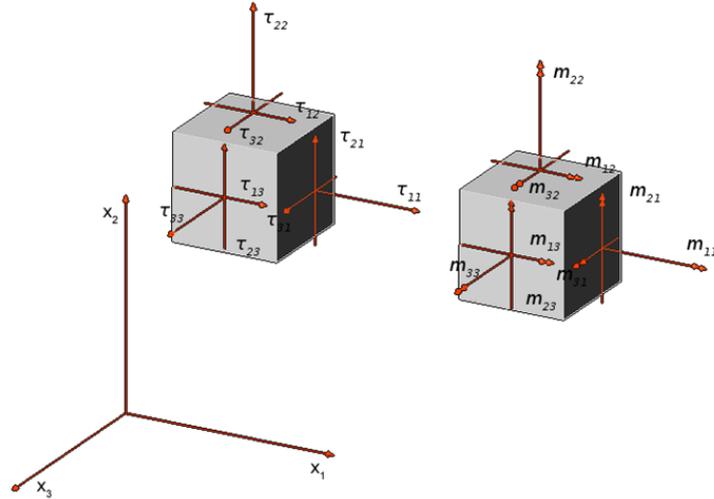


Figure 6: Stresses and couple-stresses (moments) of Cosserat continuum.

### 4.3 PVP and upscaling for deriving constitutive laws

The above continua cannot be used unless appropriate constitutive laws are used for solving engineering problems. Of particular interest are heterogeneous systems, where the higher-order continuum theories show several advantages (e.g. physically-based regularization in strain localization problems and wave dispersion). Constitutive models can be derived either experimentally, by trial and error, or by explicitly considering the microstructure using upscaling techniques.

Upscaling (or homogenization) is a class of methods that aim at deriving an equivalent continuum theory that describes the macroscopic behavior of heterogeneous systems. (Asymptotic) Homogenization is a mathematically rigorous, well established theory for performing this task [BP89, SP86, SP80, PdCOTD09, Cha10, CN84, TC97, ABG09]. This method is based on the asymptotic expansion of the various state fields (displacements, deformations, stresses) in terms of a small quantity  $\varepsilon$ , which represents the ratio of the characteristic size of the elementary volume over the overall size of the structure, and provides an equivalent to the heterogeneous system homogeneous continuum for  $\varepsilon \rightarrow 0$ . Besides the rigorous mathematical formulation of this approach, its main advantage is the ability to determine error estimators of the derived continuum for finite values of  $\varepsilon$ . However, when it comes to generalized continua, such as the Cosserat continuum, that possess internal lengths, the asymptotic limit  $\varepsilon \rightarrow 0$  loses interest as it cancels out the internal lengths [FS98, FPS01]. In other words, by this pass to the limit, asymptotic homogenization erases any internal lengths related to the material's micro-structure, which higher order continuum theories, such as Cosserat are in principle able to capture. The separation of scales,

intrinsic in asymptotic homogenization theory, does not hold anymore.

To overcome this problem several alternative schemes have been proposed for upscaling heterogeneous systems (see for example [AL94, FS98, BV01, SSV08, SSV10, BG12, GSSS16, RC16], among others). The majority of these schemes is based upon the “homogeneous equivalent continuum” concept [Cha10], in the sense that the derived higher order continuum shares a) the same power (internal) and b) the same kinematics with the heterogeneous medium for any generalized virtual velocity field. This approach reminds us the PVP, applied on a reduced space of kinematics and averaging. The classical asymptotic homogenization expansion Ansatz that leads to a Cauchy continuum as the ratio of the size of the unit cell over the overall structure tends to zero is not followed in this case. Therefore, these heuristic approaches remain applicable even when the size of the microstructure is not infinitesimal as compared to the overall size of the system, or in other words, when scale-separation is not satisfied.

A typical example for applying and testing these upscaling methods are masonry-like structures. Masonry can be seen as a geomaterial whose building blocks are often quasi-periodically arranged in space. Moreover, the building blocks are at the human scale, which makes them an ideal toy-model, contrary to granular media whose microstructure is small, shows topological complexity and has to be statistically described. When the upscaling scheme is correctly formulated, it is possible to capture the wave dispersion behavior of a heterogeneous system, even when the wave length is comparable to the block size. Notice that in this case, the classical, Cauchy continuum approach fails as it is not a dispersive medium. In figure 7 we present the modal frequencies of a masonry panel that was upscaled with Cosserat continuum versus the number of its building blocks. Even when the number of the building blocks is small, the Cosserat homogenized continuum model succeeds in representing the dynamics of the discrete heterogeneous structure. In figure 8 we present the out-of-plane-displacement contours of the first three flexural modal shapes of a homogenized masonry panel and the comparison with the flexural modal shapes provided by the Discrete Element Method (DEM) [GSSS14]. It is worth mentioning that DEM provides very satisfactory results compared to well controlled experimental tests and, therefore, it can be used as reference [GSS17]. The upscaled Cosserat continuum behaves very well, even for non-linear material behavior (see figure 9).

Nonlinear behavior, accompanied by strain softening is inherent in granular materials [Var09]. In figure 10 we show the response of a granular layer under shearing studied with the Discrete Element Method. Upscaling granular media and transferring adequate information at the macroscale (e.g. internal lengths) is a challenging topic [BV01, RC16, GSSS16] and can provide valuable information on strain localization and energy dissipation in the absence of detailed experimental data, which is often the case for fault gouges due to their complex thermo-hydro-chemo-mechanical behavior (see Figure 12, [SSV11, VSS13, SS16, RSS18a, RSS<sup>+</sup>18b]). However, Cosserat continuum is effective under shearing. In the case of pronounced extension or compaction, Cosserat kinematics have no effect and at least a complete first order micromorphic continuum has to be used or, alternatively, a second gradient model (restrained 1<sup>st</sup>

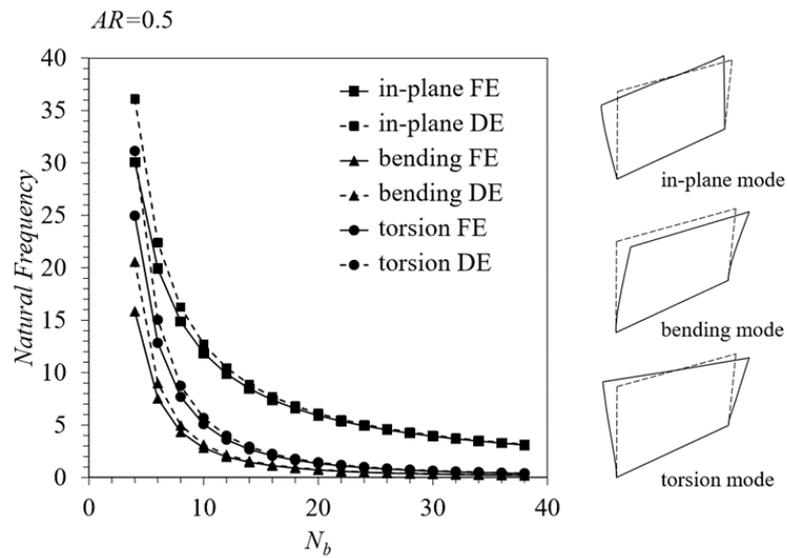


Figure 7: Modal frequencies of a masonry panel versus the number of building blocks: comparison between the results extracted by the Discrete Element Method and by the use of the Cosserat continuum [SSV08, GSSS15].

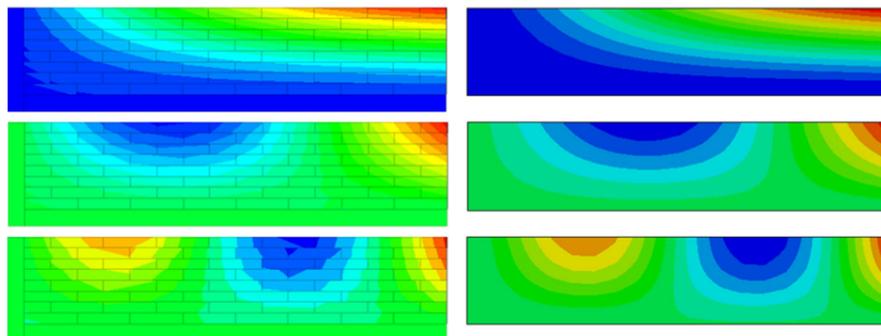


Figure 8: Out-of-plane-displacement contours of the first three flexural modal shapes. Left: Discrete Elements solution. Right: Cosserat Finite Element solution [GSSS14].

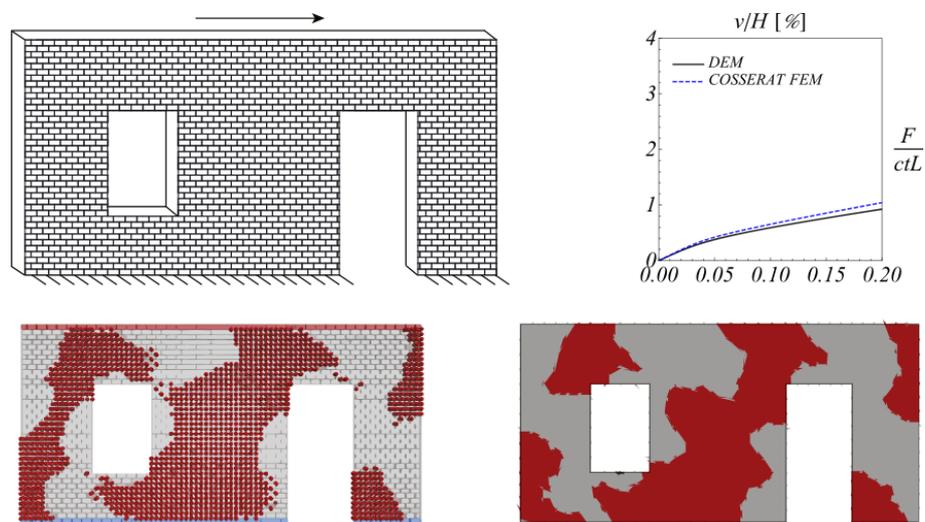


Figure 9: Numerical simulation of a confined masonry panel undergoing shear deformation. Above: tested configuration (left) and normalized force-displacement curves from DEM and Cosserat FEM analyses (right); Below: Comparison between the pattern of plastic deformation obtained from the discrete (left) and Cosserat Finite Element (right) models [GSSS16].

order micromorphic).

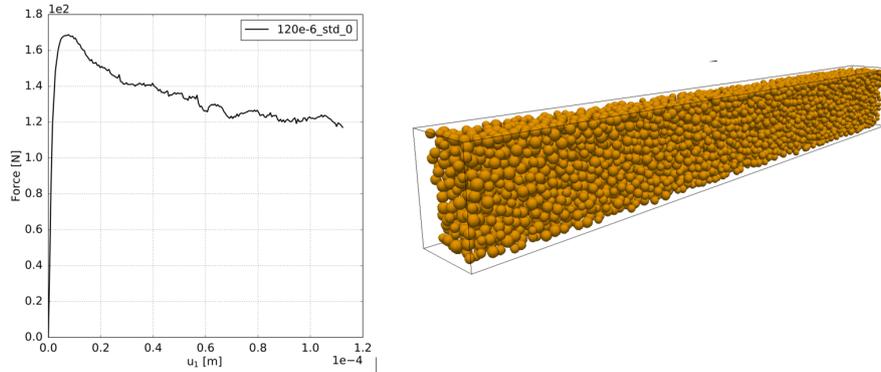


Figure 10: Example of DEM simulation of a granular layer under shearing with constant velocity and shear stress-strain response (courtesy: Efthymios Papachristos).

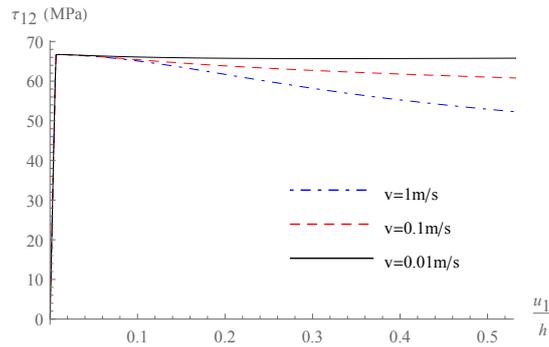


Figure 11: Apparent rate dependency of an infinite layer of a fault gouge under shearing ( $\frac{u_1}{h}$ ), modeled as a Cosserat continuum.  $h$  is the gouge thickness, often of the order of some  $\mu\text{m}$ . The slip rate has a direct impact on the shear stress-strain response, even in the ideal case of perfect plasticity (zero hardening) [RSS18a, RSS<sup>+</sup>18b].

## 5 Finite elements

The principle of virtual power, being global rather than local, is directly amenable to numerical schemes such as the Finite Element Method (FEM). Moreover, it is independent of constitutive laws, providing high degree of generality and abstraction. In this paragraph, we give two simple examples showing its direct use with the Finite El-

ement library FEniCS<sup>3</sup> [ABH<sup>+</sup>15]. We don't get into the details of the finite element method. For a consistent presentation of the method, the interested reader is referred to classical textbooks such as those of [ZT13, Hug00, Bat07].

The problem solved is that of the infinite layer of height  $h$  subjected to shearing, as shown in figure 12. Half of the layer is modeled. Both a Cauchy and a Cosserat continuum is used. Small deformations and linear elasticity is considered. The problems have analytical solutions, but their derivation and the comparison with the numerical solution is left to the reader (hint: a mesh convergence analysis should always be performed).

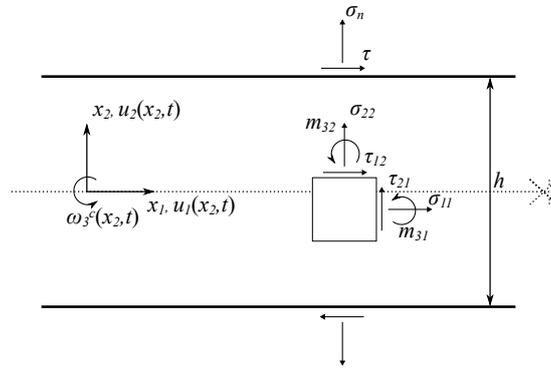


Figure 12: Infinite layer under shearing.

## 5.1 Simple shear with Cauchy continuum

Considering the invariance of the problem in  $x_1$  and  $x_3$  directions (infinite layer) and the symmetry of the stress tensor,  $\sigma_{12} = \sigma_{21}$ , equations (10) become:

$$\begin{aligned} \tilde{p}^{(int)} &= \sigma_{22} \tilde{v}_{2,2} + \sigma_{12} \tilde{v}_{1,2} \\ \tilde{p}^{(ext,t)} &= t_\alpha \tilde{v}_\alpha, \end{aligned} \quad (22)$$

with  $\alpha = 1, 2$  (repeated indices denote summation). The stresses are equal to:

$$\begin{aligned} \sigma_{22} &= M u_{2,2} \\ \sigma_{12} &= G u_{1,2}, \end{aligned} \quad (23)$$

where  $\nu$  is the Poisson ratio,  $E$  the Young's modulus,  $G = \frac{E}{2(1+\nu)}$  the shear modulus,  $M = \frac{E(1-\nu)}{(1-2\nu)(1+\nu)}$  the P-wave modulus and  $u_\alpha$  the displacements (unknowns of the problem).

<sup>3</sup>Installing FEniCS is straightforward in all operating systems: <https://fenicsproject.org>. In Windows we suggest installing the Windows subsystem for Linux: <https://docs.microsoft.com/en-us/windows/wsl/about>.

The principle of virtual power (i.e. the variational/weak form of the problem) together with the above constitutive law and the boundary conditions are directly written in Python in symbolic/mathematical form and the library FEniCS takes on for formulating the finite element problem and solving the linear system of equations, automatically! The Python code is given below, annotated. The numerical results of this simple problem are given in figure 13.

```

1 # Import FEniCS library
2 from dolfin import *
3 # Parameters
4 h=2.           #height of the layer
5 s12_bc=.5     #applied shear stress at boundary
6 s22_bc=.25   #applied normal stress at boundary
7 E_el=1000.   #Young modulus
8 nu_el=0.     #Poisson coefficient
9 ny=3         #number of elements – FE discretization
10 G_el=E_el/(2.*(1+nu_el))
11 M_el=E_el*(1.-nu_el)/((1.-2.*nu_el)*(1.+nu_el))
12
13 # Generate mesh
14 mesh=IntervalMesh(ny,0.,h/2.)
15 # Define element topology (1D)
16 cell=interval
17 # Defines a Lagrangian FE of degree 1 and two unknowns in each node (
    vector)
18 element=VectorElement("Lagrange", cell , degree=1, dim=2)
19 # Assign the element to the mesh
20 V=FunctionSpace(mesh, element)
21
22 # Define test function (virtual velocities)
23 v=TestFunction(V)
24 # Define trial (unknown) function
25 u=TrialFunction(V)
26 # Store the solution to sol
27 sol=Function(V)
28
29 # Define boundary conditions
30 def middle(x, on_boundary):
31     return on_boundary and near(x[0],0.)
32 bc=DirichletBC(V,(0.,0.), middle)
33 # Define traction vector
34 ti=Constant((s22_bc, s12_bc))
35
36 # Define internal virtual power
37 Pint=(
38     M_el*Dx(u[0],0)*Dx(v[0],0) + #sigma_22*v_2,2
39     G_el*Dx(u[1],0)*Dx(v[1],0)   #sigma_12*v_1,2
40 )*dx
41 # Define external virtual power
42 Pext=dot(ti, v)*ds
43
44 # Solve the problem (does the FE formulation, matrix assembly and
    linear solve)
45 solve(Pint == Pext, sol, bc)
46
47 # Plot solution

```

```

48 import matplotlib.pyplot as plt
49 plot(sol[0], label='$u_2$')
50 plot(sol[1], label='$u_1$')
51 plt.xlabel("$x_2$")
52 plt.legend(loc='upper left')
53 plt.ylim(0, .001)
54 font = {'size' : 18}
55 plt.rc('font', **font)
56 plt.show()

```

Listing 1: FEniCS Python code for Cauchy shearing.

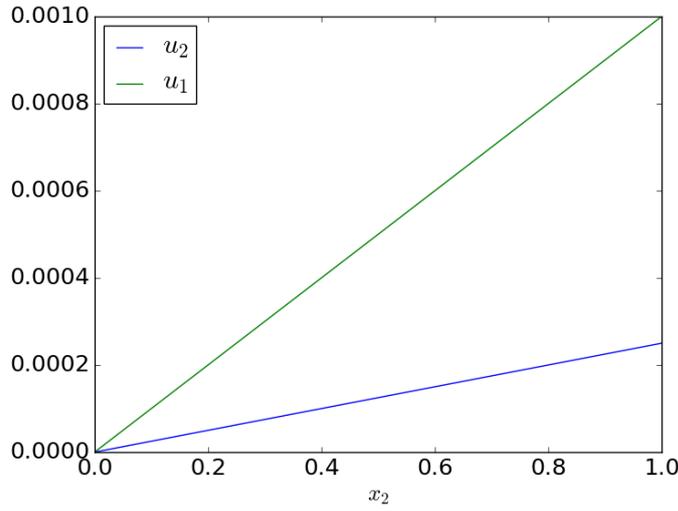


Figure 13: Calculated displacements of the sheared layer using Cauchy continuum.

## 5.2 Simple shear with Cosserat theory

Considering, again, the invariance of the problem in  $x_1$  and  $x_3$  directions equations (20) become:

$$\begin{aligned}
 \tilde{p}^{int} &= \tau_{22}\tilde{\gamma}_{22} + \tau_{12}\tilde{\gamma}_{12} + \tau_{21}\tilde{\gamma}_{21} + m_{32}\tilde{k}_{32} \\
 \tilde{p}^{(ext,t)} &= t_\alpha\tilde{v}_\alpha + \mu_3\tilde{\omega}_3^c
 \end{aligned} \tag{24}$$

As far as it concerns the constitutive law, the symmetric part of  $\tau_{ij}$  is equal to the Cauchy stress tensor, i.e.  $\tau_{(ij)} = \sigma_{ij}$ , and therefore equations (23) can be used. The antisymmetric part of  $\tau_{ij}$  is equal to  $\tau_{[ij]} = 2G_c\gamma_{[ij]}$ , where  $G_c = \eta_1 G$ ; see [Var09]. Regarding the couple stresses,  $m_{32} = M_c k_{32}$ , where  $M_c = 2\eta_3 G l^2$ .  $\eta_1$  and  $\eta_3$  are

material parameters and  $l$  is the Cosserat internal length. Therefore, we have:

$$\begin{aligned}\tau_{22} &= \sigma_{22} = Mu_{2,2} \\ \tau_{12} &= (G + G_c) u_{1,2} + 2G_c \omega_3^c \\ \tau_{21} &= (G - G_c) u_{1,2} - 2G_c \omega_3^c \\ m_{32} &= M_c \omega_{3,2}^c\end{aligned}\tag{25}$$

Similarly to the previous paragraph, the principle of virtual power (i.e. the variational/weak form of the problem) together with the above constitutive law and the boundary conditions are directly written in Python in symbolic/mathematical form and the library FEniCS takes on the rest. Quadratic interpolation was chosen for the displacement and linear for the Cosserat rotation field. The Python code is given below and the numerical results in figure 14.

```

1 # Import FEniCS library
2 from dolfin import *
3 # Parameters
4 h=2.           #height of the layer
5 s12_bc=.5     #applied shear stress at boundary
6 s22_bc=.25   #applied normal stress at boundary
7 m3_bc=.0     #applied Cosserat moment at boundary
8 E_el=1000.   #Young modulus
9 nu_el=0.     #Poisson ratio
10 eta_1=.8    #Cosserat coefficient for Gc
11 eta_3=5./2. #Cosserat coefficient for Mc
12 lc=h/3.    #Cosserat length
13
14 ny=10       #number of elements – FE discretization
15
16 G_el=E_el/(2.*(1+nu_el))
17 M_el=E_el*(1.-nu_el)/((1.-2.*nu_el)*(1.+nu_el))
18 Gc_el=eta_1*G_el
19 Mc_el=2*eta_3*G_el*lc**2
20
21 # Generate mesh
22 mesh=IntervalMesh(ny,0.,h/2.)
23 # Define element topology (1D)
24 cell=interval
25 # Defines a Lagrangian FE of degree 2 for the displacements
26 element_disp=VectorElement("Lagrange", cell, degree=2,dim=2)
27 # Defines a Lagrangian FE of degree 1 for the rotations
28 element_rot=FiniteElement("Lagrange", cell, degree=1)
29 # Creates a mixed element
30 element=element_disp*element_rot
31 # Assign the element to the mesh
32 V=FunctionSpace(mesh, element)
33 # Define test functions (virtual velocities)
34 v=TestFunction(V)
35 # Define trial functions (unknown displacements and Cos. rotation)
36 u=TrialFunction(V)
37
38 # Store the solution to sol
39 sol=Function(V)
40

```

```

41 # Define boundary conditions
42 def middle(x, on_boundary):
43     return on_boundary and near(x[0],0.)
44 bc=DirichletBC(V,(0.,0.,0.),middle)
45 # Define traction vector
46 timui=Constant((s22_bc, s12_bc, m3_bc))
47
48 # Define internal virtual power
49 Pint=(
50     M_e1*Dx(u[0],0)*Dx(v[0],0) + #
51     tau_22*gamma_22 #
52     ((G_e1+Gc_e1)*Dx(u[1],0)+2*Gc_e1*u[2])*(Dx(v[1],0)+v[2]) + #
53     tau_12*gamma_12 #
54     ((G_e1-Gc_e1)*Dx(u[1],0)-2*Gc_e1*u[2])*(-v[2]) + #
55     tau_21*gamma_21 #
56     Mc_e1*Dx(u[2],0)*Dx(v[2],0) #
57     m_32*k_32 #
58     )*dx
59 # Define external virtual power
60 Pext=dot(timui,v)*ds
61
62
63 # Solve the problem (does the FE formulation, matrix assembly and
64 # linear solve)
65 solve(Pint == Pext, sol, bc)
66
67 # Plot solution
68 import matplotlib.pyplot as plt
69 plot(sol[0], label='$u_2$')
70 plot(sol[1], label='$u_1$')
71 plot(-sol[2], label='$-\omega_3$')
72 plt.xlabel("$x_2$")
73 plt.legend(loc='upper left')
74 plt.ylim(0, .001)
75 font = {'size' : 18}
76 plt.rc('font', **font)
77 plt.show()

```

Listing 2: FEniCS Python code for Cosserat shearing.

## 6 Summary

The target of the present chapter was to give the basic ideas and intuition behind the principle of virtual power. A short historical review was made, aiming at clarifying the fundamental ideas of the principle and its connection with the equilibrium equations. After providing the statement of the principle, several examples were presented for showing its application to simple problems involving discrete systems of one and several degrees of freedom. The equivalence of the principle with the equilibrium equations was shown.

Focus was placed then on continuum systems and the generalization of the principle for deriving the differential equilibrium equations of the Cauchy continuum. The

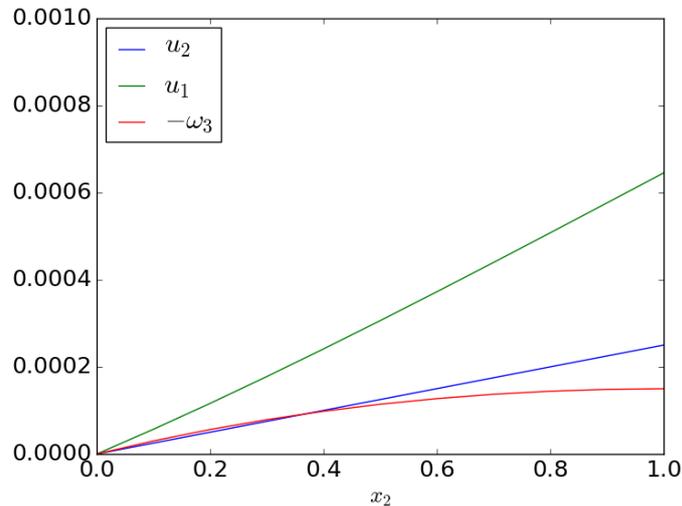


Figure 14: Calculated displacements of the sheared layer using Cosserat continuum.

principle of the virtual power provides a systematic and rigorous way for going further and deriving the equations of more advanced continuum theories as well, i.e. the micromorphic continuum theory. The equations of the micromorphic theory were presented in a general form and the hierarchical structure of the theory was illustrated. It was shown that Cosserat and strain gradient theories are special cases of this more general framework. Some applications were also given for showing the advantages of these continuum theories and the use of the principle of virtual power for upscaling.

The principle of virtual power, being global rather than local, is directly amenable to numerical schemes such as the Finite Element Method. The application of the method provides directly the weak/variational form of the equilibrium equations, which is the starting point in any finite element formulation. Nowadays, several codes exist that allow the user performing FEM analyses by simply entering the variational form of the mathematical problem that (s)he wants to solve. One of these is FEniCS open source FEM library. The example of shearing of an infinite layer modeled as a Cauchy and a Cosserat continuum was presented for showing the FEniCS finite element implementation and the use of the principle of virtual power.

## **Acknowledgements**

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

## References

- [ABG09] JL Auriault, C Boutin, and Christian Geindreau. *Homogenization of Coupled Phenomena in Heterogenous Media*. Wiley, New York, 2009.
- [ABH<sup>+</sup>15] Martin Alnæs, Jan Blechta, Johan Hake, August Johansson, Benjamin Kehlet, Anders Logg, Chris Richardson, Johannes Ring, Marie Rognes, and Garth Wells. The fenics project version 1.5. *Archive of Numerical Software*, 3(100), 2015.
- [AL94] W. B. Anderson and R. S. Lakes. Size effects due to Cosserat elasticity and surface damage in closed-cell polymethacrylimide foam. *Journal of Materials Science*, 29(24):6413–6419, 1994.
- [Bat07] Klaus-Jürgen Bathe. *Finite Element Procedures*. Klaus-Jurgen Bathe, 2007.
- [BG12] Andrea Bacigalupo and Luigi Gambarotta. Computational two-scale homogenization of periodic masonry: Characteristic lengths and dispersive waves. *Computer Methods in Applied Mechanics and Engineering*, 213-216:16–28, mar 2012.
- [BP89] N Bakhvalov and G. Panasenko. *Homogenisation: Averaging Processes in Periodic Media: Mathematical Problems in the Mechanics of Composite Materials*, 1989.
- [BV01] J. P. Bardet and Ioannis Vardoulakis. The asymmetry of stress in granular media. *International Journal of Solids and Structures*, 38(2):353–367, jan 2001.
- [CCC06] F. Collin, R. Chambon, and R. Charlier. A finite element method for poro mechanical modelling of geotechnical problems using local second gradient models. *International Journal for Numerical Methods in Engineering*, 65(11):1749–1772, mar 2006.
- [CCE98] R. Chambon, D. Caillerie, and N. El Hassan. One-dimensional localisation studied with a second grade model. *European Journal of Mechanics, A/Solids*, 17(4):637–656, 1998.
- [Cha10] Nicolas Charalambakis. Homogenization Techniques and Micromechanics. A Survey and Perspectives. *Applied Mechanics Reviews*, 63(3):030803, 2010.
- [CN84] Denis Caillerie and J. C. Nedelec. Thin elastic and periodic plates. *Mathematical Methods in the Applied Sciences*, 6(1):159–191, 1984.
- [DAD<sup>+</sup>17] Jacques Desrues, A Argilaga, S. Dal Pont, G Combe, Denis Caillerie, and T. kein Nguyen. Restoring Mesh Independency in FEM-DEM Multi-scale Modelling of Strain Localization Using Second Gradient Regularization. In E. Papamichos, P. Papanastasiou, E. Pasternak, and

- A. Dyskin, editors, *Bifurcation and Degradation of Geomaterials with Engineering Applications. IWBDG 2017. Springer Series in Geomechanics and Geoengineering*, pages 453–457. Springer, Cham, 2017.
- [DSMP93] R. De Borst, Lambertus J Sluys, H.B. Mühlhaus, and J. Pamin. Fundamental issues in finite element analyses of localization of deformation. *Engineering Computations*, 10(2):99–121, feb 1993.
- [Eri99] A. Cemal Eringen. *Microcontinuum Field Theories*. Springer Verlag, 1999.
- [FPS01] Samuel Forest, Francis Pradel, and Karam Sab. Asymptotic analysis of heterogeneous Cosserat media. *International Journal of Solids and Structures*, 38(26-27):4585–4608, jul 2001.
- [Fré02] Michel Frémond. *Non-Smooth Thermomechanics*. Springer-Verlag, Berlin Heidelberg, 2002.
- [FS98] Samuel Forest and Karam Sab. Cosserat overall modeling of heterogeneous materials. *Mechanics Research Communications*, 25(4):449–454, jul 1998.
- [Fun65] Y. C. Fung. *Foundations of Solid Mechanics*. Prentice Hall, Canada, 1965.
- [GDL98] C. Geuzaine, P. Dular, and W. Legros. A general environment for the treatment of discrete problems and its application to coupled finite element and boundary integral methods. In *Proceedings of the 8th International IGTE Symposium on Numerical Field Calculation in Electrical Engineering*, Graz (Austria), September 1998.
- [Ger73] P. Germain. The Method of Virtual Power in Continuum Mechanics. Part 2: Microstructure. *SIAM Journal on Applied Mathematics*, 25(3):556–575, nov 1973.
- [GFA10] Morton E. Gurtin, Eliot Fried, and Lallit Anand. *The Mechanics and Thermodynamics of continua*. Cambridge University Press, 2010.
- [GSS17] Michele Godio, Ioannis Stefanou, and Karam Sab. Effects of the dilatancy of joints and of the size of the building blocks on the mechanical behavior of masonry structures. *Meccanica*, may 2017.
- [GSSS14] Michele Godio, Ioannis Stefanou, Karam Sab, and Jean Sulem. Cosserat Elastoplastic Finite Elements for Masonry Structures. *Key Engineering Materials*, 624:131–138, 2014.
- [GSSS15] Michele Godio, Ioannis Stefanou, Karam Sab, and Jean Sulem. Dynamic finite element formulation for Cosserat elastic plates. *International Journal for Numerical Methods in Engineering*, 101(13):992–1018, mar 2015.

- [GSSS16] Michele Godio, Ioannis Stefanou, Karam Sab, and Jean Sulem. Multi-surface plasticity for Cosserat materials: Plate element implementation and validation. *International Journal for Numerical Methods in Engineering*, 108(5):456–484, nov 2016.
- [Hec12] F. Hecht. New development in freefem++. *J. Numer. Math.*, 20(3-4):251–265, 2012.
- [Hug00] Thomas J. R. Hughes. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover Publications, 2000.
- [KABC08] P. Kotronis, S. Al Holo, P. Bésuelle, and R. Chambon. Shear softening and localization: Modelling the evolution of the width of the shear zone. *Acta Geotechnica*, 3(2):85–97, mar 2008.
- [Lag88] Joseph-Louis Lagrange. *Méchanique Analytique*. Jacques Gabay, Paris, 1788.
- [Mau14] Gérard A. Maugin. *Continuum Mechanics Through the Eighteenth and Nineteenth Centuries*, volume 214. 2014.
- [Min64] R.D. Mindlin. Micro-structure in linear elasticity. *Archive for Rational Mechanics and Analysis*, 16(1), 1964.
- [PdCOTD09] J. Pinho-da Cruz, J.a. Oliveira, and F. Teixeira-Dias. Asymptotic homogenisation in linear elasticity. Part I: Mathematical formulation and finite element modelling. *Computational Materials Science*, 45(4):1073–1080, jun 2009.
- [PZ16] Panos C. Papanastasiou and Antonios Zervos. Numerical modelling of strain localization. In *Modelling of instabilities and bifurcation in Geomechanics, ALERT geomaterials doctoral school*, 2016.
- [RC16] Roozbeh Rezakhani and Gianluca Cusatis. Asymptotic expansion homogenization of discrete fine-scale models with rotational degrees of freedom for the simulation of quasi-brittle materials. *Journal of the Mechanics and Physics of Solids*, 88:320–345, 2016.
- [RSS18a] Hadrien Rattiez, Ioannis Stefanou, and Jean Sulem. The importance of Thermo-Hydro-Mechanical couplings and microstructure to strain localization in 3D continua with application to seismic faults. Part I: Theory and linear stability analysis. *Journal of the Mechanics and Physics of Solids*, 115:54–76, jun 2018.
- [RSS<sup>+</sup>18b] Hadrien Rattiez, Ioannis Stefanou, Jean Sulem, Manolis Veveakis, and Thomas Poulet. The importance of Thermo-Hydro-Mechanical couplings and microstructure to strain localization in 3D continua with application to seismic faults. Part II: Numerical implementation and post-bifurcation analysis. *Journal of the Mechanics and Physics of Solids*, 115:1–29, jun 2018.

- [SP80] E. Sanchez-Palencia. Non homogeneous media and vibration theory. In *Lecture notes in physics*. 1980.
- [SP86] E Sánchez-Palencia. Homogenization in mechanics. A survey of solved and open problems. *Rend. Sem. Mat. Univ. . . .*, 44:1–46, 1986.
- [SS16] Jean Sulem and Ioannis Stefanou. Thermal and chemical effects in shear and compaction bands. *Geomechanics for Energy and the Environment*, 6:4–21, jun 2016.
- [SSV08] Ioannis Stefanou, Jean Sulem, and Ioannis Vardoulakis. Three-dimensional Cosserat homogenization of masonry structures: elasticity. *Acta Geotechnica*, 3(1):71–83, feb 2008.
- [SSV10] Ioannis Stefanou, Jean Sulem, and Ioannis Vardoulakis. Homogenization of interlocking masonry structures using a generalized differential expansion technique. *International Journal of Solids and Structures*, 47(11-12):1522–1536, jun 2010.
- [SSV11] Jean Sulem, Ioannis Stefanou, and Emmanuil Veveakis. Stability analysis of undrained adiabatic shearing of a rock layer with Cosserat microstructure. *Granular Matter*, 13(3):261–268, feb 2011.
- [TC97] H Tollenaere and Denis Caillerie. Continuous modeling of lattice structures by homogenization. *Advances in Engineering Software*, pages 699–705, 1997.
- [Var09] Ioannis Vardoulakis. *Lecture notes on Cosserat continuum mechanics with application to the mechanics of granular media*. <http://geolab.mechan.ntua.gr/teaching/lectnotes/CCM2009.pdf>, 2009.
- [VSS13] Emmanuil Veveakis, Ioannis Stefanou, and Jean Sulem. Failure in shear bands for granular materials: thermo-hydro-chemo-mechanical effects. *Géotechnique Letters*, 3(April-June):31–36, may 2013.
- [ZPV01] Antonios Zervos, Panos C. Papanastasiou, and Ioannis Vardoulakis. A finite element displacement formulation for gradient elastoplasticity. *International Journal for Numerical Methods in Engineering*, 50(6):1369–1388, 2001.
- [ZT13] Olek C Zienkiewicz and Robert L. Taylor. *The Finite Element Method*. Butterworth Heinemann, 2013.



---

# Energetics in Discrete Element Modelling

**François Guillard**

*School of Civil Engineering, The University of Sydney*

---

*Discrete Element Modelling (DEM) is a widely used simulation method that provides direct insight into the behavior and internal properties of particulate materials, by simulating individual particles interacting through a chosen contact model. Here we discuss the use of DEM to investigate the energy transfers occurring in a granular material. A multiscale approach will be followed, where the energy balance of the entire system is first studied, before describing the grain-level mechanisms that are able to store or dissipate energy. Finally, the coarse-graining method, allowing mesoscopic local measurements of the granular properties, will be presented. The continuum fields extracted in this way could be directly compared with model predictions. The relevance of the various results will be illustrated with DEM simulations of a granular medium undergoing simple shear.*

## 1 Introduction

Discrete Element Modelling (DEM) is ubiquitous to study numerically the behavior of particulate materials, by simulating the displacement of every individual particle of a medium. This method has been widely used in the industry owing to the omnipresence of granular materials, for example to simulate hopper flows [GE05], mixing of pharmaceuticals [KaEH09], and fluidised beds [TKT93]. DEM is also used to investigate the properties of geological and geotechnical systems such as rockfalls [NCM01], and soil anchors [VC04].

The versatility and wide applicability of DEM has led to its use in most research and industrial fields dealing with particulate materials, and beyond [LP91, KH93]. However, several limitations remain that should not be overlooked. DEM relies heavily on computing power, and therefore is limited to the simulation of small systems, for short physical time, and typically with softer particles than in reality. More fundamentally, the modelling assumptions for the contact laws between the particles are controlling the dynamics of the system, and therefore have to be chosen very carefully.

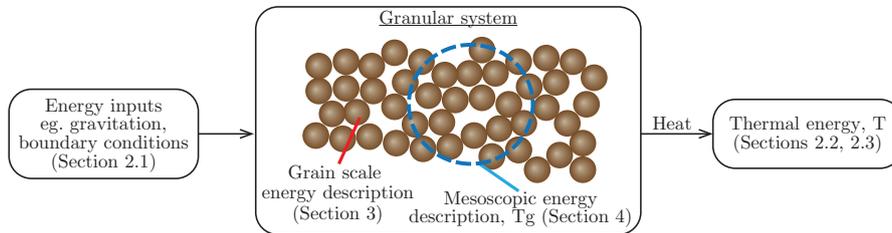


Figure 1: Energy transfer from the macroscopic to the microscopic scale in granular materials, with reference to the relevant sections in this chapter.

The continuum modelling of materials allows to overcome these limitations, by providing constitutive laws that apply to the material seen as a continuum, and which do not require the knowledge of the microscopic details at grain level. Such models have been very successful to explain and predict the behavior of granular media [JFP06, HK13] or soils [MD97, Ein07] for example, while providing a local continuum description of the material and insights into its behavior. However, the physical foundations of some assumptions can sometimes be questioned, and the upscaling from the particle level to the continuum level is not straightforward. The comparison of continuum models with experimental results and simulation outcomes is therefore crucial to assess the relevance of the proposed laws, and DEM provides invaluable information at grain level that are challenging to obtain experimentally.

As models develop toward a deeper understanding of the thermodynamics of granular material [JL09, JEL17], following the success at the atom scale of statistical physics and classical thermodynamics, it seems appropriate to investigate the energetic properties of granular media in DEM. This should provide a better understanding of the exact assumptions made when a DEM simulation is performed, and develop the appropriate tools to allow comparison of discrete numerical results with continuum modelling predictions.

Figure 1 provides a graphical summary of the content of this chapter. Section 2 will be dedicated to the energy balance at the global scale for a granular system, and the transfer of energy to heat. Section 3 will delve into the microscopic properties of grains in DEM, and the energy storage, transfer and dissipation in such system. Finally, section 4 will bridge the scales and provide the tools to extract continuum mechanical and energetic fields representing the material at an intermediate scale between the grain size and the full system, suitable for comparison with constitutive models.

## 2 Macroscopic energy balance

### 2.1 Energy input

Granular materials are frictional materials, which means that a large part of their macroscopic behaviour is governed by the fact that the interaction between the individual grains is frictional. Consequently, they are dissipative materials and cannot sustain motion without constant energy input. Figure 1 shows that energy is first injected to the granular system by the boundaries or by external forces such as gravity, and, if the power provided is large enough, the grains move and reorganise. The energy of the motion is then dissipated in heat through the friction between the grains, which transforms the grain kinetic energy to thermal energy, effectively increasing the temperature of the material.

In some situations it is quite straightforward to estimate the energy injected to a granular material. Two examples are given below.

**Simple shear** In this situation a granular material of thickness  $H$  and depth and width  $L$ , is subjected to shearing by a top plate moving at velocity  $V$ , applying a pressure  $P$ , without gravity, a configuration similar to the one discussed in section 3.5. The power injected by the top plate to the grains is  $\mathcal{P} = FVdt$  with  $F$  the horizontal force applied by the top plate. To reach a steady state, this power has to be dissipated by the material. Let us assume a simple Coulomb friction law, with the a local viscosity equal to  $\mu P$  with  $\mu$  the friction coefficient and  $P$  the pressure. The total power dissipated in the material is  $\mathcal{P}_g = L^2 \int_0^H \mu P \frac{\partial v}{\partial z} dz dt = \mu P L^2 \frac{V}{H}$  assuming a linear velocity profile through the material (simple shear), leading to  $F = \mu P L^2$ . Therefore, by applying a given  $F$  and  $P$  to a granular material, the only way to reach a steady state is for the material to adapt its internal friction to dissipate enough energy.

**Flow down an inclined plane** A granular material flows down a plane inclined with an angle  $\theta$  from the horizontal. In that case, energy is injected to the material from the change in gravitational potential:  $\mathcal{P} = mg \langle v \rangle \sin \theta dt$  where  $\langle v \rangle$  is the average velocity of the material across the depth,  $m$  the mass of material on a certain slice of width  $dx$  and height  $h$ ,  $g$  the acceleration of gravity. The dissipation can be calculated as  $\mathcal{P}_g = \int_0^h dx \mu(z) P(z) \frac{\partial v}{\partial z} dt dz$ . Assuming a constant friction coefficient  $\mu$  across the depth and a hydrostatic pressure profile gives  $\mathcal{P}_g = dx \mu \rho g \cos \theta dt \langle v \rangle h$ . By equating the injected and dissipated energy,  $\mu = \tan \theta$ .

Since experimentally the friction coefficient is bounded by  $\mu_s < \mu < \mu_2$  for the dense flow regime [MiD04], with  $\mu_s$  the static friction coefficient, and  $\mu_2$  the maximum friction at high shear rate, there is a limited range of energy injection rates, ie. angle, where granular material can sustain steady flow down an incline [AFP13]. If the power is lower than the minimum power, the medium will dissipate all the energy

through elastic wave absorption or localisation of deformation in shear bands. If the rate of energy input is too high, the flow accelerates and eventually enters a different regime of behavior, eg. collisional or gaseous.

## 2.2 Heat

The energy dissipated by the friction does not disappear but is converted in thermal energy, which increases the temperature of the material. Most of the time, in experiment and in DEM, the temperature increase is not considered. However, if the power provided to the system is high enough, the temperature increase cannot be ignored, as the material properties of the grains change, and may even lead to the partial melting of the grains if the temperature reached is too high [VVDT07, ERMS18]. The internal motion of the grains, and the diffusion of heat between grains and through the contacts, can generate additional spatial inhomogeneities in the material properties, whose relevance should be assessed.

## 2.3 Neglecting thermal fluctuations

Although the increase of thermal energy due to the friction can modify the mechanical properties of the grains, direct conversion of thermal energy in kinetic energy is never considered, as it is always negligible in granular materials (ie. granular media are athermal). To demonstrate that, we can calculate the kinetic energy of the grains due to their temperature using the Boltzmann equation. Each degree of freedom contributes with a factor  $\frac{1}{2}k_B T$  to the kinetic energy, with  $T$  the temperature and  $k_B = 1.410^{-23} \text{ J.K}^{-1}$  the Boltzmann's constant. This leads to  $\frac{1}{2}m^i v^{i2} = \frac{3}{2}k_B T$  for a grain  $i$  of mass  $m^i$  and velocity  $v^i$ , which translates to a velocity of  $3.5 \cdot 10^{-8} \text{ m.s}^{-1}$  for 1 mm glass grains. Consequently, for grains larger than a few microns in diameter, the grain motion due to their temperature is completely negligible.

Note that this temperature effect is the motion of the grains due to *thermal* temperature  $T$ , which is negligible. This is very different to the concept of *granular* temperature  $T_g$ , which is the kinetic energy of the fluctuating motion of the grains from the bulk flow, and is in general far from being negligible. There are different ways to define the granular temperature [JS83, JL09, BP10], but the simplest one is to use the kinetic theory of granular gases,  $T_g = \langle \tilde{\mathbf{v}}^{i2} \rangle$  where  $\tilde{\mathbf{v}}^i$  is the velocity fluctuation of particle  $i$  with respect to the bulk flow, and  $\langle \rangle$  is the average over all particles.

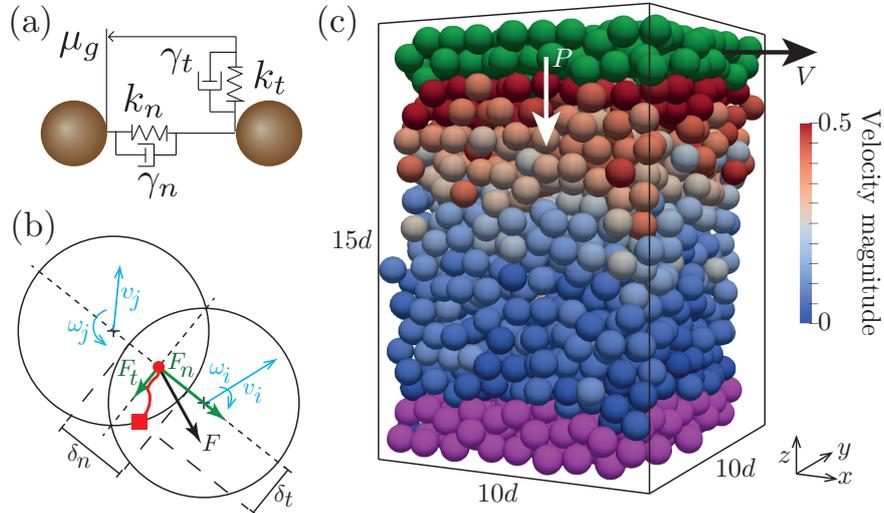


Figure 2: (a) Example of normal and tangential contact law used in DEM. (b) Definition and notation of the particle kinematic properties and contact properties. The red line is the path followed by the contact point since the initiation of the contact (red square). (c) Simple shear flow DEM configuration. The purple grains at the bottom are fixed, the green grains at the top move as rigid body with a horizontal velocity  $V$ , and apply a constant pressure  $P$ . The grains are coloured by their instantaneous velocity.

### 3 Grain-level energy balance

#### 3.1 Particulate description in DEM

DEM is based on the simulation of every individual particle. At each time-step, the forces on the particles are computed, and the equations of motion for the momentum and angular momentum are integrated in time, for example using a Verlet algorithm. The forces on a particle have two sources. The first contribution comes from the external forces such as gravity or fluid drag force. The second contribution emerges from the contacts between particles, which are solved based on a particular contact law. The choice of the contact law determines the type of material that is simulated, since DEM is used to simulate atoms as well as granular materials. We are focussing on the simulation of particulate media with frictional interaction in the following [Lud08, SEG<sup>+</sup>01]. The contact law is used to relate the positions of two particles in contact at a given time to the force that these particles apply on each other (Figure 2b).

Here we are considering a normal contact force between two grains idealised as the

sum of a spring contribution, representing the contribution of the elastic deformation of the grain material, and a damper to account of all the dissipation mechanisms during a collision (figure 2a). The normal force follows a Hertz law in three dimension [Her82, CS79], and a Hooke law (linear spring) in two dimensions<sup>1</sup>, leading to a dependency on the normal interpenetration  $\delta_n$  with a power 3/2 for the former and 1 for the latter.

The normal force  $F_n$  is therefore given by:

$$F_n = k_n \delta_n^{3/2} + \gamma_n v_n \quad \text{for Hertz contacts,} \quad (1)$$

$$F_n = k_n \delta_n + \gamma_n v_n \quad \text{for Hooke contacts,} \quad (2)$$

where  $v_n$  is the relative normal velocity between the particles in contact (cf. equation 3), and  $k_n$  and  $\gamma_n$  expressed from the material properties are given in table 1 for identical particles in contact. The relative velocity between two particles  $i$  and  $j$ , with  $\mathbf{r}^i$  and  $\mathbf{r}^j$  the vector between their respective center and the contact point, velocities  $\mathbf{v}^i$  and  $\mathbf{v}^j$ , and angular velocities  $\boldsymbol{\omega}^i$  and  $\boldsymbol{\omega}^j$  (figure 2b) is given by:

$$\mathbf{v}_{\text{rel}} = \mathbf{v}^i + \mathbf{r}^i \times \boldsymbol{\omega}^i - \mathbf{v}^j - \mathbf{r}^j \times \boldsymbol{\omega}^j \quad (3)$$

with  $\times$  the vector cross product. This allows to define  $v_n = \mathbf{v}_{\text{rel}} \cdot \mathbf{n}$  the normal relative velocity, with  $\mathbf{n}$  the vector normal to the contact; and  $\mathbf{v}_t = v_t \mathbf{t} = \mathbf{v}_{\text{rel}} - v_n \mathbf{n}$  the tangential relative velocity.

A large part of the richness and complexity of granular materials emerge from the frictional forces between the particles, that create local rotations of the grains. Normal forces between grains only affect the translational motion, while tangential forces modify both the translation and rotation of the grains *via* the torques that they generate. The tangential contact law is usually modelled by a spring-dashpot-slider law (figure 2a). The tangential force  $F_t$  increases with the tangential interpenetration – the tangential displacement since the beginning of the contact (figure 2b) – and dissipates energy from the tangential viscosity, until the force reaches the Coulomb friction threshold which caps it to  $\mu_g F_n$ , where  $\mu_g$  is the friction coefficient.

$$\begin{aligned} F_t &= k_t \delta_t + \gamma_t v_t & \text{if } |F_t| < \mu_g |F_n| \\ &= \mu_g |F_n| & \text{otherwise.} \end{aligned} \quad (4)$$

The coefficients  $k_t$ ,  $\gamma_t$  and  $\mu_g$  are given in table 1.

The normal and tangential contact laws covered above are the basis of almost every DEM simulation in the context of granular materials. Obviously, more advanced laws can be implemented to better describe the interactions between the particles [BM14]. For example, an additional short range attractive force can be added to simulate cohesive particles, or long range interactions can be integrated to take into account electrostatic effects. Common extensions of the contact law include the addition of rolling or twisting friction, which are torque-only contributions arising from the difference in relative rotational velocity between the particles in contact.

<sup>1</sup>Note that for performance purposes the Hooke law can also be used in 3D, with little effect on the qualitative behaviour of the system

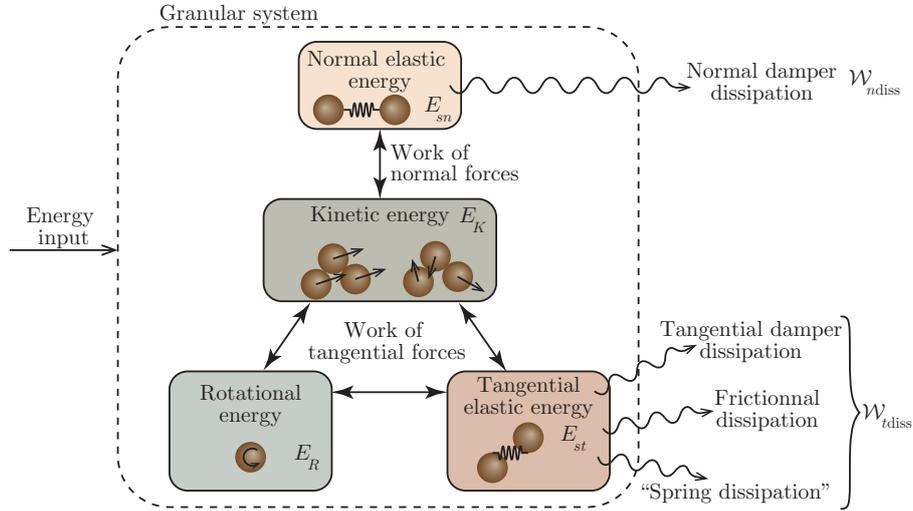


Figure 3: Conserved energies and transfers at the microscopic scale.

### 3.2 Conserved energies

Figure 3 gives a schematic view of the energy balance at the microscopic level in a granular material as described in DEM simulations. The energy in the material is stored in four different ways.

**Kinetic energy** The translational motion of the particles stores the kinetic energy of the system. For a given particle  $i$  of mass  $m^i$  and velocity  $\mathbf{v}^i$ , the kinetic energy  $E_K^i$  is classically given by  $E_K^i = \frac{1}{2} m^i |\mathbf{v}^i|^2$ .

**Rotational energy** The rotation of every particle also stores energy in the motion. Although this form of energy is not very often thought as a means of storage, it makes a non negligible contribution to the total energy stored in the granular system. The expression for this energy is very similar to the definition of the kinetic energy, where the mass is replaced by the moment of inertia of the particle, and the translational velocity by the rotational velocity. For a spherical particle<sup>2</sup> with a moment of inertia  $I^i$  and a rotational velocity  $\boldsymbol{\omega}^i$ , the rotational energy  $E_R^i$  is  $E_R^i = \frac{1}{2} I^i |\boldsymbol{\omega}^i|^2$ .

**Elastic energy** The spring components of the contact forces are also able to store and release energy. The normal and tangential springs have quite different properties

<sup>2</sup>Non spherical particles have a moment of inertia which is a second order tensor,  $\mathbf{I}$ . For such particles, the rotational energy is given by  $E_R^i = \frac{1}{2} \boldsymbol{\omega}^i \cdot \mathbf{I}^i \cdot \boldsymbol{\omega}^i$

Table 1: Coefficient of Hertzian and Hookean contact law expressed from the material properties of two identical particles in contact, and values used for the simulations section 3.5 (units are SI).  $E$  the Young's modulus,  $\nu$  the Poisson's ration,  $m$  the mass,  $e$  restitution coefficient,  $v$  the characteristic velocity,  $\beta = \ln(e)/\sqrt{\ln^2(e) + \pi^2}$ .

	Hertz	Hooke	Simulations: Hooke
$F_n$	$k_n \delta_n^{3/2} + \gamma_n v_n$	$k_n \delta_n + \gamma_n v_n$	
$F_t,  F_t  < \mu_g  F_n $	$k_t \delta_t + \gamma_t v_t$	$k_t \delta_t + \gamma_t v_t$	
$k_n$	$\frac{E}{3(1-\nu^2)} \sqrt{d}$	$\left( \frac{4E\sqrt{d}}{15(1-\nu^2)} \right)^{\frac{4}{5}} (mv)^{\frac{1}{5}}$	$8 \cdot 10^5$
$k_t$	$\frac{E}{(2-\nu)(1+\nu)} \sqrt{d}$		$2.2 \cdot 10^5$
$\gamma_n$	$-\beta \sqrt{\frac{5Em}{6(1-\nu^2)}} \sqrt[4]{d\delta_n}$	$\beta \sqrt{2mk_n}$	50
$\gamma_t$	$-\beta \sqrt{\frac{20E}{6(2-\nu)(1+\nu)}} \sqrt[4]{d\delta_n}$		50
$\mu_g$			0.5

though, as the tangential elastic component is non-conservative and is able to dissipate energy, unlike idealised springs. This effect will be discussed in detail in section 3.4. On the other hand, the normal spring component is a regular conservative force, albeit non-linear in the case of Hertzian contact. A force is conservative when the work it produces does not depends on the path of the displacement. Such a force derives from a potential, ie. there is  $\phi$ , a scalar function of space, such that  $F_n = \nabla \phi$ , with  $\nabla$  the spatial gradient. If one expresses the gradient function attached to the contact, the potential function of the normal force is therefore  $\frac{1}{2} k_n \delta_n^2$  for the linear hookean spring contact, and  $\frac{2}{5} k_n \delta_n^{5/2}$  for an Hertzian spring. Therefore, both cases of a linear or non-linear normal contact spring lead to the ability to store and release energy in the contact.

The case of the tangential spring energy storage is more intricate and will be discussed in detail in section 3.4. For the purpose of simplifying the discussion of the energy balance at grain level we will consider linear tangential springs that store an energy  $E_{st}^{ij} = \frac{1}{2} k_t \delta_t^2$  for the contact between particles  $i$  and  $j$ . To be coherent we also limit ourselves to Hookean normal contact storing an energy  $E_{sn}^{ij} = \frac{1}{2} k_n \delta_n^2$ .

Since the energy is extensive, it is straightforward to go from a single particle or a single contact energy to the energy stored in the full system for a given conserved energy type, by summing over the particles or the contacts:

$$E_K = \sum_{i=1}^N E_K^i ; E_R = \sum_{i=1}^N E_R^i ; E_{sn} = \sum_{i=1}^N \sum_{j>i}^N E_{sn}^{ij} ; E_{st} = \sum_{i=1}^N \sum_{j>i}^N E_{st}^{ij} \quad (5)$$

### 3.3 Energy transfers

The contact forces transfer conserved energy between the different types (kinetic, potential elastic etc.). The work  $\mathcal{W}$  of a contact force  $\mathbf{f}$  over a certain displacement  $l$  is given by  $\mathcal{W} = \int_l \mathbf{f} \cdot d\mathbf{l}$ , which can also be expressed in terms of the time  $T$  needed to achieve the displacement at velocity  $\mathbf{v}(\tau)$ :  $\mathcal{W} = \int_T \mathbf{f} \cdot \mathbf{v} d\tau$ .

As shown in figure 3, the normal forces only affect the translational motion of the particles, and therefore only modify the balance between the kinetic energy and the normal elastic potential energy. The spring and damper components of the normal force can directly be separated, leading to a change in normal elastic energy  $\delta E_{sn}$  and a dissipated energy  $\mathcal{W}_{ndiss}$  during a time  $\delta\tau$ :

$$\delta E_{sn} = \mathcal{W}_{kn} = k_n \delta n v_n \delta\tau \quad (6)$$

$$\mathcal{W}_{ndiss} = -\gamma_n v_n^2 \delta\tau < 0 \quad (7)$$

The work of the tangential forces converts energy between the translational and rotational kinetic energy in the tangential potential elastic energy. However, the dissipative and non-dissipative parts of the work cannot be distinguished as easily as in the case of the normal forces, but the dissipation  $\mathcal{W}_{diss}$  can still be evaluated from the total work of the tangential force  $\mathcal{W}_t$  and the change in tangential elastic energy  $\delta E_{st}$ :

$$\delta E_{st} = \mathcal{W}_t - \mathcal{W}_{tdiss} = F_t v_t \delta\tau - \mathcal{W}_{tdiss} \quad (8)$$

Finally, at the macroscopic scale, the energy  $\mathcal{W}_{in} > 0$  supplied over a time  $\delta t$  to the system is related to the change in energy between the conserved energy types and to the dissipated energy:

$$\mathcal{W}_{in} = \delta E_K + \delta E_R + \delta E_{sn} + \delta E_{st} - \mathcal{W}_{ndiss} - \mathcal{W}_{tdiss} \quad (9)$$

Note that in steady state, on average, the quantity of each type of conserved energy is constant, so all the terms in  $\delta E$  in the equation 9 are zero, leading to the expected expression that the energy supplied to the system is equal to the energy dissipated.

### 3.4 Tangential dissipative mechanisms

As mentioned earlier, the energy dissipation mechanism of the tangential contact can be quite intricate. First of all, in the case of a Hertz contact, the law mentioned in table 1 gives rise to a tangential elastic component proportional to  $\delta_n^{1/2} \delta_t$ . This expression mixes the normal and tangential interpenetration at the contact, and cannot be integrated to extract a potential that would also be compatible with the normal elastic potential. Using the previous law therefore leads to a non-conservative tangential elastic force.

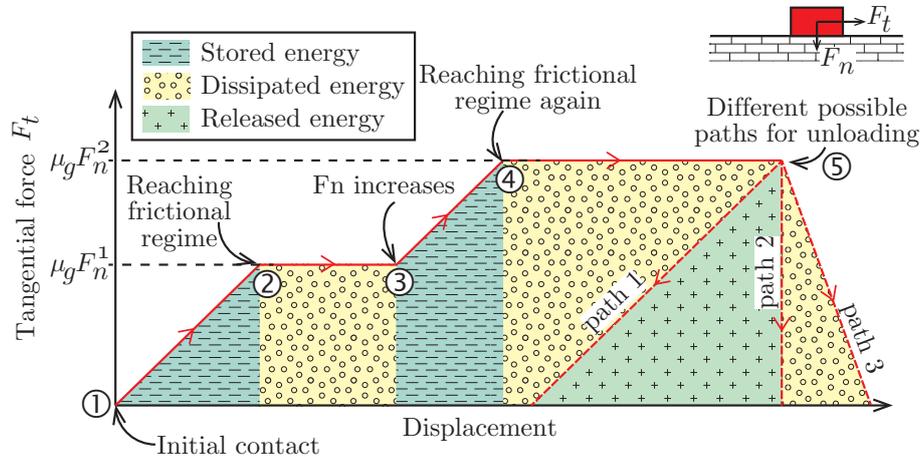


Figure 4: Complex one-dimensional loading and unloading path, and related work, for a block in elastic-frictional contact with a surface.

Moreover, even in the case of a Hooke tangential contact law, where the tangential springs are linear and independent of the normal interpenetration, the energy stored in the elastic contact is not necessarily released to the system when the contact opens, effectively leading to the dissipation of this energy in heat. This effect is better understood with the thought experiment of a block in frictional contact with a plane, with only one direction of motion, as shown in figure 4 (circled numbers in the following refer to the numbers in that figure).

At the beginning of the contact, the block is pressed against the plane, with a normal force  $F_n$ . There is no tangential force yet, as the block has not moved horizontally. As this motion starts ①, the tangential force increases (according to  $k_t$ ), and energy is stored. When the block reaches the Coulomb friction criterion ②, that is  $F_t = \mu_g F_n$ , the force stops growing and all the work of the tangential force is dissipated. If at this point the normal force is increased ③, the block re-enters the elastic regime and the stored energy grows again until the friction criterion is met once again ④. The unloading part of the path is where the complexity arises, because multiple different paths can be followed from ⑤. One unloading path (path 1) is simply to move back the block horizontally while maintaining the normal force constant. In that case all the energy that was stored in the tangential spring during the loading phase is released. On the other hand, if the normal force is decreased while the block is kept static horizontally (path 2), no work is performed by the tangential forces until the contact is broken, effectively corresponding to the full dissipation of the energy that was stored in the tangential spring. An even more extreme situation happens if the block keeps moving horizontally in the same direction as the normal force is reduced (path 3): not only is the stored energy fully lost, but additional dissipation occurs from the frictional forces.

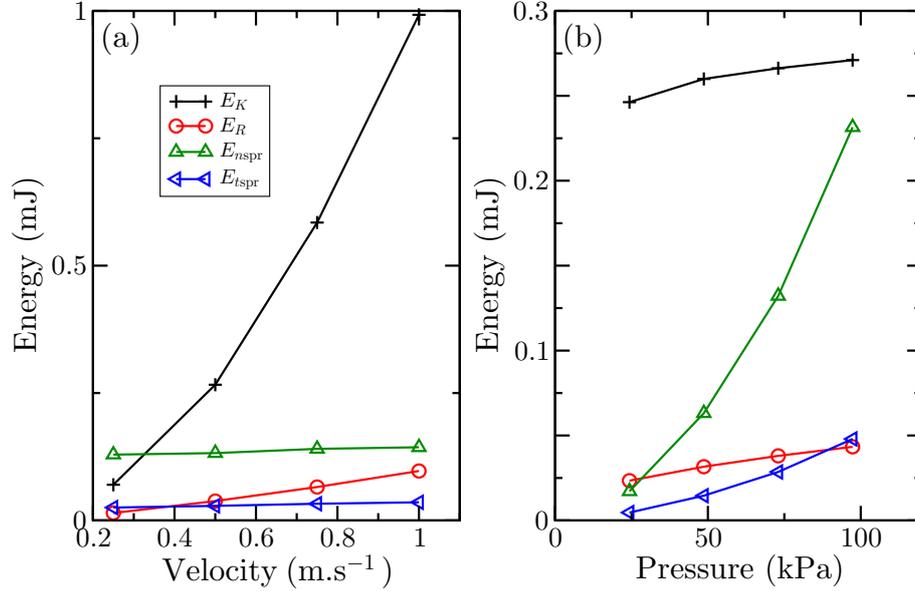


Figure 5: Conserved energies in DEM. (a) Various shearing velocity  $V$  and constant pressure  $P = 73$  kPa. (b) Various pressure  $P$  and constant velocity  $V = 0.5$  m.s<sup>-1</sup>.

Consequently, depending on the unloading path of the contact, the energy stored in the tangential spring can be recovered in full, partially or not at all. The energy stored in the tangential spring is only an upper bound of the energy that can actually be released during the unloading phase.

### 3.5 Example: DEM of three dimensional simple shear

To illustrate the previous discussion we perform some DEM simulations of simple shear using LIGGGHTS [KGGH<sup>+</sup>12]. The simulated configuration, shown in figure 2c, is a collection of  $N = 1300$  grains of diameter  $d = 0.0015$  mm and density  $\rho_g = 2500$  kg.m<sup>-3</sup>. The simulation is periodic in the horizontal directions  $x$  and  $y$ . The grains are sheared from the top by a rough plate moving at constant horizontal velocity  $V$ , which is free to move vertically to apply a constant confining pressure  $P$ . The contact between the particles is following a Hooke contact law, with the parameters given in table 1. The simulations are performed dimensionally to simplify the comparison with physical systems, but could be non-dimensionalised by using  $d$  as a unit length,  $\rho_g/(\pi d^3/6)$  as unit mass, and  $d/V$  as unit time. The time-step used is  $10^{-6}$  s. The inertial number of the system is quite low,  $I = \dot{\gamma}d/\sqrt{P/\rho_g} \simeq 1.2 \cdot 10^{-2}$  with  $\dot{\gamma}$  the shear rate, leading to a constant friction coefficient  $\mu \sim 0.35$  [MiD04].

Figure 5 shows the different types of energy stored in the system at steady state. As

expected, the kinetic energy is proportional to the velocity squared. In simple shear, the velocity is directly related to the shear rate (the symmetric part of the velocity gradient), which is equal to the vorticity (the antisymmetric part of the velocity gradient). Therefore the rotational energy follows the same trend as the kinetic energy, but is quite smaller. The normal elastic energy is related to the pressure applied to the grains and is therefore essentially constant when the pressure is kept constant (figure 5a) but increases with increasing pressure (figure 5b), as the interpenetration between the particles increases to accommodate the higher confining pressure.

## 4 Energetic Coarse-Graining

The previous microscopic energy balance can be used to understand exactly the mechanisms of energy transfer at grain level, or to study the full particulate system. However, the microscopic information is usually not that useful, as the knowledge of every particle energetic properties should not be necessary to describe a granular material made of millions of grains. On the other hand, the summation of the grain contribution to perform the energy balance at the level of the whole system is now too coarse, as all the spatial inhomogeneities and temporal transients have been erased. In the following, we will develop the necessary tools to extract continuum fields from the discrete grain and contact information obtained from DEM. These fields will give important insights into the material response and could be compared to experimental results and continuum modelling outcome.

The process of extracting continuum fields from the particle locations, masses, velocities etc. and from the contact forces is called coarse-graining. It corresponds to the spatial and temporal averaging of the grain and contact properties. However, this procedure is not as straightforward as one might expect, and should be examined in detail. First of all, it is clear that the size of the averaging window will have to be chosen carefully, to average enough grains to obtain a smooth field, while being local enough to reflect the spatial and temporal variations in the system. Secondly, and more subtly, the continuum fields obtained must themselves verify conservation equations to be useful. Since the energy balance involves all the properties of the grains and of the contacts as we have seen in the previous section, the mass, momentum and angular momentum conservation laws will have to be verified by the computed continuum fields as well as the energy balance.

### 4.1 Generalities on the coarse-graining procedure

We first consider the spatial coarse-graining of the grain properties. The coarse-graining procedure typically treats the particles properties as point-like values located at the center of mass of the grain. To get the value of a certain continuum field at a given point  $\mathbf{x}$  in space, all the particle properties are weighted according to their distance from  $\mathbf{x}$  using a suitable window function  $w$ . For example, the density  $\rho$  at point

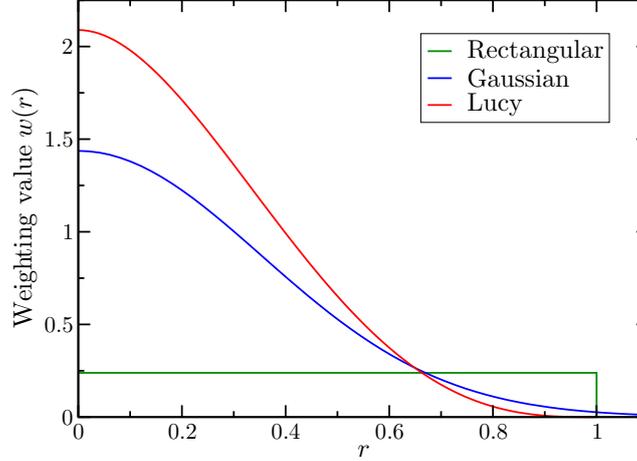


Figure 6: Examples of window functions with  $c = 1$  (rectangular, Lucy), and  $c = 0.5$  (gaussian). The Lucy function combines the advantages of being differentiable, with finite support and computationally efficient [WHTL13].

Rectangular window:  $w(r) = \frac{\mathcal{H}(c-r)}{4/3\pi c^3}$  with  $\mathcal{H}$  the Heaviside function.

Gaussian window:  $w(r) = \frac{1}{\pi^{3/2}c^3} \exp\left(-\left(\frac{r}{c}\right)^2\right)$ .

Lucy window:  $w(r) = \frac{105}{16\pi c^3} \left(-3(r/c)^4 + 8(r/c)^3 - 6(r/c)^2 + 1\right)$ .

$\mathbf{x}$  is computed from all the masses  $m^i$  of the grains at location  $\mathbf{x}^i$ , with  $i \in \{1 \dots N\}$ , by:

$$\rho(\mathbf{x}) = \sum_{i=1}^N w^i m^i \text{ with } w^i \equiv w(|\mathbf{x}^i - \mathbf{x}|) \quad (10)$$

The choice of the averaging window function  $w(r)$ , which depends only on the distance  $r \equiv |\mathbf{x}^i - \mathbf{x}|$  from a given grain to the point where the continuum field is estimated, is very important. This function has to be normalised, ie. its volume integral over the entire space must be equal to one, to ensure that the average of a constant field is the value of the field. The window function must also be monotonically decreasing with the distance to the averaging point, and preferably go to zero at a finite distance for computing efficiency. Some examples of window functions are given in figure 6.

However, the most crucial parameter for the choice of the averaging window is not the function itself, but its width. The width of the window in space will determine how many grains are averaged at each location. If the window is too small, sub-grain details will be visible on the computed continuum field, which is usually not desirable. On the other hand, if the window is too large, the field will be less sensitive to the medium scale inhomogeneities in the material, and may be affected by the boundaries of the system. Consequently it is important to try different sizes of averaging window to ensure the measurements are accurate. From previous studies [WHTL13, GAC<sup>+</sup>06],

a window size  $c \gtrsim d$  seems a reasonable choice.

When coarse-graining of the contact information is required, for example to compute the stress, a similar procedure will be followed. However, the contacts themselves are not physical entities, therefore the weighting coefficient that will have to be applied to the contact information will incorporate the locations of the two particles  $i$  and  $j$  in contact. Depending on the precise field of interest, the weighting coefficient will be either  $\check{w}^{ij}$  the integral of the window function along the contact vector  $\mathbf{r}^{ij} \equiv \mathbf{x}^j - \mathbf{x}^i$ ; or  $\bar{w}^{ij}$  the average of the window at the two particles in contact.

$$\check{w}^{ij} = \int_0^1 w(|\mathbf{x}^i + s\mathbf{r}^{ij} - \mathbf{x}|) ds \quad (11)$$

$$\bar{w}^{ij} = \frac{w(|\mathbf{x}^i - \mathbf{x}|) + w(|\mathbf{x}^j - \mathbf{x}|)}{2} \quad (12)$$

Note that for wide coarse-graining windows,  $\check{w}^{ij} \simeq \bar{w}^{ij}$ , which is the form we will use for simplicity in the implementation section 4.3.

Finally, the continuum field obtained at a particular time-step is usually averaged over time to provide smoother information. A process similar to the one followed for the spatial average can be implemented, with an averaging window in the temporal space. However, it is usually sufficient to use a rectangular window in that case, corresponding to a moving window temporal average. Once again, the size of the time window should be chosen carefully, in particular for the study the transient behaviours of the system.

## 4.2 Balance laws

The balance equations for the coarse-grained field have been derived multiple times for the mass and momentum conservation [GG01, WTLB12], angular momentum conservation [Bab97], and energy conservation [Bab97, AR15]. We will not reproduce these derivation here, but only describe the salient results and important conclusions on the matter. Most of the following results have been derived by [Bab97].

The general form form of the balance equation for a conserved quantity  $\Psi$  is:

$$\underbrace{\frac{\partial \Psi}{\partial t}}_{\text{material derivative}} + \underbrace{\nabla(\bar{\mathbf{v}}\Psi)}_{\text{advection}} = \underbrace{\nabla \cdot \mathbf{q}}_{\text{fluxes}} + \underbrace{S}_{\text{sources}} - \underbrace{s}_{\text{sinks}} \quad (13)$$

where  $\nabla$  is the gradient operator and  $\nabla \cdot$  the divergence operator. All the balance equations have the same form. The left term describes the material derivative of the conserved quantity, which includes a contribution from the change of the quantity in the considered infinitesimal volume, and another contribution from the advection of

the quantity due to the flow. The terms on the right handside of the balance equation include the fluxes from the surrounding material, as well as sources and sinks from external contributions.

Note that the advective term depends on the mass weighted average velocity  $\bar{\mathbf{v}}$ . It can be computed from the particle velocities using

$$\rho \bar{\mathbf{v}} = \sum_{i=1}^N w^i m^i \mathbf{v}^i \quad (14)$$

where  $\rho$  is given by equation 10.

The mass and momentum balance therefore read respectively:

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \bar{\mathbf{v}}) = 0 \quad (15)$$

$$\frac{\partial \rho \bar{\mathbf{v}}}{\partial t} + \nabla(\rho \bar{\mathbf{v}} \bar{\mathbf{v}}) = \nabla \cdot (\boldsymbol{\sigma}^C + \boldsymbol{\sigma}^K) + \rho \mathbf{g} \quad (16)$$

with  $\mathbf{g}$  the gravitational acceleration. As is clear in equation 16, the stress tensor has two distinct contributions, respectively  $\boldsymbol{\sigma}^C$  the contact stress, and  $\boldsymbol{\sigma}^K$  the kinetic stress, defined from the contact force  $\mathbf{f}^{ij}$  and contact vector  $\mathbf{r}^{ij} = \mathbf{x}^i - \mathbf{x}^j$  between particles  $i$  and  $j$ , with  $\otimes$  the tensor product:

$$\boldsymbol{\sigma}^C = \sum_{i=1}^N \sum_{j>i}^N \tilde{w}^{ij} \mathbf{r}^{ij} \otimes \mathbf{f}^{ij} \quad (17)$$

$$\boldsymbol{\sigma}^K = \sum_{i=1}^N w^i m^i \tilde{\mathbf{v}}^i \otimes \tilde{\mathbf{v}}^i \quad (18)$$

The kinetic stress is related to the velocity fluctuations of the grains  $\tilde{\mathbf{v}}^i = \mathbf{v}^i - \bar{\mathbf{v}}$ : the more the grains fluctuate around the average bulk velocity, the more internal pressure will be created. However, defining what is actually the fluctuating velocity can be subtle [AR15]. In particular, choices must be made on where and when the bulk velocity  $\bar{\mathbf{v}}$  should be measured (at the coarse graining point or at the particle location, averaged in time or not etc.).

The balance equation for the angular momentum is given below, and leads to the definition of the couple stress tensor  $\boldsymbol{\mu}^C + \boldsymbol{\mu}^K$  and the rate of supply of internal spin from the contacts  $\boldsymbol{\mu}^{cc}$ :

$$\frac{\partial \rho I \bar{\boldsymbol{\omega}}}{\partial t} + \nabla(\rho \bar{\mathbf{v}} I \bar{\boldsymbol{\omega}}) = \nabla \cdot (\boldsymbol{\mu}^C + \boldsymbol{\mu}^K) + \boldsymbol{\mu}^{cc} \quad (19)$$

with

$$\tilde{\omega}^i = \omega^i - \bar{\omega}; \quad \rho I = \sum_{i=1}^N w^i m^i I^i; \quad \rho \bar{\omega} = \sum_{i=1}^N w^i m^i \omega^i \quad (20)$$

$$\boldsymbol{\mu}^K = \sum_{i=1}^N w^i m^i I^i \tilde{\omega}^i \otimes \tilde{\omega}^i; \quad \boldsymbol{\mu}^C = \sum_{i=1}^N \sum_{j>i}^N \tilde{w}^{ij} \mathbf{r}^{ij} \otimes (\mathbf{m}^{ij} - \mathbf{m}^{ji}) \quad (21)$$

$$\boldsymbol{\mu}^{cc} = \sum_{i=1}^N \sum_{j>i}^N \tilde{w}^{ij} \mathbf{r}^{ij} \times \mathbf{f}^{ij} \quad (22)$$

Note that  $\boldsymbol{\mu}^C$  vanishes if particles  $i$  and  $j$  have the same radius, since in that case the torque  $\mathbf{m}^{ij}$  created by  $i$  over  $j$  is equal to  $\mathbf{m}^{ji}$  the torque of  $j$  on  $i$ . Finally  $\boldsymbol{\mu}^{cc} = \sum_{i=1}^N \sum_{j>i}^N \tilde{w}^{ij} \mathbf{r}^{ij} \times \mathbf{f}^{ij}$ . We will not go into more details on the balance of rotational energy, except to mention that it is at the center of Cosserat theories, which explicitly describe the elastic property alterations due to the couple stress tensor and the grain micro-rotation [Lak95, SSV08].

We now turn our attention to the energy balance equation. We will consider separately the balance law for the kinetic (translational) energy, the rotational energy and the thermal energy, as these are the ones giving the most insight into the system. The global energy balance law could be recovered simply by summing those 3 equations. The thermal energy acts as an energy sink, since there is no mechanism to convert back the thermal energy into mechanical energy (cf. section 2.3). Moreover we do not consider any feedback of the (thermal) temperature on the mechanical properties of the grains, which would require to use the thermal energy balance to compute the temperature. Therefore we will not consider the thermal energy balance here, although the law can be found in [Bab97].

The translational kinetic energy and the rotational kinetic energy balance laws are the following:

$$\frac{\partial \rho(\bar{E}_K + \tilde{E}_K)}{\partial t} + \nabla(\rho(\bar{E}_K + \tilde{E}_K)) = \nabla \cdot ((\boldsymbol{\sigma}^C + \boldsymbol{\sigma}^K) \bar{\mathbf{v}} + \mathbf{q}_K^C + \mathbf{q}_K^K) - s_K + \rho \mathbf{g} \bar{\mathbf{v}} \quad (23)$$

$$\frac{\partial \rho(\bar{E}_R + \tilde{E}_R)}{\partial t} + \nabla(\rho(\bar{E}_R + \tilde{E}_R)) = \nabla \cdot ((\boldsymbol{\mu}^C + \boldsymbol{\mu}^K) \bar{\boldsymbol{\omega}} + \mathbf{q}_R^C + \mathbf{q}_R^K) - s_R \quad (24)$$

In the same way that we have split translational and rotational velocities in their average values and fluctuating parts, the energy density  $E_K = \bar{E}_K + \tilde{E}_K$  and  $E_R =$

$\bar{E}_R + \tilde{E}_R$  are split such as:

$$\bar{E}_K = \frac{1}{2} \bar{\mathbf{v}} \bar{\mathbf{v}}; \quad \rho \tilde{E}_K = \frac{1}{2} \sum_{i=0}^N w^i m^i \tilde{\mathbf{v}}^i \tilde{\mathbf{v}}^i \quad (25)$$

$$\bar{E}_R = \frac{1}{2} I \bar{\boldsymbol{\omega}} \bar{\boldsymbol{\omega}}; \quad \rho \tilde{E}_R = \frac{1}{2} \sum_{i=0}^N w^i m^i I^i \tilde{\boldsymbol{\omega}}^i \tilde{\boldsymbol{\omega}}^i \quad (26)$$

The fluxes of energy involve the work produced by the stress and moment stress tensor on the displacement due to the average velocity, as well as some additional contributions from the fluctuating velocities defined as:

$$\mathbf{q}_K^C = \frac{1}{2} \sum_{i=1}^N \sum_{j>i}^N \tilde{w}^{ij} \mathbf{r}^{ij} \mathbf{f}^{ij} \cdot (\tilde{\mathbf{v}}^i + \tilde{\mathbf{v}}^j); \quad \mathbf{q}_K^K = -\frac{1}{2} \sum_{i=1}^N w^i m^i I^i \tilde{\mathbf{v}}^i \cdot \tilde{\mathbf{v}}^i \tilde{\mathbf{v}}^i \quad (27)$$

$$\mathbf{q}_R^C = \frac{1}{2} \sum_{i=1}^N \sum_{j>i}^N \tilde{w}^{ij} \mathbf{r}^{ij} (\mathbf{m}^{ij} \cdot \tilde{\boldsymbol{\omega}}^i - \mathbf{m}^{ji} \cdot \tilde{\boldsymbol{\omega}}^j); \quad \mathbf{q}_R^K = -\frac{1}{2} \sum_{i=1}^N w^i m^i I^i \tilde{\boldsymbol{\omega}}^i \cdot \tilde{\boldsymbol{\omega}}^i \tilde{\mathbf{v}}^i \quad (28)$$

Finally, as discussed before, some of the energy is dissipated in thermal energy, through the terms  $s_K$  and  $s_R$ :

$$s_K = - \sum_{i=1}^N \sum_{j>i}^N \bar{w}^{ij} \mathbf{f}^{ij} \cdot (\mathbf{v}^i - \mathbf{v}^j) \quad (29)$$

$$s_R = - \sum_{i=1}^N \sum_{j>i}^N \bar{w}^{ij} (\mathbf{m}^{ij} \cdot \boldsymbol{\omega}^i + \mathbf{m}^{ji} \cdot \boldsymbol{\omega}^j) \quad (30)$$

### 4.3 Example: Coarse-graining in simple shear

To give an example on how the coarse graining procedure can be implemented and to verify the validity of the previous equations, we are computing the coarse-grained field from the simulations detailed in section 3.5. The coarse graining is performed using a Lucy function of width  $c = 2d$ , and all the time-steps are averaged together since the material is supposed to be in steady state. Note that no care is taken to handle the boundaries of the system, therefore the fields will be incorrect near the base, free surface and periodic boundaries of the system. A proper treatment of coarse-graining near boundaries is available in [WTLB12]. For simplicity we will only focus on the balance of the translational kinetic energy, as the rotational part is quite small in this system.

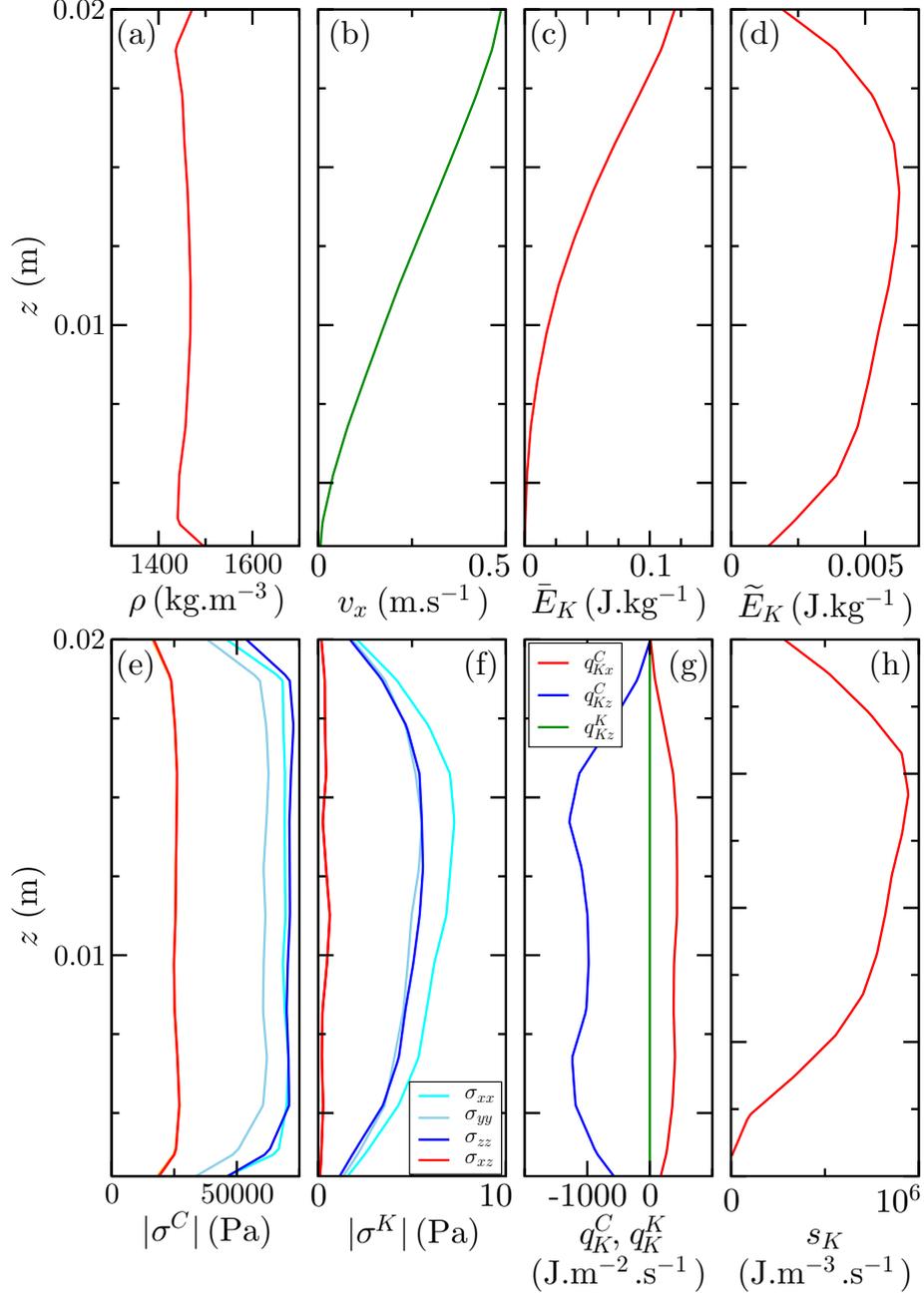


Figure 7: Vertical coarse-grained profile at the center of the simulation box of various quantities related to translational energy conservation.  $V = 0.5 \text{ m}\cdot\text{s}^{-1}$ ,  $P = 73 \text{ kPa}$ . (a) density, (b) horizontal velocity, (c) kinetic energy of the mean translational velocity, (d) kinetic energy of the fluctuating velocity, (e) contact stress tensor components, (f) kinetic stress tensor components, (g) contact and kinetic flux component of the kinetic energy, (h) translational kinetic energy density converted into thermal energy. Only non-negligible components are displayed. The legend in (f) also applies to (e).

Figure 7 shows the non-negligible terms entering equation 23. The density profile figure 7a is essentially constant, as expected at such a low inertial number, corresponding to a grain volume fraction of about  $\frac{\rho}{\rho_g} \simeq 0.58$ . The velocity profile figure 7b shows a linear profile with slight inflections near the wall, coherent with the simulated configuration [MRME13]. The kinetic energy of the mean translational flow figure 7c increases with the square of the depth as expected from the velocity profile, and is coherent with the total kinetic energy shown in figure 3a when scaled by the total mass of the system  $Nm_g$ . The fluctuating translational energy figure 7d is two orders of magnitude smaller than  $\bar{E}_K$  but is expected to play an important role as it is closely related to the granular temperature concept [JL09].

Figure 7e-f show respectively the contact stress and the kinetic stress components. By symmetry, only the diagonal components (normal stresses) and the  $\sigma_{xz}$  tangential stress component are non-zero. The contact stress is several orders of magnitude higher than the kinetic stress, which means that the confining pressure is dominating over the pressure due to the fluctuations of the particles. The contact stress also shows that there is little difference in the normal stress components ( $\sigma_{xx} \simeq \sigma_{yy} \simeq \sigma_{zz}$ ). The shear stress is in agreement with a friction coefficient of about 0.35, once again a reasonable value. The flux terms shown in figure 7g also show a much smaller contribution from the velocity fluctuations than from the contact, and are essentially constant in space, except near the boundaries. Finally figure 7h shows the energy dissipated by volume and by second. One can check that the value of  $s_K$  is reasonable by comparing it to the energy provided by the moving top wall. The former dissipates a total energy in the system of  $\iiint s_K \sim 2.5 \text{ J.s}^{-1}$  while the plate provides a power of  $\iint \sigma_{xz} V \sim 3 \text{ J.s}^{-1}$ , which is a rough agreement considering the inaccuracy in the measurement of  $\sigma_{xz}$  near the boundary with the currently implemented coarse-graining method.

The steady state simple shear configuration implies that equation 23 can be simplified quite extensively. In particular, the material derivative is 0, leaving only the flux and source terms in the equation. Moreover, the system is invariant by translation along  $x$  and  $y$  thanks to the periodic boundary conditions, therefore any derivatives along these directions are zero ( $\partial_x = \partial_y = 0$ ). Finally, the velocity is non-zero only along  $x$  and there is no shear stress along the direction  $y$ . Eq. 23 can therefore be simplified to:

$$\frac{\partial(\sigma_{xz}v_x + q_z)}{\partial z} = s_K \quad (31)$$

Figure 8 shows that equation 31 is well verified for two different pressures in our system, despite the inaccuracy of the coarse-graining near the boundary and the spatial gradients that lead to a quite staggered flux term. Equation 31 seems therefore reasonably verified using the developed coarse-graining method in this simple shear system.

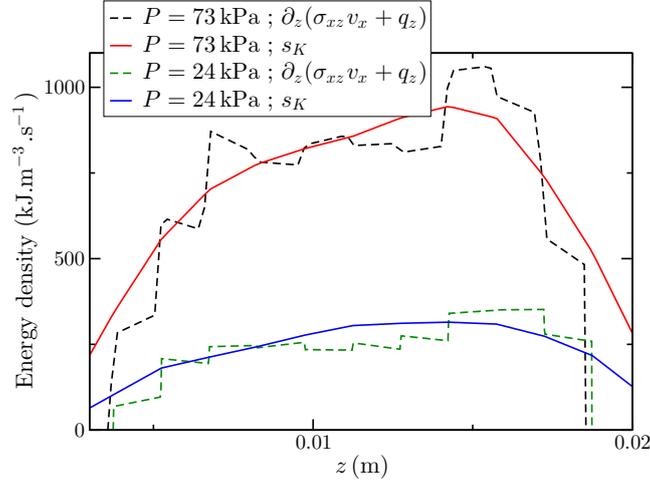


Figure 8: Comparison of the translational energy flux with the dissipated translational energy, for 2 confining pressure.  $V = 0.5 \text{ m.s}^{-1}$ .

## 5 Conclusions

DEM provides the information necessary to verify the viability of proposed continuum mechanical models for granular materials. In particular, the development of constitutive relationship provides expressions for the stress tensors and dissipation. Following the coarse-graining procedure developed in section 4, it is then possible to compare the continuum fields extracted from DEM to the fields hypothesised as constitutive laws in continuum modelling. The closure provided by the constitutive laws then allows to describe the full dynamic of the system using the balance equations, and opens the opportunity for continuum modeling of the material to overcome the limitations of DEM.

The use of DEM gives access to informations at grain level, that can be coarse-grained to obtain mesoscopic fields. However, it is important to remember that both representations carry important assumptions. At the grain level, the contact law models the grain collisions is a idealised way, which can be far from the actual physics of the contact between grains in the material. At the mesoscopic level, the definition of a coarse-grained field can be challenging. In particular, the length-scale of the spatial inhomogeneities in the system must be large enough to be able to even define a continuum. Moreover, identifying some quantities such as the fluctuating velocity and the granular temperature relies on strong hypothesis that should be questioned and tested.

Finally, DEM simulations can be extended to account for other mechanisms, such as cohesion between grains, grain breakage [RE13], or interactions with fluid [GMBE13]. The choice of how such processes are implemented will modify how the energy is

dissipated in the material, and therefore has to be considered carefully to describe numerically the physical processes in a reasonable manner.

## References

- [AFP13] Bruno Andreotti, Yoël Forterre, and Olivier Pouliquen. *Granular media: between fluid and solid*. Cambridge University Press, 2013.
- [AR15] Riccardo Artoni and Patrick Richard. Average balance equations, scale dependence, and energy cascade for granular materials. *Physical Review E*, 91(3):032202, 2015.
- [Bab97] Marijan Babić. Average balance equations for granular materials. *International journal of engineering science*, 35(5):523–548, 1997.
- [BM14] R Balevičius and Z Mróz. A finite sliding model of two identical spheres under displacement and force control. part ii: dynamic analysis. *Acta Mechanica*, 225(6):1735–1759, 2014.
- [BP10] Nikolai V Brilliantov and Thorsten Pöschel. *Kinetic theory of granular gases*. Oxford University Press, 2010.
- [CS79] Peter A Cundall and Otto DL Strack. A discrete numerical model for granular assemblies. *geotechnique*, 29(1):47–65, 1979.
- [Ein07] Itai Einav. Breakage mechanics—part ii: Modelling granular materials. *Journal of the Mechanics and Physics of Solids*, 55(6):1298–1320, 2007.
- [ERMS18] I Einav, P Rognon, T Miller, and J Sulem. Faults get colder through transient granular vortices. *Geophysical Research Letters*, 45(6):2625–2632, 2018.
- [GAC<sup>+</sup>06] Chay Goldenberg, Allbens PF Atman, Philippe Claudin, Gael Combe, and Isaac Goldhirsch. Scale separation in granular packings: stress plateaus and fluctuations. *Physical review letters*, 96(16):168001, 2006.
- [GE05] Tibor J Goda and Fritz Ebert. Three-dimensional discrete element simulations in hoppers and silos. *Powder technology*, 158(1-3):58–68, 2005.
- [GG01] BJ Glasser and I Goldhirsch. Scale dependence, correlations, and fluctuations of stresses in rapid granular flows. *Physics of Fluids*, 13(2):407–420, 2001.
- [GMBE13] Yixiang Gan, Federico Maggi, Giuseppe Buscarnera, and Itai Einav. A particle-water based model for water retention hysteresis. *Géotechnique Letters*, 3:152, 2013.

- [Her82] Heinrich Rudolf Hertz. Über die berührung fester elastischer körper und über die harte. *Verhandlung des Vereins zur Beförderung des Gewerbefleißes, Berlin*, page 449, 1882.
- [HK13] David L Henann and Ken Kamrin. A predictive, size-dependent continuum model for dense granular flows. *Proceedings of the National Academy of Sciences*, 110(17):6730–6735, 2013.
- [JEL17] Yimin Jiang, Itai Einav, and Mario Liu. A thermodynamic treatment of partially saturated soils revealing the structure of effective stress. *Journal of the Mechanics and Physics of Solids*, 100:131–146, 2017.
- [JFP06] Pierre Jop, Yoël Forterre, and Olivier Pouliquen. A constitutive law for dense granular flows. *Nature*, 441(7094):727, 2006.
- [JL09] Yimin Jiang and Mario Liu. Granular solid hydrodynamics. *Granular Matter*, 11(3):139, 2009.
- [JS83] James T Jenkins and Stuart B Savage. A theory for the rapid flow of identical, smooth, nearly elastic, spherical particles. *Journal of fluid mechanics*, 130:187–202, 1983.
- [KaEH09] William R Ketterhagen, Mary T am Ende, and Bruno C Hancock. Process modeling in the pharmaceutical industry using the discrete element method. *Journal of pharmaceutical sciences*, 98(2):442–470, 2009.
- [KGH<sup>+</sup>12] Christoph Kloss, Christoph Goniva, Alice Hager, Stefan Amberger, and Stefan Pirker. Models, algorithms and validation for opensource dem and cfd-dem. *Progress in Computational Fluid Dynamics, an International Journal*, 12(2-3):140–152, 2012.
- [KH93] Georg Kresse and Jürgen Hafner. Ab initio molecular dynamics for liquid metals. *Physical Review B*, 47(1):558, 1993.
- [Lak95] Roderic Lakes. Experimental methods for study of cosserat elastic solids and other generalized elastic continua. *Continuum models for materials with microstructure*, pages 1–25, 1995.
- [LP91] Larry D Libersky and Albert G Petschek. Smooth particle hydrodynamics with strength of materials. In *Advances in the free-Lagrange method including contributions on adaptive gridding and the smooth particle hydrodynamics method*, pages 248–257. Springer, 1991.
- [Lud08] Stefan Luding. Introduction to discrete element methods: basic of contact force models and how to perform the micro-macro transition to continuum theory. *European Journal of Environmental and Civil Engineering*, 12(7-8):785–826, 2008.
- [MD97] Majid T Manzari and Yannis F Dafalias. A critical state two-surface plasticity model for sands. *Geotechnique*, 47(2):255–272, 1997.

- [MiD04] GDR MiDi. On dense granular flows. *The European Physical Journal E*, 14(4):341–365, 2004.
- [MRME13] Thomas Miller, P Rognon, B Metzger, and I Einav. Eddy viscosity in dense granular flows. *Physical review letters*, 111(5):058002, 2013.
- [NCM01] F Nicot, B Cambou, and G Mazzoleni. Design of rockfall restraining nets from a discrete element modelling. *Rock Mechanics and Rock Engineering*, 34(2):99–118, 2001.
- [RE13] Adrian R Russell and Itai Einav. Energy dissipation from particulate systems undergoing a single particle crushing event. *Granular Matter*, 15(3):299–314, 2013.
- [SEG<sup>+</sup>01] Leonardo E Silbert, Deniz Ertas, Gary S Grest, Thomas C Halsey, Dov Levine, and Steven J Plimpton. Granular flow down an inclined plane: Bagnold scaling and rheology. *Physical Review E*, 64(5):051302, 2001.
- [SSV08] Ioannis Stefanou, Jean Sulem, and Ioannis Vardoulakis. Three-dimensional cosserat homogenization of masonry structures: elasticity. *Acta Geotechnica*, 3(1):71–83, 2008.
- [TKT93] Yutaka Tsuji, Toshihiro Kawaguchi, and Toshitsugu Tanaka. Discrete particle simulation of two-dimensional fluidized bed. *Powder technology*, 77(1):79–87, 1993.
- [VC04] Pascal Villard and Bruno Chareyre. Design methods for geosynthetic anchor trenches on the basis of true scale experiments and discrete element modelling. *Canadian Geotechnical Journal*, 41(6):1193–1205, 2004.
- [VVDT07] Emmanuil Veveakis, Ioannis Vardoulakis, and Giulio Di Toro. Thermo-poromechanics of creeping landslides: The 1963 vaiont slide, northern italy. *Journal of Geophysical Research: Earth Surface*, 112(F3), 2007.
- [WHTL13] Thomas Weinhart, Remco Hartkamp, Anthony R Thornton, and Stefan Luding. Coarse-grained local and objective continuum description of three-dimensional granular flows down an inclined surface. *Physics of fluids*, 25(7):070605, 2013.
- [WTLB12] Thomas Weinhart, Anthony R Thornton, Stefan Luding, and Onno Bokhove. From discrete particles to continuum fields near a boundary. *Granular Matter*, 14(2):289–294, 2012.



---

# Hierarchical guide for constructing thermodynamically admissible constitutive models

**Itai Einav**

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

---

*The number of ways one can employ for constructing constitutive models is staggeringly boundless, but the response of a given material does not care about this subjective freedom – it will behave as it wants. Therefore, constraints must be placed for guiding the process of constitutive modelling, with the laws of thermodynamics joining mass and momentum conservation laws as the most prominent and universal examples. Considering those laws, the freedom in modelling is greatly reduced, but not entirely. Therefore, what could further be done (as shown in this chapter) is to classify the mathematical conditions thermodynamics place on well-known mathematical examples of classes of models, building up hierarchically rate-independent models from elasticity to hypo-plasticity, hyper-plasticity and  $h^2$ plasticity. In parallel we discuss issues related to rate-dependency.*

## 1 Introduction

Mass, momentum and energy conservation laws, in addition to entropy balance and the non-negativeness of entropy productions constitute the set of Hydrodynamic Equations that any material must satisfy as it moves and deforms. Let it be solid, fluid, gas or their mixture, granular material, copper or water, those equations universally apply. The complete Hydrodynamic Procedure, which was described by many scholars in physics ([LL80],[LL87],[Kha65],[dGP93],[JL09]), and also so described in Mario Liu's book chapter, represents all those equations at the material level, and thus automatically satisfies Rational Mechanics's form of macroscopic balance laws, so described in Ioannis Stefanou's book chapter. In this chapter we first briefly review the strict connection of those two formulations through the property of work input. Next, we summarise the general mathematical restrictions on constitutive models imposed by the Hydrodynamic Equations. Knowing how the internal energy explicitly de-

depends on internal variables, it is possible to study those restrictions further. Therefore, in the second part of this chapter we demonstrate how to construct several examples for the mathematical structure of thermodynamically admissible models in a hierarchical way. We start from simple viscous fluid and elasticity, and gradually build up to construct rate independent and rate-dependent hypoplasticity, hyperplasticity and  $h^2$ plasticity.

## 2 Rational Mechanics approach

In Rational Mechanics, the work input is identified by localizing a global energy balance into a local form (see Gurtin et al. [GFA10] and the Ioannis Stefanou's chapter in this book). Accordingly, the global energy balance for regular solid continua belonging to spatial region  $\mathcal{P}$  convecting with the body is (here  $\mathcal{P}$  is used for  $\mathcal{P}_t$  in [GFA10]):

$$d_t \mathcal{E}(\mathcal{P}) + d_t \mathcal{K}(\mathcal{P}) = \mathcal{Q}(\mathcal{P}) + \mathcal{W}(\mathcal{P}), \quad d_t \equiv \partial_t + v_k \nabla_k, \quad (1)$$

where  $\mathcal{E}$  denotes the global net internal energy of  $\mathcal{P}$ ;  $\mathcal{K}$  the kinetic energy of  $\mathcal{P}$ ;  $\mathcal{Q}$  the rate of heat flow transferred to  $\mathcal{P}$  through its surface  $\partial\mathcal{P}$ ; and  $\mathcal{W}$  the conventional external power in the moving frame (here  $\mathcal{W}$  is used for  $\mathcal{W}^0$  in [GFA10]); also,  $d_t$  is the material derivative and  $\nabla_k v_k$  the gradient of the barycentric velocity vector  $v_k$ . The four energetic terms above are specified as follows

$$\mathcal{E}(\mathcal{P}) = \int_{\mathcal{P}} \rho e dV, \quad \mathcal{W}(\mathcal{P}) = \int_{\mathcal{P}} b_i v_i dV + \int_{\partial\mathcal{P}} T_{ij} n_i v_j da. \quad (2)$$

$$\mathcal{K}(\mathcal{P}) = \int_{\mathcal{P}} \frac{1}{2} \rho v_k v_k dV, \quad \mathcal{Q}(\mathcal{P}) = - \int_{\partial\mathcal{P}} q_k n_k da, \quad (3)$$

while any scalar heat supply, for instance through radiation, is excluded from  $\mathcal{Q}(\mathcal{P})$  for simplicity of exposition. In the above,  $dV$  and  $da$  are the volume and area increments of  $\mathcal{P}$  and  $\partial\mathcal{P}$ , respectively;  $\rho$  is the material density;  $e$  the internal energy per mass;  $b_i$  the body force with a classical example being due to gravity (here  $b_i$  denotes the  $b_i^0$  term in [GFA10]);  $n_i$  the unit vector;  $q_k$  the dissipative heat flux; and  $T_{ij}$  the Cauchy stress tensor that is positive during tension. However, in Soil Mechanics, it is more customary to define the Cauchy stress tensor to be positive in compression

$$\sigma_{ij} = -T_{ij}. \quad (4)$$

Similarly, whereas the stretching rate tensor  $D_{ij}$  is used as the symmetric component of the velocity gradient  $L_{ij} = \nabla_i v_j$ , which is being positive under tension, in Soil Mechanics it is more common to adopt the symmetric strain rate tensor  $\dot{\epsilon}_{ij}$  that is positive under compression:

$$\dot{\epsilon}_{ij} = -D_{ij}, \quad D_{ij} \equiv \frac{1}{2} (\nabla_i v_j + \nabla_j v_i). \quad (5)$$

In addition, the anti-symmetric part of the velocity gradient can be identified as

$$\Omega_{ij} \equiv \frac{1}{2}(\nabla_i v_j - \nabla_j v_i), \quad (6)$$

such that  $L_{ij} = D_{ij} + \Omega_{ij}$ .

## 2.1 Work input in Rational Mechanics

Since the region  $\mathcal{P}$  convects with the body such that its volume remains constant, and in light of the divergence theorem, the local balance of linear momentum, and the symmetry of stress, equation (2) can be rewritten as [GFA10]:

$$d_t \mathcal{E}(\mathcal{P}) = \int_{\mathcal{P}} \rho d_t e dV, \quad \mathcal{W}(\mathcal{P}) = \int_{\mathcal{P}} \sigma_{ij} \dot{\epsilon}_{ij} dV + d_t \mathcal{K}(\mathcal{P}), \quad (7)$$

where  $T_{ij} D_{ij} = \sigma_{ij} \dot{\epsilon}_{ij}$  was used thanks to equations (4,5). By inserting equations (3,7) into equation (1), we obtain the global form of the energy conservation:

$$\int_{\mathcal{P}} (\rho d_t e - \sigma_{ij} \dot{\epsilon}_{ij} + \nabla_k q_k) dV = 0, \quad (8)$$

which upon localization gives the local energy balance for a unit volume of the media [GFA10]:

$$w \equiv \rho d_t e + \nabla_k q_k, \quad \text{with } w = \sigma_{ij} \dot{\epsilon}_{ij}. \quad (9)$$

Extensions for multi-component continua, such as saturated and partially saturated soils are discussed in [GFA10, EL18].

## 3 Hydrodynamics approach

The purpose of the following section is to arrive at an exposition that shows the full consistency between the Hydrodynamics and Rational Mechanics formulations adopted in this ALERT book chapters by Mario Liu and Ioannis Stefanou, respectively. A more complete demonstration could be found in [EL18]. For that purpose, we describe the building steps of the Hydrodynamic Procedure, which begins with the first law of thermodynamics (conservation of energy) as:

$$\partial_t U + \nabla_i E_i = \rho v_i G_i, \quad (10)$$

where  $U = U(\dots, v_i)$  and  $u = u(\dots) = \rho e$  are the conserved energy densities in the moving and rest ( $v_i = 0$ ) frames, respectively, with  $v_i$  being the velocity ( $u$  also known as the internal energy density);  $\rho G_i$  is the gravitational force density; and  $E_i$  is the energy flux, whose form will be determined below. We note that the conserved and internal free energy densities are linked through the kinetic energy  $k = g_i^2/2\rho$

using  $U = u + k$ , with  $g_i = \rho v_i$  being the momentum and  $\rho$  the density. In other words, that  $U = u + g_i^2/2\rho$  and  $\partial_t U = \partial_t u + v_i \partial_t g_i - v_i^2 \partial_t \rho/2$ .

Also note the second law of thermodynamics (entropy balance):

$$\partial_t s + \nabla_i (s v_i - f_i) = R/T \geq 0, \quad (11)$$

with  $s v_i$  and  $f_i$  as the convective and dissipative entropy currents, respectively; where  $T$  is the temperature, and  $R$  is the total rate of dissipation, whose form will be determined below.

In addition, any material must also satisfy the conservation laws of mass and momentum<sup>1</sup>,

$$\partial_t \rho + \nabla_i (\rho v_i) = 0, \quad (12)$$

$$\partial_t g_i + \nabla_j (g_i v_j + \sigma_{ij}) = \rho G_i, \quad (13)$$

It is also possible to identify a conserved elastic strain  $\varepsilon_{ij}^e$  for all materials, even for conventional fluids where it vanishes. For that purpose, we have to describe an addition conservation equation for the elastic strain, whose rate is most generally given as<sup>2</sup>

$$(\partial_t + v_k \nabla_k) \varepsilon_{ij}^e + \Omega_{ik} \varepsilon_{kj}^e - \varepsilon_{ik}^e \Omega_{kj} = \dot{\varepsilon}_{ij} - \dot{\varepsilon}_{ij}^p, \quad (14)$$

where the term  $\Omega_{ik} \varepsilon_{kj}^e - \varepsilon_{ik}^e \Omega_{kj}$  above accounts for solid body rotation of the elastic strain  $\varepsilon_{ij}^e$ , while  $\dot{\varepsilon}_{ij}$  accounts for the actual deformation. Also, the dissipative contribution to the above equation  $\dot{\varepsilon}_{ij}^p$  is called the plastic strain rate. See further discussion on the above equation in [EL18].

We summarise that irrespective of their structure all materials carry their conserved energy, entropy, density, momentum and elastic strain ( $U$ ,  $s$ ,  $\rho$ ,  $g_i$  and  $\varepsilon_{ij}^e$ ), each receiving their own equation (equations 10-14). More generally, the conserved energy describes a scalar quantity that links all those state variables with  $U = U(s, \rho, g_i, \varepsilon_{ij}^e, \dots)$ , with  $\dots$  denoting the possibility of having additional internal variables. For now, we ignore the possibility of additional internal variables, although soils certainly depends on other state variables such as granular temperature and grainsize distribution / breakage (see Mario Liu's, as well as Yida Zhang & Giuseppe Buscarnera's book chapters). Therefore, this chapter restricts attention to materials whose conserved energy is fully described by  $U = U(s, \rho, g_i, \varepsilon_{ij}^e)$ .

We now have all that is required to do some mathematical manipulations and retrieve a more explicit set of constraints for our rather limited  $U = U(s, \rho, g_i, \varepsilon_{ij}^e)$  material.

<sup>1</sup>The plus sign ahead of  $\sigma_{ij}$  in equation (13) is consistent with Soil Mechanics sign convention of positive stresses under compression. For positive stresses under tension one should use the negative sign instead.

<sup>2</sup>This equation is specified to be consistent with Soil Mechanics sign convention with positive symmetric strain rate  $\dot{\varepsilon}_{ij}$  and elastic strain  $\varepsilon_{ij}^e$  tensors under compression. For positive symmetric strain rate under tension  $D_{ij}$ , the terms  $\dot{\varepsilon}_{ij}$  and  $\dot{\varepsilon}_{ij}^p$  should correspondingly be replaced by  $D_{ij}$  and  $D_{ij}^p$ , in which case  $\varepsilon_{ij}^e$  will also be positive under tension.

Given the general dependence of the conserved energy on four internal variables, four conjugate variables (temperature, chemical potential, velocity and elastic stress) can be defined as:

$$T \equiv \frac{\partial U}{\partial s}, \quad \mu \equiv \frac{\partial U}{\partial \varrho}, \quad v_i \equiv \frac{\partial U}{\partial g_i}, \quad \pi_{ij} \equiv \frac{\partial U}{\partial \varepsilon_{ij}^e}. \quad (15)$$

Skipping some details that could be found in [EL18], it is also possible to identify the thermodynamic pressure from the differentiation of the internal energy per unit mass ( $u/\varrho$ ) by the inverse of density

$$P_T = - \left. \frac{\partial(u/\varrho)}{\partial(1/\varrho)} \right|_{s/\rho, \varepsilon_{ij}^e} = Ts + \mu^0 \varrho - u = \mu^0 \varrho - \psi, \quad (16)$$

which could be seen as the Legendre conjugate of the Helmholtz free energy  $\psi = \psi(\varrho, \dots)$ , replacing the density  $\varrho$  with the chemical potential in the rest system  $\mu^0 = \mu - v_i^2/2$ .

It is also convenient to define the following stress measure,

$$\sigma_{ij}^D \equiv -\sigma_{ij} + \pi_{ij} + P_T \delta_{ij}, \quad (17)$$

whose meaning as the viscous stress will be discussed below. Here,  $\delta_{ij}$  denotes the Kronecker delta, returning 0 when  $i \neq j$  and 1 when  $i = j$ .

In the absence of viscous stress (realistic close to equilibrium), the above becomes a generalised version of Terzaghi's equation for effective stress [Ter43]  $\sigma_{ij}^{eff} \equiv \sigma_{ij} - P_T \delta_{ij}$ , with the elastic stress taking the role of the effective stress  $\sigma_{ij}^{eff} \equiv \pi_{ij}$  and  $P_T$  dependent on the media. This idea has been developed in depth by Jiang et al. [JEL17], with specific expressions resolved for  $P_T$  in fully and partially saturated soils.

The fact that the energy does not change under rigid body rotation can be expressed in terms of the *rotation identity*,

$$\pi_{ji} \varepsilon_{ik}^e - \pi_{ki} \varepsilon_{ij}^e = 0. \quad (18)$$

After a slightly subtle derivation using all of the above Hydrodynamic Equations above, one can find (see [EL18])

$$\nabla_i E_i = \nabla_i ([U + P_T] v_i - f_i T + [\pi_{ij} - \sigma_{ij}^D] v_j) - R + f_i \nabla_i T - \sigma_{ij}^D \dot{\varepsilon}_{ij} + \pi_{ij} \dot{\varepsilon}_{ij}^p.$$

The energy flux  $E_i$  is then identified with all terms appearing within the gradient, while the total dissipation  $R$  with the other terms:

$$E_i = (U + P_T) v_i + (\pi_{ij} - \sigma_{ij}^D) v_j - T f_j, \quad (19)$$

$$R = f_i \nabla_i T + \pi_{ij} \dot{\varepsilon}_{ij}^p - \sigma_{ij}^D \dot{\varepsilon}_{ij} \geq 0. \quad (20)$$

The first term in  $R$  represents thermal dissipation; the second represents dissipation through plasticity, or irreversible straining; and the last term is viscous dissipation most familiar in fluids. In soils it is generally negligible for small strain rates, and thus coming back to equation (17), Terzaghi's equation could be explained. However, under very high rates, it should not be forgotten, and then represents a physical correction to Terzaghi's principle.

Generally speaking, for materials with internal free energy dependent on other internal variables than just the primary ones ( $s$ ,  $\varepsilon_{ij}^e$  and  $\varrho$ ), the total rate of dissipation may depend on further dissipative contributions to  $R$ . For example, in breakage mechanics (eg., see Zhang & Buscarnera's ALERT book chapter) an additional term will involve the product of the breakage energy by the rate of breakage.

### 3.1 Work input in Hydrodynamics

It is possible to show that according to the Hydrodynamic Equations above, one can find exactly the same explicit expression for the work input equation (9b) as found by Rational Mechanics.

Considering the definition of the work input in equation (9a), it is possible to redefine it using the original Hydrodynamic equations (10-13):

$$w = d_t u - T d_t s + R - f_k \nabla_k T + \psi \nabla_k v_k. \quad (21)$$

Since  $du = T ds + \mu^0 d\varrho + \pi_{ij} d\varepsilon_{ij}^e$  and given  $R$  in equation (20), the above becomes

$$w = \mu^0 d_t \varrho + \pi_{ij} (d_t \varepsilon_{ij}^e + \dot{\varepsilon}_{ij}^p) - \sigma_{ij}^D \dot{\varepsilon}_{ij} + \psi \nabla_k v_k. \quad (22)$$

Then, by employing the mass conservation in equation (12) with the balance law for the elastic strain in equation (14), the consequence of the rotation identity in equation (18), the Soil Mechanics convention of strain rate in equation (5), and the relation between the thermodynamic potential and Helmholtz free energy in equation (16) we obtain:

$$\begin{aligned} w &= (\pi_{ij} - \sigma_{ij}^D) \dot{\varepsilon}_{ij} - (\mu^0 \varrho - \psi) \nabla_k v_k \\ &= (\pi_{ij} - \sigma_{ij}^D + P_T \delta_{ij}) \dot{\varepsilon}_{ij} \\ &= \sigma_{ij} \dot{\varepsilon}_{ij}, \end{aligned} \quad (23)$$

which is fully consistent with the result in Rational Mechanics by localizing the global energy balance (see equation (9b)).

Note that taking the energy to depend on both the density and elastic strain is more general than specific cases of solids independent on density (where  $\sigma_{ij}^D = P_T \delta_{ij} = 0$  and  $\sigma_{ij} = \pi_{ij}$ ) or fluids independent on elastic strain (where  $\pi_{ij} = 0$  and  $\sigma_{ij} =$

$-\sigma_{ij}^D + P_T \delta_{ij}$ ). Also, till now we have established the consistency of the Hydrodynamic Procedure and Rational Mechanics, by showing the consistency in their quantity of work input. The demonstration was done for materials with  $u = u(s, \rho, \varepsilon_{ij}^e)$ , yet could be easily established for more general materials with more internal variables.

## 4 Summary of the thermodynamic constraints

Here, we summarise the constitutive restrictions any constitutive model with  $u = u(s, \rho, \varepsilon_{ij}^e)$  should obey based on the above derivation. Considering the Legendre transformation of the internal free energy  $u = \psi + Ts$ , and since the temperature  $T$  is often the measured quantity rather than the entropy  $s$ , it is often more convenient to represent materials using the Helmholtz free energy:

$$\psi = \psi(T, \rho, \varepsilon_{ij}^e).$$

Using equation (15) we therefore find the following relationships between all the state variables:

$$s(T, \rho, \varepsilon_{ij}^e) \equiv -\frac{\partial \psi}{\partial T}, \quad \mu(T, \rho, \varepsilon_{ij}^e) \equiv \frac{\partial \psi}{\partial \rho}, \quad \pi_{ij}(T, \rho, \varepsilon_{ij}^e) \equiv \frac{\partial \psi}{\partial \varepsilon_{ij}^e}. \quad (24)$$

In addition, it is also important to recall the general constraint on the total dissipation rate in equation (20):

$$R = f_i \nabla_i T + \dot{\varepsilon}_{ij}^p \pi_{ij} - \sigma_{ij}^D \dot{\varepsilon}_{ij} \geq 0. \quad (25)$$

In summary, from the viewpoint of thermodynamics, a fully admissible constitutive model of an  $\psi = \psi(T, \rho, \varepsilon_{ij}^e)$  type should therefore satisfy the two equations above. Then, using the definitions of the viscous stress  $\sigma_{ij}^D$  in equation (17) and the thermodynamic pressure  $P_T$  in equation (16) we can work out the total Cauchy stress  $\sigma_{ij}$ . What we need in order to complete the construction of a new model is to assume:

1. an equation for  $\psi = \psi(T, \rho, \varepsilon_{ij}^e)$  from which we can find the entropy  $s$ , chemical potential  $\mu$  and elastic stress  $\pi_{ij}$ , as well as
2. equations for the dissipative entropy current  $f_i$ , the plastic strain rate  $\dot{\varepsilon}_{ij}^p$  and viscous stress  $\sigma_{ij}^D$  that ensure  $R \geq 0$  for any conceivable loading path.

Starting with an explicit expression for  $\psi = \psi(T, \rho, \varepsilon_{ij}^e)$ , Step 1 above is normally not too difficult to respect, although the choice of  $\psi$  should represent the material at hand. The more difficult challenge and freedom is normally embedded in establishing Step 2 above. A few ways from the literature are discussed below.

#### 4.1 Ways to satisfy the dissipation inequality $R \geq 0$

There are many ways to chose  $f_i$ ,  $\varepsilon_{ij}^p$  and  $\sigma_{ij}^D$  that will ensure  $R \geq 0$ . A most general way is based on Onsager's reciprocal principle [Ons31] that allows for cross couplings between the  $R$  terms [DGM13] (see also Mario Liu's chapter in this ALERT book),

$$\begin{pmatrix} f_i \\ \varepsilon_{ij}^p \\ \sigma_{ij}^D \end{pmatrix} = \begin{pmatrix} \kappa_{ik} & b_{ikl} & c_{ikl} \\ b_{kli} & e_{ijkl} & h_{ijkl} \\ -c_{kli} & -h_{klij} & \eta_{ijkl} \end{pmatrix} \cdot \begin{pmatrix} \nabla_k T \\ \pi_{kl} \\ -\dot{\varepsilon}_{kl} \end{pmatrix}, \quad (26)$$

where the non-negativeness of  $R$  is ensured provided the following conditions are satisfied: First, taking both  $\sigma_{ij}^D$  and  $\dot{\varepsilon}_{kl}$  as a six-tuple vector, and writing  $\eta_{ijkl}$  as a 6x6 matrix,  $\eta_{\alpha,\beta}$  ( $\alpha, \beta$  going from 1 to 6), we require it to have only positive Eigenvalues. Next, taking  $(f_i, \varepsilon_{ij}^p)$  and  $(\nabla_k T, \pi_{kl})$  as two nine-tuple vectors, we again require the 9x9 matrix of coefficients connecting them to have positive Eigenvalues. Note that  $c_{kli}$  and  $h_{klij}$  are reactive coefficients that do not contribute to  $R$ , and thus there are no constraints for them.

Please note that the various coefficients in the matrix above are often taken as function of the different state variables. For example, using  $\kappa_{ik} \equiv \kappa_{ik}(T, \varrho, \varepsilon_{ij}^e)$ ,  $b_{ikl} \equiv b_{ikl}(T, \varrho, \varepsilon_{ij}^e)$ , etc. It is also often convenient to think of these coefficients in terms of the elastic stress rather than the elastic strain ( $\kappa_{ik} \equiv \kappa_{ik}(T, \varrho, \pi_{ij}^e)$ ,  $b_{ikl} \equiv b_{ikl}(T, \varrho, \pi_{ij}^e)$ , etc), as we note that transformations between those forms is possible thanks to equation (24).

A very limited subset of the above setup, which satisfies the non-negativeness of  $R$ , is often followed by ignoring all the off-diagonal terms in equation (26), such that:

$$f_i = \kappa_{ij} \nabla_j T, \quad \varepsilon_{ij}^p = e_{ijkl} \pi_{kl}, \quad \sigma_{ij}^D = -\eta_{ijkl} \dot{\varepsilon}_{kl}. \quad (27)$$

However, the equations above are often too limiting because they set up all the various additive terms in  $R$  of equation (20) to be non-negative, which goes beyond what is actually required from equation (20). In fact, based on equation (20)  $R$  may still be possible even if one/two of the three additive terms is/are not actually positive, as the other two/one terms may be, whose sum should have a larger absolute value. This situation of permitting one or two negative terms in equation (20) is enabled through the use of the Onsager relationships in equation (26).

A comment – there are other ways of satisfying the non-negativeness of  $R$ . For example, in (hyper)elastoplasticity to be discussed later, it is common to introduce a ‘plastic flow potential’ function  $g$  that is not a thermodynamic potential, but a function from which the unknown variables  $f_i$ ,  $\varepsilon_{ij}^p$  and  $\sigma_{ij}^D$  could be found through its derivation. For example, it is possible to use the following relationships:

$$f_i = \kappa_{ij}^* \frac{\partial g}{\partial \nabla_j T}, \quad \varepsilon_{ij}^p = e_{ijkl}^* \frac{\partial g}{\partial \pi_{kl}}, \quad \sigma_{ij}^D = -\eta_{ijkl}^* \frac{\partial g}{\partial \dot{\varepsilon}_{kl}}, \quad (28)$$

with certain convexity restrictions on the function  $g$ , to be discussed later in this chapter under subsection (hyper)elastoplasticity.

## 5 Setting up your model – Hierarchical development

### 5.1 Viscous fluid

In conventional compressible fluids we ignore the dependence of the free energy on the elastic strains, and consider only its dependence on density  $\psi = \psi(\varrho)$ . In this case we find:

$$\pi_{ij} = 0, \quad \sigma_{ij} = \sigma_{ij}^D + P_T \delta_{ij}, \quad (29)$$

so the remaining assumptions require one to define  $P_T$  and  $\sigma_{ij}^D$ .

The relationship between the thermodynamic pressure  $P_T$  and density  $\varrho$  comes from the density dependence of the free energy  $\psi$ , and is known in fluid mechanics as the equation of state.

On the other hand, the expression for  $\sigma_{ij}^D$  comes from the restriction on  $R$ , as most generally represented by the Onsager relationships (26). A rather simple choice is ignoring all the coefficients in the corresponding big matrix, apart from the values of the  $\eta_{ijkl}$  coefficients, in which case  $R = \eta_{ijkl} \dot{\varepsilon}_{ij} \dot{\varepsilon}_{kl} \geq 0$ . Further demanding the fluid to be isotropic with only one viscosity coefficient  $\eta$  one would typically employ  $\sigma_{ij}^D = -\eta \dot{\varepsilon}_{ij}$  with  $R = \eta \dot{\varepsilon}_{ij} \dot{\varepsilon}_{ij} \geq 0$ .

### 5.2 (Hyper)elasticity

As a second example, consider the case of non-dissipative ( $R = 0$ ) materials whose free energy depends only on the elastic strain  $\psi = \psi(\varepsilon_{ij}^e)$ . In other words, a thermodynamically admissible elastic material (known as a hyper-elastic material) is fully defined by specifying an explicit form for  $\psi = \psi(\varepsilon^e)$ .

Given  $R = 0$ , the viscous stress is zero by definition ( $\sigma_{ij}^D = 0$ ), and given the density independence of the free energy, thus the thermodynamic pressure as well ( $P_T = 0$ ). Therefore, in this case, in the absence of plastic strain rate, we have:

$$\sigma_{ij} = \pi_{ij} \equiv \frac{\partial \psi(\varepsilon_{ij}^e)}{\partial \varepsilon_{ij}^e}. \quad (30)$$

We can then retrieve the general rate equation for the stress in elastic media:

$$\dot{\sigma}_{ij} = E_{ijkl} \dot{\varepsilon}_{kl}^e, \quad \text{with } E_{ijkl} = \frac{\partial^2 \psi}{\partial \varepsilon_{ij}^e \partial \varepsilon_{kl}^e}, \quad (31)$$

where the general stiffness coefficient may depend on the (elastic) strain  $E_{ijkl} = E_{ijkl}(\varepsilon_{kl}^e)$ , or thanks to equation (24) on the (elastic) stress  $E_{ijkl} = E_{ijkl}(\sigma_{kl})$ .

The term ‘hyperelastic’ is often used to designate such materials, as to highlight that they are fully consistent with thermodynamics, and in this case deducible from energy potentials (such as the Helmholtz free energy,  $\psi$ ).

### 5.2.1 Linear (hyper)elasticity

In the case of linear elasticity, the stiffness coefficients in  $E_{ijkl}$  are all constants and independent on the (elastic) stress (or the (elastic) strain). A famous example for an explicit form of the energy potential is given for linear elastic, homogeneous and isotropic materials:

$$\psi(\varepsilon_{ij}^e) = \frac{1}{2}K(\varepsilon_v^e)^2 + G\varepsilon_{ij}^e\varepsilon_{ij}^e, \quad (32)$$

where  $K$  and  $G$  are the bulk and shear moduli, respectively;  $\varepsilon_{ij}^e = \varepsilon_{ij}^e - \frac{1}{3}\delta_{ij}\varepsilon_v^e$  is the deviator of the strain tensor; and the volumetric strain  $\varepsilon_v^e = \varepsilon_{ii}^e$  is taken positive in compression. Using equation (31) we find the final relations between the total Cauchy stress and strain, and between their rates, in such idealised linear elastic materials:

$$\sigma_{ij} = K\varepsilon_v^e\delta_{ij} + 2G\varepsilon_{ij}^e. \quad (33)$$

$$\dot{\sigma}_{ij} = \left[ \left( K - \frac{2}{3}G \right) \delta_{ij}\delta_{kl} + 2G\delta_{ik}\delta_{jl} \right] \dot{\varepsilon}_{ij}^e. \quad (34)$$

There is of course, nothing new in the above linear elastic equation, which was proposed well before the laws of thermodynamics were well articulated. However, in showing that the final equation could be derived from the thermodynamic potential, the obtained result can also be called hyperelastic.

### 5.2.2 Nonlinear (hyper)elasticity

In this case, the stiffness coefficients  $E_{ijkl}$  are no longer independent of the stress, and in general  $E_{ijkl} = E_{ijkl}(\sigma_{kl})$ . For simplicity, a simple useful example is demonstrated in the following by considering only two out of three invariants of the elastic strains, the volumetric and shear parts  $\varepsilon_v^e$  and  $\varepsilon_s^e$ , as the conjugates of the mean and shear stresses  $p$  and  $q$ , respectively.

Consider the Hertz contact between two spherical grains of diameter  $d$ , whose deflection at the contact is  $\delta$ . According to Hertz’s theory, the typical stress by the contact force  $F_n$  scales as

$$\frac{F_n}{d^2} \propto \left( \frac{\delta}{d} \right)^{3/2}.$$

Next assume that this typical mean stress scales linearly with the pressure  $p \propto F_n$  and that the volumetric elastic strain scales as  $\varepsilon_v^e \propto \delta/d$ . Accordingly, from Hertz's theory we may expect  $p \propto (\varepsilon_v^e)^{3/2}$ . As the bulk stiffness  $K$  scales as  $K \propto \partial p / \partial \varepsilon_v^e$ , and since the same considerations may follow for the shear part, according to Hertz's theory it is reasonable to expect the bulk and shear moduli with:

$$K \propto (\varepsilon_v^e)^{1/2} \propto p^{1/3}, \quad \text{and} \quad G \propto K.$$

Alternatively, following Goddard [God90], one may consider the interaction problem between a sphere of diameter  $d$  and a cone, for the same deflection  $\delta$  at the contact we find in this case

$$\frac{F_n}{d^2} \propto \left(\frac{\delta}{d}\right)^2,$$

that is,  $p \propto (\varepsilon_v^e)^2$  and thus

$$K \propto \varepsilon_v^e \propto p^{1/2}, \quad \text{and} \quad G \propto K.$$

This solution should be arguably more realistic for natural sand with asperities, in which case contacts are most frequently between a flat curved surface and a cone reflective of one asperity.

Therefore, a more general model can be specified with

$$K \propto (\varepsilon_v^e)^n \propto p^{n/(n+1)}, \quad \text{and} \quad G \propto K, \quad (35)$$

with  $n = 1/2$  and  $n = 1$  representative of the sphere-to-sphere and cone-to-sphere contact problems, respectively.

Note– the above scaling laws were derived without any thermodynamic considerations. For example, inputting the above relations into  $K$  and  $G$  in equation (34) is thermodynamically wrong, with many people taking this step while ignoring the inconsistencies. Take for example the scaling  $G \propto (\varepsilon_v^e)^n$ . Integrating  $G = 3\partial q / \partial \varepsilon_s^e$  to get  $q$  is not simple, since its value depends on the integration path. Assuming constant  $\varepsilon_s^e$  we may expect  $q = 3\bar{G}\varepsilon_s^e(\varepsilon_v^e)^n$  and  $\psi_s = \frac{3}{2}\bar{G}(\varepsilon_s^e)^2(\varepsilon_v^e)^n + F_1(\varepsilon_v^e)$  for the shear part. Similarly, integrating  $K = \partial p / \partial \varepsilon_v^e$  under constant  $\varepsilon_s^e$  would give  $\psi_v = \frac{\bar{K}}{(n+1)(n+2)}(\varepsilon_v^e)^{n+2} + F_2(\varepsilon_s^e)$ . Comparing  $\psi_s$  with  $\psi_v$ , one can consider the following energy potential:

$$\psi \equiv \psi(\varepsilon_v^e, \varepsilon_s^e) = \frac{\bar{K}}{(n+1)(n+2)}(\varepsilon_v^e)^{n+2} + \frac{3}{2}\bar{G}(\varepsilon_s^e)^2(\varepsilon_v^e)^n. \quad (36)$$

Now that we have a Helmholtz free energy, we can get the explicit relationships (beyond just scaling rules). Starting with the shear stress, we have

$$q \equiv \frac{\partial \psi}{\partial \varepsilon_s^e} = 3\bar{G}\varepsilon_s^e(\varepsilon_v^e)^n, \quad (37)$$

which precisely matches the expected scaling above. However, compared with the initially assumed scaling, the mean stress is slightly more complicated:

$$p \equiv \frac{\partial \psi}{\partial \varepsilon_v^e} = \frac{\bar{K}}{(n+1)} (\varepsilon_v^e)^{n+1} + \frac{3n}{2} \bar{G} (\varepsilon_s^e)^2 (\varepsilon_v^e)^{n-1}. \quad (38)$$

While the first term matches the expected scaling, the second one is new, and reflects the elastic cross-coupling of the volumetric deformations with the shear deformations. The rate equation for the stress becomes:

$$\begin{pmatrix} \dot{p} \\ \dot{q} \end{pmatrix} = \begin{pmatrix} \boxed{\bar{K} (\varepsilon_v^e)^n} + 3n \bar{G} (\varepsilon_s^e)^2 (\varepsilon_v^e)^{n-1} & 3n \bar{G} (\varepsilon_s^e) (\varepsilon_v^e)^{n-1} \\ 3n \bar{G} (\varepsilon_s^e) (\varepsilon_v^e)^{n-1} & \boxed{3 \bar{G} (\varepsilon_v^e)^n} \end{pmatrix} \cdot \begin{pmatrix} \dot{\varepsilon}_v^e \\ \dot{\varepsilon}_s^e \end{pmatrix}, \quad (39)$$

where the boxed terms reflect the initial scalings based on which this law was developed, while the other terms are correction terms required by thermodynamics (the existence of energy potential  $\psi$ ).

Similar derivations and conclusions on the implications of the pressure dependence of the shear modulus on the full elastic constitutive law for granular/sand media can be found in [JL03, HAR05, NE09], as well as discussion on the elastic stability of such equations. The terms not highlighted by a box in the stiffness matrix above represents the dependence of the volumetric response on the shear stress as the latter increases, and may represent the reorientation of force chains upon shear.

### 5.2.3 Density dependent nonlinear elasticity

Next, we discuss the case of a non-dissipative elastic medium whose energy potential does not only depend on the elastic strain, but also on the density  $\psi = \psi(\varepsilon_{ij}^e)$ . Let us also limit the discussion to functions satisfying the multiplicity in the form

$$\psi = f_\psi(\varrho) \psi_e(\varepsilon_{ij}^e). \quad (40)$$

Given  $R = 0$ , the viscous stress  $\sigma_{ij}^D$  is zero by definition. Therefore, from equation (17) we find

$$\sigma_{ij} = f_\psi(\varrho) \frac{\partial \psi_e}{\partial \varepsilon_{ij}^e} - \psi_e(\varepsilon_{ij}^e) \frac{\partial (f_\psi/\varrho)}{\partial (1/\varrho)} \delta_{ij}. \quad (41)$$

This equation is rather involved. One can continue by carrying all the terms, yet for the sake of clarity, it is useful to consider the following approximation under small elastic deformation. Under such a condition, since  $\psi \propto (\varepsilon_v^e)^m$  with  $3 \geq m \geq 2$  and  $\psi \propto (\varepsilon_s^e)^2$  (see the subsection on nonlinear elasticity) one finds  $\psi_e \ll \frac{\partial \psi_e}{\partial \varepsilon_{ij}^e}$ , such that we can neglect the last term (similar to taking  $P_T \rightarrow 0$ ):

$$\sigma_{ij} \approx f_\psi(\varrho) \frac{\partial \psi_e}{\partial \varepsilon_{ij}^e}. \quad (42)$$

### 5.2.4 Viscoelasticity of Kelvin-Voigt type

For simplicity, we return to the limited discussion on idealised density independent materials with  $P_T = 0$ . To capture scenarios with non-negligible viscous stress  $\sigma_{ij}^D$ , we find from equation (17) that for density independent materials with  $P_T \rightarrow 0$

$$\sigma_{ij} = \pi_{ij} - \sigma_{ij}^D, \quad \pi_{ij} = \frac{\partial \psi(\varepsilon_{ij}^e)}{\partial \varepsilon_{ij}^e}, \quad \sigma_{ij}^D = -\eta_{ijkl} \dot{\varepsilon}_{kl}, \quad (43)$$

where the elastic stress  $\pi_{ij}$  is defined from the energy potential  $\psi$  using equation (31) and the evolution law for the viscous stress  $\sigma_{ij}^D$  is chosen to ensure  $R \geq 0$  using equation (27). In such a way, the model automatically satisfies thermodynamics. Furthermore, we note that in this model the total Cauchy stress  $\sigma_{ij}$  is the sum of the elastic and viscous stresses, the celebrated model known as the Kelvin-Voigt viscoelasticity. The model is often schematically represented as a viscous damper and elastic spring connected in parallel.

## 5.3 Thermodynamically admissible hypoplastic media

Hypoplastic models are models that are developed without regard to thermodynamic considerations and are specified directly in an incrementally non-linear irreversible form. Such models are often written to take the following structure [Kol91, CDHC94]:

$$\dot{\sigma}_{ij} = E_{ijkl} (\dot{\varepsilon}_{kl} - \dot{\varepsilon}_{kl}^p), \quad (44)$$

with

$$\begin{aligned} \dot{\varepsilon}_{kl}^p &= B_{kl} \|\dot{\varepsilon}'\|, \quad \|\dot{\varepsilon}'\| = \sqrt{\dot{\varepsilon}'_{mn} \dot{\varepsilon}'_{mn}}, \\ E_{ijkl} &\equiv E_{ijkl}(\sigma_{ij}), \quad \text{and} \quad B_{kl} \equiv B_{kl}(\sigma_{ij}). \end{aligned}$$

Surprisingly, however, unlike the intention behind the adjective "hypo", it is quite simple to construct a thermodynamically admissible hypoplastic constitutive law, probably more than a hyperplastic law (whose adjective "hyper" came to highlight thermodynamic consistency).

Firstly, following from equation (31), and assuming for simplicity a close to equilibrium system with  $\sigma_{ij}^D = 0$  and a density independent energy potential (meaning  $P_T = 0$  and  $\pi_{ij} = \sigma_{ij}$ ), we find:

$$E_{ijkl}(\sigma_{ij}) = \frac{\partial^2 \psi}{\partial \varepsilon_{ij}^e \partial \varepsilon_{kl}^e}. \quad (45)$$

Secondly, using equation (14) (ignoring for simplicity the rotational parts) we may identify the plastic strain rate in hypoplasticity with  $\dot{\varepsilon}_{kl}^p = B_{kl}(\sigma_{ij}) \|\dot{\varepsilon}'_{mn}\|$ . For this case, it follows from the thermodynamic dissipation inequality in equation (20) that

such models should be constrained by  $\dot{\varepsilon}_{ij}^p \pi_{ij} \geq 0$ . Since  $\|\dot{\varepsilon}'_{mn}\|$  is positive by definition, and in this case  $\pi_{ij} = \sigma_{ij}$ , thermodynamics further require:

$$B_{ij} \sigma_{ij} \geq 0. \quad (46)$$

To summarise, a hypoplastic constitutive law of the general form of equation (44) is thermodynamically admissible as long as the  $E_{ijkl}$  can be derived from an energy potential  $\psi$  and that the structure of  $B_{kl}$  is taken to satisfy the inequality above given any state of stress  $\sigma_{ij}$ . Where the issue of finding explicit  $E_{ijkl}$  from  $\psi$  has already been discussed in the context of hyperelasticity, the remaining challenge is to motivate a structure for  $B_{kl} \equiv B_{kl}(\sigma_{ij})$  that works well against physical experiments. At the moment, although constrained by equation (46) above, there is no systematic method for determining this function.

In summary, a thermodynamically admissible hypoplastic constitutive relationships may take the following form:

$$\dot{\sigma}_{ij} = \frac{\partial^2 \psi}{\partial \varepsilon_{ij}^e \partial \varepsilon_{kl}^e} (\dot{\varepsilon}_{kl} - B_{kl} \|\dot{\varepsilon}'\|), \quad \text{with } B_{ij} \sigma_{ij} \geq 0. \quad (47)$$

Note – the above relation is rate-independent given the property  $\dot{\sigma}_{ij}(\dot{\varepsilon}_{ij} dt) \equiv \dot{\sigma}_{ij}(\varepsilon_{ij}) dt$  for any arbitrary positive increment of time  $dt$ .

### 5.3.1 Thermodynamically admissible rate-dependent hypoplasticity

There are several ways to construct a rate-dependent hypoplastic model that is admissible from the viewpoint of thermodynamics. One way is to simply combine the above model with the Kelvin-Voigt type of viscoelasticity. First, recall that the relation in equation (47) so-written for  $\dot{\sigma}_{ij}$  is actually representing  $\dot{\pi}_{ij}$ . Therefore, using this relation to update  $\pi_{ij}$  in addition to the last relation in equation (43) for  $\sigma_{ij}^D$  provides the ingredients to calculate the rate-dependent equation for  $\sigma_{ij}$  using the first relation in equation (43). This model can be schematically represented as a viscous damper element and hypoplastic element connected in parallel. The rate dependency is driven from the last term of  $R$  in equation (25), that is bringing rate-dependency while requiring  $\sigma_{ij}^D \dot{\varepsilon}_{ij} \geq 0$ .

Another way to construct a rate-dependent hypoplastic model can be made in light of the second term of  $R$  in equation (25), that is bringing rate-dependency while requiring  $\pi_{ij} \dot{\varepsilon}_{ij}^p \geq 0$ . Recalling that in the model above  $\dot{\varepsilon}_{ij}^p = B_{ij} \|\dot{\varepsilon}'\|$ , one may simply introduce rate dependency into  $B_{ij} = B_{ij}(\dot{\varepsilon}_{ij})$ , while still demanding that  $B_{ij} \sigma_{ij} \geq 0$ .

## 5.4 (Hyper)elastoplastic media

A more complicated mathematical structure of an inelastic constitutive law is given by the elastoplastic model:

$$\dot{\sigma}_{ij} = E_{ijkl} (\dot{\epsilon}_{kl} - \dot{\epsilon}_{kl}^p), \quad (48)$$

with

$$\dot{\epsilon}_{kl}^p = \lambda C_{kl}(\sigma_{ij}), \quad \lambda(\sigma_{ij}, \dot{\epsilon}_{ij}) \geq 0, \quad y(\sigma_{ij}) \leq 0.$$

The value of the multiplier  $\lambda$  is defined as zero as long as  $y < 0$ , and when ‘yielding’, ie.,  $y = 0$ , it is defined so that the yield function  $y$  never becomes positive. This condition sets  $\lambda$  upon yield by demanding  $\dot{y} = \frac{\partial y}{\partial \sigma_{ij}} \dot{\sigma}_{ij} = 0$ , from which:

$$\lambda = \left( \frac{\frac{\partial y}{\partial \sigma_{ij}} E_{ijkl}}{\frac{\partial y}{\partial \sigma_{ij}} E_{ijkl} C_{kl}} \right) \dot{\epsilon}_{kl}. \quad (49)$$

Considering all of the conditions above, it is possible to rewrite the elastoplastic model using a single equation:

$$\dot{\sigma}_{ij} = E_{ijkl} (\dot{\epsilon}_{kl} - \langle \lambda \rangle C_{ij} I(y)), \quad (50)$$

where we introduced two functions. The function  $I(y)$  is defined as 0 for  $y < 0$  and 1 for  $y = 0$ . The function  $\langle x \rangle$  is  $x$  for  $x \geq 0$  and 0 for  $x < 0$ .

One can introduce additional internal variables into the function  $y$ , in which case the solution for  $\lambda$  will change, but the logic of the derivation above will remain.

Next, we ask the question – which conditions should the various functions ( $E_{ijkl}$ ,  $y$ ,  $C_{ij}$ , and  $\lambda$ ) follow such that the model will satisfy the laws of thermodynamics?

Firstly, we require that the stiffness tensor  $E_{ijkl}$  will follow the logic of hyperelasticity, by requiring it to be derivable from the energy potential.

$$E_{ijkl}(\sigma_{ij}) = \frac{\partial^2 \psi}{\partial \epsilon_{ij}^e \partial \epsilon_{kl}^e}. \quad (51)$$

Secondly, using equation (14) (ignoring for simplicity the rotational parts) we may identify the plastic strain rate in hypoplasticity with  $\dot{\epsilon}_{ij}^p = \langle \lambda \rangle C_{ij} I(y)$ . For this case, since  $\langle \lambda \rangle$  and  $I(y)$  are positive by definition, it follows from the thermodynamic dissipation inequality in equation (20) that such models should be constrained by:

$$C_{ij} \sigma_{ij} \geq 0. \quad (52)$$

Conveniently, this could be satisfied by considering  $C_{ij} = \frac{\partial y}{\partial \sigma_{ij}}$  through the derivative of the yield surface  $y$  with respect to the stress  $\sigma_{ij}$ , and then demanding  $y$  to be a

convex function containing the origin in the stress space. A more general way, is to define a different function, a so-called ‘plastic potential’  $g(\sigma_{ij})$ , which must also be defined as a convex function that contains the origin in the stress space, with  $C_{ij} = \frac{\partial g}{\partial \sigma_{ij}}$ , and

$$C_{ij} = \frac{\partial g}{\partial \sigma_{ij}}, \quad \text{and} \quad \frac{\partial g}{\partial \sigma_{ij}} \sigma_{ij} \geq 0. \quad (53)$$

Setting  $g$  to be different than  $y$  defines a so-called non-associated flow rule. It is also possible to satisfy  $g = 0$  in the limit  $y = 0$ , and  $g < 0$  for  $y < 0$ . This choice of  $g$  does not necessarily satisfy  $g = y$  under  $g < 0$ , yet it satisfies  $I(y) = I(g)$ . Therefore, under those setups, one can define a structure of (hyper)elastoplastic constitutive laws that follows:

$$\dot{\sigma}_{ij} = E_{ijkl} \left( \dot{\epsilon}_{kl} - \langle \lambda \rangle \frac{\partial g}{\partial \sigma_{ij}} I(g) \right), \quad \lambda = \left( \frac{\frac{\partial y}{\partial \sigma_{ij}} E_{ijkl}}{\frac{\partial y}{\partial \sigma_{ij}} E_{ijkl} \frac{\partial g}{\partial \sigma_{kl}}} \right) \dot{\epsilon}_{kl}. \quad (54)$$

Note – it is easy to produce an exact connection between the above formulation and the hyperplastic formulation explored in [Zie77, CH97, HP07]. In either of these hyperplastic formulations, the adjective ‘hyper’ comes to reflect the fact that such relations satisfy the thermodynamic laws. However, as we mentioned in the previous subsection, hypoplastic models may too be structured in such a way that satisfies thermodynamics laws, which motivates the construction of the more encompassing  $h^2$ plasticity formulation discussed in the following.

Also note – the above relation is rate-independent given the property  $\dot{\sigma}_{ij}(\dot{\epsilon}_{ij} dt) \equiv \dot{\sigma}_{ij}(\epsilon_{ij}) dt$  for any arbitrary positive increment of time  $dt$ .

#### 5.4.1 Thermodynamically admissible rate-dependent hyperplasticity

There are several ways to construct a rate-dependent hyperplastic model. One way is to simply combine the above model with the Kelvin-Voigt type of viscoelasticity, directly as was demonstrated earlier for obtaining a rate-dependent hypoplastic model of the same type.

Another way to construct a rate-dependent hypoplastic model can be made in light of the second term of  $R$  in equation (25), that is bringing rate-dependency while requiring  $\pi_{ij} \dot{\epsilon}_{ij}^p \geq 0$ . Recalling that in the model above  $\dot{\epsilon}_{ij}^p = \langle \lambda \rangle \frac{\partial g}{\partial \sigma_{ij}} I(g)$ , one way to introduce rate dependency is to simply replace the linearity in time embedded in  $\lambda$  with a nonlinear function in time. One way is commonly followed through Perzyna’s formulation [Per63] that replaces the role of  $\lambda \geq 0$  with  $\langle y \rangle / \tau \geq 0$ , with  $\tau$  designating a typical relaxation time as a function of the overshooting of yield. As  $\langle y \rangle$  gets bigger the relaxation is stronger, and given a smaller typical time  $\tau$  the relaxation happens faster to satisfy  $y = 0$ , in which case capturing rate-independency as alternative to what has been done before through the solution for  $\lambda$ . Interested readers are

invited to review other ideas in Sari et al.'s book chapter, where the relaxation is also temperature dependent.

## 5.5 $H^2$ plastic media

The only difference between the hyperplastic and the thermodynamically admissible hypoplastic formulations above could be seen to be in the choice of the plastic strain rate. Unlike the hypoplastic formulation, the behaviour of those models experiences step changes in stiffness due to the presence of the functions  $\langle \cdot \rangle$  and  $I(y)$  above, which do not exist in the hypoplastic relationships in the previous section. This is a behavioural advantage of hypoplasticity over hyperplasticity. However, as we also highlighted above, unlike hyperplasticity, the choice of  $B_{ij}$  in the plastic strain rate of hypoplasticity is not very structured. Therefore, in the current example we simply combine the corresponding advantages of hyper- and hypo- plasticity to establish a unifying formulation, the so-called  $h^2$ plasticity (see [Ein12]).

$$\dot{\sigma}_{ij} = E_{ijkl} (\dot{\epsilon}_{kl} - \dot{\epsilon}_{kl}^p), \quad (55)$$

where

$$E_{ijkl}(\sigma_{ij}) = \frac{\partial^2 \psi}{\partial \epsilon_{ij}^e \partial \epsilon_{kl}^e}, \quad \text{and} \quad \dot{\epsilon}_{ij}^p = \langle \lambda \rangle \frac{\partial A(g)}{\partial \sigma_{ij}} = \frac{\partial A(g)}{\partial g} \frac{\partial g}{\partial \sigma_{ij}},$$

with  $\frac{\partial A(g)}{\partial g} \geq 0$  and  $\frac{\partial g}{\partial \sigma_{ij}} \sigma_{ij} \geq 0$ .

Accordingly, here too the rate of plastic dissipation is guaranteed to be non-negative, since

$$\sigma_{ij} \dot{\epsilon}_{ij}^p = \langle \lambda \rangle \frac{\partial A(g)}{\partial g} \frac{\partial g}{\partial \sigma_{ij}} \sigma_{ij} \geq 0. \quad (56)$$

With  $\lambda$  taking from (hyper)elastoplasticity, the  $h^2$ plastic relationship for the stress rate becomes

$$\dot{\sigma}_{ij} = E_{ijkl} \left( \dot{\epsilon}_{kl} - \langle \lambda \rangle \frac{\partial g}{\partial \sigma_{ij}} \frac{\partial A(g)}{\partial g} \right), \quad \lambda = \left( \frac{\frac{\partial y}{\partial \sigma_{ij}} E_{ijkl}}{\frac{\partial y}{\partial \sigma_{ij}} E_{ijkl} \frac{\partial g}{\partial \sigma_{kl}}} \right) \dot{\epsilon}_{kl}. \quad (57)$$

Compare the above equation with equation (54) of (hyper)elastoplasticity, the two formulations are identical only for a choice of  $A$  such that its derivative with respect to  $g$  satisfies  $\frac{\partial A(g)}{\partial g} = I(g)$ . Take for example the following function:

$$A(g) = \frac{1}{s+1} (1+g)^{s+1}, \quad (58)$$

with  $s > 0$  and  $-1 \leq g \leq 0$  with  $g = 0$  when  $y = 0$ . In this case, equation (57) becomes

$$\dot{\sigma}_{ij} = E_{ijkl} \left( \dot{\epsilon}_{kl} - \langle \lambda \rangle \frac{\partial g}{\partial \sigma_{ij}} (1 + g)^s \right), \quad \lambda = \left( \frac{\frac{\partial y}{\partial \sigma_{ij}} E_{ijkl}}{\frac{\partial y}{\partial \sigma_{ij}} E_{ijkl} \frac{\partial g}{\partial \sigma_{kl}}} \right) \dot{\epsilon}_{kl}.$$

Note that in the limit of  $s \rightarrow \infty$  the above model identically becomes that of hyperplasticity, since at this limit the function  $(1 + g)^s$  becomes  $I(g)$ . Also note, where in hyperplasticity there is a switch in the tangent stiffness from a state with  $y = g = 0$  to a state with  $y < 0$  (and  $g < 0$ ), this is not the case in  $h^2$ plasticity. However, in the limit of  $g = y = 0$ , the equation above reduces identically to the hyperplastic relation in equation (54). In this way, the structure of hyperplasticity to recover phenomenological yielding and approach to critical state is preserved in  $h^2$ plasticity, but the yielding in  $h^2$ plasticity is gradual since  $(1 + g)^s$  is not zero for  $g < 0$ . There is still however a certain switch in the loading direction given by the positiveness of  $\lambda$  contained within  $\langle \lambda \rangle$ . This switch can be eliminated by replacing  $\langle \lambda \rangle$  with  $|\lambda|$ . The choice between these two options, or any linear combination is constitutive (for further details see [Ein12]).

## 6 Conclusions

This Chapter provided a guide for the construction of thermodynamically admissible constitutive models. We first reviewed the general conditions imposed by the thermodynamic laws through the eyes of both Rational Mechanics and the Hydrodynamic Procedure. We demonstrated that the end implications of these two formulations are actually identical, which should not be surprising, given that they both came to represent the same physical reality. We continued with the Hydrodynamic Procedure as a concise way to study materials at the scale of a representative volume element, and established a few examples for various classes of constitutive models. The demonstration of the constitutive models was presented in a hierarchical way, starting from simple viscous fluid, then moving on to study rate-independent and rate-dependent elasticity and elastoplasticity (or hyperelasticity and hyperplasticity, respectively, by virtue of satisfying thermodynamics). Next, we discussed the structure of hypoplastic models. Finally, by comparing the mathematical structure of the thermodynamically admissible hypoplasticity and hyperplasticity, we were able to construct a unification called  $h^2$ plasticity formulation that captures the advantages of both models.

## References

- [CDHC94] René Chambon, Jacques Desrues, Walid Hammad, and Robert Charlier. Cloe, a new rate-type constitutive model for geomaterials theoretical basis

- and implementation. *International Journal for Numerical and Analytical Methods in Geomechanics*, 18(4):253–278, 1994.
- [CH97] I.F. Collins and G.T. Houlsby. Application of thermomechanical principles to the modelling of geotechnical materials. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 453, pages 1975–2001. The Royal Society, 1997.
- [DGM13] S.R. De Groot and P. Mazur. *Non-Equilibrium Thermodynamics*. Dover Books on Physics. Dover Publications, 2013.
- [dGP93] P.G. de Gennes and J. Prost. *Continuum Mechanics*. Clarendon Press, Oxford, 1993.
- [Ein12] Itai Einav. The unification of hypo-plastic and elasto-plastic theories. *International Journal of Solids and Structures*, 49(11-12):1305–1315, 2012.
- [EL18] Itai Einav and Mario Liu. Hydrodynamic derivation of the work input to fully and partially saturated soils. *Journal of the Mechanics and Physics of Solids*, 110:205–217, 2018.
- [GFA10] M.E. Gurtin, E. Fried, and L. Anand. *The Mechanics and Thermodynamics of Continua*. Cambridge University Press, 2010.
- [God90] JD Goddard. Nonlinear elasticity and pressure-dependent wave speeds in granular media. *Proc. R. Soc. Lond. A*, 430(1878):105–131, 1990.
- [HAR05] GT Houlsby, A Amorosi, and E Rojas. Elastic moduli of soils dependent on pressure: a hyperelastic formulation. *Géotechnique*, 55(5):383–392, 2005.
- [HP07] G.T. Houlsby and A.M. Puzrin. *Principles of Hyperplasticity: An Approach to Plasticity Theory Based on Thermodynamic Principles*. Springer London, 2007.
- [JEL17] Y. Jiang, I. Einav, and M. Liu. A thermodynamic treatment of partially saturated soils revealing the structure of effective stress. *Journal of the Mechanics and Physics of Solids*, 100:131–146, 2017.
- [JL03] Yimin Jiang and Mario Liu. Granular elasticity without the coulomb condition. *Physical review letters*, 91(14):144301, 2003.
- [JL09] Y. Jiang and M. Liu. Granular solid hydrodynamics. *Granular Matter*, 11:139, 2009.
- [Kha65] I.M. Khalatnikov. *Introduction to the Theory of Superfluidity*. Benjamin, New York, 1965.
- [Kol91] Dimitrios Kolymbas. An outline of hypoplasticity. *Archive of applied mechanics*, 61(3):143–151, 1991.

- [LL80] L.D. Landau and E.M. Lifshitz. *Statistical Physics*. Butterworth-Heinemann, 1980.
- [LL87] L.D. Landau and E.M. Lifshitz. *Fluid Mechanics*. Butterworth-Heinemann, 1987.
- [NE09] Giang D Nguyen and Itai Einav. The energetics of cataclasis based on breakage mechanics. *Pure and applied geophysics*, 166(10-11):1693–1724, 2009.
- [Ons31] L. Onsager. Reciprocal relations in irreversible processes. i. *Physical review*, 37(4):405, 1931.
- [Per63] Piotr Perzyna. The constitutive equations for rate sensitive plastic materials. *Quarterly of applied mathematics*, 20(4):321–332, 1963.
- [Ter43] K. Terzaghi. *Theory of consolidation*. Wiley Online Library, 1943.
- [Zie77] H. Ziegler. *An introduction to thermomechanics*. North-Holland series in applied mathematics and mechanics. North-Holland Pub. Co., 1977.

---

# Energetical background of common approaches in geomechanics

**Eleni Gerolymatou**

*Chalmers University of Technology*

---

*As is well known, mechanics in general is based on three basic conservation laws, those of mass, momentum and energy. Of those, the first two can be derived from the last, when one considers the objectivity with respect to the observer as a prerequisite. A further energy related concept significant for modeling is the second law of thermodynamics. This chapter discusses several common methods and concepts used in geomechanics, with respect to their relationship to the first and second law of thermodynamics. The aim is to provide the reader with a better understanding of the underlying assumptions of these approaches. The various methodologies presented are subdivided into two categories, principles based on the concept of energy conservation and variational principles. In the first category the derivation of the mass, momentum etc. balance equations from the energy balance by means of assuming a change of observer and the derivation of the micromechanical formulation of stress measures are discussed, as well as a brief discussion on energy based upscaling. In the category of variational principles, the minimum potential energy, the maximum plastic work and the second order work are outlined. The virtual work method is omitted, as it has been discussed in more detail elsewhere in this book.*

## 1 Introduction

Though the concept of energy is a rather new one for humanity, the idea that Nature has some goal appeared a long time ago. Aristoteles claimed in Physics that Nature follows the easiest path or, equivalently, the one requiring the smallest amount of effort in all its manifestations. The idea may be even older, since Euler [Eul53] in turn claims that Aristoteles most likely borrowed this dogma from his predecessors, rather than developed it independently. The same principle today remains at the basis of variational principles [Ber09]. The notion of a quantity that is eternally conserved was first introduced apparently by Heraclitus in the 6th century B.C., who denoted this quantity as *fire* [Sch68], though not in the literal sense. Several centuries later,

Leonardo da Vinci [Mac41] and Girolamo Cardano [Car51] both remarked on the fact that it is impossible in terrestrial phenomena to get something for nothing, implying again the idea of a quantity that must be conserved.

Which quantity is conserved has not always been obvious. René Descartes, on the basis of his experiments on the impacts of bodies, formulated what he called the conservation of the quantity of motion, known today as momentum conservation [Lag88]. Leibniz [Lei86] disagreed with him and expressed the view that the quantity conserved is the product of the mass and the second power of the velocity, a quantity he denoted as *vis viva*, or living force. To account for slowing due to friction, Leibniz claimed that heat consisted of the random motion of the constituent parts of matter. A controversy was sparked that lasted more than a century. By many the controversy is thought to have been resolved by Lagrange [Lag88], who endeavored to express mechanics in a mathematical manner analogous to a geometry of four dimensions, the time and the three space coordinates. He derived the conservation of the *vis viva* from the laws of Newton.

The term 'energy' was first introduced by Thomas Young [You07], who defined in 1807 the energy of a body as the product of its mass into the square of its velocity. Gustav-Gaspard Coriolis introduced the term 'kinetic energy' in 1829, while William Rankine introduced 'potential energy' as a concept in 1853. The laws of thermodynamics were finally formulated by William Thomson, Lord Kelvin, in 1851 [Tho51], who claimed that there is an entity (energy) which is conserved, and a different one (motivity -concentration of energy) which is lost. This is the great generalization expressed as the 'universal tendency in nature to the dissipation of energy' and led to the mathematical formulation of the concept of energy by Rudolf Clausius. The interested reader may find more details on the history of the concept of energy in [Lin71].

Even in our days, with the conservation of energy firmly established, the concept remains hazy. Richard Feynman stated [Fey64]

That is a most abstract idea, it is not a description of a mechanism, or anything concrete; it is just a strange fact that we can calculate some number and when we finish watching nature go through her tricks and calculate the number again, it is the same.

In fact, both the first and the second law of thermodynamics can be classified as empirical, as they are based on observations. Even so, no exceptions to their validity are known and they form the basis for a wide range of concepts and methods in a wide variety of fields, among which mechanics are also to be found.

## 2 Conservation principles

### 2.1 Derivation of balance equations

The notion of using the energy rate invariance as a tool to extract the basic principles of mechanics dates back to 1959 and was first introduced by W. Noll [Nol59], though Lagrange [Lag88] showed already in 1788 that the opposite was also possible. The idea was later used also by other authors, see [GR64, PG97]. The scope here is to apply the energy rate invariance to the domain of granular assemblies, in order to extract the basic balance equations in a natural way. Note that it is the energy rate that is invariant to non-inertial changes of observer, rather than the energy. The derivation of the basic conservation laws of mechanics from the energy rate balance is here illustrated using the simple example of a single grain.

The power balance equation in its general form reads

$$\dot{E} = \dot{W}_F + \dot{W}_C + \dot{W}_q - \dot{E}_{kin} - \dot{E}_{el} - \dot{Q} - D = 0 \quad (1)$$

where  $W_F$  is the work of the forces,  $W_C$  is the work of the couples,  $W_q$  is the amount of heat inflow to the system,  $E_{kin}$  is the kinetic energy of the grain,  $E_{el}$  is the stored in it elastic energy,  $Q$  is the heat of it and  $D$  is the rate of energy dissipation. In the present work heat will be neglected. As a result the above can be rewritten as

$$\dot{E} = \dot{W}_F + \dot{W}_C - \dot{E}_{kin} - \dot{E}_{el} - D = 0 \quad (2)$$

Note that, as observed in [Nol59], the kinetic energy rate corresponds to the work rate originating in inertial forces, which can be viewed as the result of the interaction of our system with the rest of the bodies in the universe. In that sense, the above representation of the energy rate is not the only one admissible. This point is extensively discussed in [PG97], where the form of each of the energy terms is also discussed.

The energy rate that should be maintained between observers is

$$\dot{\hat{E}} = \dot{\hat{W}}_F + \dot{\hat{W}}_C - \dot{\hat{E}}_{kin} \quad (3)$$

When considering a single grain, this is characterized by its mass, a scalar measure  $m$ , the position of its center of mass  $\mathbf{x}$ , its linear velocity  $\mathbf{v}$ , its rotational velocity  $\mathbf{w}$  and its tensor of angular inertia  $\underline{\theta}$ . Since rigid grains are discussed, there is a single force and a single couple to which all forces and couples acting on the single grain can be reduced, with their center of mass as a reference point. Therefore

$$\dot{\hat{W}}_F + \dot{\hat{W}}_C = \mathbf{f} \cdot \mathbf{v} + \mathbf{m} \cdot \mathbf{w} \quad (4)$$

For the kinetic energy rate, on the other hand, the following holds

$$\dot{\hat{E}}_{kin} = \frac{1}{2} \dot{m} \mathbf{v} \cdot \mathbf{v} + m \mathbf{v} \cdot \dot{\mathbf{v}} + \frac{1}{2} \mathbf{w}^T \dot{\underline{\theta}} \mathbf{w} + \mathbf{w}^T \underline{\theta} \dot{\mathbf{w}} \quad (5)$$

Two changes of observer are considered. The first moves at a constant linear velocity with respect to the initial system, while the second moves at a constant rotational velocity with respect to the initial system.

### 2.1.1 Galilean change of observer

An observer moving at velocity  $\mathbf{a}$  with respect to the initial system is considered. The rotational velocity and the rotational inertia observed by the second observer are the same as those observed by the first one. The same holds for the moments:

$$\underline{\boldsymbol{\theta}}' = \underline{\boldsymbol{\theta}} \quad \& \quad \mathbf{w}' = \mathbf{w} \quad \& \quad \mathbf{m}' = \mathbf{m} \quad (6)$$

The transformation rule for the force remains unknown, while that of the velocity reads

$$\mathbf{v}' = \mathbf{v} - \mathbf{a} \quad (7)$$

Then the invariance requirement yields

$$\begin{aligned} \mathbf{f} \cdot \mathbf{v} + \mathbf{m} \cdot \mathbf{w} - \mathbf{f}' \cdot \mathbf{v}' - \mathbf{m} \cdot \mathbf{w} &= \frac{1}{2} \dot{m} \mathbf{v} \cdot \mathbf{v} + m \mathbf{v} \cdot \dot{\mathbf{v}} + \\ &+ \frac{1}{2} \mathbf{w}^T \underline{\boldsymbol{\theta}} \dot{\mathbf{w}} + \mathbf{w}^T \underline{\boldsymbol{\theta}} \dot{\mathbf{w}} - \frac{1}{2} \dot{m} \mathbf{v}' \cdot \mathbf{v}' - m \mathbf{v}' \cdot \dot{\mathbf{v}}' - \frac{1}{2} \mathbf{w}'^T \underline{\boldsymbol{\theta}} \dot{\mathbf{w}} - \mathbf{w}'^T \underline{\boldsymbol{\theta}} \dot{\mathbf{w}} \Rightarrow \\ \mathbf{f} \cdot \mathbf{v} - \mathbf{f}' \cdot \mathbf{v}' &= \frac{1}{2} \dot{m} \mathbf{v} \cdot \mathbf{v} + m \mathbf{v} \cdot \dot{\mathbf{v}} - \frac{1}{2} \dot{m} \mathbf{v}' \cdot \mathbf{v}' - m \mathbf{v}' \cdot \dot{\mathbf{v}}' \Rightarrow \end{aligned}$$

$$\begin{aligned} \mathbf{f} \cdot \mathbf{v} - \mathbf{f}' \cdot \mathbf{v}' + \mathbf{f}' \cdot \mathbf{a} &= \frac{1}{2} \dot{m} \mathbf{v} \cdot \mathbf{v} + m \mathbf{v} \cdot \dot{\mathbf{v}} + \\ &- \frac{1}{2} \dot{m} (\mathbf{v} - \mathbf{a}) \cdot (\mathbf{v} - \mathbf{a}) + \\ &- m (\mathbf{v} - \mathbf{a}) \cdot (\dot{\mathbf{v}} - \mathbf{a}) \Rightarrow \end{aligned}$$

$$\begin{aligned} \mathbf{f}' \cdot \mathbf{a} &= (\mathbf{f}' - \mathbf{f}) \cdot \mathbf{v} + \\ &+ \dot{m} \mathbf{a} \cdot \mathbf{v} + m \mathbf{a} \cdot \dot{\mathbf{v}} - \frac{1}{2} \dot{m} \mathbf{a} \cdot \mathbf{a} \end{aligned}$$

Since  $\mathbf{a}$  is arbitrarily selected, the above equation can be decomposed as follows:

$$\mathbf{f}' = \mathbf{f} \quad (8)$$

$$\mathbf{f} = \dot{m} \mathbf{v} + m \dot{\mathbf{v}} \quad (9)$$

$$\dot{m} = 0 \quad (10)$$

yielding from top to bottom the force balance, the momentum balance and the mass balance.

### 2.1.2 Leibniz change of observer

The Leibniz change of observer refers to an observer moving at constant rotational speed and zero translational speed with respect to the first one. An observer moving at rotational speed  $\mathbf{b}$  with respect to the initial system is now considered. Then, at given

time  $t$ , the axes of the second system will be rotated with respect to those of the initial one by a rotation  $\underline{\mathbf{O}}(t)$ , to be signified from now on as  $\underline{\mathbf{O}}$  for simplicity.

The rate of change in rotation will be

$$\begin{aligned}\dot{\underline{\mathbf{O}}} &= \lim_{dt \rightarrow 0} \frac{\underline{\mathbf{O}}(t+dt) - \underline{\mathbf{O}}(t)}{dt} \Rightarrow \\ \dot{\underline{\mathbf{O}}} &= \lim_{dt \rightarrow 0} \frac{d\underline{\mathbf{O}} - \underline{\mathbf{I}}}{dt} \underline{\mathbf{O}} \Rightarrow \\ \dot{\underline{\mathbf{O}}} &= \underline{\mathbf{W}} \underline{\mathbf{O}} \Rightarrow \\ \underline{\mathbf{W}} &= \dot{\underline{\mathbf{O}}} \underline{\mathbf{O}}^T\end{aligned}$$

where in the specific case  $\underline{\mathbf{W}}$  is a rotational velocity tensor such that

$$\underline{\mathbf{W}}\underline{\mathbf{R}} = \mathbf{b} \times \underline{\mathbf{R}}$$

for any vector  $\underline{\mathbf{R}}$ . Note that between  $\underline{\mathbf{W}}$  and  $\mathbf{b}$  the following property holds:

$$\underline{\mathbf{W}} = -\underline{\mathbf{E}}^T \mathbf{b}, \quad \mathbf{b} = -\frac{1}{2} \underline{\mathbf{E}} \underline{\mathbf{W}}$$

where  $\underline{\mathbf{E}}$  is the third order Ricci tensor.

The second derivative of  $\underline{\mathbf{O}}$  is now considered.

$$\begin{aligned}\ddot{\underline{\mathbf{O}}} &= \lim_{dt \rightarrow 0} \frac{\dot{\underline{\mathbf{O}}}(t+dt) - \dot{\underline{\mathbf{O}}}(t)}{dt} \Rightarrow \\ \ddot{\underline{\mathbf{O}}} &= \lim_{dt \rightarrow 0} \frac{\underline{\mathbf{W}}(t+dt)\underline{\mathbf{O}}(t+dt) - \underline{\mathbf{W}}(t)\underline{\mathbf{O}}(t)}{dt}\end{aligned}$$

In the specific case  $\underline{\mathbf{W}}$  is constant in time, so

$$\begin{aligned}\ddot{\underline{\mathbf{O}}} &= \lim_{dt \rightarrow 0} \frac{\underline{\mathbf{W}}\underline{\mathbf{O}}(t+dt) - \underline{\mathbf{W}}\underline{\mathbf{O}}(t)}{dt} \Rightarrow \\ \ddot{\underline{\mathbf{O}}} &= \underline{\mathbf{W}} \lim_{dt \rightarrow 0} \frac{\underline{\mathbf{O}}(t+dt) - \underline{\mathbf{O}}(t)}{dt} \Rightarrow \\ \ddot{\underline{\mathbf{O}}} &= \underline{\mathbf{W}} \dot{\underline{\mathbf{O}}}\end{aligned}$$

Additionally, as is well known for any rotation tensor

$$\begin{aligned}\underline{\mathbf{O}} \underline{\mathbf{O}}^T &= \underline{\mathbf{I}} \Rightarrow \\ \dot{\underline{\mathbf{O}}} \underline{\mathbf{O}}^T + \underline{\mathbf{O}} \dot{\underline{\mathbf{O}}}^T &= \underline{\mathbf{0}} \Rightarrow \\ \underline{\mathbf{W}} + \underline{\mathbf{W}}^T &= \underline{\mathbf{0}} \Rightarrow \\ \underline{\mathbf{W}} &= -\underline{\mathbf{W}}^T \Rightarrow \\ \underline{\mathbf{W}} &\in \mathbb{S}k\mathbf{w}\end{aligned}\tag{11}$$

As far as the relationship between the quantities of the first system and those of the second system is concerned, the following is known:

$$m' = m, \quad \dot{m}' = \dot{m} \quad (12)$$

$$\mathbf{x}' = \underline{\mathbf{Q}}^T \mathbf{x} \quad (13)$$

$$\dot{\mathbf{x}}' = \underline{\dot{\mathbf{Q}}}^T \mathbf{x} + \underline{\mathbf{Q}}^T \dot{\mathbf{x}} \Rightarrow \mathbf{v}' = \underline{\dot{\mathbf{Q}}}^T \mathbf{x} + \underline{\mathbf{Q}}^T \mathbf{v} \quad (14)$$

$$\dot{\mathbf{v}}' = \underline{\ddot{\mathbf{Q}}}^T \mathbf{x} + \underline{\dot{\mathbf{Q}}}^T \dot{\mathbf{x}} + \underline{\dot{\mathbf{Q}}}^T \mathbf{v} + \underline{\mathbf{Q}}^T \dot{\mathbf{v}} \Rightarrow$$

$$\dot{\mathbf{v}}' = \underline{\dot{\mathbf{Q}}}^T \underline{\mathbf{W}}^T \mathbf{x} + 2\underline{\dot{\mathbf{Q}}}^T \mathbf{v} + \underline{\mathbf{Q}}^T \dot{\mathbf{v}} \quad (15)$$

$$\mathbf{w}' = \underline{\mathbf{Q}}^T (\mathbf{w} - \mathbf{b}) \quad (16)$$

$$\dot{\mathbf{w}}' = \underline{\dot{\mathbf{Q}}}^T (\mathbf{w} - \mathbf{b}) + \underline{\mathbf{Q}}^T \dot{\mathbf{w}} \quad (17)$$

For the transformation of the forces and the moments the rule for vectors is used:

$$\mathbf{f}' = \underline{\mathbf{Q}}^T \mathbf{f}, \quad \mathbf{m}' = \underline{\mathbf{Q}}^T \mathbf{m} \quad (18)$$

For  $\underline{\boldsymbol{\theta}}$ ,  $\underline{\boldsymbol{\theta}}'$ ,  $\underline{\dot{\boldsymbol{\theta}}}$  and  $\underline{\dot{\boldsymbol{\theta}}}'$  the assumption is made that they are symmetric.

To begin with, the form of the kinetic energy in the second system is evaluated. Introducing in the kinetic energy rate equation the above equations one gets

$$\begin{aligned} \dot{E}'_{kin} &= \frac{1}{2} \dot{m} \left[ \underline{\dot{\mathbf{Q}}}^T \mathbf{x} + \underline{\mathbf{Q}}^T \mathbf{v} \right]^T \left[ \underline{\dot{\mathbf{Q}}}^T \mathbf{x} + \underline{\mathbf{Q}}^T \mathbf{v} \right] + \\ &+ m \left[ \underline{\dot{\mathbf{Q}}}^T \mathbf{x} + \underline{\mathbf{Q}}^T \mathbf{v} \right]^T \left[ \underline{\dot{\mathbf{Q}}}^T \underline{\mathbf{W}}^T \mathbf{x} + 2\underline{\dot{\mathbf{Q}}}^T \mathbf{v} + \underline{\mathbf{Q}}^T \dot{\mathbf{v}} \right] + \\ &+ \frac{1}{2} \left[ \mathbf{w}^T - \mathbf{b}^T \right] \underline{\mathbf{Q}} \underline{\dot{\boldsymbol{\theta}}}' \underline{\mathbf{Q}}^T \left[ \mathbf{w} - \mathbf{b} \right] + \\ &+ \frac{1}{2} \left[ \mathbf{w}^T - \mathbf{b}^T \right] \underline{\mathbf{Q}} \underline{\boldsymbol{\theta}}' \left[ \underline{\dot{\mathbf{Q}}}^T (\mathbf{w} - \mathbf{b}) + \underline{\mathbf{Q}}^T \dot{\mathbf{w}} \right] + \\ &+ \frac{1}{2} \left[ (\mathbf{w}^T - \mathbf{b}^T) \underline{\dot{\mathbf{Q}}} + \dot{\mathbf{w}}^T \underline{\mathbf{Q}} \right] \underline{\boldsymbol{\theta}}' \underline{\mathbf{Q}}^T \left[ \mathbf{w} - \mathbf{b} \right] \end{aligned}$$

Using the above and setting the energy rate in the new system equal to the energy rate in the reference system yields the following balance equations

$$\underline{\dot{\boldsymbol{\theta}}} = \underline{\mathbf{W}} \underline{\boldsymbol{\theta}} + \underline{\boldsymbol{\theta}} \underline{\mathbf{W}} \quad (19)$$

$$\underline{\mathbf{m}} = \underline{\dot{\boldsymbol{\theta}}} \underline{\mathbf{w}} + \underline{\boldsymbol{\theta}} \underline{\dot{\mathbf{w}}} \quad (20)$$

for the angular inertia tensor and the angular momentum respectively, where  $\underline{\mathbf{W}}$  is the rotational velocity tensor corresponding to the vector  $\mathbf{w}$ .

Though the example given here is a very simple one, the same principle can be used to derive the balance equations in significantly more complicated cases, such as multiphase media.

## 2.2 Micromechanical stress tensors

In several works [CMNN81, CL90, BV01], the sum of the work rates of the contacts is set equal to the total work rate, in an attempt to derive the micromechanical formulation of the stress, the couple stress and the hyper stress tensors in granular media on the basis of the inter-granular forces. The methodology is briefly outlined and critically discussed here.

For the sake of simplicity, the present analysis is restricted to a Cosserat continuum. A more extended analysis may be found in [FG10]. The same approach has been applied to so called non-simple materials [Nol72] to derive the formulation of the corresponding higher order stress measures, for example for micropolar media [CL90, BV01, TW02] and couple stress tensors. An assembly of grains is considered, with a set  $\mathcal{C}$  of ordered contacts  $a$ . The forces and couples at the contacts are denoted as  $\mathbf{f}^a$  and  $\mathbf{m}^a$  respectively. A reference point  $O$  is defined. The assumption is made that the displacement and rotation rates are affine, which means the displacement rate of the grain ( $k$ ) can be expressed in the form of a Taylor expansion about the center  $O$  of the assembly as

$$\mathbf{u}^{(k)} = \mathbf{u}^O + \nabla \mathbf{u}^O \cdot \mathbf{x}^{(k)} \quad (21)$$

for the displacement rates and

$$\mathbf{w}^{(k)} = \mathbf{w}^O + \nabla \mathbf{w}^O \cdot \mathbf{x}^{(k)} \quad (22)$$

for the rotation rates. The relative velocity of the grains ( $i$ ) and ( $j$ ) at their contact  $a$  will then be

$$\mathbf{v}^a = \mathbf{v}^{(i),c} - \mathbf{v}^{(j),c} \quad (23)$$

$$\begin{aligned} \mathbf{v}^a &= \nabla \mathbf{u}^O \cdot (\mathbf{x}^{(i)} - \mathbf{x}^{(j)}) + \mathbf{w}^O \times (\mathbf{x}^{(j)} - \mathbf{x}^{(i)}) + \\ &+ \nabla \mathbf{w}^O \cdot (\mathbf{x}^{(i)} - \mathbf{x}^O) \times (\mathbf{x}^a - \mathbf{x}^{(i)}) + \\ &- \nabla \mathbf{w}^O \cdot (\mathbf{x}^{(j)} - \mathbf{x}^O) \times (\mathbf{x}^a - \mathbf{x}^{(j)}) \end{aligned} \quad (24)$$

Since the internal power is invariant to the selection of the contact point, no loss in generality results from assuming it to be located at mid-distance between the centers of the two grains. As a result

$$\mathbf{v}^a = \nabla \mathbf{u}^O \cdot \mathbf{l}^a - \mathbf{w}^O \times \mathbf{l}^a + \nabla \mathbf{w}^O \cdot (\mathbf{x}^a - \mathbf{x}^O) \times \mathbf{l}^a \quad (25)$$

where

$$\mathbf{l}^a = \mathbf{x}^{(i)} - \mathbf{x}^{(j)} \quad (26)$$

The relative rotation is simply

$$\mathbf{w}^a = \nabla \mathbf{w}^O \cdot (\mathbf{x}^{(i)} - \mathbf{x}^{(j)}) \quad (27)$$

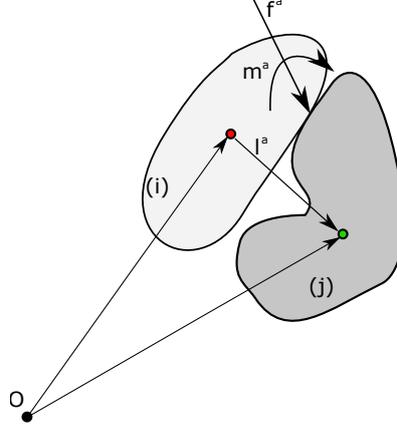


Figure 1: Contact between two grains.

This means that the rate of the internal work may be written as

$$\begin{aligned}
 P_{int} &= \sum_{a \in \mathcal{C}} \left( (\mathbf{f}^a \otimes \mathbf{l}^a) : (\nabla \mathbf{v}^O - \underline{\mathbf{W}}^O) \right) \\
 &+ \sum_{a \in \mathcal{C}} \left( (\mathbf{f}^a \times \mathbf{l}^a) \otimes (\mathbf{x}^a - \mathbf{x}^O) : \nabla \mathbf{w} \right) \\
 &+ \sum_{a \in \mathcal{C}} (\mathbf{m}^a \otimes \mathbf{l}^a : \nabla \mathbf{w}^O)
 \end{aligned}$$

where  $\underline{\mathbf{W}}^{(0)}$  is the tensorial form of the rotation vector  $\mathbf{w}^{(0)}$ .

For a continuum the internal power of a Cosserat medium is known to be

$$\bar{P}_{int} = V \left( \underline{\boldsymbol{\sigma}} : \underline{\dot{\boldsymbol{\Gamma}}} + \underline{\boldsymbol{\mu}} : \underline{\dot{\boldsymbol{\kappa}}} \right) \quad (28)$$

where  $\underline{\boldsymbol{\Gamma}}$  is the generalized strain,  $\underline{\boldsymbol{\mu}}$  is the couple stress tensor and  $\underline{\boldsymbol{\kappa}}$  is the curvature. This would lead one to define

$$\begin{aligned}
 \underline{\dot{\boldsymbol{\Gamma}}} &= \nabla \mathbf{v} - \underline{\mathbf{W}} \\
 \underline{\boldsymbol{\sigma}} &= \sum_{(i,j) \in \mathcal{C}} \mathbf{f}^a \otimes \mathbf{l}^a \\
 \underline{\dot{\boldsymbol{\kappa}}} &= \nabla \mathbf{w}
 \end{aligned}$$

and

$$\underline{\boldsymbol{\mu}} = \frac{1}{V} \sum_{a \in \mathcal{C}} (\mathbf{m}^a \otimes \mathbf{l}^a) + \frac{1}{V} \sum_{a \in \mathcal{C}} ((\mathbf{f}^a \times \mathbf{l}^a) \otimes \mathbf{x}^a) \quad (29)$$

In the above, the dependence of the couple stress tensor on the selection of the reference point is clear. This result is in agreement with the work of different authors

[CL90, BV01]. It is however counter-intuitive, in that it shows a strong dependence of the result on the choice of the reference point. In fact, contacts further from the reference point seem to contribute more significantly to the final result.

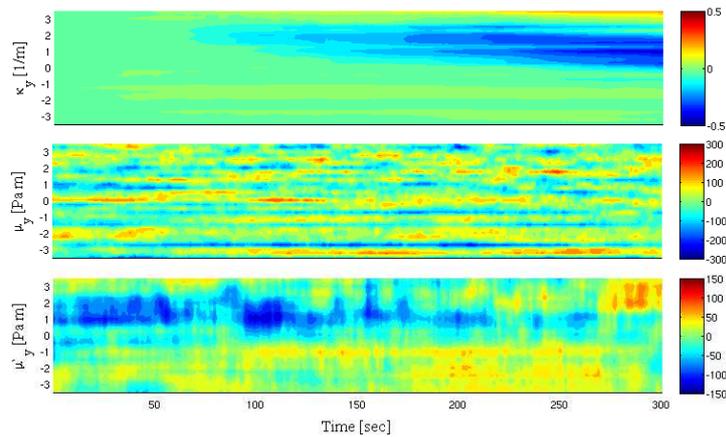


Figure 2: Comparison of different couple stress tensor formulations [Ger11].

An alternative using only the first neighbors of each grain was suggested by Tordesillas and Walsch [TW02]. This eliminates the dependence on the reference point, as the expansion is now performed around a single grain. A comparison of the two different measures is given in figure 2 for a simple shear test on a specimen of hexagons with periodic boundary conditions. The code LMG90 was used. The curvature in the vertical direction is shown at the top, while the standard couple stress formulation given here [CL90, BV01] and that of a variation of the one by Tordesillas and Walsch [TW02, Ger11] follow. As may be observed, the results are very different. While part of the difference may be attributed to fluctuations due to the discrete character of the simulation, there can be no doubt that the two formulations provide very different results. There is no doubt that such differences will also arise when considering other stress measures, such as the hyper stress tensor.

The assumption of affinity is of course not a valid one, but rather only an approximation. It has been shown from DEM simulations [AMVH<sup>+</sup>06, Kru14] and experiments [CCL97] that in reality displacements deviate significantly from this idealization. Even so, the discrepancies in this case do not arise from it. The well known formula for the stress tensor that has resulted from this consideration, under the assumption that the displacement field is affine is used widely and has been found to be in excellent agreement with the stress tensor evaluated from the forces applied at the boundary in discrete element simulations. Moreover, similar approaches using instead the elastic energy at the contacts show no dependence on the reference point [Ger14].

The variations seem to arise mostly from the process of equating a micromechanical

quantity, such as the virtual power of the contacts, to a continuum quantity, in this case the virtual power of a Cosserat continuum. Extreme care should be taken when using this approach to make sure that the quantities equated to the micromechanical ones truly correspond to their continuum equivalents. A more detailed discussion on the comparison between micromechanical and continuum measures may be found in [FTV06]. So, while there is no doubt of the correctness of the concept of setting the internal work equal to the external work, care needs to be exercised in the details, especially when passing from a discrete to a continuum formulation.

### 3 Variational methods

A variational principle is an assertion stating that some quantity, that is defined for all possible processes that could potentially take place, reaches its minimum (or maximum, or stationary) value for the real process [Ber09]. The notion stems from the idea, that the observable events are extreme in their character and that the general principles sought are variational, i.e. they assert that certain parameters obtain their maximum or minimum values in realizable physical processes. A simple example is the intuitive understanding that Nature, given the chance, will follow a path of least resistance to the final state, expressed loosely as the ‘principle of least action’.

Variational principles are not directly linked to either the first or the second law of thermodynamics and are usually empirical in nature. Some of the more commonly used ones in the field of mechanics will be discussed here.

#### 3.1 Minimum potential energy

The potential energy of a body with respect to a reference configuration can be defined as the work done in moving the body from the reference configuration to its new one. It has the following characteristics [Kel05]:

1. a force field exists
2. to move something in the force field, work must be done
3. the force field is conservative
4. the force field itself does negative work when another force is moving something against it
5. it is recoverable energy

An example is shown in figure 3. Taking into account gravity, the potential energy of the first sphere is larger than that of the second one. In the absence of gravity, both spheres have the same potential energy. In the same way, considering internal energy, one may consider a very loose granular assembly in a box, with the particles at a significant distance from each other. Under this definition, reducing the size of

the box does not lead to a change in internal energy, as no work has been expended for the change in the configuration. On the other hand, if the particles of the granular assembly were acting as a gas with a degree of kinetic energy, an increase in internal energy would follow the reduction of the size of the box. Correspondingly, the poten-

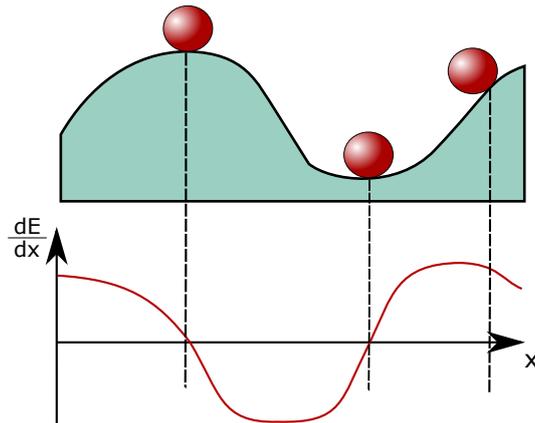


Figure 3: Equilibrium and potential energy gradient.

tial energy of a deformable body may be viewed as its elastic energy, which results from the field of the internal stresses.

It is a well known principle, that a body will be in equilibrium, when it is located at a position corresponding to zero gradient of the potential energy. Examples in figure 3 are the first sphere (located at a local maximum) and the second sphere (located at a local minimum), while the third sphere is not in equilibrium. The bodies located at local minima, like the second sphere, are in addition stable. Thus, one speaks of stable and unstable equilibrium points.

The principle of minimum potential energy follows directly from the principle of virtual work for elastic materials. It is used to obtain approximate solutions to problems and is also known as the Rayleigh-Ritz method. The underlying idea is exceedingly simple. A body under a certain load will reach equilibrium and, provided nothing prohibits it from it, it will endeavor to reach a stable equilibrium. If the loads are known, the field of displacements will be the one that maximizes or minimizes the potential energy. It is usually assumed that the equilibrium is stable and that the potential energy will be minimized. For the field of displacements, trial functions are used. As a result, the solution evaluated with the Rayleigh-Ritz method is an approximation of the actual solution.

Examples are not included here, as such may be found in the chapter of this book that is treating virtual work.

### 3.2 Maximum plastic work principle

The domain in which the deformations of geomaterials are reversible is at best limited. To predict the mechanical response of such materials, the most widely used approach is elasto-plasticity. The strains, and also their increments, are most often decomposed additively in an elastic and a plastic part

$$\dot{\epsilon}_{ij} = \dot{\epsilon}_{ij}^e + \dot{\epsilon}_{ij}^p \quad (30)$$

The yield function  $f(\underline{\sigma}, :)$  is used to determine whether the deformation rate is elastic. Elastic (reversible) strains only occur when

$$f(\underline{\sigma}) < 0 \quad (31)$$

whereas plastic strain increments become non-zero when the corresponding equality holds.

Within the frame of elasto-plastic theories, it is common to assume that the plastic strain is analogous to the gradient of the plastic potential function in the stress space:

$$\dot{\epsilon}_{ij}^p = \dot{\lambda} \frac{\partial g}{\partial \sigma_{ij}} \quad (32)$$

where  $g(\underline{\sigma}, :)$  stands for the plastic potential function [Yu06]. When the functions  $f(\underline{\sigma}, :)$  and  $g(\underline{\sigma}, :)$  coincide, one speaks of an associated flow rule or of the normality rule, because the strain increment is normal to the yield surface. This plastic flow rule was based on the observation in 1870 by de Saint-Venant [dSV71] that for metals the principal axes of the plastic strain rate coincide with those of the stress. This is the so-called coaxial assumption, which has been the foundation of almost all the plasticity models for metals. An example after [IaH11] is shown in figure 4 for biaxial tests on rolled titanium sheets. The yield surface has been interpolated using splines, while for the strain directions the normality rule has been used.

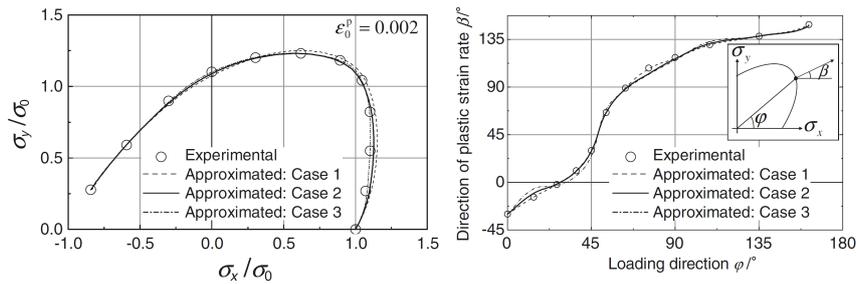


Figure 4: Yield surface (left) and flow directions (right) after [IaH11].

The normality rule represents a maximum for the plastic work, as shown by Hill [Hil48]. The proof is briefly outlined here. Let us assume that the plastic strain rate

$\dot{\epsilon}_{ij}$  is given and is determined from the normality rule and the yield criterion at the corresponding stress state,  $\sigma_{ij}$ , represented by a point P in the stress space, figure 5. If  $\sigma_{ij}^*$  is an arbitrary state of stress represented by a point P\* on or inside the yield surface, then the difference between the incremental plastic works done by the two stress states on the actual plastic strain rate is

$$\dot{W}_p = (\sigma_{ij} - \sigma_{ij}^*) \dot{\epsilon}_{ij}^p \quad (33)$$

The above represents the scalar product of the vector P\*P and the plastic strain rate. If the yield surface is strictly convex, the angle between these vectors is acute and the scalar product is positive. Therefore

$$\dot{W}_p = (\sigma_{ij} - \sigma_{ij}^*) \dot{\epsilon}_{ij}^p \geq 0 \quad (34)$$

which means that the plastic work has a maximum for associated flow rule, under the condition that the yield surface is convex. This is known as the maximum plastic work principle or theorem. Though presented here as given by Hill [Hil48], it seems to have been originally proven by von Mises [vM28].

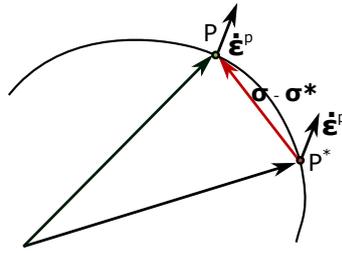


Figure 5: Potential energy gradient.

In short, the maximum plastic work principle is a mathematical statement of the following two important ideas: (a) The yield surface is convex; (b) The plastic strain rate (or increment) is normal to the yield surface.

It must be noted that a large amount of experimental data suggests that the coaxial assumption is generally not valid for soils and rocks. The discrepancy is especially large when it comes to stress rotations. An example after [GI00] is shown in figure 6. Efforts have been made in recent years to formulate the elasto-plastic response in such a way as to satisfy the maximum plastic work principle, for example in [CH97, Kra09, Fra18] and it has been shown that it is possible to recover associativity at the cost of introducing a dependence on internal variables of the functions describing the internal energy or the dissipation.

It should be remarked at this point, that, though associativity is an empirical observation for metals, the requirement that the plastic work be maximum is not one resulting

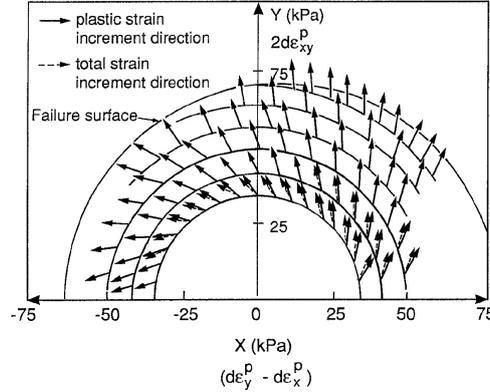


Figure 6: Unit plastic strain increment vectors obtained from pure principal rotation tests on dense Toyoura sand after [GI00].

from the first or the second thermodynamic law. It is rather an assumption that has proven to be very useful, but is currently being debated. The intuitive justification may be seen as follows: a body on which external work is imposed will, in agreement with the requirement for energy balance, tend to either store this energy in the form of internal energy or to dissipate it. Maximizing the dissipation is thus equivalent to assuming that the body will strive to reach a minimum internal energy, or potential energy, in the absence of kinetic energy, reaching thus a stable equilibrium as fast as possible. Minima in potential energy are empirically accepted as attractors, but there is no proof that the path to these attractors is the shortest possible.

From the point of view of the second law of thermodynamics, the restriction on the direction of the plastic strains is much more relaxed, requiring only the dissipation to remain non-negative at any given time. This would mean that the direction of the plastic strains cannot deviate by more than  $90^\circ$  from the direction of the stresses.

## 4 Second order work

Hill [Hil58] stated that, under certain assumptions, the equilibrium of the body occupying the volume  $V$  is stable if

$$\int_V \dot{\sigma}_{ij} \frac{\partial \dot{u}_i}{\partial x_j} dV > 0, \quad \forall \dot{u}_i(x_j) \quad (35)$$

The underlying assumptions are mainly the small strain assumption and the assumption of quasi static deformation. This gave rise to the so called second order work

criterion, which is expressed as follows: If

$$d^2W = \dot{\sigma}_{ij}\dot{\epsilon}_{ij} > 0, \quad \forall \dot{\epsilon}_{ij} \quad (36)$$

then the equilibrium is stable.

It is obvious that the local and the global formulations are not equivalent. The global form, equation (35), can still be satisfied, if there exist points not satisfying the local form, equation (36), as long as they are few enough. The satisfaction of the local form at one point of the body occupying the volume  $V$  does not provide information on the global form, equation (35). If the stress, strain and their increments are uniform in the volume  $V$ , the two formulations are equivalent. The distinction is ignored here.

The second work criterion is classified as a stability criterion. Stability is however an ambiguous term. Lyapunov [Lya92] defined a system as stable when small perturbations in the initial conditions will lead to finite perturbations of the final solutions. Let us consider once more figure 3. The second sphere is obviously in a stable state. The first sphere from the right on the other hand is stable, if the length of the slopes on its right and on its left are bounded, or, more simply, if the peak it finds itself at is a local but not a global maximum of the height. From the point of view of potential energy, the sphere is of course in a state of unstable equilibrium. Thus a more suitable definition for the case of continua would be to require the perturbation of the final solution to a continuous function of the perturbation of the initial condition. A more extensive discussion on the topic of stability definition, validity of the criterion and comparison to other formulations may be found in the extensive literature, for example [IN98, CCV04, NDK07]. The present section is limited to the connection of the formulation to energetic considerations.

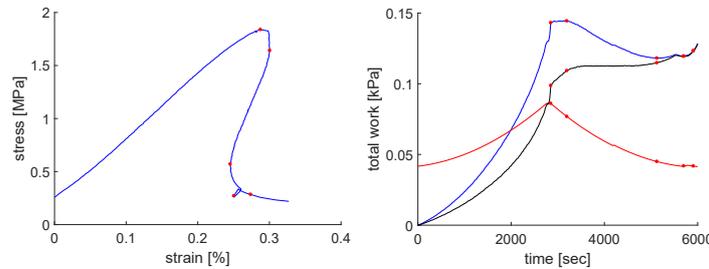


Figure 7: Uniaxial compression test (left) and corresponding work measures (right).

The work rate, or power, for the stresses reads

$$\dot{W} = \sigma_{ij}\dot{\epsilon}_{ij} \quad (37)$$

Considering the rate of the above equation yields

$$\ddot{W} = \dot{\sigma}_{ij}\dot{\epsilon}_{ij} + \sigma_{ij}\ddot{\epsilon}_{ij} = d^2W + \sigma_{ij}\ddot{\epsilon}_{ij} \quad (38)$$

It is obvious that the second order work is the rate of the power, if one neglects accelerations. Since the application is as a rule restricted to the quasi-static regime, accelerations are usually negligible. The requirement for a positive second order work is thus equivalent to the requirement for a strictly increasing work rate. As the second order work criterion requires the second order work to be positive for every possible strain rate direction, it requires the work rate to be increasing also for every possible strain rate direction. It should be remarked at this point that the criterion states that the equilibrium is stable, when equation (36) is satisfied. This does not mean that the equilibrium is necessarily unstable when equation (36) is not satisfied, meaning that a non positive second order work is a necessary but not sufficient condition for unstable equilibrium.

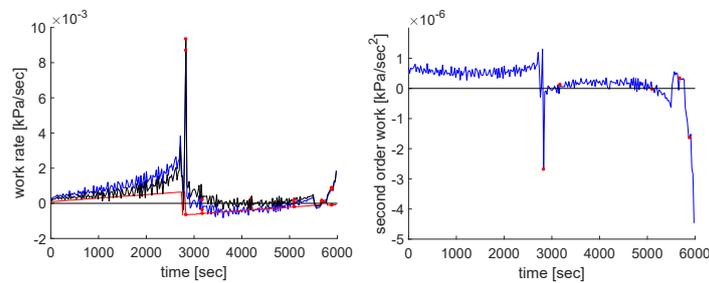


Figure 8: Work rates (left) and second order work (right).

The question that arises, is what the physical meaning of the second order work principle is. A negative second order work is linked in a sense to a plastic unloading: while the strain rate is positive, the stress rate is negative. The opposite is of course also possible, depending on sign conventions. In such a situation, the rate of the reversible changes in the internal energy will still be non negative. Since changes in the kinetic energy are neglected, a positive value of the second order work would mean that the rate of dissipation changes faster than the rate of the internal power can change.

We will try to illustrate this here, using a simple example. On the left of figure 7 the axial stress versus the axial strain of a uniaxial test on a soft calcarenite is shown. The test was performed with a constant radial strain rate as the controlling variable. On the right of the same figure the total work, the elastic energy and the dissipated energy are plotted. A constant elastic modulus was assumed, for the sake of simplicity, even if, strictly speaking, the elastic modulus decreases in the course of the test. The total, elastic and dissipated work are given in the same figure in blue, red and black respectively, while their rates are given on the right of figure 8. The second order work is given on the left of the same figure. Consider the red points in the figures in the order of time as denoted by A, B, C, D and E, respectively.

From the beginning of the test to A the stress and the strain rate are both positive and so is the second order work. It would also be positive in the case of a reversal of sign

of the strain rate. All work rates are also positive. According to the second order work criterion, the test is stable.

Between the points A and B the second order work is negative. The test is however stable, in the sense that strain rate perturbations lead to continuous changes in the stress. This is not in conflict with the second order work criterion, as it does not claim that a negative second order work is a definite sign of instability. It should also be remarked, that stress control at these points is not possible, something that is however more a question of controllability than a question of stability, see [IN98].

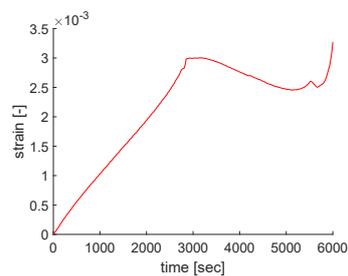


Figure 9: Strain rate versus time.

Between points B and C the second order work is positive and the specimen is unstable, since a positive strain rate would lead to a jump in the axial stress. This situation is also in agreement with the second order work criterion, as it requires the inequality to hold not for the actual, but for all possible displacement rates.

At point E one can easily observe that, while the rate of work is clearly increasing, and has thus a positive time derivative, the second order work is negative. This is a result of the increasing strain rate, as may be seen in figure 9, which means that the change in the kinetic energy is not negligible, even in the case of a clearly quasi-static uniaxial test.

Summarizing, it should be noted that, when applying the second order work criterion, attention should be paid to the definition of stability and its differences to controllability. Care should also be taken with respect to the kinetic energy, which was explicitly excluded by Hill, but may be present even in quasi-static settings. A relevant discussion may be found in [CCV04]. Finally, it should be stressed that the criterion requires the inequality to hold for all possible displacement increments, rather than just the actual one.

## 5 Conclusions

In this chapter different methods and formulations, more or less closely related to the concept of energy were outlined. The list is by no means exhaustive and several

significant and commonly used methods were omitted. An example is the use of concepts linked to energy to regularize ill-posed problems. The goal was to provide the reader with a somewhat better understanding of the underlying assumptions of each of these methods. It is hoped that this will render it easier to avoid common pitfalls in application, that may lead to physically unrealistic results.

## References

- [AMVH<sup>+</sup>06] F. Alonso-Marroquín, I. Vardoulakis, H. Herrmann, D. Weatherley, and P. Mora. Effect of rolling on dissipation in fault gouges. *Physical Review E*, 74(3):1–10, 2006.
- [Ber09] V.L. Berdichevsky. *Variational principles I*. Springer Verlag, 2009.
- [BV01] J.P. Bardet and I. Vardoulakis. The asymmetry of stress in granular media. *International Journal of Solids and Structures*, 38(2):353–367, 2001.
- [Car51] G. Cardano. *De subtilitate*. Milan, 1551.
- [CCL97] F. Calvetti, G. Combe, and J. Lanier. Experimental micromechanical analysis of a 2D granular material: relation between structure evolution and loading path. *Journal of Cohesive-Frictional Materials*, 2:121–163, 1997.
- [CCV04] R. Chambon, D. Caillerie, and G. Viggiani. Loss of uniqueness and bifurcation vs instability: some remarks. *Revue Francaise de Gcnie Civil*, 8(5-6):517–535, 2004.
- [CH97] I.F. Collins and G.T. Houlsby. Application of thermomechanical principles to the modelling of geotechnical materials. *Proceedings of the Royal Society A*, 453(1964):1975–2001, 1997.
- [CL90] C.S. Chang and C.L. Liao. Constitutive relation for a particulate medium with the effect of particle rotation. *International Journal of Solids and Structures*, 26(4):437–453, 1990.
- [CMNN81] J. Christoffersen, M.M. Mehrabadi, and S. Nemat-Nasser. A micromechanical description of granular material behavior. *Journal of Applied Mechanics*, 48:339–344, 1981.
- [dSV71] M. de Saint-Venant. Mémoire sur l'établissement des équations différentielles des mouvements intérieurs opérés dans les corps solides ductiles au delà des limites où l'élasticité pourrait les ramener à leur premier état. *Journal de mathématiques pures et appliquées*, 16:103–124, 1871.

- [Eul53] L. Euler. *Dissertatio de principio minimae actionis una cum examine objectionis*. Cl. Prof. Koenigii contra hoc principium factarum, 1753.
- [Fey64] R. Feynman. *The Feynman lectures on physics*. Addison-Wesley, 1964.
- [FG10] F. Froio and E. Gerolymatou. Stress tensors for granular materials as non simple continua. *European Journal of Environmental and Civil Engineering*, 14(8-9):1167–1183, 2010.
- [Fra18] G.A. Francfort. Recovering convexity in non-associated plasticity. *Comptes Rendus Mecanique*, 346:198–205, 2018.
- [FTV06] F. Froio, G. Tomassetti, and I. Vardoulakis. Mechanics of granular materials: The discrete and the continuum descriptions juxtaposed. *International Journal of Solids and Structures*, 43:7684–7720, 2006.
- [Ger11] E. Gerolymatou. *Micropolar continuum theories as applied to the description of the deformation of granular media*. PhD thesis, National Technical University of Athens, 2011.
- [Ger14] E. Gerolymatou. A micromechanically derived anisotropic micropolar constitutive law for granular media: Elasticity. *International Journal for Numerical and Analytical Methods in Geomechanics*, 38(17):1761–1775, 2014.
- [GI00] M. Gutierrez and K. Ishihara. Non-coaxiality and energy dissipation in granular materials. *Soils and Foundations*, 40(2):49–59, 2000.
- [GR64] A.E. Green and R.S. Rivlin. On Cauchy's equations of motion. *Zeitschrift für Angewandte Mathematik und Physik*, 15:290–292, 1964.
- [Hil48] R. Hill. A variational principle of maximum plastic work in classical plasticity. *The Quarterly Journal of Mechanics and Applied Mathematics*, 1(1):18–28, 1948.
- [Hil58] R. Hill. A general theory of uniqueness and stability in elastic–plastic solids. *Journal of the Mechanics and Physics of Solids*, 6:239–249, 1958.
- [IaH11] M. Ishiki and T. Kubawara and Y. Hayashida. Measurement and analysis of differential work hardening behavior of pure titanium sheet using spline function. *International Journal of Material Forming*, 4(2):193–204, 2011.
- [IN98] S. Imposimato and R. Nova. An investigation on the uniqueness of the incremental response of elastoplastic models for virgin sand. *Mechanics of Cohesive-frictional Materials*, 3:65–87, 1998.
- [Kel05] P. Kelly. *Solid Mechanics Part I*, chapter Energy and virtual work, pages 231–274. Cambridge University Press, 2005.

- [Kra09] K. Krabbenhoft. A variational principle of elastoplasticity and its application to the modeling of frictional materials. *International Journal of Solids and Structures*, 46(3–4):464–479, 2009.
- [Kru14] N.P. Kruyt. Micromechanical study of elastic moduli of three-dimensional granular assemblies. *International Journal of Solids and Structures*, 51(13):2336–2344, 2014.
- [Lag88] J.L. Lagrange. *Mécanique analytique*. Paris, 1788.
- [Lei86] G.W. Leipzig. *Acta eruditorum*. Leipzig, 1686.
- [Lin71] R.B. Lindsay. The concept of energy and its early historical development. *Foundations of Physics*, I(4):384–393, 1971.
- [Lya92] A.M. Lyapunov. *The general problem of stability of motion*. PhD thesis, University of Kharkov, 1892.
- [Mac41] E. MacCurdy. *The notebooks of Leonardo da Vinci*. Garden City Publishing Company, 1941.
- [NDK07] F. Nicot, F. Darve, and H.D.V. Khoa. Bifurcation and second-order work in geomaterials. *International Journal for Numerical and Analytical Methods in Geomechanics*, 31:1007–1032, 2007.
- [Nol59] W. Noll. La mécanique classique, basée sur un axiome d’objectivité. In *Colloque International sur la Méthode Axiomatique dans les Mécaniques Classiques et Nouvelles, 1963*, pages 47–56, Paris, 1959.
- [Nol72] W. Noll. A new mathematical theory of simple materials. *Archive for Rational Mechanics and Analysis*, 48:1–50, 1972.
- [PG97] P. Podio-Guidugli. Inertia and invariance. *Annali di Matematica Pura ed Applicata*, 172:103–124, 1997.
- [Sch68] A. Schwegler. *Handbook of the history of philosophy*. Edmonston & Douglas, 1868.
- [Tho51] W. Thomson. On the dynamical theory of heat, with numerical results deduced from Mr. Joule’s equivalent of a thermal unit, and M. regnault’s observations on steam. *Transactions of the Royal Society of Edinburgh*, 20:261–268; 289–298, 1851.
- [TW02] A. Tordesillas and D.C.S. Walsh. Incorporating rolling resistance and contact anisotropy in micromechanical models of granular media. *Powder Technology*, 124(1-2):106–111, 2002.
- [vM28] R. von Mises. Mechanik der plastischen Formänderung von Kristallen. *Zeitschrift der angewandten Mathematik und Mechanik*, 8:161–185, 1928.

- [You07] T. Young. *A course of lectures on natural philosophy and the mechanical arts*. London, 1807.
- [Yu06] H.-S. Yu. *Plasticity and Geotechnics*. Springer, 2006.



# An Energy Based Constitutive framework for Multiphysics Geomechanics

Mustafa Sari<sup>1</sup>, Jack Lin<sup>1</sup>, Sotiris Alevizos<sup>1</sup>, Thomas Poulet<sup>2</sup>, Manolis Veveakis<sup>3,‡</sup>

<sup>1</sup>UNSW Australia, <sup>2</sup>CSIRO Australia, <sup>3</sup>Duke University  
(‡ [manolis.veveakis@duke.edu](mailto:manolis.veveakis@duke.edu))

---

*Geomechanics as a field has increasingly being expanding towards including information from different scales as well as physical mechanisms. The need for additional information has brought elaborate requirements into the constitutive models required to represent such rich information. In this work we present a constitutive framework for including multi-physical mechanisms, relying on explicitly allowing for the energy balance and the second law of thermodynamics to evolve and lead the system to its point(s) of self organisation. Through theoretical examples from the fault mechanics community and practical applications from Geomechanics, we show that different self-organisation regimes can be achieved for different loading conditions and physical mechanisms included.*

## 1 Introduction

Constitutive models for multiphysical processes in Geomechanics have the key responsibility of describing processes that are happening over several different length and time-scales, yet affect the apparent response of materials related to all the spectrum of civil engineering applications. The challenging component in including multiphysical processes in constitutive models, is that the systems become extremely rich in energetics and may self organize away from the maximum dissipation limit. Such systems and cases are very well known in chemical engineering, as showcased by the nobel-price winning work of Prigogine and his co-workers [Pri90].

Geomaterials, unlike chemical engineering materials, have the particularity of changing their behaviour from solid-like to fluid-like upon failure. Modelling Geomaterials has therefore the additional challenge of capturing such transitions. In this work we will define failure and lay down the fundamental principles of multiphysics modelling

in Geomaterials. After doing so, we will present a constitutive law based on these principles. Examples given in this work include cases from fault mechanics as well as laboratory experiments.

## 1.1 Failure and fracture in Geomaterials

Failure and fracture in geomaterials may have different manifestations, such as the propagation of single fractures in otherwise intact rock, crack growth in presence of multiple interacting fractures, initiation and percolation of damage zones under quasi-static and dynamic loading, strain localization into shear- or compression bands.

Localization theory is a natural extension of Mohr's strength of materials theory. The mathematical formulation of bifurcation and post-bifurcation phenomena and related instabilities constitutes the basis of a contemporary continuum theory of failure of geomaterials [VS95]. From the theoretical point of view, the so-called Thomas-Hill-Mandel shear-band model ([Hil62], [Man66]) was introduced in the early '60s and it was widely publicized by the paper of Rudnicki and Rice [RJ75].

Localized failure, exhibited as shear banding, has been recently claimed as the primary failure pattern for faults [Ric06] and landslides [Var2a]. However, field evidence from exhumed faults reveals that seismic events take place in even narrower zones within shear-bands, formed from post-failure evolution of the fault zone, dominated by weakening mechanisms that are considered to be mainly thermal in origin [Ric06]. Note that in ductile, visco-thermoplastic materials like metals, shear banding is commonly observed during high strain-rate inelastic deformation.

## 1.2 The role of energy in shear localization of Elasto-Plasticity

These concepts of shear banding as material bifurcation of an elasto-plastic skeleton have led to the identification of a material length scale defining the width of shear bands. The critical stress, as well as the orientation and the thickness of the localized shear failure planes (shear zones) are calculated through the eigenvalues of the elasto-plastic stiffness modulus  $C_{ijkl}^{ep}$  of the material [RJ75] obeying a rate-independent rheology,  $\dot{\sigma}'_{ij} = C_{ijkl}^{ep} \dot{\epsilon}_{kl}$  (the prime denoting effective stress). Within the framework of the solid mechanical instabilities the shear (fault) zone thickness emerges as a solution depending on the microstructure [MV87, PV11].

This rate-independent regime is typically used as an adequate description of brittle processes, thus placing the formation of the shear zones (onset of localization) near the maximum deviatoric stress [RJ75]. For example, in linear elastic fracture mechanics brittle fracture occurs without thermal activation when a critical stress level is reached to split the bonds. At a critical energy threshold an elasto-dynamic fast time-scale instability ensues where a variety of dissipative processes kick in such as grain/particle rotations which release heat in extremely fast time-scales. The fast time

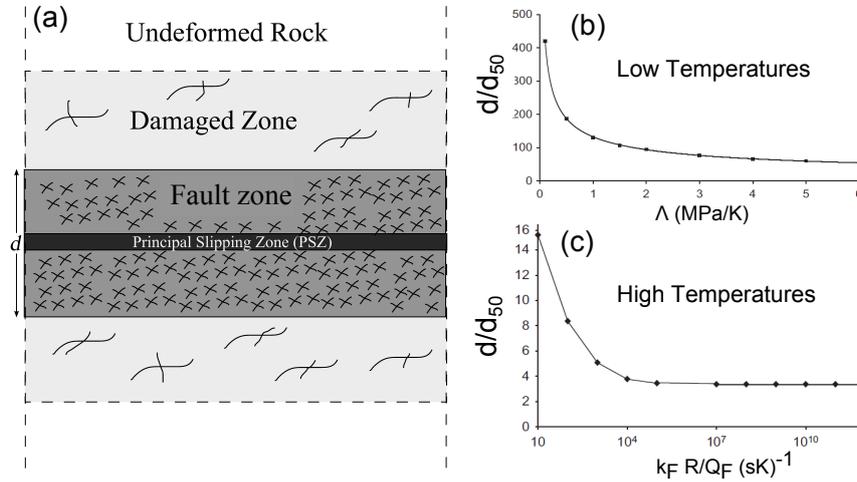


Figure 1: (a) Conceptual model of the internal structure of a fault described by two scales: the Principal Slip Zone (PSZ) in the centre surrounded by the fault zone. (b) using the geothermal temperature as a thermal condition for the boundary of the fault zone we obtain the thickness of the fault zone as a function of its internal structure (average particle size) and the thermal pressurisation. (c) Using the higher temperature required for chemical reactions as recorded in the PSZ a much length scale down to 3 times the average grain/particle size is obtained.

scale instability simplifies the processes as temperature and fluids do not have time to diffuse. This regime is called undrained-adiabatic and is characterising extremely fast co-seismic slip [Ric06, SSV11] of brittle failure events.

Within this brittle, solid mechanical framework recent studies [SSV11, VSS12, VSS13] derived different levels of localisation corresponding to different energy (temperature) regimes. Their results showcase the role of different mechanisms operating at different scales and energy (temperature) levels. Even by adopting classical, rate-independent, elasto-plastic theory for a material incorporating chemical, hydraulic and thermal sensitivity we obtain some strong dependence of the shear band thickness with pressure and temperature conditions.

This is shown by calculating the extent (thickness) of the brittle fault zone of Figure 1 (a) at different temperature regimes, marked by: (1) the temperature inferred from the geothermal gradient for the boundary of the fault zone, (2) the activation temperature of the dominant chemical reaction observed in the Principal Slipping Zone (PSZ). The details of the approach can be found in [SSV11, VSS12, VSS13] and the results are summarised in Figure 1 (b) and (c).

### 1.3 Rate sensitivity

A vast category of geomaterials exhibit rate-dependent behavior during shear, thus changing their mechanical properties and state with velocity. Recently, it has been shown by Veveakis *et al.* [VVD07] that a fully saturated clay material may change its properties past a critical temperature and slip velocity, when sheared under constant load. In particular, a rapid solid-liquid phase transition may take place when a normally consolidated clay expels water due to heat and becomes liquefied. This phenomenon is called thermal pressurization and is usually related to the difference of the thermal expansion coefficients of the two phases at low hydraulic diffusivities and/or to the dehydration or “fugacity” of water bound to the mineral particles. Thermal pressurization is considered as a major weakening mechanism both for landslides ([Var00], [LA07]) and seismic slip ([Ric06], [SF09]). Recently, Fialko and Khazan [FK05] used a similar framework of shear heating and thermodynamics of phase transitions for the effect of melting. They showed that chemical transformations may exhibit the same behavior with thermal pressurization, and provided the conditions at which melting may reduce rapidly the strength of a fault. Since these weakening processes are thermal in origin [Ric06], it is rational to anticipate that triggering one of them could be due to thermally active, prolonged slip and in particular due to the heat produced by friction at the base of a slipping fault.

In order to model this kind of behaviors (such as prolonged slip) it is common to account for rate and state dependency in the frictional behavior of the geomaterials, after the works of Dieterich ([Die72], [Die78]). In fact, the need to account for rate dependent friction laws in order to model the stick-slip motions that characterize an earthquake was brought into light by Gu *et al.* [GRRT84]. Scholz [Sch98] remarked that the development of a full constitutive law for rock friction is crucial to understand earthquake phenomena such as seismogenesis and seismic coupling, pre- and post-seismic phenomena, and the insensitivity of earthquakes to stress transients. Later, Marone [Mar98] indicated that the differences in the behavior of bare rock surfaces as compared to shear within granular fault gouge can be attributed to dilation within fault gouge. Following this study, Garagash and Rudnicki ([Gar12]) retrieved stable and unstable regimes concerning the shear heating of a fluid-saturated slip-weakening dilatant fault zone. They verified that shear heating tends to increase pore pressure and decrease the effective compressive stress and the resistance to slip and consequently tends to promote instability [VAV10, APV14, VPA14, VRL14].

Chester and Chester [CC98], based on field observations from the Punchbowl fault, San Andreas, suggested that unstable seismic slip is an extremely localized phenomenon, that occurs primarily in an ultra thin shear zone of the order of millimeter (the so-called “Primary Slip Surface” or PSS), which lies within a finely granulated (ultracataclastic) fault-core of typically tens to hundreds millimeter. Note also that the Aigion fault core of clay size particles consisting of finely crushed radiolarites, extended to about 1m and was intercepted by a “fresh” distinct slip surface of sub-millimeter size [SJP05]. This structure of an ultrathin shear zone within the ultracataclastic core is observed only in the case of a wet gouge, under hot conditions [BMJ95].

It is to be noted also that under these hot-wet conditions little grain size reduction was observed in the ultracataclastic core, and that a temperature weakening and velocity strengthening frictional behavior is observed, which is dominated by solution transfer processes, such as pressure solution.

In this study we will initially study a simple 1D shear problem, considering the ultracataclastic core to be filled with a fully saturated, visco-plastic, porous, soil material. The solid phase is assumed to have permanent contacts between the “grains” so that a solid skeleton is formed and a normal effective stress can be defined. Based on basic principles of thermodynamics we formulate the corresponding governing equations for the problem of simple shear of a biphasic granular material. The non-Newtonian viscous behavior of these materials is then studied for each assumed particular friction law to the physics of the problem, emphasizing temperature weakening and velocity strengthening behavior. We characterize each of these friction laws in terms of mathematical properties, stability and physical behavior, to extract criteria on the range of the validity of each model. Through this analysis, we show that for all the friction laws studied, when hardening fails to counterbalance softening, localization instability occurs thus showing that shear heating of active faults is a post-failure localizing mechanism, in a narrow shear band within the initial fault core.

The significant temperature increase due to this instability could have the potential to set in effects like pressure solution and decomposition reactions ([APV14]). To explore this possibility we consider the effect of chemical reactions that could be triggered at elevated temperatures and are claimed to dominate the later, unstable stage of the slip evolution that is taking place in seismic velocity range ( $> 1$  m/s), in extremely localized zones.

## 2 Thermodynamical Aspects

In this section we will construct the governing heat diffusion equation for a biphasic material, consisting of a solid granular phase and a fluid phase, under constant volume from first principles of thermodynamics. In multiphase materials partial temperatures for each phase are usually defined. In this study, we assume that the specific surface of the grains is large enough, so that the system’s phases are always in thermal equilibrium. This assumption allows us to define a unique temperature field for both phases  $\theta(x_k, t)$ , and restricts the validity of the model presented in this study.

Let  $e = e(x_k, t)$ ,  $P^{(m)} = P^{(m)}(x_k, t)$  and  $Q_i = Q_i(x_k, t)$  be the specific internal energy, the stress power and the heat flux, respectively. The local form of the energy balance equation as applied to the mixture reads as follows ([VS95])

$$\rho \frac{D^{(m)}e}{Dt} = P^{(m)} - \frac{\partial Q_k}{\partial x_k} \quad (1)$$

where  $\frac{D^{(m)}}{Dt}$  is the barycentric material time derivative of the mixture and  $\rho = (1 -$

$\phi\rho_s + \phi\rho_w$  is the mixture's density,  $\phi$  being the porosity and  $\rho_s, \rho_w$  the densities of the solid and fluid phase, respectively.

One of the most common expressions for the 2<sup>nd</sup> law of Thermodynamics is given by Clausius -Duhem inequality

$$\rho \frac{D^{(m)}s}{Dt} \geq -\frac{\partial}{\partial x_k} \left( \frac{Q_k}{\theta} \right) \quad (2)$$

where  $\frac{D^{(m)}s}{Dt}$  is the specific entropy production.

By substituting the Energy Balance Equation (Eq. (1)) into Eq. (2) we can identify the expression

$$D_{loc} = \rho\theta \frac{D^{(m)}s}{Dt} - \rho \frac{D^{(m)}e}{Dt} + P^{(m)} \quad (3)$$

as the local dissipation of the mixture ([Tru69]). This equation (Eq. (3)), along with the Energy Balance one (Eq. (1)), yield

$$\rho\theta \frac{D^{(m)}s}{Dt} = D_{loc} - \frac{\partial Q_k}{\partial x_k} \quad (4)$$

The last equation is known as the local entropy production balance law. Based on this equation, one can derive an evolution equation for the temperature by expressing the local entropy production as a function of temperature and adopting Fourier's law for the heat flux.

Indeed, we may use the definition of the specific entropy through Maxwell's relations,  $s = -\left(\frac{\partial\psi}{\partial\theta}\right)_V$ , where  $\psi = e - \theta s$  is the Helmholtz free energy, to the first term of Eq. (4),

$$\rho\theta \frac{D^{(m)}s}{Dt} = -\rho\theta \frac{\partial}{\partial\theta} \left( \frac{\partial\psi}{\partial\theta} \frac{D^{(m)}\theta}{Dt} \right) \quad (5)$$

If we expand the derivative of the r.h.s. of Eq. (5) and set  $-\theta \frac{\partial^2\psi}{\partial\theta^2} = jC_m$ , we obtain the following expression for the specific entropy production

$$\rho\theta \frac{D^{(m)}s}{Dt} = j(\rho C)_m \frac{D^{(m)}\theta}{Dt} \quad (6)$$

In the above expressions  $C_m$  is the specific heat capacity of the mixture and  $j = 4.2J/cal$  is the mechanical equivalent of heat. The specific heat capacity of the mixture is calculated in reference to the partial masses of the constituents, through mixture's theory,  $(\rho C)_m = (1 - \phi)\rho_s C_s + \phi\rho_w C_w$ , where  $C_s$  and  $C_w$  is the specific heat of the solids and of the water, respectively.

In order to determine the third term of Eq. (4) we assume that the heat flux vector obeys Fourier's Law for thermally isotropic media:

$$Q_k = -jk_m \frac{\partial\theta}{\partial x_k} \quad (7)$$

where  $k_m = (1 - \phi)k_s + \phi k_w$  is the thermal conductivity of the soil-water mixture and  $k_s, k_w$  is Fourier's thermal conductivity of the solid and water, respectively. The negative sign in Eq. (7) represents that heat flows from the hotter areas to the colder ones.

Thus, Eq. (4) can be written, in view of Eqs. (6) and (7), as a heat diffusion equation ([VS95], [VVD07])

$$\frac{D^{(m)}\theta}{Dt} = \kappa_m \frac{\partial^2 \theta}{\partial x_k^2} + \frac{1}{j(\rho C)_m} D_{loc} \quad (8)$$

where  $\kappa_m = \frac{k_m}{(\rho C)_m}$  is Kelvin's thermal diffusion coefficient of the mixture. In order to evaluate it we usually use the expressions of heat capacity and thermal conductivity from mixture's theory formulation ( $k_m = \phi k_w + (1 - \phi)k_s$ ). However, the experimentally calculated thermal conductivity of the mixture,  $k_m^{eff}$ , is in general lower than the theoretically predicted one,  $k_m$ , being comparable only for small values.

In the derived heat diffusion equation for a biphasic mixture under constant volume (Eq. (8)) the dissipation function,  $D_{loc}$ , remains undefined. It should be carefully evaluated for the material under study, depending on the specific problem (simple shear, triaxial compression etc) through experiments and proper constitutive modeling.

### 3 Shear Heating Formulation

In this section we will apply the governing heat diffusion equation, derived in the previous section, in a specific geomechanical problem. Therefore we consider the deformation of a long shear-zone of water-saturated clayey material. We assume that this shear-zone has a thickness  $d$  and that the various mechanical fields vary only along the short  $z$ -direction, (Fig. 2). The maximum value of the temperature profile is achieved at the center of the shear-zone,  $z = 0$ . In addition, we assume that the shear-zone material is at critical state (in a geotechnical sense), deforming thus under constant volume and load. The only non-zero velocity component is  $v = v(z, t)$ .

At the shear-zone's boundary a constant normal stress  $\sigma_n$  is assumed to act, together with a shear stress  $\tau_d$ . Following Terzaghi's effective stress principle the effective stress tensor acting on the solid skeleton in our case (Fig. 2) consists of two components,  $\sigma' = \sigma_n - p$  (where  $p$  is the pore fluid pressure) and  $\tau$ . The latter, is commonly assumed to obey a Coulomb friction law,  $\tau = \mu \sigma'$ , where  $\mu$  is the friction coefficient at critical state (remaining constant for the considered case).

Following recent studies of Rice [Ric06] and Sulem and Famin [SF09] we neglect inertia inside the shear band,

$$\frac{\partial \tau}{\partial z} = 0 \quad (9)$$

This requirement is valid when the shear band thickness  $d$  is negligible. Indeed, one

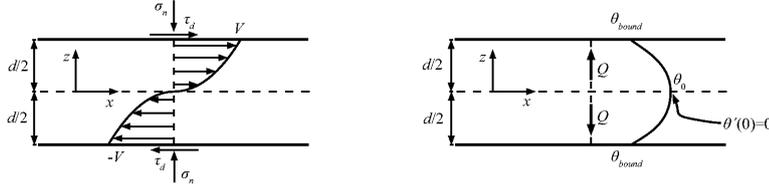


Figure 2: Steady shearing of a shear-zone. The steady-state profiles of (a) the loading and boundary conditions and (b) the thermal conditions are depicted.

may write

$$\frac{\partial \tau}{\partial z} = \rho \frac{\partial v}{\partial t} \Rightarrow \frac{\Delta \tau}{d} = \rho \eta g \quad (10)$$

where  $\Delta \tau = \tau(z = d/2) - \tau(z = -d/2)$  is the shear stress variance in the shear-band,  $g = 9.81 \text{ m/s}^2$  is gravity's acceleration and  $\eta$  is a proportionality constant. In order to have a significant shear stress difference across the shear-zone the product  $\Delta \tau = \rho \eta g d$  must be significant. Assuming an acceleration of about 10 times the acceleration of gravity (i.e  $\eta = 10$ ) and a typical value for the soil density of about  $2500 \text{ kg/m}^3$ , we obtain  $\Delta \tau \simeq 25 \cdot 10^{-2} \cdot d \text{ MPa/m}$ . In order to have significant variations across the shear-band,  $d$  should be of the order of tens of meters. However, as noted also in the Introduction, since  $d \sim 10^{-1} - 10^{-3} \text{ m}$ ,  $\Delta \tau$  is negligible, so that  $\tau = \tau_d(t)$  remains constant across the shear-zone [SF09]. It is evident that, since  $\mu = \text{const.}$  (critical state friction coefficient) the pore pressure is constant across the shear zone as well ( $p = p(t)$ ). Since no pore pressure variations are assumed, apart from the excess pore pressure generation due to the various chemical processes,  $p = \text{const.}$  during creep (where  $\sigma' = \text{const.}$ ), and thus  $\tau_d = \text{const.}$

### 3.1 Viscosity and Dissipation Function

Assuming that the material is fully plastic, in the one-dimensional framework the dissipation function of Eq. (8) is always positive and can be written for the problem of simple shear as [Var2a]:

$$D_{loc} \simeq \beta_T \tau_d \dot{\gamma} \geq 0 \quad (11)$$

where  $\beta_T$  is the Taylor-Quinney coefficient [TQ34], a fraction of the mechanical work that converts into heat during simple shear, and  $\dot{\gamma} = \dot{\gamma}^p \approx V/d$  is the plastic shear strain rate, where  $V$  is the interface velocity ([Var2a], [VVD07]). Under the above definitions, Eq. (8) can be written in its final form:

$$\frac{D^{(m)} \theta}{Dt} = \kappa_m \frac{\partial^2 \theta}{\partial z^2} + \frac{\beta_T \tau_d}{j(\rho C)_m} \dot{\gamma} \quad (12)$$

In order to proceed with our analysis, we are obliged to further investigate the nature of the friction coefficient, introduced in the Coulomb's law. Without loss of generality, we introduce indicative reference quantities for the temperature and strain rate,  $\theta_{ref}$  and  $\dot{\gamma}_{ref}$  respectively, along with the dimensionless quantities

$$z^* = \frac{z}{(d/2)}, t^* = \frac{\kappa_m}{(d/2)^2} t, \theta^* = \frac{\theta}{\theta_{ref}}, \dot{\gamma}^* = \frac{\dot{\gamma}}{\dot{\gamma}_{ref}} \quad (13)$$

and reduce equation (12) to a simpler, normalized heat equation,

$$\frac{D^{(m)}\theta^*}{Dt^*} = \frac{\partial^2 \theta^*}{\partial z^{*2}} + \frac{\beta_T \dot{\gamma}_{ref} \tau_d}{\theta_{ref} (j k_m)} \left(\frac{d}{2}\right)^2 \dot{\gamma}^* \quad (14)$$

In this study we will assume that the friction coefficient is a function of rate and state variables. State variables according to Thermodynamics are the macroscopic quantities that result from the microscopic interaction of particles in a material (such as temperature, pressure, internal energy, enthalpy, entropy, volume, density, mass). Therefore, we assume that the friction coefficient depends on the dimensionless strain rate and temperature, in a usual multiplicative sense,

$$\mu(\dot{\gamma}^*, \theta^*) = f(\dot{\gamma}^*)g(\theta^*) \quad (15)$$

In order for the friction coefficient to remain constant at critical state, the two functions of strain-rate and temperature should exhibit an antagonistic behavior. Thus, rate effects should counterbalance thermal effects (and vice versa) and determine the rheological behavior of the material. Furthermore, we assume that the two functions  $f(\dot{\gamma}^*)$ ,  $g(\theta^*)$  are monotonous. This issue is further addressed in the next paragraph.

It is easily understood under the above assumption that the material exhibits a non-Newtonian rheology ([Var02]), since

$$\tau_d = \left[ \sigma'_n \frac{f(\dot{\gamma}^*)}{\dot{\gamma}^*} g(\theta^*) \right] \dot{\gamma}^* \quad (16)$$

As shear stress is assumed to be constant and  $f, g$  are monotonous, one may express the strain rate as a function of temperature,

$$f(\dot{\gamma}^*) = \frac{\tau_d}{\sigma'_n g(\theta^*)} \Rightarrow \dot{\gamma}^* = f^{-1} \left( \frac{\tau_d}{\sigma'_n g(\theta^*)} \right) = \phi(\theta^*) \quad (17)$$

In view of Eq. (17), the heat equation, Eq. (12), becomes

$$\frac{D^{(m)}\theta^*}{Dt^*} = \frac{\partial^2 \theta^*}{\partial z^{*2}} + Gr \phi(\theta^*) \quad (18)$$

where the parameter,

$$Gr = \frac{\beta_T \dot{\gamma}_{ref} \tau_d}{\theta_{ref} (j k_m)} \left(\frac{d}{2}\right)^2 \quad (19)$$

is called the Gruntfest number [Gru63] and is the ratio of the two characteristic time scales of the problem, namely the time scale of the heat produced due to mechanical work over the time scale that the material conducts heat away. When  $Gr \rightarrow 0$  the system is “diffusive”, since it has the capability to conduct all the produced energy, and thus remain stable. On the contrary, when  $Gr \rightarrow \infty$  the system is “adiabatic”, since the heat produced doesn’t have the time to be conducted away. This limit is characterized by high strain rates and temperatures, having the tendency to promote instability [Gar12].

### 3.2 Temperature-Dependent Friction Law

The problem at hand is a viscoplastic flow corresponding to an exothermal, temperature dependent process. Obviously, given the assumption of monotonous antagonistic mechanisms, stated in the previous paragraphs, either rate hardening and thermal softening or vice versa will provide the same equation [VAV10]. Following the discussion on the behavior of faults under hot-wet conditions [BMJ95], we restrict ourselves to the former case, which is more pronounced also for temperature sensitive clays [HB90, HP02, CL04, LLC08, CZV11].

Hicher [Hic74] reported experimentally a decrease of the critical state friction coefficient with temperature of kaolin clay. His results are summarized in Fig. 3, where several possible fits are also presented. In this section we will derive the heat diffusion equation for three of these fits. In particular, we will study an exponential, a logarithmic and an Arrhenius-like dependency of the friction coefficient with temperature.

#### 3.2.1 Exponential-power law dependency:

Following Vardoulakis [Var02] and Veveakis *et al.* [VVD07] we consider the exponential dependency of the critical state friction coefficient with temperature, which is shown in Fig. 3 as the thermal softening mechanism. Motivated by the experiments of Leinenkugel [Lei76] on kaolin clay we also assume a power-law dependency of the critical state friction coefficient to strain-rate, to obtain finally that [Var02]

$$\mu = \mu_{ref} \left( \frac{\dot{\gamma}}{\dot{\gamma}_{ref}} \right)^N e^{-M(\theta - \theta_1)} \quad (20)$$

where  $M$  and  $N$  are the frictional thermal-sensitivity coefficient and the frictional rate-sensitivity coefficient, respectively, while  $\mu_{ref}$  and  $\dot{\gamma}_{ref}$  are reference quantities of the friction coefficient and the strain rate and  $\theta_1$  a reference temperature. This temperature is usually identified as either the temperature at the boundaries of the shear zone, or an ambient temperature at which material properties, such as friction coefficient, are measured. In this study we assume the latter, in order to study the effect of thermal boundary conditions to the system.

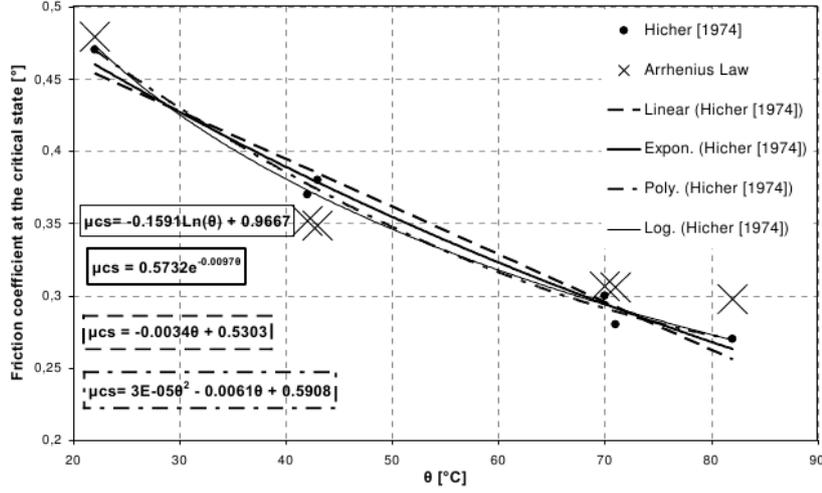


Figure 3: Critical state friction coefficient of a black (kaolin) clay. Various fitting curves are depicted, on the experimental data provided by Hicher. The best fit for the Arrhenius law,  $\mu = \mu_0 \exp\left(\frac{E_d}{R\theta}\right)$ , is achieved for  $\mu_0 = 0.25$  and  $E_d/R = 14.3^\circ\text{C}$  ([Hic74]).

Setting in this case  $\theta^* = m(\theta - \theta_1)$ , where  $m = M/N$ , the corresponding dimensionless diffusion equation is obtained from Eq. (18), for  $\phi(\theta^*) = m\dot{\gamma}_0 e^{\theta^*}$ ,  $Gr = m \frac{\beta_T \tau_d \dot{\gamma}_0}{j k_m} \left(\frac{d}{2}\right)^2$ , and  $\dot{\gamma}_0 = \dot{\gamma}_{ref} \left(\frac{\tau_d}{\sigma'_n \mu_{ref}}\right)^{\frac{1}{N}}$  [Var02]

$$\frac{D^{(m)}\theta^*}{Dt^*} = \frac{\partial^2 \theta^*}{\partial z^{*2}} + Gr e^{\theta^*} \quad (21)$$

Notice that in this law the material sensitivity to temperature and strain rate, expressed through  $m$ , is incorporated to the expression of the Gruntfest number.

### 3.2.2 Log-log dependency:

If we assume a logarithmic dependency of  $\mu$  with temperature and strain rate, then we may consider the widely accepted *Dieterich-Ruina* friction law ([Die78]) for the case of shearing of a finite-thickness shear-band ([VVD07])

$$\mu = \mu_0 + A \ln\left(\frac{\dot{\gamma}}{\dot{\gamma}_0}\right) + B \ln\left(\frac{\theta}{\theta_{ref}}\right) \quad (22)$$

where  $A$  and  $B$  are the frictional thermal-sensitivity coefficient and the frictional rate-sensitivity coefficient, respectively. Notice that the original law is frequently used

for rate sensitive materials, where  $\theta$  is a state-parameter, conjugated to the asperity of rock-to-rock contacts. Therefore adopting this law for thermally and rate sensitive materials, one should thoroughly study experimentally the range of the sensitivity coefficients. As we will see in the following section and unlike the exponential law, this type of model is sensitive to the ratio of these two coefficients.

By accepting the dimensionless quantities of Eq. (13) for  $\dot{\gamma}_{ref} = \dot{\gamma}_0$ , the corresponding diffusion equation is obtained from Eq. (18), for  $\phi(\theta^*) = \dot{\gamma}_0 \theta^{*m}$ , where  $m = -B/A$ , and  $Gr = \frac{\beta_T \tau_d \dot{\gamma}_0}{\theta_{ref} (j k_m)} \left(\frac{d}{2}\right)^2$

$$\frac{D^{(m)}\theta^*}{Dt^*} = \frac{\partial^2 \theta^*}{\partial z^{*2}} + Gr \theta^{*m} \quad (23)$$

### 3.2.3 Arrhenius-power law dependency:

The most known among all the friction laws (rheological models) is the Arrhenius law,

$$\mu = \mu_{ref} \left( \frac{\dot{\gamma}}{\dot{\gamma}_{ref}} \right)^N e^{\frac{E_d}{R\theta}} \quad (24)$$

where  $E_d$  is the activation energy of the frictional heat production and  $R$  the universal gas constant ( $8.3JK^{-1}mole^{-1}$ ). Setting  $\theta^* = m(\theta - \theta_1)$  the corresponding dimensionless diffusion equation is obtained from Eq. (18), for  $\phi(\theta^*) = m\dot{\gamma}_0 e^{\left(-\frac{Ar_d}{1+\delta\theta^*}\right)}$ ,  $Ar_d = E_d/RN\theta_1$ ,  $\delta = 1/(m\theta_1)$ ,  $Gr = m \frac{\beta_T \tau_d \dot{\gamma}_0}{j k_m} \left(\frac{d}{2}\right)^2$ , and  $\dot{\gamma}_0 = \dot{\gamma}_{ref} \left(\frac{\tau_d}{\sigma_n \mu_{ref}}\right)^{\frac{1}{N}}$

$$\frac{D^{(m)}\theta^*}{Dt^*} = \frac{\partial^2 \theta^*}{\partial z^{*2}} + Gr e^{\left(-\frac{Ar_d}{1+\delta\theta^*}\right)} \quad (25)$$

Notice that the ratio of the thermal to rate sensitivity of the material is also present in this law, expressed by the quantity  $Ar_d$ . In a chemical engineering context, this is defined as the activation energy of the friction. In the present formalism,  $Ar_d$  is interpreted as the energy threshold at which the friction coefficient becomes thermally and rate sensitive. We need to notice that for rate hardening, for  $0 < N < 1$  the rate response of the friction coefficient is concave, for  $N = 1$  linear and for  $N > 1$  convex. Since we anticipate that the friction coefficient remains finite even at large strain rates, we require a concave response,  $0 < N < 1$ .

## 4 Characterization of Rate- and State- Friction Laws

### 4.1 Characterization in terms of mathematical properties

From previous works [VVD07] we know that Eqs. (21),(23) may present a property that is called finite-time blow up (i.e. finite time singularity in their solution). On the

other hand, it can be shown that Eq. (25) does not present such kind of a behavior. The physical properties of the different behaviors are not directly obvious and call for further investigation. Therefore, we will proceed by performing in the next paragraph the numerical bifurcation analyses of the three different rheological (frictional) models presented in section 3.

## 4.2 Characterization in terms of Stability and Bifurcation Analysis

The steady state of the three equations (21),(23),(25) are studied, together with Dirichlet and Neumann boundary conditions,

$$\frac{d^2\theta^*}{dz^{*2}} + Gre^{\theta^*} = 0, \{\theta^*(1) = \theta_{bound}, \theta^{*'}(0) = 0\} \quad (26)$$

$$\frac{d^2\theta^*}{dz^{*2}} + Gr\theta^{*m} = 0, \{\theta^*(1) = \theta_{bound}, \theta^{*'}(0) = 0\} \quad (27)$$

$$\frac{d^2\theta^*}{dz^{*2}} + Gre^{-\frac{Ar_d}{(1+\delta\theta^*)}} = 0, \{\theta^*(1) = \theta_{bound}, \theta^{*'}(0) = 0\} \quad (28)$$

Notice that since we are studying the steady state of the heat equations, we may neglect convective terms, as shown by Vardoulakis [Var02]. The temperature profile is assumed to satisfy a standard symmetry condition [CDMM89], being maximum at the core of the shear-zone ( $z^* = 0$ ). The value of the temperature at the boundary of the shear-zone ( $z^* = 1$ ) is assumed to be constant. This choice is dictated by the fact that we examine the behavior of the shear zone in steady state and will not be valid in any time-dependent framework, since in that case the problem of the heat diffusion is driven by a characteristic time scale which should be compared to the characteristic time scale of slip evolution [Gar12], expressed in this model by the Gruntfest number. For example in Veveakis *et al.* [VVD07] it was shown that during quasi-static creep, the boundary temperature evolved along the variations of the core temperature.

For determining the stability of the three models, we solve each equation using a finite differences scheme along with an arc-length continuation method ([CK91]). The eigenvalues of the Jacobian matrix of the numerical system are also calculated and determine the stability regimes of the system (according to the lines of the arc-length continuation method, thoroughly presented in [CK91]). As shown in Fig. 4 ((a) and (b)) the first two equations (26),(27) present a turning point at a critical value of the Gruntfest number,  $Gr_c$ : (Eq. 26) presents it always, while Eq. (27) only when  $m > 1$ . This turning point defines through the sign of the eigenvalues a stable (lower) and an unstable (upper) branch of the solution.

The behavior depicted in Fig. 4 ((a), (b) and (c)) represents the celebrated phenomenon of thermal runaway [Fow97]. Depending on the value of the Gruntfest number, the problem has two ( $Gr < Gr_c$ ), one ( $Gr = Gr_c$ ) or zero ( $Gr > Gr_c$ ) steady state

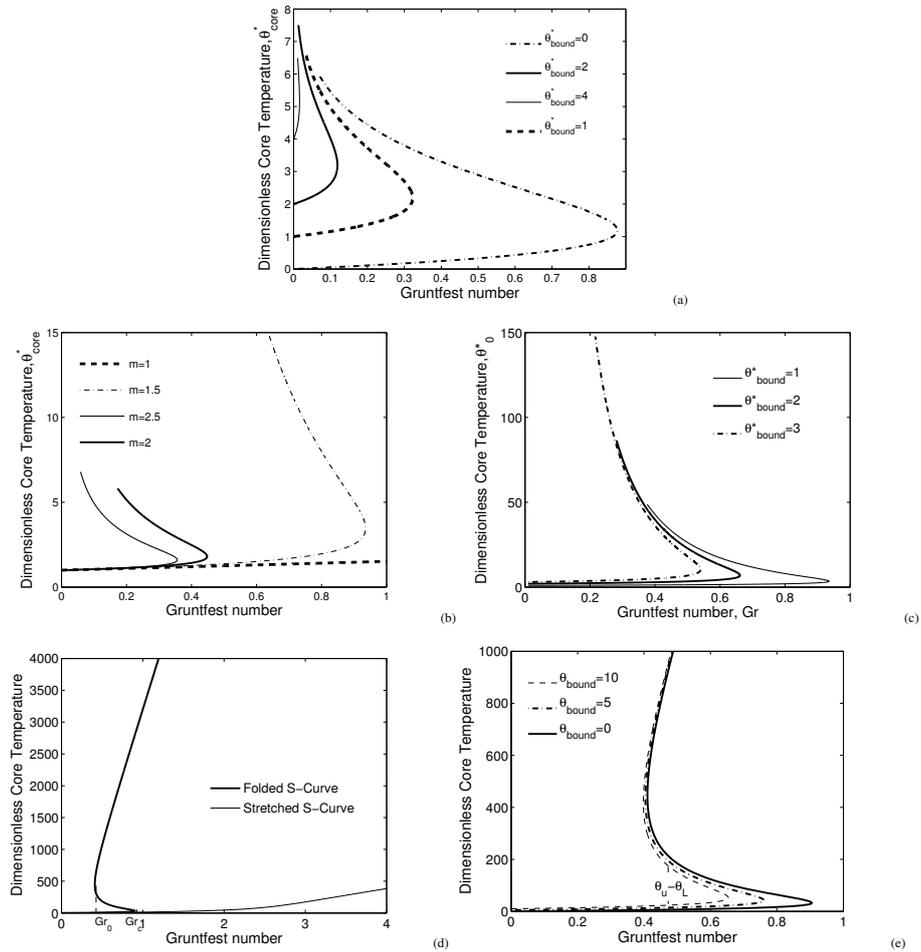


Figure 4: Bifurcation Diagrams (or response curves) of: (a) the Bratu problem. Results are plotted for different values of the boundary temperature. All the profiles present a turning point, for  $Gr = Gr_c$ . The lower branch is stable, while the upper branch is unstable; (b) the log-log (or *Dieterich-Ruina*) law for various values of the power coefficient,  $m$ . It is to be noted that for  $m > 1$  the model presents a turning point, exactly like the Bratu problem, while for  $m < 1$  the solution is stable; (c) the same law for  $m = 1.5$ , plotted for different boundary temperatures; (d) the Arrhenius law model, plotted for  $\theta_{bound}^* = 0$ ; (e) the lower branch of a folded S-curve of the Arrhenius law model, plotted for different boundary temperatures.

solutions. The effect of the boundary conditions is also depicted in this figure. As it can be seen in Fig. 4(a) and (c), while boundary temperature increases the critical value for the Gruntfest number,  $Gr_c$ , decreases. Due to the assumption of constant shear stress during creep (see Section 3) and the definition of the Gruntfest number (see Section 3.2.1), the latter implies a decrease in the critical strain-rate value,  $\dot{\gamma}_0$ . From the previous analysis we may conclude that, the exponential and the logarithmic law, for  $m > 1$ , present the same qualitative behavior.

Similarly to the logarithmic law, the number of steady state solutions of the Arrhenius one may vary (Fig. 4(d)). As discussed by Law ([Law06], pp. 313-317) one may approximate the condition for which the Arrhenius law loses steady state solutions by the inequality

$$Ar_d \leq 4(1 + \delta \cdot \theta_{bound}) \quad (29)$$

As long as this inequality holds, the system has one stable steady state solution for all Gruntfest numbers, a response frequently addressed as “stretched S-curve” [Law06]. As we observe in Fig. 4(e), by increasing the boundary temperature (imposing a higher energy level to the system), the lower stable branch shrinks with respect to the loading conditions, as expressed by  $Gr$ . At the same time, the range of temperatures defining the stable area ( $\theta_u - \theta_L$ ) for a given  $Gr$  shrinks accordingly. The same inequality implies that, when the friction has a small  $Ar_d$ , and therefore the heat production inside the shear zone is significant at lower temperatures, the system response tends to that of a stretched S-curve. Again, for a given boundary temperature (energy level) the ability of the system to remain at relatively low temperatures is diminished the faster the frictional effect takes place (Fig. 4(d)). This discussion highlights also the necessity for detailed experiments on the rate and thermal response of the material under study, given the fact that  $m$  and  $Ar_d$  depend on the thermal to rate sensitivity of the material.

When inequality (29) fails, the bifurcation diagram of the steady state problem is that of a folded S-curve as depicted in Fig. 4(d). In this case, although the Arrhenius law does not provide the property of thermal runaway, its behavior is identical to the exponential and logarithmic one for lower temperatures; a lower, stable branch is followed by an intermediate, unstable one. As shown by the asymptotic analysis of Fowler ([Fow97], pp.181-190) on the problem of Eq. (25), the Arrhenius law coincides with the exponential one for temperatures of order  $\sim O(1)$ . The difference however is that at elevated temperatures the Arrhenius law presents a third, stable, branch (Fig. 4(d)) which means that for higher orders of temperature the system will tend to a rather high, but finite temperature [MLB9a]. Indeed, again from this asymptotic analysis it is shown that the upper stable branch is reached for temperatures greater than the order of  $\sim O(Ar_d^2)$ , raising the question as whether these temperatures required to be on this branch of the Arrhenius-power law are realistic. Finally, we can observe from Fig. 4(e) the same decreasing behavior of the critical Gruntfest number of the lower turning point,  $Gr_c$ , with increasing boundary temperature.

### 4.3 Characterization in terms of physical behavior

In Fig. 4 the unstable branches of the diagrams represent a significant process; by plotting the dimensionless dissipation function (Fig. 5) we may observe that, after the turning point, the curvature of the dissipation profiles changes, leading to its localization towards the center of the shear-zone. Thus, the unstable branch of the system may correspond to thermal localization, in a time-dependent, quasi-static evolution. Since the shear strain-rate is conjugated with temperature through Eq. (17), it is evident that -similarly to strain softening- thermal softening corresponds to deformation localization as well ([VVD07]). This localization signifies the formation of an even thinner shear-band within the initial shear-zone. This physical behavior is identical for all the laws under certain conditions ( $m > 1$ ,  $Ar_d > 4(1 + \delta \cdot \theta_{bound})$ ), since they all present an unstable branch. The latter highlights an important difference between the three laws; the response of the exponential is independent of the material parameters and boundary conditions having a single free parameter ( $Gr$ ), whereas the ones of the logarithmic and the Arrhenius law depend also on the sensitivity ratio  $m$ , and the sensitivity ratio  $Ar_d$  and the boundary temperature, respectively. It remains to be verified whether these additional stability diagrams correspond to an observable manifestation of the real geophysical problem.

In real geomechanical processes, in quasi-static conditions, the system will assume a specific Gruntfest number. Depending on the initial condition, the evolution of the temperature is evidently characterized by the equilibrium points. For all the models and for all initial temperatures below the unstable steady state, the system will asymptotically tend to the low, stable steady state temperature for the given Gruntfest number. If the initial condition is above the unstable steady state temperature, the first two laws predict an infinite temperature growth. This effect, along with the mathematical property of blow-up (that the two laws present), are not physically admissible properties, since infinite temperature growth is not allowed in nature. On the other hand, the Arrhenius law presents qualitatively the same mathematical and physical behavior at lower temperatures, fits the experimental data (from experiments at ambient temperatures) equally well (Fig. 3) and leads the system to rather elevated, but finite, temperatures. Therefore, from a modeling point of view, it is reasonable to prefer it from the other two models.

Concerning the behavior of the shear zone at a higher regime, one could assume that the temperature is sufficiently high to trigger a chemical reaction within the shear-band. Such an effect could counterbalance the localization effect and provide a new, stable, steady state at different orders of temperatures, depending on the nature of the aforementioned reaction. To clarify this claim, in the following section we present the coupling of frictional heating of the Arrhenius type with chemical reactions triggered at high temperatures.

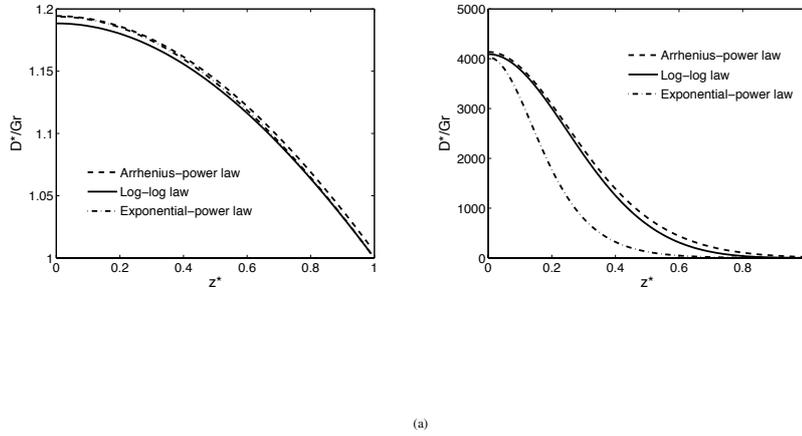


Figure 5: Plot of the profiles of the dissipation functions (1)  $D^*/Gr = e^{\theta^*}$ , (2)  $D^*/Gr = \theta^{*m}$ ,  $m = 5$ , (3)  $D^*/Gr = e^{-\frac{Ar_d}{1+\delta\theta^*}}$ ,  $\delta = 10^{-2}$  and for an arbitrary value of  $Ar_d$ , across the shear zone. Since  $Gr > 0$ , the figure depicts qualitatively the shape of each dissipation function (a) at the lower stable and (b) at the unstable branch of the response curves of Fig. 4. We notice that localization of the dissipation occurs after the bifurcation point  $Gr_c$ , i.e. at the unstable branch of the response curves of the three laws.

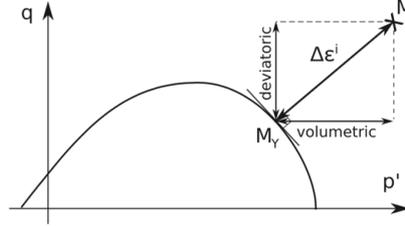


Figure 6: Schematic for the decomposition of the plastic flow rule in volumetric and Deviatoric components. A point  $M$  in the  $p' - q$  space (mean effective stress – shear stress) is shown with its corresponding point  $M_Y$  on the yield envelope.

## 5 Generalization for a 3D Constitutive model

Based on the work of the previous sections, we identified the Arrhenius law as the most generic and representative in order to describe thermal sensitivity of geomaterials. However, the Arrhenius law can be generalized to include activation energies for interface and internal dissipative mechanisms. In this section, we will generalize the Arrhenius dependency previously described, to include such effects, and cast it into a 3D constitutive law of elastoplasticity able to describe multiphysical effects.

The underlying physical model for this exercise is based on the principles of overstress plasticity [Per66], used in a novel elasto-viscoplastic approach [PV16]. Following the classical considerations of mechanics, the total strain rate is decomposed in an elastic (reversible) part and a plastic (irreversible) part. The reversible component is assumed to obey a linear thermo-elastic relationship,

$$\dot{\epsilon}_{ij}^r = C_{ijkl} \dot{\sigma}'_{kl} - \lambda_s \Delta\theta \quad (30)$$

where  $C_{ijkl}$  is the compliance modulus. The irreversible element of the strain rate obeys an associative visco-plastic flow law of the form  $\dot{\epsilon}_{ij}^i = \dot{\Pi} n_{ij}$ , where  $n_{ij} = \frac{\partial f}{\partial \sigma'_{ij}} / \left\| \frac{\partial f}{\partial \sigma'_{ij}} \right\|$  the unit vector of the plastic potential and  $f$  is the yield function and  $\dot{\Pi}$  is a (scalar) plastic multiplier (see Fig. 6).

The plastic multiplier describes the overstress function, here taken to be the Euclidean norm of the distance of the current stress state vs the original yield surface:  $\dot{\Pi} = \sqrt{\dot{\epsilon}_d^i{}^2 + \dot{\epsilon}_v^i{}^2}$ . In this expression,  $\dot{\epsilon}_d^i$  and  $\dot{\epsilon}_v^i$  are the deviatoric and volumetric parts of the strain rate tensor (see Fig. 6), respectively, following the incremental relations

$$\dot{\epsilon}_d^i = \dot{\epsilon}_0 \left\langle \frac{q - q_Y}{\sigma_{ref}} \right\rangle^m \exp \left( -\frac{Q_{mech}^d}{R\theta} \right), \quad (31a)$$

$$\dot{\epsilon}_v^i = \dot{\epsilon}_0 \left\langle \frac{p' - p_Y}{\sigma_{ref}} \right\rangle^m \exp \left( -\frac{Q_{mech}^V}{R\theta} \right), \quad (31b)$$

where  $\dot{\epsilon}_0$  is a reference strain rate,  $q$  is the equivalent deviatoric (or von Mises) stress,  $p'$  is the volumetric mean effective stress,  $q_Y$  and  $p_Y$  are the respective effective stresses at yield,  $\sigma_{ref}$  is a reference stress,  $R$  is the universal gas constant,  $\theta$  is the temperature field and  $\langle \cdot \rangle$  denote the Macaulay brackets.

These expressions imply that the material is admitting thermal sensitivity expressed through the activation enthalpies for the deviatoric ( $Q_{mech}^d$ ) and the volumetric ( $Q_{mech}^v$ ) components. This activation enthalpy incorporates the activation energies of all the micromechanical mechanisms, like frictional initiation [Ric06] or volumetric mass diffusion mechanisms [VRL15] like pore collapse [PV16] and can be a function of all the global and internal state variables of the problem at hand. In essence, these exponential expressions act like the hardening laws of classical plasticity allowing for information to be transferred across the scales. An example of such information will be illustrated in the following paragraphs, concerning the mechanism of pore collapse.

The activation enthalpy is in principle expressed in the generalised form  $Q = E + P^*V_{act}$  where  $P^*$  is a measure of pressure responsible for driving the internal mechanism, while  $E$  and  $V_{act}$  are the activation energy and volume of the internal mechanism considered (here pore collapse) and are yet to be determined. Following the recent work of Poulet and Veveakis [PV16], we assume isotropic response of the activation enthalpies ( $Q_{mech}^d = Q_{mech}^v = Q_{mech}$ ), and seek the most appropriate expressions for  $P^*$  and  $V_{act}$ . This will be achieved in the next sections, through fitting data for different tests (in particular triaxial and isotropic compression) and materials (sandstone and mudstone). Following this exercise, the final form of the activation enthalpy will be retrieved.

## 5.1 Modelling Isotropic compression

We start by considering isotropic compression tests conducted on Bleuswiller sandstone, for wet and dry conditions [FGS07]. The specimens used were 80 mm in length and 40 mm in diameter, having initial porosity 25%, grain size 80 - 180  $\mu m$  and permeability  $2 \times 10^{-14} m^2$ . The experiment have conducted one dry and wet conditions under hydrostatic loading at confining pressure up to 280 MPa.

We modelled both dry and wet conditions under hydrostatic loading. The results are shown in the fig. 7. We normalised the stress for the numerical experiments using the yield values of 190 MPa for the dry sample and 132 MPa for the wet sample. Since we have drained isotropic tests, the excess pore pressure was calculated to be zero, therefore suggesting that the activation enthalpy is a linear function of the confinement pressure, obeying:

$$Q_{mech} = E + P'_c OCR \alpha_1 R \theta \quad (32)$$

where in this case the activation volume was expressed as  $V_{act} = \alpha_1 R \theta$ , with  $\alpha_1$  a constant,  $OCR = P'_{c(max)}/P'_c$  the Overconsolidation ratio expressed as the ratio

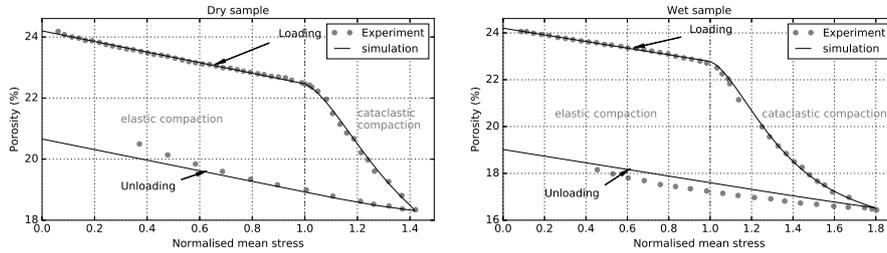


Figure 7: Matching the loading phases of isotropic compression experiments from [FGS07]. The model parameters are calibrated to match the experimental results for the wet case (left) with a reference stress of 132 MPa. The same model is then used with a reference stress of 190 MPa and matches the dry case as well.

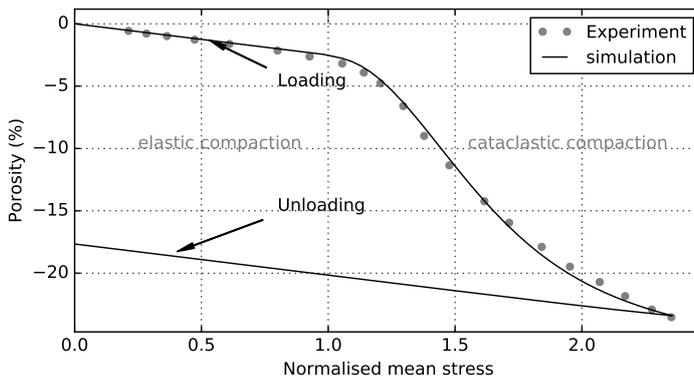


Figure 8: The porosity evolution with increase the effective pressure

between maximum effective confinement and  $P'_{c(max)}$  and the current confining effective pressure  $P'_c$ .

Using this expression, we were also able to match isotropic compression data from tests that have been performed on diatomaceous mudstone [OKH<sup>+</sup>11]. The authors of this experimental work [OKH<sup>+</sup>11] used a prismatic specimen with 8 cm high and 4 cm side. We have modeled this behavior and there is an excellent agreement of the simulation with the experimental data as shown in figure 8.

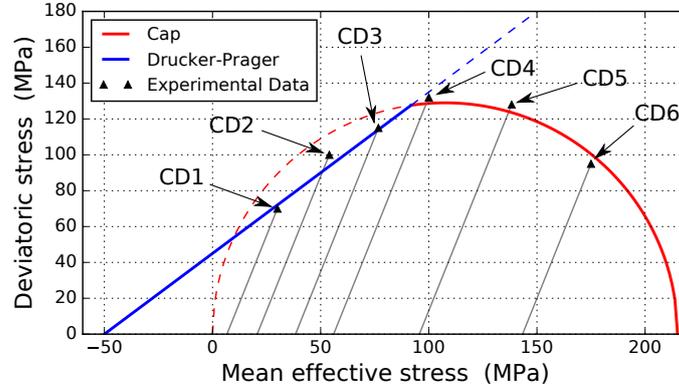


Figure 9: The experimentally derived yield points (circles) for experiments performed at 6 different confinements (CD 1-6) by [WDZ97] and the modified cam-clay yield surface used in modelling the tests (solid line).

## 5.2 Modelling Triaxial compression

To show how is the model is responding against real data from the experiments, we have chosen a very notable set of experiments in sedimentary rocks. The experiments were conducted by [WDZ97] on Bleurswiller sandstone at different confinements. The samples had initial porosity of 22.6 % and were cored parallel to the bedding. Cylindrical sample have been used for this test ( 18.4 mm diameter and 38.1 mm length ), under a fix loading rate of  $5 * 10^{-5}$ /s. The levels of confinement pressure to observe different deformation patterns were (5, 20, 40, 60, 100, 150) Mpa. This sandstone materials has a cap yield surface as shown in figure (9) which is a combination of Drucker prager and modified cam clay as a cap .

In order to model the response of sandstone, we run a numerical experiment to model the real experiment to reproduce the stress strain curves with different confining pressure and the methodolgy has been described by [PV16]. We fit the experimental curves reported for these materials by using the material properties listed on table 1 and varying only  $V_{act}$  at different confining pressures. The optimal fit acquired is shown in Fig. 10, and the formula of  $Q_{mech}$  was inverted to be:

$$Q_{mech} = E_0 + P^* V_{act} \quad (33)$$

$$P^* = P'_c OCR - \frac{\alpha_2}{\alpha_1} P_f \quad (34)$$

$$V_{act} = \alpha_1 \left( 1 - \frac{\ln P'_c OCR}{\ln P_{cs}} \right) R\theta \quad (35)$$

This expression implies that the energy term  $P^* V_{act}$  represents the effect of interface strength of cemented materials (whether by solid cement, capillary bridges etc),

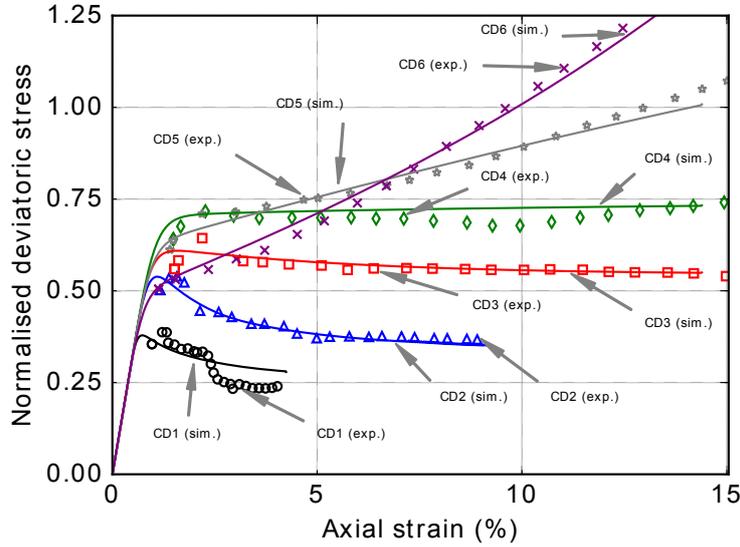


Figure 10: The deviatoric stress ( $\tau$ ) vs. axial strain results for the experiments (symbols) and the numerical simulations (lines).

with the driving pressure  $P^*$  being a Bishop-like effective pressure representing the maximum effective pressure the interface has ever experienced. The activation volume required to describe the strength of this interface follows a stress path dependent Laplace-like law, in which  $P_{cs}$  represents the initial effective confinement required for the material to end in critical state while being loaded at a given stress path.

Having inverted for the final formula of  $Q_{mech}$ , we are extending the work by fitting a different material. To this end we recall the triaxial compression experiments that have been conducted on diatomaceous mudstone by [OKH<sup>+</sup>11]. They performed a series of triaxial tests on six rectangular shaped-prismatic specimens with 8 cm high and 4 cm side. In order to avoid the effect of the initial anisotropy the specimens was made with longitudinal direction perpendicular to the plane of sedimentation. They used different levels of confining pressure to observe different deformation patterns and the scenarios are listed in Tab. 2. All tests were pursued to a maximum axial strain of 20% approximately.

To model the response of this soft material and its behavior for this specific test, we have used the inverted formula of  $Q_{mech}$  of Eq. 33, combined with a modify cam clay as a yield surface (Fig.11), defined as:

$$\left(\frac{q}{M}\right)^2 + p(p - p_c) = 0 \quad (36)$$

where  $p_c$  is the pre-consolidation pressure and  $M$  the slope of the critical state line.

Parameter	Sandstone
$c_{th}$ [ $m^2/s$ ]	$10^{-6}$
$k_{\pi}$ [ $m^2$ ]	$3 \times 10^{-16}$
$\mu_f$ [ $Pa.s$ ]	$10^{-3}$
$\beta_m$ [ $Pa^{-1}$ ]	$10^{-7}$
$\dot{\epsilon}_0$ [ $s^{-1}$ ]	1
$E_{mech}$ [ $Jmol^{-1}K^{-1}$ ]	25
$\lambda_m$ [ $K^{-1}$ ]	$10^{-3}$
$\sigma'_{ref}$ [ $MPa$ ]	189
$T_{ref}$ [ $K$ ]	300
$x_{ref}$ [ $m$ ]	0.02
$\chi$ [—]	1

Table 1: Parameters used in order to fit the experimental data for sandstone. We have used the expression  $c_{hy} = k_{\pi}/\mu_f\beta_m$  for the hydraulic diffusivity where  $k_{\pi}$  the permeability and  $\mu_f$  the fluid viscosity.

Table 2: Confinement pressures used in experiments from [OKH<sup>+</sup>11]

Case No.	CD1	CD3	CD3	CD4	CD5	CD6
Effective confining pressure (MPa)	0.25	0.5	0.75	1.0	1.5	2.0

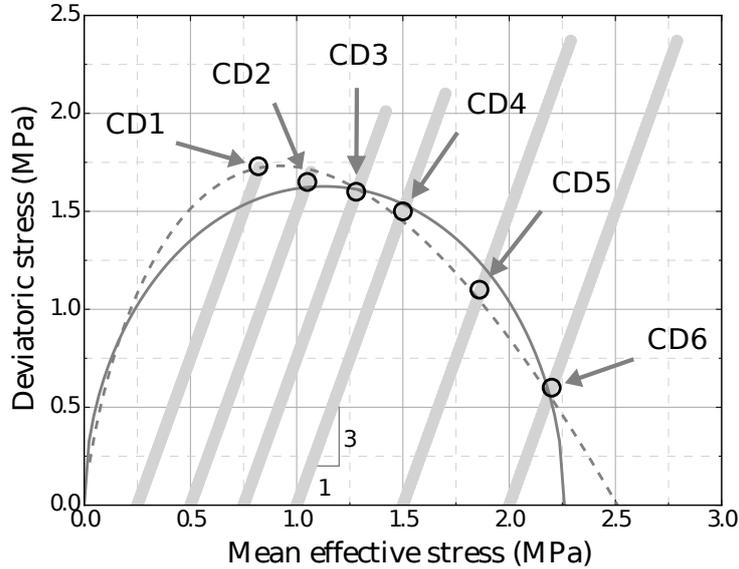


Figure 11: Best Fits for Modified Cam Clay (plain line) and Original Cam Clay (dashed line) yield surfaces are plotted with experimental stress paths and yield points (dots) digitised from Fig. 5 in [OKH<sup>+</sup>11]

We fit the experimental curves reported for these materials by using the material properties listed in Tab. 5.2. The results are shown in figure 12.

It is to be noted that both materials were fitted without significant temperature increase inside the sample (less than one degree). Also, the optimal values of  $\alpha_1$  and  $\alpha_2$  in the expressions of Eqs. 33 -characterizing the driving pressure  $P^*$  and the activation volume  $V_{act}$  respectively- are such that the system's dissipation rate self organizes along a global maximum.

## 6 Conclusions

In this work we have presented an energy-based constitutive framework that can accommodate multi-physical mechanisms. We have shown that within the theory of multiplicative visco-plasticity we may explicitly account for the temperature sensitivity of materials and that all commonly used laws for thermal sensitivity produce the same qualitative response. We have therefore used the most generic one to generalize for a 3D constitutive law of thermo-poro-elasto-viscoplasticity and showed that by considering internal interface mechanisms we may reproduce laboratory results for different rocks and tests with the same model. Following these promising results,

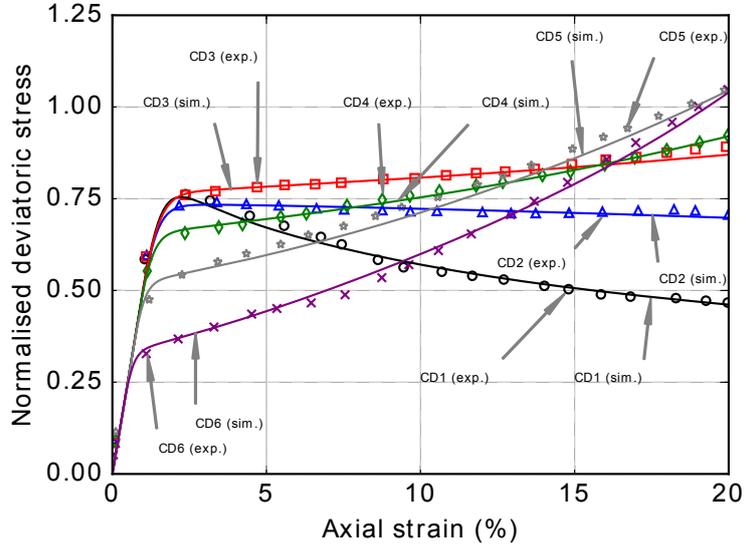


Figure 12: Matching CD1-CD6 experiments with simulation results. Note that the results are shown in a normalised stress space, where the preconsolidation pressure  $\sigma_{ref} = 2.26$  MPa is used as the reference value of the stress used for normalising the experimental data. In the absence of information from [OKH<sup>+</sup> 11] on the value of the pore fluid pressure during the experiments, the numerical results are normalised using the same value, requiring that the maximum pore fluid pressure during the experiments is around 1 MPa.

Parameter	Mudstone
$c_{th}$ [m <sup>2</sup> /s]	10 <sup>-6</sup>
$k_{\pi}$ [m <sup>2</sup> ]	10 <sup>-16</sup>
$\mu_f$ [Pa.s]	10 <sup>-3</sup>
$\beta_m$ [Pa <sup>-1</sup> ]	10 <sup>-7</sup>
$\dot{\epsilon}_0$ [s <sup>-1</sup> ]	1
$E_{mech}$ [Jmol <sup>-1</sup> K <sup>-1</sup> ]	18
$\lambda_m$ [K <sup>-1</sup> ]	10 <sup>-6</sup>
$\sigma'_{ref}$ [MPa]	2.26
$T_{ref}$ [K]	300
$x_{ref}$ [m]	0.02
$\chi$ [—]	0.6

Table 3: Parameters used in order to fit the experimental data for mudstone. We have used the expression  $c_{hy} = k_{\pi}/\mu_f\beta_m$  for the hydraulic diffusivity where  $k_{\pi}$  the permeability and  $\mu_f$  the fluid viscosity.

more work is required to further constrain the open parameters used in this model and reduce the number of experiments required to attain the same amount of information.

## References

- [APV14] S. Alevizos, T. Poulet, and E. Veveakis. Thermo-poro-mechanics of chemically active creeping faults. 1: Theory and steady state considerations. *Journal of Geophysical Research: Solid Earth*, 119(6):4558–4582, 2014.
- [BMJ95] Lockner D.A. Blanpied M.L. and Byerlee J.D. Frictional slip of granite at hydrothermal conditions. *J. Geophys. Res. - Solid Earth*, 100:13045–13064., 1995.
- [CC98] F. M. Chester and J. S. Chester. Ultracataclasite structure and friction processes of the punchbowl fault, san andreas system, california. *Tectonophysics*, 295:199–221, 1998.
- [CDMM89] H.T.Z. Chen, A. S. Douglas, and R. Malek-Madani. An asymptotic stability condition for inhomogeneous simple shear. *Q. Appl. Math.*, 47:247–262, 1989.
- [CK91] TFC. Chan and HB. Keller. Arc-length continuation and multi-grid techniques for nonlinear elliptic eigenvalue problems. *SIAM J. Sci. Stat. Comput.*, 3(2):173 – 194, 1991.
- [CL04] C. Cekerevac and L. Laloui. Experimental study of thermal effects on the mechanical behaviour of a clay,. *Int. J. Num. Anal. Meth. Geomech.*, 28:209–228, 2004.
- [CZV11] Francesco Cecinato, Antonis Zervos, and Emmanuil Veveakis. A thermo-mechanical model for the catastrophic collapse of large landslides. *International Journal for Numerical and Analytical Methods in Geomechanics*, 35(14):1507–1535, 2011.
- [Die72] J.H. Dieterich. Time-dependent friction in rocks. *J. Geophys. Res.*, 377:3690–3697, 1972.
- [Die78] J.H. Dieterich. Time-dependent friction and the mechanics of stick-slip. *Pure Appl. Geophys.*, 116:790–806, 1978.
- [FGS07] Jérôme Fortin, Yves Guéguen, and Alexandre Schubnel. Effects of pore collapse and grain crushing on ultrasonic velocities and vp/vs. *ournal of Geophysical Research: Solid Earth*, page B08207, 2007.
- [FK05] Y. Fialko and Y. Khazan. Fusion by earthquake fault friction: Stick or slip? *J. Geophys. Res.*, 110:B12407, 2005.

- [Fow97] A.C. Fowler, editor. *Mathematical Models in the Applied Sciences*. Cambridge University Press, 2 edition, 1997.
- [Gar12] D. I. Garagash. Seismic and aseismic slip pulses driven by thermal pressurization of pore fluid. *Journal of Geophysical Research: Solid Earth*, 117(B4), 2012.
- [GRRT84] J.C. Gu, J. R. Rice, A. L. Ruina, and S. T. Tse. Slip motion and stability of a single degree of freedom elastic system with rate and state dependent friction. *J.Mech.Phys.Solids*, 32(3):167–196, 1984.
- [Gru63] I. Grunfest. Thermal feedback in liquid flow: Plane shear at constant stress. *Trans. Soc. Rheol.*, 7:95–207., 1963.
- [HB90] T. Hueckel and G. Baldi. Thermoplastic behavior of saturated clays: an experimental constitutive study. *J. Geotech. Eng.*, 116(12):1778–1796, 1990.
- [Hic74] P. Hicher. Etude des proprietes mecaniques des argiles a l' aide d'essais triaxiaux, in uence de la vitesse et de la temperature. *Soil Mech. Lab., Ecole Cent. de Paris, Paris*, 1974.
- [Hil62] R. Hill. Acceleration waves in solids. *J. Mech. Phys. Solids*, 10:1–16, 1962.
- [HP02] T. Hueckel and R. Pellegrini. Reactive plasticity for clays: application to a natural analog of long-term geomechanical effects of nuclear waste disposal. *Eng. Geology*, 64:195–215, 2002.
- [LA07] Goren L. and E. Aharonov. Long runout landslides: The role of frictional heating and hydraulic diffusivity. *Geophys. Res. Lett.*, 34:L07301, 2007.
- [Law06] C.K. Law, editor. *Combustion Physics*. Cambridge University Press, 2006.
- [Lei76] H.J. Leinenkugel. Deformations und festigkeitverhalten bindiger erdstoffe: Experimental ergebnisse und ihre physikalische deutung. Doctoral Dissertation, Univ. Karlsruhe, Karlsruhe, Germany, 1976.
- [LLC08] L. Laloui, S. Leroueil, and S. Chalindar. Modelling the combined effects of strain rate and temperature on one-dimensional compression of soils,. *Can. Geotech. J.*, 45:173 – 194, 2008.
- [Man66] J. Mandel. Conditions de stabilite et postulate de drucker. *Rheology and Soil Mechanics*, pages 58–67, 1966.
- [Mar98] C. Marone. Laboratory-derived friction laws and their application to seismic faulting. *Annu. Rev. Erth Planet. Sci.*, 26:643–696, 1998.
- [MLB9a] M. Matalon, G.S.S Ludford, and J. Buckmaster. Diffusion flames in a chamber. *Acta Astronautica*, 6:943–959, 1979a.

- [MV87] H. Muhlhaus and I. Vardoulakis. Thickness of shear bands in granular materials. *Geotechnique*, 37(3):271–283, 1987.
- [OKH<sup>+</sup>11] F. Oka, S. Kimoto, Y. Higo, H. Ohta, T. Sanagawa, and T. Kodaka. An elasto-viscoplastic model for diatomaceous mudstone and numerical simulation of compaction bands. *International Journal for Numerical and Analytical Methods in Geomechanics*, 35(2):244–263, 2011.
- [Per66] P. Perzyna. Fundamental problems in viscoplasticity. *Adv. Appl. Mech.*, 9:243–377, 1966.
- [Pri90] I. Prigogine. Time, dynamics and chaos: Integrating poicare’s ‘non-integrable systems’. In *Nobel Conference XXVI*, volume CONF-9010321-1, pages 1–27. U.S. Department of Energy, Office of Energy Research (DOE/ER); Commission of the European Communities (CEC), 1990.
- [PV11] S. Papanicolopoulos and E. Veveakis. Sliding and rolling dissipation in cosserat plasticity. *Granular Matter*, 13(3):197–204, 2011.
- [PV16] T. Poulet and M. Veveakis. A viscoplastic approach for pore collapse in saturated soft rocks using redback: an open-source parallel simulator for rock mechanics with dissipative feedbacks. *Computers and Geotechnics*, 74:211–221, 2016.
- [Ric06] J. R. Rice. Heating and weakening of faults during earthquake slip. *J. Geophys. Res.*, 111:B05311, 2006.
- [RJ75] Rice JR. Rudnicki JW. Conditions for the localization of deformation in pressure sensitive dilatant materials. *J. Mech. Phys. Solids*, 23:371–394, 1975.
- [Sch98] C.H. Scholz. Earthquakes and friction laws. *Nature*, 391:37–42, 1998.
- [SF09] J. Sulem and V. Famin. Thermal decomposition of carbonates in fault zones: slip-weakening and temperature-limiting effects. *J. Geophys. Res.*, 114:B03309, 2009.
- [SJP05] Ouffroukh H. Sulem J., Vardoulakis I. and V. Perdikatsis. Thermo-poro-mechanical properties of the aigion fault clayey gouge - application to the analysis of shear heating and fluid pressurization. *Soils and Foundations*, 45:97–108, 2005.
- [SSV11] J. Sulem, I. Stefanou, and E. Veveakis. Stability analysis of undrained adiabatic shearing of a rock layer with cosserat microstructure. *Granular Matter*, 13(3):261–268, 2011.
- [TQ34] G. Taylor and H. Quinney. The latent energy remaining in a metal after cold working. *Proc. R. Soc., Ser. A.*, 143:307 – 326., 1934.

- [Tru69] C. Truesdell, editor. *Rational Thermodynamics*. McGraw-Hill Series in Modern Applied Mathematics, 1969.
- [Var2a] I. Vardoulakis. Dynamic thermo-poro-mechanical analysis of catastrophic landslides. *Geotechnique*, 52:157–171, 2002a.
- [Var00] I. Vardoulakis. Catastrophic landslides due to frictional heating of the failure plane. *Mech. Cohesive Frict. Mater.*, 5:443–467, 2000.
- [Var02] I. Vardoulakis. Steady shear and thermal run-away in clayey gouges. *Int. J. Solids Struct.*, 39:3831 – 3844, 2002.
- [VAV10] E. Veveakis, S. Alevizos, and I. Vardoulakis. Chemical reaction capping of thermal instabilities during shear of frictional faults. *J. Mech. Phys. Solids*, 58:1175 – 1194, 2010.
- [VPA14] E. Veveakis, T. Poulet, and S. Alevizos. Thermo-poro-mechanics of chemically active creeping faults: 2. transient considerations. *Journal of Geophysical Research: Solid Earth*, 119(6):4583–4605, 2014.
- [VRL14] E. Veveakis and K. Regenauer-Lieb. The fluid dynamics of solid mechanical shear zones. *Pure Appl. Geophys.*, 171:3159–3174, 2014.
- [VRL15] E. Veveakis and Regenauer-Lieb. Cnoidal waves in solids. *Journal of Mechanics and Physics of Solids*, 78:231–248, 2015.
- [VS95] I. Vardoulakis and J. Sulem, editors. *Bifurcation Analysis in Geomechanics*. Blackie Acc. and Professional, 1995.
- [VSS12] E. Veveakis, J. Sulem, and I. Stefanou. Modeling of fault gouges with cosserat continuum mechanics: Influence of thermal pressurization and chemical decomposition as coseismic weakening mechanisms. *J. Struct. Geology*, 38:254–264, 2012.
- [VSS13] E. Veveakis, I. Stefanou, and J. Sulem. Failure in shear bands for granular materials: thermo-hydro-chemo-mechanical effects. *Geotechnique Let.*, 3(2):31–36, 2013.
- [VVD07] E. Veveakis, I. Vardoulakis, and G. DiToro. Thermoporomechanics of creeping landslides: The 1963 vaiont slide, northern italy. *J. Geophys. Res.*, 112:F03026, 2007.
- [WDZ97] T. Wong, C. David, and W. Zhu. The transition from brittle faulting to cataclastic flow in porous sandstones: Mechanical deformation. *Journal of Geophysical Research: Solid Earth*, 102(B2):3009–3025, 1997.



---

## **Energetics of Crushable Granular Materials: from Particle Fracture to Breakage Mechanics**

**Yida Zhang<sup>1</sup> and Giuseppe Buscarnera<sup>2</sup>**

<sup>1</sup> *Department of Civil, Environmental and Architectural engineering, University of Colorado Boulder, Colorado, USA*

<sup>2</sup> *Department of Civil and Environmental Engineering, Northwestern University, Illinois, USA.*

---

*Much of the complexity of the macroscopic behaviour of granular materials originates from interactions among their constituents, as well as from the propagation of these interactions across scales. Examples of these multi-scale processes are mechanisms such as grain rearrangement and breakage, which are often driven by the simultaneous action of mechanical and environmental phenomena. Understanding these processes is important to engineer geotechnical structures, design infrastructure materials, and optimize material processing technologies. This chapter reviews the application of energy methods to crushable granular materials subjected to multiphysical loading. First, the energetics and kinetics of fracture growth in brittle solids will be reviewed in the context of the thermodynamics of internal variables (TIV). The same methodology will then be extended to a granular system consisting of brittle particles. By defining the work input and energy storage for the system under study, it will be shown that the multiphysical response of particulate materials can be modelled in close analogy to the mathematics of fracture in brittle solids. The goal of this chapter is thus to show with reference to the particular case of crushable granular materials how TIV can be used to interpret a variety of multi-physical processes taking place at very different length scales.*

## 1 Introduction

The mechanics of granular geomaterials has been studied in the context of various disciplines, such as geotechnical engineering, petroleum engineering, geoscience and physics due to their pervasive presence in nature and the wide range of relevant technological and industrial applications. At elevated pressure, particle breakage is one of the crucial mechanisms governing mechanical and hydraulic properties of granular assemblies. For example, grain breakage is responsible for excessive settlement during dam construction [Old01], subsidence due to depletion of hydrocarbon reservoirs [Ver17], and long-runout behaviour of landslides [Wan02]. Grain crushing is often exacerbated in the presence of environmental loadings. Evidence includes water vapor assisted creep of rockfill [Old07], saturation-induced crushing of sands [Ova13] and chemical weakening of granular rocks [Bau00]. Since geomaterials are exposed to constantly changing environmental conditions, studying the interplay between grain crushing and hydro-chemo-mechanical agents becomes crucial to assess the short and long-term behaviour of many geotechnical and geological systems.

The understanding of grain crushing has been significantly improved thanks to the numerous experimental, numerical and theoretical studies conducted during the past decades [Har85, Lad96, McD96, Nak01a, Ova13, Cil14]. Some key conclusions from these studies are summarized below:

- 1) The yielding and hardening in high pressure compression tests correspond to the onset and accumulation of grain breakage. The terms ‘elastic yielding’ and ‘elastic hardening’ are often used to highlight the role of grain breakage in these processes [McD98].
- 2) Crushing of grain assemblies depends on individual particle characteristics, packing condition and grain size distribution (GSD) [Nak99, Nak01b, Nak01a]. Large values of yielding stress are often observed in specimens composed of tougher particles and smaller grains; Well-rounded particles give sharper yielding transition than angular particles; Densely packed specimens display higher yielding stress than loosely packed ones [Nak01a, McD02c].
- 3) Granular assemblies monotonically loaded (e.g. via 1D compression and simple shear) to sufficiently high compressive stresses tend to achieve an ultimate fractal GSD [Sam87, McD96]. The fractal dimension of such GSD inferred from laboratory and numerical experiments is about 2.5~2.7 [Sam87, Ben10, McD13].
- 4) The creation of new surface area through particle fragmentation must be considered as a source of energy dissipation in addition to frictional dissipation [McD98]. It has also been recognized that there is an additional amount of energy dissipated through the redistribution of the contact forces across the soil matrix upon each breakage event [Ngu09, Rus13].

Existing models which tackle the mechanics of crushable granular soils can be categorized into three groups: phenomenological approaches based on elastoplasticity, micromechanical models based on fractal mathematics, and thermomechanical models based on breakage mechanics theory. In elastoplasticity approaches, the primary

focus is on the stress-strain response, and the evolution of soil gradation is a not explicitly considered or modelled. For example the model by [Pes95] describes the high pressure 1D compression response of sand via an incremental stress-strain relation that implicitly incorporated the onset of particle breakage by approaching to the so-called Limiting Compression Curve (LCC). The elastoplastic model proposed by [Kik10] introduces gradation-dependent critical state. The yield mechanism associated with grain crushing is described via a separate yield function (i.e. a cap surface) in addition to the Mohr-Coulomb yield surface. These models often exhibit good agreement with the stress-strain responses, but overlook the microscopic mechanisms and the energetics associated with the creation of new surface area. On the other hand, fractal models rely on the grain size distribution, the pore size distribution and the single-particle failure probability to describe the evolution of gradation and porosity under isotropic compression [McD96, McD98, Rus11]. The energy function of [Ros63] is often revised to account for the energy dissipation due to creation of new surface area, to establish the macroscopic compressive response of the system. These approaches are physically appealing, considering their potential to link with the microscopic statistics of granular soils. However, they have been so far limited to only the 1D case and have not been generalized to triaxial compression or simple shear paths.

The Breakage Mechanics theory [Ein07a] offers an energy-based continuum framework to model crushable granular materials, while providing analytical tools to up-scale microstructural processes at the continuum level. The theory is established within the framework of thermodynamics with internal variables (TIV) by using the grain size distribution to define a new internal variable referred to as *Breakage*. Such enrichment allows the coupling of the energy, microstructure, and stress-strain response of the granular assembly. In this paper, the intimate connection between the particle-scale energetics of grain fracture and the mechanics of continuum breakage in granular solids is discussed. In particular, it will be shown how this framework represents a versatile platform to model the rate dependency and environmental effect associated with grain crushing.

## 2 Energetics of brittle fracture

Since the pioneering work of [Gri21], fracture mechanics has become a discipline that is intimately connected with energy principles and methodologies. The Griffith criterion states that cracks propagate if the energy available to extend their length (i.e. the energy release rate) is equal to the surface energy of the newly created area, which is essentially a special case of the principle of energy balance. This energy-based approach circumvents the difficulty of dealing with infinite tensile stresses at the crack tip and accounts for the most crucial factors controlling the growth of cracks, such as material properties, crack geometry, and stress level. Since its proposition, numerous generalizations have been developed, e.g. for nonlinear solids, non-isothermal conditions, dynamic propagations, and fatigue under cyclic and environmental loadings.

Others seek for reinterpretation of Griffith theory in the context of modern thermodynamic theories. One of the major achievements is the J-integral [Ric68]. When studying the energetics of fracture propagation, Rice identified a path-independent integral that characterizes the variation of elastic energy during an infinitesimal advancement of the crack front, which is later proven to be the identical to the energy release rate in the Griffith theory. Such development has been extensively used in computational methods to calculate the energy release rate (e.g., [Moe99, Sta03]). Another important work by [Ric78] is the recast of the Griffith theory in the framework of TIV. Such treatment has provided a unifying perspective of the quasi-static and rate-dependent propagation of cracks, as well as a framework to incorporate environmental factors. Here we summarize Rice's approach adapted for crack growth in a uniaxially stressed particle, and show how the same results can be achieved through the Hyperplasticity formalism [Hou07].

Consider a disk-shaped particle (2D loading configuration) that contains a pre-existing flaw. The state of the selected system can be characterized by an external variable  $\Delta$  (representing the reversible displacement of the loading plate) and an internal variable  $l$  (representing the length of the crack, here assumed to be monotonically increasing). Isothermal conditions are assumed, thus including the possibility of heat exchanges with the environment (e.g., the heat flow  $Q$  in figure 1).

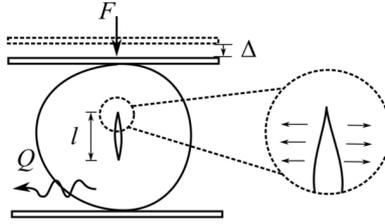


Figure 1: Schematic diagram of fracturing process in disk-shaped particle with centre crack. A tensile stress field is assumed to prevail near the crack tip, thus considering a scenario of mode I crack propagation.

## 2.1 Rice's approach

The first law of thermodynamics states:

$$\dot{Q} + F\dot{\Delta} = \dot{U} \quad (1)$$

where  $U$  is the internal energy of the system which depends on  $\Delta$  and  $S$  (external variables) and crack length  $l$  (an internal variable). The second law can be written in the following form:

$$\dot{S} = \dot{Q}/T + \Lambda_c, \quad \Lambda_c \geq 0 \quad (2)$$

where  $S$  is entropy;  $\Lambda_c$  is the entropy production rate during cracking. Combining equations (1) and (2) to eliminate  $\dot{Q}$ , one obtains the free energy balance equation:

$$F\dot{\Delta} - \dot{\Psi} = T\Lambda_c \geq 0 \quad (3)$$

where  $\Psi$  is the Helmholtz free energy of the entire system (per unit thickness) defined as:

$$\Psi(\Delta, l, T) = U(\Delta, l, S) - TS \quad (4)$$

Considering isothermal conditions (i.e.  $\dot{T} = 0$ ), equation (3) can be rewritten as:

$$\left(F - \frac{\partial \Psi}{\partial \Delta}\right)\dot{\Delta} - \frac{\partial \Psi}{\partial l}\dot{l} = T\Lambda_c \geq 0 \quad (5)$$

The system's Helmholtz free energy can be decomposed as energy stored in the solid as elastic strain energy  $\Psi_s$  and on the crack surface in the form of surface tension:

$$\Psi = \Psi_s(\Delta, l) + 2\gamma l \quad (6)$$

Here the surface is treated as an independent phase in the Gibb's sense, i.e. a surface phase can carry mass, momentum, energy and entropy but with zero thickness.

According to the definition of elastic strain energy, the following relations can be derived:

$$F = \frac{\partial \Psi_s}{\partial \Delta} = \frac{\partial \Psi}{\partial \Delta}, \quad G_I = -\frac{\partial \Psi_s}{\partial l} \quad (7)$$

where  $G_I$  is the type-I (tension) energy release rate. Substituting equations (6) and (7) into (5) yields:

$$(G_I - 2\gamma)\dot{l} = T\Lambda_c \geq 0 \quad (8)$$

equation (8) recovers the classical Griffith criterion for the case in which only crack growth is allowed (i.e.,  $\dot{l} \geq 0$ ), having that:

$$\begin{cases} \dot{l} = 0 & \text{for } 0 \leq G_I < 2\gamma \\ \dot{l} \geq 0 & \text{for } G_I \geq 2\gamma \end{cases} \quad (9)$$

## 2.2 Hyperplasticity formalism

An alternative way of deriving equation (8) is to adopt the Hyperplasticity formalism [Hou07]. This methodology defines the constitutive response of the system via two scalar functions, namely the Helmholtz free energy and the dissipation rate function. Under isothermal conditions, the first and second laws can be written as:

$$W = \dot{\Psi} + \Phi, \text{ with } \Phi \geq 0 \quad (10)$$

where  $W$  is the work input rate and  $\Phi$  is the dissipation rate function. For the system depicted in figure 1, the only source of work input is external loading,  $W = F\dot{\Delta}$ . For  $\Psi$  and  $\Phi$ , we take a slightly different view compared to Rice: the Helmholtz free energy is identical to the elastic strain energy and all the energy transferred to the surface is considered as “dissipated”. This is because we focus on crack growth only and the energy that enters in the surface phase is regarded as irreversible. Thus, one can write

$$\Psi = \Psi_s(\Delta, l) \quad (11)$$

$$\Phi = 2\gamma\dot{l} + T\Lambda_c \geq 0 \quad (12)$$

Substituting equation (11) into (10) gives:

$$\left(F - \frac{\partial \Psi_s}{\partial \Delta}\right)\dot{\Delta} + \left(-\frac{\partial \Psi_s}{\partial l} - 2\gamma\right)\dot{l} = T\Lambda_c \geq 0 \quad (13)$$

Substituting equation (7) into (13) one obtains again the Rice’s criterion (8). The new insight provided by this formalism is the view of fracture as a yielding phenomenon. For example, assuming that energy is dissipated solely by surface area creation (i.e.  $\Lambda_c = 0$ ),  $\Phi$  becomes a homogenous function of  $\dot{l}$  of degree 1 and its degenerated Legendre transformation represents a yield surface in the stress space [Hou07]:

$$y_c = \frac{G_I}{2\gamma} - 1 = 0 \quad (14)$$

Such a criterion bounds the reversible regime for cracked solid and represents a yield surface for fracture expressed in energy terms. Note that this yield surface does not contain any hardening parameters, the predicted fracture phenomenon resembles the elastic-perfectly-plastic scenario. This means the crack velocity becomes unbounded once the critical energy release rate is reached ( $G_I = G_{IC} = 2\gamma$ ) (figure 2).

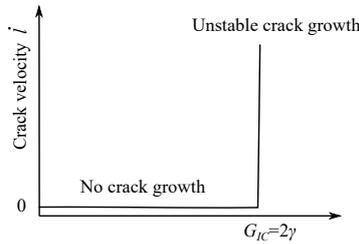


Figure 2: Griffith crack model in  $G_I - \dot{l}$  plane.

### 3 Kinetics of crack growth

The predicted unbounded crack growth velocity at  $G_{IC}$  and the lack of crack growth below  $G_{IC}$  does not agree with experimental observations of brittle solids in natural environments. A common phenomenon is the so-called subcritical crack growth, referring to the extension of cracks at a finite velocity below the critical threshold  $G_{IC}$  [Atk87]. This phenomenon has been observed in glasses, ceramics, quartz and various types of natural rocks, and is highly sensitive to environmental factors such as relative humidity and temperature [Wie67, Atk82]. Figure 3a shows a typical plot of crack propagation velocity,  $\dot{l}$ , as a function of the stress intensity factor  $K_I$  (related to energy release rate via the celebrated Irwin's relation  $G_I = K_I^2 / E$ ), in the presence of various degrees of water vapour concentration. Three crack growth regimes can be identified, all located at values of  $K_I$  greater than  $K_{I0}$  (often referred to as *stress corrosion limit*). Region I, is the regime where the crack velocity increases exponentially, and it is assumed to be controlled by the reaction rate between environmental species and strained bonds in proximity of the crack tip (thus making  $\dot{l}$  dependent on both  $K_I$  and chemical conditions). At higher values of  $K_I$  (region II), where  $\dot{l}$  displays a much weaker dependence on  $K_I$ , the diffusion of the reactive species towards the crack tip is often hypothesized to govern the crack propagation velocity. Finally, in region III  $K_I$  approaches the critical value  $K_{IC}$  in vacuum, at which the crack propagates catastrophically without displaying any noticeable dependence on the environmental agents. Based on such observations, many kinetic models have been proposed (see the review by [Atk87]). Among them, the Charles equation [Cha58] has been widely applied to interpret subcritical crack growth data. A possible law to express such a simple kinetics law is given as follows [Old07]:

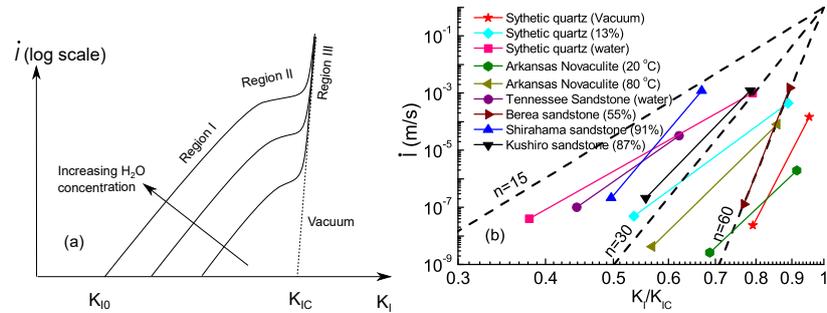


Figure 3: a) Typical stress corrosion curves represented in the  $K_I - \dot{l}$  plane (after [Atk87]); b)  $K_I - \dot{l}$  diagrams for quartz-rich rocks exposed to various environmental conditions. The dash lines represent the Charles equation used with  $K_{Iref} = K_{IC}$  and  $\dot{l}_0 = 1$  m/s (data from [Atk80, Atk84, Nar12]).

$$\dot{i} = \dot{i}_0 \left( \frac{K_I}{K_{Iref}} \right)^n \quad (15)$$

where  $n$  is the stress corrosion index (i.e., the constant controlling the rate of crack advancement);  $\dot{i}_0$  is a reference crack propagation velocity; and  $K_{Iref}$  is a stress intensity factor used for normalization purposes (which depending on convenience may be either the fracture toughness under vacuum,  $K_{IC}$ , or the minimum stress intensity factor for fracture growth under the prevailing environmental conditions,  $K_{I0}$ ). The Charles equation satisfactorily models the crack growth behaviour observed in region I for most minerals and rocks. For example, figure 3b collects  $K_I - \dot{i}$  diagrams for synthetic quartz and quartz-rich rocks subjected to different environments, and it shows that the values of  $n$  may vary between 15 and 60.

A limitation of the Charles equation is that it implies crack extension in the presence of any magnitude of stress concentration. In other words, it does not have a stress corrosion limit below which the growth of a crack ceases. This limit, however, is necessary to explain the rate of crack growth in glasses at low values of stress intensity factor [Wie70], and its existence has also been hypothesized for rocks [Atk87]. An alternative model that includes a stress corrosion limit was proposed by [Mau85], and it can be expressed through the following relation:

$$\dot{i} = \frac{1}{\alpha^{1/n'}} \left\langle \frac{K_I^2}{K_{I0}^2} - 1 \right\rangle^{1/n'} \quad (16)$$

where  $\alpha$  and  $n'$  are two model parameters and the Macauley brackets are used to limit the length evolution to crack growth only (i.e.,  $\dot{i} > 0$ ). It is apparent that such a model imposes a threshold value of  $K_I$  below which the growth of a crack is not possible (i.e.,  $K_I < K_{I0}$ ). By specifying the following relations between the parameters of equations (15) and (16):

$$n' = \frac{2}{n} ; \alpha = \frac{1}{\dot{i}_0^{2/n}} \quad (17)$$

it is possible to show that the Charles and Maugis equations have the same asymptotic behaviour at large values of  $K_I$ :

$$\lim_{K_I \rightarrow \infty} \left[ \dot{i}_0 \left( \frac{K_I}{K_{I0}} \right)^n \right] / \left[ \frac{1}{\alpha^{1/n'}} \left\langle \frac{K_I^2}{K_{I0}^2} - 1 \right\rangle^{1/n'} \right] = 1 \quad (18)$$

where equation (1) has been used with  $K_{Iref} = K_{I0}$ . Thus, the Maugis equation can be expressed in terms of the widely used parameters of the Charles law, as follows:

$$\dot{i} = \dot{i}_0 \left\langle \frac{K_I^2}{K_{I0}^2} - 1 \right\rangle^{n/2} = \dot{i}_0 \left\langle \frac{G_I}{2\gamma} - 1 \right\rangle^{n/2} \quad (19)$$

Figure 4a shows a plot of the Charles and Maugis equations on a double-logarithmic scale, from which the convergence of the two models at large values of  $K_I$  is readily apparent. While both laws depict a subcritical crack growth with the same corrosion index ( $n=10$ ), the Maugis equation also imposes a stress corrosion limit ( $K_{I0}=0.1 \text{ MPa}\cdot\text{m}^{0.5}$ ). Figure 4b illustrates the importance of such limit with reference to data reported by [Wie70]. It can be noted that the non-linearity of the crack growth process in proximity of the stress corrosion limit is accurately captured by equation (19), which therefore offers both conceptual and practical advantages.

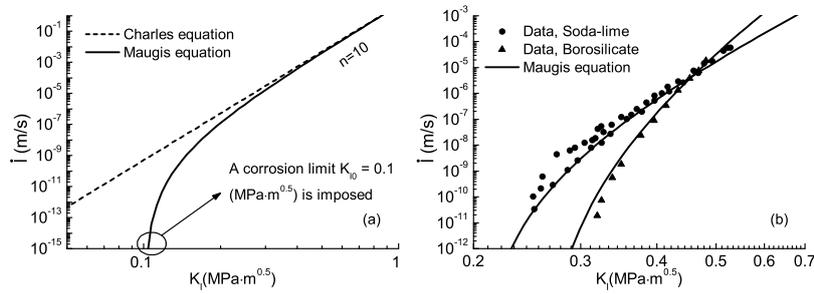


Figure 4: a) Typical plot of Charles and Maugis crack propagation equation ( $K_{I0}=0.1 \text{ MPa}\cdot\text{m}^{0.5}$ ,  $n=10$  and  $\dot{i}_0=5\times 10^{-10} \text{ m/s}$ ); b) Interpretation data for soda-lime and borosilicate glass [Wie70] through the Maugis equation ( $K_{I0}=0.2 \text{ MPa}\cdot\text{m}^{0.5}$ ,  $n=12$  and  $\dot{i}_0=8\times 10^{-10} \text{ m/s}$  for soda-lime glass;  $K_{I0}=0.25 \text{ MPa}\cdot\text{m}^{0.5}$ ,  $n=16$  and  $\dot{i}_0=5\times 10^{-9} \text{ m/s}$  for borosilicate glass).

### 3.1 Thermally admissible kinetic relations

In the previous section we introduced two kinetic relations for crack growth. These relations appear phenomenological in nature and are not evidently connected with thermodynamic principles. Indeed, equation (8) defines the constraints that any kinetic laws must satisfy to comply with the second law of thermodynamics. Substituting the Charles law equation (15) into equation (8) gives

$$T\Lambda_c = \dot{i}_0 (G_I - 2\gamma) \left( \frac{G_I}{2\gamma} \right)^{n/2} \quad (20)$$

It is apparent that the Charles equation only satisfies equation (8) when  $G_I \geq 2\gamma$ . For  $0 \leq G_I < 2\gamma$ , the Charles law gives a negative entropy production rate which cannot be justified without resorting to microscale processes (i.e., crystal structures of the solid, chemical reactions at crack tip). By contrast, by substituting the Maugis law (19) into (8) it follows

$$T\Lambda_c = \dot{i}_0 (G_I - 2\gamma) \left\langle \frac{G_I - 2\gamma}{2\gamma} \right\rangle^{n/2} \quad (21)$$

which satisfies equation (8) for all values of  $G_I$ .

By using the Hyperplasticity formalism, it is possible to construct kinetic relations that automatically satisfy the thermodynamic constraints. This can be done by specifying a non-negative dissipation rate function that is homogeneous of degree higher than one or nonhomogeneous [Hou02]. For example, equation (12) becomes a non-homogeneous function by specifying

$$\Phi(\dot{i}) = 2\gamma\dot{i} + \overbrace{2\gamma\dot{i} \left( \frac{\dot{i}}{\dot{i}_0} \right)^{2/n}}^{T\Lambda_c} \geq 0 \quad (22)$$

Comparing equation (22) with the definition of dissipation rate  $\Phi(\dot{i}) \equiv G_I \dot{i}$ , the following relation immediately follows:

$$G_I = 2\gamma \left[ 1 + \left( \frac{\dot{i}}{\dot{i}_0} \right)^{2/n} \right] \geq 0 \quad (23)$$

The inversion of equation (23) leads to a fracture kinetics:

$$\dot{i} = \dot{i}_0 \left\langle \frac{G_I}{2\gamma} - 1 \right\rangle^{n/2} = \dot{i}_0 \xi_c^{n/2} \quad (24)$$

which is identical to the Maugis law (19). The dimensionless term

$$\xi_c = \left\langle \frac{G_I}{2\gamma} - 1 \right\rangle \quad (25)$$

can be interpreted as an overstress function similar to those widely used in Perzyna-type viscoplasticity models, since  $\xi_c = 0$  represents the rate-independent fracture function  $\gamma_c$  defined in equation (14) and positive values of  $\xi_c$  produce higher crack growth velocity  $\dot{i}$ . For this reason,  $\xi_c$  will be hereafter referred to as *fracture overstress function*.

Now let us revisit the physical meaning of the new term introduced in equation (22):

$$T\Lambda_c = 2\gamma\dot{i} \left( \frac{\dot{i}}{\dot{i}_0} \right)^{2/n} \quad (26)$$

This term describes the dependency of the entropy production rate on the crack velocity. It accounts for all the additional energy dissipation mechanisms during fracturing. If  $T\Lambda_c = 0$ , the loss of elastic strain energy in the solid is completely converted to

the surface energy of newly created surface area and equation (22) leads to Rice's fracture criterion. In reality, the loss of energy by surface area creation is always accompanied by additional dissipation mechanisms, which may include wave emission [Bou96], heat generation [Dol84], solid-environment reaction at crack surface [Mic82] or viscoelastic/plastic deformation of the solid in the vicinity of crack tip [Wei74]. Therefore, it is necessary to have  $T\Lambda_c > 0$  if any of the above-mentioned processes take place during fracturing. Deriving the particular expression of  $T\Lambda_c$  can, in principle, be achieved by examining the physical origins of each dissipation mechanism. For example, given a crack geometry, one can compute the energy consumption during diffusion of reactive species from the ambient environment to the newly generated surfaces. Energy dissipation upon spontaneous chemical reaction of fluid molecules with strained solid bonds depends on the rate at which such processes take place and on the involved reactants. Dissipation through viscoelastic/inelastic deformation near the crack tip can be evaluated based on the knowledge of the near-tip stress field as well as the material properties of the solid. In this chapter, these mechanisms are not differentiated, and are collectively accounted for via the expression (26). As a result, the Maugis law is recovered.

### 3.2 Dissipation function and potentials

To provide a context for the derivations which will be presented in the following sections, it is useful to introduce a more general way of obtaining equation (23) from (22) by distinguishing the definition of dissipation rate function and dissipation potentials. A comprehensive overview of this topic is given by [Hou14]. In general, the energy dissipation rate caused by the thermodynamic velocities  $\dot{\mathbf{a}}$  can be expressed as:

$$\Phi(\mathbf{A}, \dot{\mathbf{a}}) \equiv \boldsymbol{\chi} : \dot{\mathbf{a}} \geq 0 \quad (27)$$

where  $\boldsymbol{\chi} = \{\chi_1, \chi_2, \dots, \chi_n\}$  are thermodynamic forces conjugated to  $\dot{\mathbf{a}}$ ;  $\mathbf{A}$  is a list of external variables. In describing the behaviour of materials, the thermodynamic velocities and forces generally correspond to the rate of change of internal variables (e.g. plastic strain  $\dot{\boldsymbol{\epsilon}}^p$ , crack length  $\dot{l}$ ) and their driving forces (e.g. stress  $\boldsymbol{\sigma}$ , energy release rate  $G_I$ ), respectively. To define the constitutive relations between  $\boldsymbol{\chi}$  and  $\dot{\mathbf{a}}$ , the existence of a *force potential*  $z(\mathbf{A}, \dot{\mathbf{a}})$  and *flow potential*  $w(\mathbf{A}, \boldsymbol{\chi})$  can be postulated [Hou14]:

$$\boldsymbol{\chi} = \frac{\partial z}{\partial \dot{\mathbf{a}}}; \quad \dot{\mathbf{a}} = \frac{\partial w}{\partial \boldsymbol{\chi}} \quad (28)$$

where  $w$  and  $z$  are related by Legendre transformation:

$$w(\mathbf{A}, \boldsymbol{\chi}) = \boldsymbol{\chi} : \dot{\mathbf{a}} - z = \Phi - z \quad (29)$$

The  $\Phi$ ,  $z$  and  $w$  are not independent of each other. In particular, for the case of  $\Phi$  homogenous of degree one in  $\dot{\mathbf{a}}$  (i.e.  $\Phi = (\partial\Phi / \partial\dot{\mathbf{a}}) : \dot{\mathbf{a}}$ ), the thermodynamic forces  $\chi$  can be expressed as  $\chi = \partial\Phi / \partial\dot{\mathbf{a}}$ . Comparing it with equation (28) it is apparent that  $\Phi$  and  $z$  are identical. Another consequence of having both  $\Phi$  and  $z$  homogenous of degree one in  $\dot{\mathbf{a}}$  is that  $w$  collapses to zero (i.e.,  $w=0$ ), thus defining a locus that bounds the reversible region of the space of the external state variables (the so called yield locus) and outside which the inelastic response of the system is modelled as a rate-independent process [Ric71, Lub72]. This particular case is the basis of many rate-independent thermomechanical models for geomaterials [Col97, Col03, Ein07d]. In section 2.2, equation (14) is a consequence of a first-order homogeneous dissipation  $\Phi(\dot{l}) = 2\gamma\dot{l}$ .

For the more general case of  $\Phi$  homogeneous of degree  $N > 1$  or non-homogenous function of  $\dot{\mathbf{a}}$ , it can be shown that the dissipation process becomes rate-dependent. For example, the case of  $N=2$  corresponds to a linear viscoplastic behaviour [Per66], while  $N>2$  or more complex non-homogenous functions lead to nonlinear viscoplastic theories [Hou02]. In section 3.1, Maugis' kinetic law is a consequence of a non-homogeneous dissipation  $\Phi(\dot{l}) = 2\gamma\dot{l} + 2\gamma\dot{l}(\dot{l} / \dot{l}_0)^{2/n}$ .

In the latter case, the force potential  $z$  can be constructed from  $\Phi$  by using the following relation [Hou07, Hac08]:

$$z = \int_0^{\dot{l}} \frac{\Phi(\mathbf{A}, \tau\dot{\mathbf{a}})}{\tau} d\tau \quad (30)$$

where  $\tau$  is a dummy variable. Conversely, if the force potential  $z$  is known,  $\Phi$  can always be derived from the following relation:

$$\Phi(\mathbf{A}, \dot{\mathbf{a}}) = \frac{\partial z(\mathbf{A}, \dot{\mathbf{a}})}{\partial \dot{\mathbf{a}}} : \dot{\mathbf{a}} \quad (31)$$

In both cases, the flow potential  $w$  can be obtained from equation (29). More generally, since  $\Phi$ ,  $z$  and  $w$  can be converted to each other via Equations (29), (30) and (31), it is sufficient to specify one of them to define the dissipative constitutive response [Hou14].

We can immediately use equations (30) and (29) to derive the dissipation potentials for subcritical crack growth:

$$z = 2\gamma\dot{l} + \frac{2\gamma\dot{l}}{2/n+1} \left( \frac{\dot{l}}{\dot{l}_0} \right)^{2/n} ; w(G_I) = \Phi(\dot{l}) - z(\dot{l}) = \frac{2\gamma\dot{l}_0 \xi_c^{n/2+1}}{n/2+1} \quad (32)$$

The kinetic relations (23) and (24) can be readily derived via equation (28). Therefore, we have shown that the Maugis law can be derived by postulating either one of the potentials  $\Phi$ ,  $z$  and  $w$ . This conclusion is useful in deriving other kinetic models where not all of the dissipation potentials can be written explicitly like this case.

## 4 Energetics of continuum breakage

Let us now focus on the collective crushing of a granular assembly under mechanical loading. First, some basic concepts of continuum Breakage Mechanics will be introduced without referring to microscale fracturing processes. Then the similarities and connections of breakage mechanics with fracture mechanics will be reviewed. Finally, a simple upscaling strategy to bridge single particle fracture mechanisms to the macroscale breakage behaviour of granular assemblies will be discussed.

### 4.1 Definition of breakage

As summarized in section 1, granular assemblies that are monotonically loaded to sufficiently high stresses tend to approach an ultimate, self-similar GSD. This phenomenon can be explained from a statistical perspective: *“If particles fracture such that the smallest particles are in geometrically self-similar configurations under increasing macroscopic stress, with a constant probability of fracture, a fractal geometry evolves with the successive fracture of the smallest grains, in agreement with the available data.”* [McD98]. The fractal dimension of such GSD inferred from these concepts is about 2.5~2.7 (figure 5).

Based on these observations, [Ein07a] proposed a definition of breakage index based on the current location of GSD relative to the initial and ultimate fractal GSD, as shown in figure 6. Geometrically, the shaded area (the total breakage  $B_t$ ) quantifies how much grain crushing has occurred in the system, and the cross-hatched area (breakage potential  $B_p$ ) represents the total breakage the system can possibly undergo. Relative breakage  $B_r$  is defined as the ratio between the two, so that  $B_r=0$  represents the uncrushed specimen and  $B_r=1$  represents the completely crushed state. Mathematically, Einav's breakage index can be defined as:

$$B_r = \frac{\int_{d_m}^{d_M} (F(x) - F_0(x)) d(\log x)}{\int_{d_m}^{d_M} (F_u(x) - F_0(x)) d(\log x)} \quad (33)$$

where  $F$ ,  $F_0$  and  $F_u$  are the current, initial and ultimate cumulative GSD by mass;  $x$  is the grain diameter;  $d_m$  and  $d_M$  are the minimum and maximum grain sizes.

Given an initial and ultimate GSD and the relative breakage  $B_r$ , infinite examples of current GSDs can be found to satisfy equation (33) since the calculation of area does not restrict the shape of the GSD. This limits the direct use of  $B_r$  as an internal variable in continuum theories. A reasonable constraint to the shape of the current GSD is the assumption of fractional breakage [Ein07a], which hypothesizes that the measure of  $B_r$  is fractionally independent:

$$B = \frac{F(x) - F_0(x)}{F_u(x) - F_0(x)} = \frac{g(x) - g_0(x)}{g_u(x) - g_0(x)} \quad (34)$$

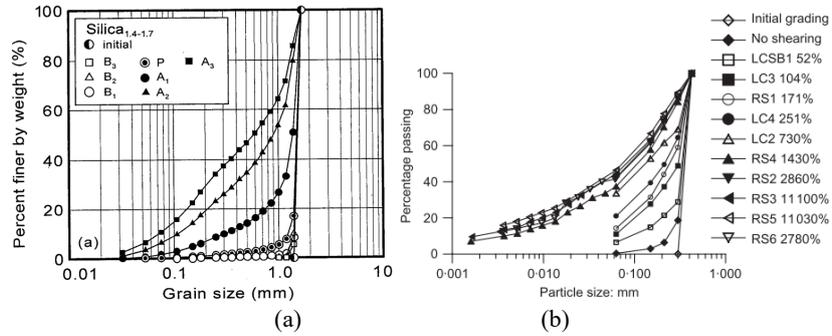


Figure 5: The development of an ultimate fractal GSD during (a) 1D compression test [Nak01b]; (b) Ring shear test [Coo04]. Please refer to the original papers for the correspondence in the legend.

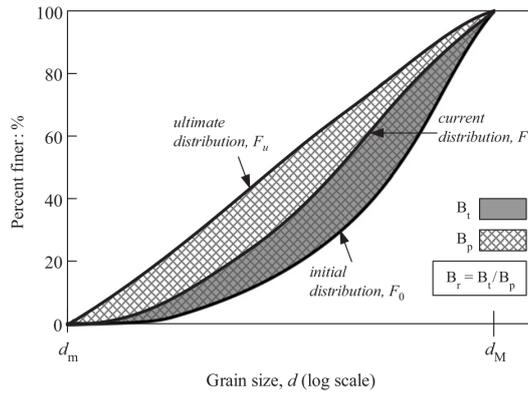


Figure 6: Einav's definition of breakage (after [Ein07a])

where  $g(x) = \partial F(x) / \partial x$  is the probability GSD;  $B$  is the fractional breakage (or simply breakage). Equation (34) is a direct outcome of (33) by assuming fractional independency. In this way, the breakage  $B$  is one-to-one linked with the current GSD of a granular system:

$$g(x, B) = (1 - B)g_0(x) + Bg_u(x) \tag{35}$$

The ultimate fractal GSD can be represented by the following function:

$$F_u(x) = \frac{x^{3-\alpha} - d_m^{3-\alpha}}{d_M^{3-\alpha} - d_m^{3-\alpha}} \quad \text{and} \quad g_u(x) = (3-\alpha) \frac{x^{2-\alpha}}{d_M^{3-\alpha} - d_m^{3-\alpha}} \quad (36)$$

with  $\alpha=2.5-2.7$ . Typically, the initial GSD can be represented by the same equation with a different fractal dimension:

$$F_0(x) = \frac{x^{3-\beta} - d_m^{3-\beta}}{d_M^{3-\beta} - d_m^{3-\beta}} \quad \text{and} \quad g_0(x) = (3-\beta) \frac{x^{2-\beta}}{d_M^{3-\beta} - d_m^{3-\beta}} \quad (37)$$

where  $\beta$  controls the coefficient of uniformity of the initial GSD. For soils with monodisperse grading, the following expression can be used:

$$F_0(x) = H(x - d_M) \quad \text{and} \quad g_0(x) = \delta(x - d_M) \quad (38)$$

where  $H$  is the Heaviside function and  $\delta$  is the Dirac delta function.

## 4.2 Breakage Mechanics

Let us start with energy balance equation (10). The system of interest is an assembly of brittle particles under mechanical loading. The work input is easily identified:

$$W = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} \quad (39)$$

where  $\boldsymbol{\sigma}$  is Cauchy stress and  $\boldsymbol{\varepsilon}$  is the infinitesimal strain tensor. To completely define the response of the system, one needs to specify the two scalar functions  $\Psi$  and  $\Phi$ . By observation, the external state variables of the system are  $\boldsymbol{\sigma}$  and  $\boldsymbol{\varepsilon}$ . The internal state variables (ISVs) include breakage  $B$ , plastic strain  $\boldsymbol{\varepsilon}^p$ , porosity  $n$  and fabric  $\mathbf{F}$ , etc. To focus on the breakage process, let us select  $B$  as the only ISV for now. The two functions are then expected to have the following forms:

$$\Psi = \Psi(\boldsymbol{\varepsilon}, B); \Phi = \Phi(\dot{B}) \quad (40)$$

[Ein07a] postulated that the elastic energy in the system characterized by GSD  $g(x, B)$  is distributed among different the grain size fractions  $\psi(\boldsymbol{\varepsilon}, x)$  which is a function of  $x$  through a power function:

$$\psi(\boldsymbol{\varepsilon}, x) = \psi_r(\boldsymbol{\varepsilon}) \left( \frac{x}{d_r} \right)^{n_M} \quad (41)$$

where  $\psi(\boldsymbol{\varepsilon}, x)$  is the grainsize specific Helmholtz free energy that satisfies

$$\Psi(\boldsymbol{\varepsilon}, B) = \int_{d_m}^{d_M} \psi(\boldsymbol{\varepsilon}, x) g(x, B) dx \quad (42)$$

;  $\psi_r$  and  $d_r$  are reference grainsize specific Helmholtz free energy and grain size respectively;  $n_M$  is a scaling parameter to be determined. [Ein07a] then conducted a

series of DEM analysis and found that  $n_M=2$  is a good approximation of the energy split in the system. Substituting equations (35) and (41) into (42) gives:

$$\Psi(\boldsymbol{\varepsilon}, B) = (1 - \mathcal{G}B)\psi_r(\boldsymbol{\varepsilon}) \quad (43)$$

where  $\mathcal{G}$  is the grading index emerged from the homogenization

$$\mathcal{G} = 1 - \frac{\int_{d_m}^{d_M} g_u(x) x^2 dx}{\int_{d_m}^{d_M} g_0(x) x^2 dx} \quad (44)$$

The thermodynamic forces defined by equation (43) can be derived as follows:

$$\boldsymbol{\sigma} = \frac{\partial \Psi(\boldsymbol{\varepsilon}, B)}{\partial \boldsymbol{\varepsilon}} = (1 - \mathcal{G}B) \frac{\partial \psi_r(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}} \quad (45)$$

$$E_B = -\frac{\partial \Psi(\boldsymbol{\varepsilon}, B)}{\partial B} = \mathcal{G}\psi_r(\boldsymbol{\varepsilon}) \quad (46)$$

Equation (45) represents a breakage-dependent elasticity model, i.e. the decrease of stiffness as gradation approach to ultimate state. Equation (46) defines a new variable named *breakage energy*. This quantity plays a similar role to that of  $G_I$  in fracture mechanics, as it quantifies the energy release rate upon an infinitesimal change of the gradation of the system  $\delta B$  (instead of  $\delta l$  for fracture). Their conceptual correspondence is illustrated in figure 7.

The breakage criterion can be defined similarly to fracture mechanics by imposing a threshold on the energy release rate  $E_B$  rather than on the stress  $\boldsymbol{\sigma}$ . Experimental evidence suggests that grain breakage progressively increases the energy required to cause further crushing [McD98], which is commonly referred to as *elastic hardening*. This leads to the postulation of the following breakage criterion:

$$E_B \leq \frac{E_c}{(1 - B)^2} \quad (47)$$

where  $E_c$  is the critical breakage energy. Compared to the Griffith criterion  $G_I \leq G_{IC}$ , the breakage criterion contains a hardening term  $1/(1 - B)^2$  so that the energy required to cause breakage approaches infinity as  $B$  approaches 1. The dissipation rate function during breakage can be derived as:

$$\Phi = E_B \dot{B} = \frac{E_c}{(1 - B)^2} \dot{B} \quad (48)$$

It is possible to examine the physical meaning of equation (48). For this purpose, let us define the *residual breakage energy*

$$E_B^* = \Psi - \Psi_u \quad (49)$$

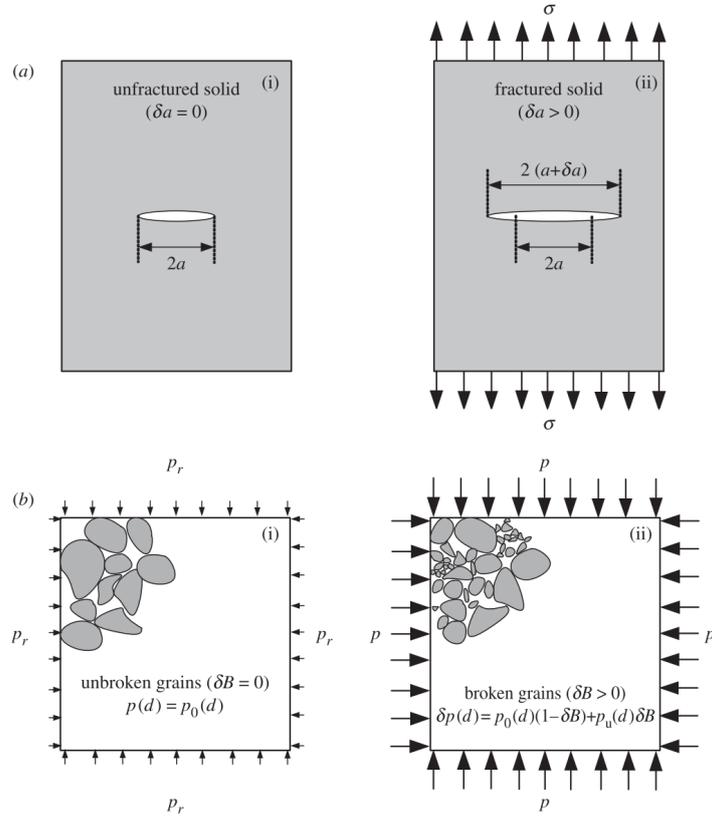


Figure 7: (a) Fracture propagation in a tensioned pre-cracked solid plate (i) before and (ii) after crack growth; (b) comminution in a confined granular assembly (i) before and (ii) after breakage growth. (after [Ein07c])

where  $\Psi_u(\boldsymbol{\varepsilon}) = \Psi(\boldsymbol{\varepsilon}, B = 1) = (1 - \mathcal{G})\psi_r(\boldsymbol{\varepsilon})$ . It represents the free energy that remains in the system for crushing grains. Substituting equation (46) into (49) gives

$$E_B^* = E_B(1 - B) \quad (50)$$

Substituting the breakage criterion (47), equation (50) finally reads:

$$\dot{E}_B^* = \frac{E_c}{(1 - B)^2} \dot{B} = \Phi \quad (51)$$

This equation reveals that the breakage dissipation is linked with the change of Helmholtz free energy due to moving the current GSD to a new position, or simply, energy dissipation from breakage is equal to the loss in residual breakage energy.

For illustration purposes, it is convenient to consider a one-dimensional (1D) scenario where the sample is subjected to isotropic compression only. Assuming the system obeys linear elasticity, the Helmholtz free energy and the corresponding elastic relations can be written as:

$$\begin{aligned}\Psi(\varepsilon_v, B) &= \frac{1}{2}(1 - \vartheta B)K\varepsilon_v^2 \\ p &= (1 - \vartheta B)K\varepsilon_v \\ E_B &= \frac{1}{2}\vartheta K\varepsilon_v^2\end{aligned}\quad (52)$$

where  $\varepsilon_v = \text{tr}(\boldsymbol{\varepsilon})$  is the volumetric strain;  $p = \text{tr}(\boldsymbol{\sigma})/3$  the mean stress;  $K$  the bulk modulus. Substituting equation (52) into (47) gives the critical breakage pressure (at  $B=0$ ):

$$p_{cr} = \sqrt{\frac{2KE_c}{\vartheta}}\quad (53)$$

Equation (53) bears numerous similarities with the well-known expression of the critical tensile stress for a brittle solid plate with a pre-existing crack:

$$\sigma_{cr} = \sqrt{\frac{EG_{Ic}}{\pi a}}\quad (54)$$

where  $E$  is the Young's modulus of the plate;  $a$  is the half-crack length. Continuum breakage mechanics, in a similar way to fracture mechanics, predicts the critical stress threshold for the growth of surface area in the system (via comminution), which depends on the material properties and the geometry of the system. For example, the effect of the stiffness is included in the bulk modulus  $K$  (thus suggesting that stiff particles and/or dense packing generate large yield pressures), while the role of the particle grading is embedded in  $\vartheta$  (which tends to zero as the initial GSD approaches its ultimate configuration, thus suggesting that well-graded soils tend to have higher yielding thresholds compared to poorly-graded soils). The picture is completed by the critical breakage energy  $E_c$  (i.e. a parameter that reflects the crushability of the material), thus playing a role conceptually similar to that of the critical energy release rate,  $G_{Ic}$ . This result, in contrast to the classical stress-based elastoplastic models, is tightly connected to the energetic principles in comminution theories and grain-scale fracture mechanics [Zha16]. It also provides tools for interpreting surface area measures during high-pressure tests [Zha18] which will be discussed in the last chapter of this paper.

A better description of the elastic response of granular materials can be achieved using the pressure-dependent elasticity:

$$\begin{aligned} \Psi(\varepsilon_v, B) &= (1 - \vartheta B) \frac{p_r}{\bar{K}(2-m)} A^{\frac{2-m}{1-m}} \\ p &= (1 - \vartheta B) p_r A^{\frac{1}{m-1}} \\ E_B &= \vartheta \frac{p_r}{K(2-m)} A^{\frac{2-m}{1-m}} \end{aligned} \tag{55}$$

where  $A = \bar{K}(1-m)\varepsilon_v + 1$ ;  $p_r$  is a reference pressure;  $m$  controls the nonlinearity of the model;  $\bar{K}$  is the dimensionless bulk modulus. The constitutive responses defined by Equation (55) and breakage criterion (47) are plotted in figure 8. This basic 1D model can already capture several main traits of crushing processes, i.e. elastic yielding and hardening as the GSD approaching the ultimate fractal distribution. Unloading does not produce reversed growth of breakage as constrained by the second law. Apparently, the model needs several extensions to capture the realistic behaviour of crushable granular materials: 1) plastic strain associated with grain breakage need to be included; 2) stress states other than isotropic need to be considered.

### 4.3 A rate-independent breakage model

A triaxial breakage model considering plastic-breakage coupling can be achieved by including the following features in the 1D framework: 1) include plastic strains to the list of ISVs; 2) include the deviatoric stress,  $q$ , in the formulation; 3) couple a frictional yielding mechanism to the breakage yielding mechanism. The step-by-step procedure in developing such model is given by [Ein07b]. Here we only present the key equations defining the final form of the model.

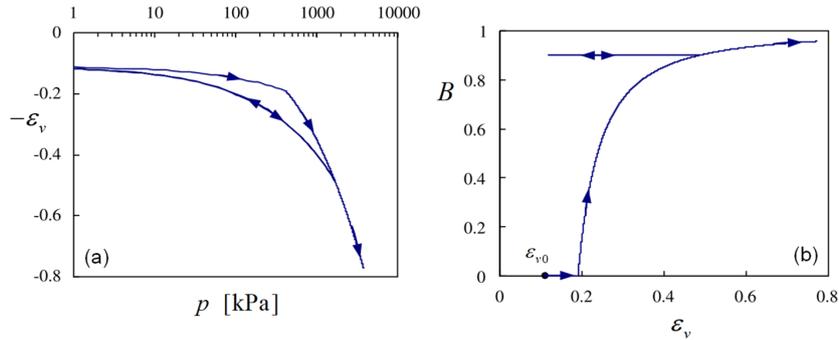


Figure 8: Compression response for a simple 1D breakage model (after [Ein07b]).

The two scalar energy functions are expressed as:

$$\Psi(\varepsilon_v^e, \varepsilon_s^e, B) = (1 - \vartheta B) \left[ \frac{1}{2} K \varepsilon_v^{e2} + \frac{3}{2} G \varepsilon_s^{e2} \right] \quad (56)$$

for linear elasticity or

$$\Psi(\varepsilon_v^e, \varepsilon_s^e, B) = (1 - \vartheta B) \left[ \frac{p_r}{\bar{K}(2-m)} A^{1-m} + \frac{3}{2} p_r \bar{G} A^{1-m} \varepsilon_s^{e2} \right] \quad (57)$$

for pressure-dependent elasticity, and

$$\Phi(\dot{B}, \dot{\varepsilon}_v^p, \dot{\varepsilon}_s^p) = \sqrt{\Phi_B(\dot{B})^2 + \Phi_p^v(\dot{\varepsilon}_v^p)^2 + \Phi_p^s(\dot{\varepsilon}_s^p)^2} \quad (58)$$

where

$$\begin{aligned} \Phi_B &= \frac{1}{(1-B)\cos\omega} \sqrt{E_B E_c} \dot{B} \\ \Phi_p^v &= \frac{p}{(1-B)\sin\omega} \sqrt{\frac{E_c}{E_B}} \dot{\varepsilon}_v^p \\ \Phi_p^s &= Mp \dot{\varepsilon}_s^p \end{aligned} \quad (59)$$

$\omega$  is the coupling angle that allocates the energy dissipation through breakage and friction;  $M$  is the critical stress ratio that controls the frictional limit during shear;  $\varepsilon_s = \sqrt{\frac{2}{3}} \mathbf{e} : \mathbf{e}$  the deviatoric strain;  $\mathbf{e} = \boldsymbol{\varepsilon} - \varepsilon_v / 3 \mathbf{I}$  the strain deviator;  $q = \sqrt{\frac{3}{2}} \mathbf{s} : \mathbf{s}$  the deviatoric stress;  $\mathbf{s} = \boldsymbol{\sigma} - p \mathbf{I}$  the stress deviator; superscripts 'e' and 'p' denote the elastic and plastic components, respectively.

Equation (58) is a homogeneous function of degree 1 with respect to its variables  $(\dot{B}, \dot{\varepsilon}_v^p, \dot{\varepsilon}_s^p)$ . Therefore, the resultant model is expected to be rate-independent and  $\Phi = z$  is the true dissipation potential of the system. Following the hyperplasticity formalism, the yield surface in true stress space can be derived as:

$$y_{pB} = \frac{E_B}{E_c} (1-B)^2 + \left( \frac{q}{Mp} \right)^2 - 1 = 0 \quad (60)$$

The flow rules are given by

$$\begin{aligned} \dot{B} &= \lambda \frac{\partial \bar{y}_{pB}}{\partial \bar{E}_B} = 2\lambda \frac{(1-B)^2 \cos^2 \omega}{E_c} \\ \dot{\varepsilon}_v^p &= \lambda \frac{\partial \bar{y}_{pB}}{\partial \bar{p}} = 2\lambda \frac{(1-B)^2 E_B \sin^2 \omega}{p E_c} \\ \dot{\varepsilon}_s^p &= \lambda \frac{\partial \bar{y}_{pB}}{\partial \bar{q}} = 2\lambda \frac{q}{M^2 p^2} \end{aligned} \quad (61)$$

where  $\lambda$  is a positive plastic multiplier;  $\bar{p}, \bar{q}, \bar{E}_B$  are dissipative stresses;  $\bar{y}_{pB}$  is the yield surface in the dissipative stress space; The performance of the triaxial breakage model has been studied by several authors [Ein07b, Ngu09, Zha13] and examples of application are shown in figure 9.

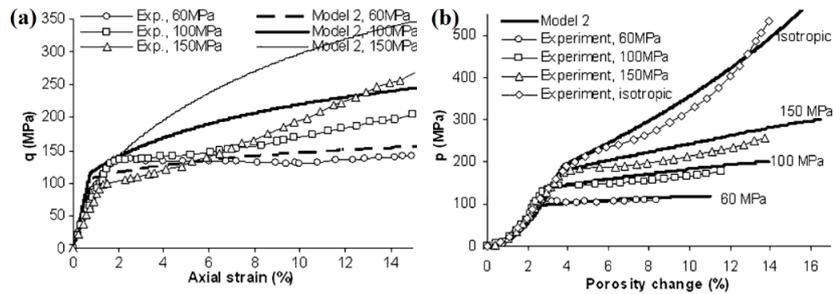


Figure 9: Model vs data for the (a) stress-strain behaviour and (b) porosity change for Adamswiller sandstone in different loading conditions. The prediction is given by pressure-dependent elasticity. (After [Ngu09])

## 5 Kinetics of breakage growth

Section 3 discussed the connection between Rice's interpretation of the Griffith's criterion and the Maugis crack growth law. The latter is shown to be a consequence of a higher-order term added to the dissipation of a rate-dependent fracture process, which augments the loss of energy via creation of new surface area by accounting for further dissipative processes. This result has inspired a strategy to convert a rate-independent initiation criterion (e.g., a yield surface) into a rate-dependent evolution law controlled by delayed fracturing (i.e., a kinetic relation), and it will be applied hereafter to the continuum modelling of delayed particle breakage.

### 5.1 Simple isotropic model for rate-dependent breakage

First, it is useful to examine the physical processes that govern the rate-dependent crushing of granular assemblies. Recent experimental data supports [Kar10, Lad10] that creep and relaxation of granular materials under high confining pressure is caused by the growth of subcritically stressed cracks in individual grains. Indeed, the time required to split a particle subjected to an external load  $P$  can be estimated by integrating the corresponding crack growth velocity with respect to time [Old07] (figure 10a). Depending on the magnitude of  $K_I$  generated by the load  $P$ , the particle can crush catastrophically (i.e.  $K_I \geq K_{IC}$ ), fracture after a finite time (i.e.  $K_{I0} < K_I < K_{IC}$ ) or remain

intact indefinitely (i.e.  $K_I \leq K_{I0}$ ). Now let us consider a sand specimen subjected to sustained confining stress (i.e., a creep test). Critically loaded particles arranged along the major force chains will be rapidly crushed. These events are reflected macroscopically as instantaneous inelastic strains developed upon the application of a load increment. The rest of the particles are either stressed with intermediate values of  $K_I$  or slightly stressed with  $K_I \leq K_{I0}$ . As time progresses, particles under sufficiently high stresses will be consecutively crushed, thus initiating grain rearrangement and force redistribution. Eventually, a stable configuration is achieved where most of the flaws within the particles are stressed below the corrosion limit and only limited strain accumulation is displayed over time (figure 10b)

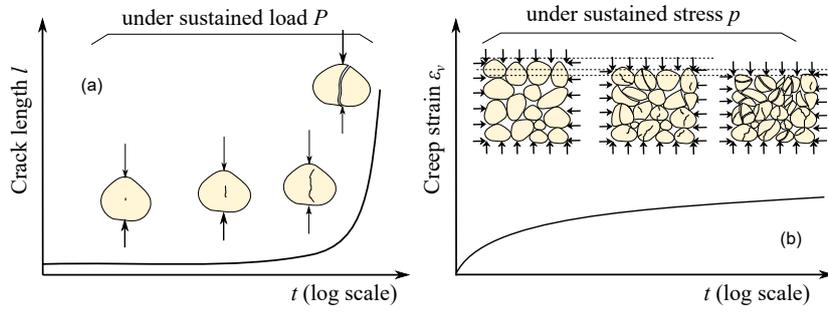


Figure 10: Schematic diagrams of a) fracture growth in a subcritically loaded particle and b) creep in a granular assembly

Based on these considerations, the energy dissipation by grain crushing should not only contain the dissipation associated with the creation of new surface area and the associated energy redistribution, but it should also include additional sources of energy loss that reflect the net dissipation due to delayed inelastic interactions at crack tips. This can be incorporated in the dissipation function (48) in a similar way as for the fracture dissipation function (22) by adding a non-negative term  $TA_B$ :

$$\Phi(\dot{B}) = \frac{E_c \dot{B}}{(1-B)^2} + \overbrace{\frac{E_c \dot{B}}{(1-B)^2} \left( \frac{\dot{B}}{\dot{B}_0} \right)^{2/n}}^{TA_B} \geq 0 \quad (62)$$

The first term on the right hand side (RHS) of (62) is the breakage dissipation accounting for the creation of new surface area and the associated loss of energy due to force redistribution in the original 1D breakage theory. The second term  $TA_B \geq 0$  collects additional dissipative terms due to stress corrosion at crack tips and delayed breakage. Apparently, this term makes equation (62) a non-homogeneous function, thus implying a rate-dependent irreversible process. Following the same procedure outlined in Section 3.2, the force and flow potentials can be derived as:

$$z(\dot{B}) = \frac{E_c \dot{B}}{(1-B)^2} \left( 1 + \frac{1}{2/n+1} \left( \frac{\dot{B}}{\dot{B}_0} \right)^{2/n} \right); w(E_B) = \frac{E_c \dot{B}_0}{(1-B)^2} \frac{\xi_B^{n/2+1}}{n/2+1} \quad (63)$$

where the dimensionless term

$$\xi_B(B, E_B) = \left\langle \frac{(1-B)^2 E_B}{E_c} - 1 \right\rangle \quad (64)$$

is the breakage overstress function. The breakage evolution law corresponding to equation (63) is derived as:

$$E_B = \frac{\partial z}{\partial \dot{B}} = \frac{E_c}{(1-B)^2} \left[ 1 + \left( \frac{\dot{B}}{\dot{B}_0} \right)^{2/n} \right] \text{ or } \dot{B} = \frac{\partial w}{\partial E_B} = \dot{B}_0 \xi_B^{n/2} \quad (65)$$

Combining equations (65) and (24), one obtains

$$\frac{\log(\dot{i}/\dot{i}_0)}{\log \xi_C} = \frac{\log(\dot{B}/\dot{B}_0)}{\log \xi_B} = \frac{n}{2} \quad (66)$$

This equation reveals the implications of having the same order of nonlinearity for the entropy production term  $TA_B$  for fracturing and breakage. In particular, according to (66) the corrosion index  $n$  not only governs the kinetics of crack propagation within the grains, but it also enters in the continuum description of the rate-dependent crushing of an assembly. This hypothesis results in the same logarithmic growth of the internal variables  $\dot{i}$  and  $\dot{B}$  when plotted against their respective overstress functions  $\xi_C$  and  $\xi_B$  (figure 11). The applicability of such hypothesis will be discussed in the subsequent sections with reference to laboratory evidences. The performance of this 1D model using linear elastic model is presented in figure 12.

## 5.2 A rate-dependent triaxial breakage model

A generalization of this 1D model to triaxial conditions can be pursued in a similar manner as in Section 4.3. The only difference is that the dissipation rate function in this case involves three thermodynamic velocities  $(\dot{B}, \dot{\epsilon}_v^p, \dot{\epsilon}_s^p)$  while the 1D case has only one  $(\dot{B})$ . The coupling of multiple dissipation mechanisms in equation (58) prevents the use of a generalization strategy based on the simple addition of a higher-order source of energy dissipation  $TA_B$ . To circumvent such difficulty, the extension to multi-dissipative mechanisms is tackled by directly postulating a flow potential  $w$  rather than one of the functions  $\Phi$  or  $z$ . For this purpose, by analogy with equation (63), the flow potential is postulated as

$$w(\bar{E}_B, \bar{p}, \bar{q}) = \frac{E_c \dot{B}_0}{2(1-B)^2} \frac{\bar{\xi}_{pB}^{n/2+1}}{n/2+1} \quad (67)$$

where  $\bar{\xi}_{pB}(\bar{E}_B, \bar{p}, \bar{q}) = \langle \bar{y}_{pB} \rangle$  is the breakage plastic overstress function for triaxial loading. Applying the definition of flow potential gives the evolution laws:

$$\begin{aligned} \dot{B} &= \frac{\partial w}{\partial \bar{E}_B} = \dot{B}_0 \cos^2 \omega \xi_{pB}^{n/2} \\ \dot{\varepsilon}_v^p &= \frac{\partial w}{\partial \bar{p}} = \dot{B}_0 \frac{E_B}{p} \sin^2 \omega \xi_{pB}^{n/2} \\ \dot{\varepsilon}_s^p &= \frac{\partial w}{\partial \bar{q}} = \frac{\dot{B}_0 E_c}{(1-B)^2} \frac{q}{M^2 p^2} \xi_{pB}^{n/2} \end{aligned} \quad (68)$$

where  $\xi_{pB}(E_B, p, q) = \langle y_{pB} \rangle$ . The relations above converge to the same breakage growth law in equation (65) whenever isotropic stress conditions are imposed ( $q=0$ ) and volumetric plastic dissipation is neglected ( $\omega=0$ ). Furthermore, by setting  $n=2$ , Equations (42) lead to an overstress-based viscous model with a linear viscous nucleus function, which has already been used in previous studies of crushing-induced compaction localization for regularization purposes [Das13]. It is shown here that such regularization can be obtained from a thermomechanical procedure based on the flow potential. Moreover, from this analysis it follows that the power law coefficient  $n$  should not be arbitrary selected, but it should rather coincide with the stress corrosion index of the constituting mineral (which is much larger than 2).

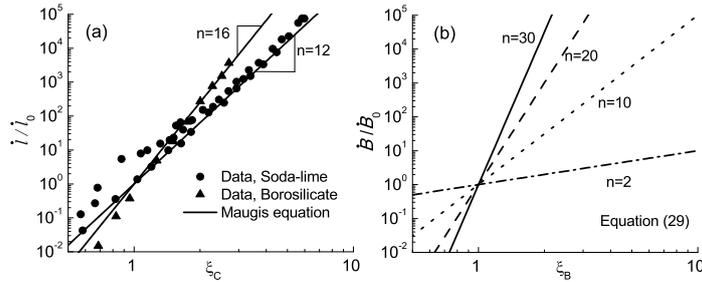


Figure 11: a) Plot of Maugis model and subcritical crack growth data on the  $\dot{l}/l_0 - \xi_c$  plane; b) visualization of equation (66) for different values of  $n$ .

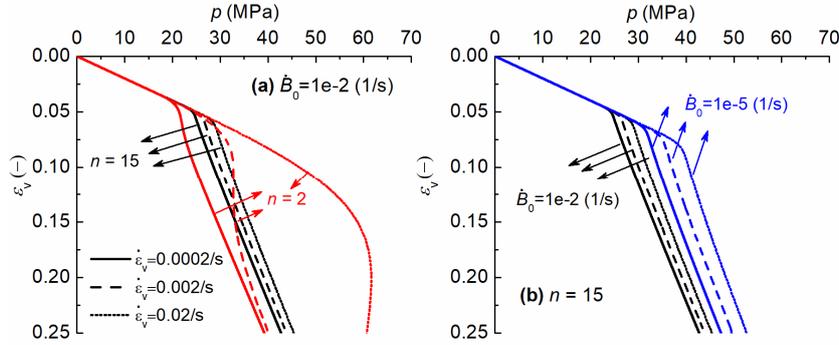


Figure 12: Examples of isotropic breakage simulations exploring the effect of the parameters  $n$  and  $\dot{B}_0$  under constant strain rate compression.

### 5.3 Model validation

Creep data from oedometric compression tests on a quartz sand reported by [Tak01] is selected for validation purposes (figure 13). The grading index has been computed by using the reported grain size ( $D_{50}=1.43$  mm) and assuming an initial uniform grading. Each load increment has been assumed to be applied in 1s, after which a creep stage was simulated. In addition, the parameters of the baseline breakage model ( $\bar{K}$ ,  $\bar{G}$ ,  $E_c$ ,  $M$  and  $\omega$ ) have been calibrated through a strategy similar to that used for the previous example, while the parameters  $n$  and  $\dot{B}_0$  have been selected to match the magnitude and rate of the measured creep response at a stress level of 20 MPa (figure 13b). It is noticed that the computed compression curves and the predicted strain histories at the other stress levels ( $\sigma_v=1.11$  MPa  $\sim$  9.95 MPa) exhibit an excellent agreement with the experiments. Such agreement has been achieved by using a power law coefficient  $n = 30$ , which is within the range of values reported for typical rock-forming minerals [Atk82]. In addition, the negligible creep measured at the lowest stress levels (i.e.  $\sigma_v=1.11$  MPa and 2.34 MPa) is correctly captured by the model, thus reflecting that under high compressive stresses and constrained kinematic conditions grain crushing plays a dominant role in the accumulation of creep. Although other time-dependent processes may influence creep rates in granular soils (e.g., at low to intermediate stress levels, where grain damage may be negligible and/or coexist with other inelastic mechanisms), the model performance in this specific example validates the hypothesized link between delayed crack growth and macroscopic creep for cases in which grain crushing is the main source of microstructural rearrangement.

Karimpour and Lade [Kar10, Kar13, Lad14] recently performed a series of triaxial compression tests on saturated Virginia Beach sand under high confining pressures (8 MPa) to study the effect of the loading rate on its stress-strain, creep and relaxation responses. Such data are here used to validate the characteristics of the proposed

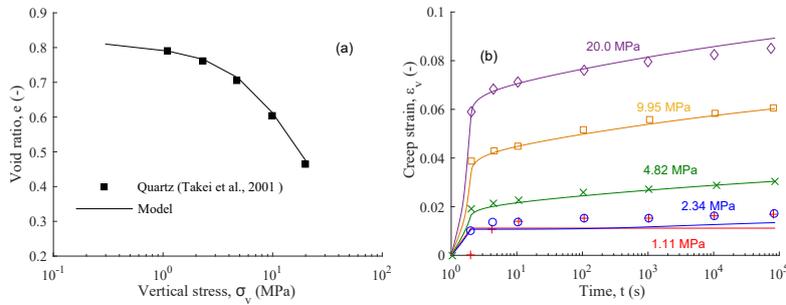


Figure 13: Data and model predictions of oedometric creep tests on quartz sand: a) compression response, b) creep response (data from Takei et al., 2001). Parameters used:  $\vartheta=0.75$ ,  $E_c=0.09$  MPa and  $\bar{K}=1500$ ,  $\bar{G}=1640$ ,  $\omega=15^\circ$ ,  $M=1.2$ ,  $n=30$  and  $\dot{B}_0=2 \times 10^{-2}$ /s.

model along stress paths that involve considerable shear strains. The strategy is to first calibrate all the model parameters based on stress-strain curves reported for various strain rates, then, using the same set of parameters, to predict stress relaxation responses to be compared with measured relaxation curves. Figure 14 presents measured and simulated stress-strain curves for constant strain rate triaxial tests, for which a calibration strategy similar to that discussed in the previous cases has been used to constrain the model parameters. Stress relaxation tests are then simulated by interrupting each test at 2.8% strain level and then maintaining constant axial strain  $\epsilon_a$  for one day, thus following exactly the procedure described by [Lad14]. The data in figure 15a suggests that upon ceasing of axial straining, a drop of  $q$  manifests on the  $q$ - $\epsilon_a$  plane, while negligible volumetric changes occur during the relaxation stage. These characteristics have been successfully predicted by the model, though there is a slight over prediction of the volumetric compaction prior to the relaxation stage. Figure 15b presents the time history of  $q$  relaxation, as well as the absolute values of decreasing stress deviator. An excellent agreement between model predictions and experimental data is observed on both planes. More interestingly, the convergence of the three  $q$ - $t$  curves after certain amount of relaxation time is successfully predicted by the proposed model.

For Virginia Beach sand a value of  $n=30$  has been used, which again falls within the range of corrosion indices summarized in figure 3b. This evidence suggests that high values of  $n$  are necessary to properly capture strain-rate sensitivity, creep and relaxation in granular soils at high pressures. Other values of  $n$  (e.g. classical Perzyna model with  $n=2$ ) would instead not be able to match neither of these effects due to the high sensitivity of the response to small changes in strain rates. Such observations indirectly validate the hypothesis in equation (66), i.e. that the power law coefficient governing the breakage evolution is linked with the corrosion index controlling the growth of cracks in particles.

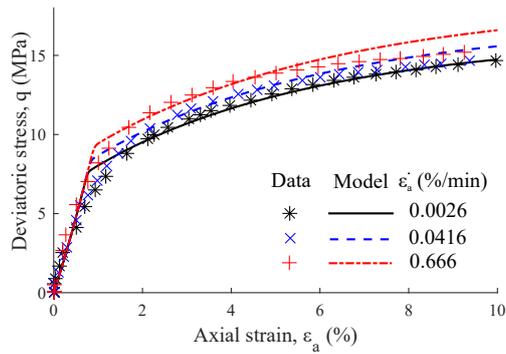


Figure 14: Experimental data and model simulations of the stress strain response of Virginia Beach sand subjected to various strain rates at a confining pressure of 8 MPa. The grading index  $\beta=0.72$  results from the reported  $D_{50}=0.6$  mm. The calibrated parameters are:  $E_c=0.12$  MPa and  $\bar{K}=4000$ ,  $\bar{G}=4360$ ,  $\omega=45^\circ$ ,  $M=1.02$ ,  $n=30$  and  $\dot{B}_0=2 \times 10^{-2}/s$ .

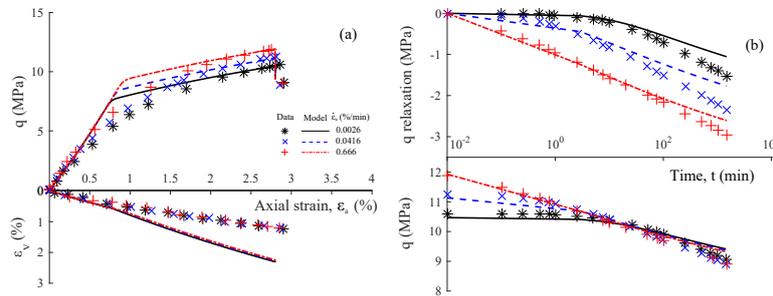


Figure 15: Experimental data and model simulations of triaxial relaxation tests on Virginia Beach sand: a)  $q$  and  $\varepsilon_v$  responses, b) relative and absolute drop of  $q$  due to relaxation.

## 6 Environmental effects

It is known that the crushing behaviour of granular materials is sensitive to the state of the fluids occupying the pore space. The previous discussion has focused on the mechanical causes of particle breakage, thus placing less emphasis on the environmental conditions at which such processes take place. Interactions between solids and

environment can indeed alter the crushing response of brittle granular materials at the scale of representative elementary volumes (REV). A common example is the effect of varying ambient relative humidity (RH), for which numerous tests on specimens made by silica and quartz particles [Che07, Brz14] or rock aggregates [Sow65] have shown that the exposure to water vapour can dramatically exacerbate crushing by lowering the macroscopic yield strength, enhancing the matrix compressibility and promoting long-term creep. For these reasons, dry sand specimens in equilibrium with high confining pressures often experience a reactivation of breakage and volumetric collapse upon water injection [Ova13]. Similarly, cemented granular rocks (e.g., sandstones) display marked water-sensitivity, in the sense that in the cataclastic regime fully saturated samples display much lower yield stress compared to nominally dry specimens [Zhu97, Bau00]. This section shows how the above phenomenon can be consistently addressed in the same thermomechanical framework by exploiting the breakage-fracture analogy at macro and microscales.

### 6.1 Environmentally enhanced fracture growth

Consider again a diametrically loaded disk with unitary thickness in the out-of-plane direction depicted in figure 16. The system is in contact with a heat reservoir with constant temperature  $T$ . The particle is immersed in a fluid mixture composed of  $k$  miscible surface-reactive species. The state of the system can be altered by controlling the compression force on the particle, as well as by changing the pressure of the fluid or modifying the number of moles of each species. The rate of work input introduced in the system through these three mechanisms can be written as:

$$W = F\dot{\Delta} + P\dot{V}_f + \sum_i^k \mu_i \dot{N}_i \quad (69)$$

where  $F$  and  $\Delta$  are the axial load and displacement, respectively;  $P$  the fluid pressure;  $V_f$  the volume of the fluid;  $\mu_i$  and  $N_i$  the chemical potential and the number of moles of species  $i$  in the system. Assuming no mass exchange between the solid phase and the other phases, the added molecules of species  $i$  can only exist in fluid form or be adsorbed on the solid surface, i.e.  $N_i$  can be decomposed into

$$N_i = N_i^f + N_i^{surf} = N_i^f + 2\Gamma_i l \quad (70)$$

where the superscript  $f$  and  $surf$  denotes fluid and surface portion, respectively;  $l$  is the crack length;  $\Gamma_i$  is the surface excess concentration representing the adsorbed number of moles of species  $i$  per unit area of the fracture. Substituting equation (69) in the energy balance (10) and assuming the same fracture dissipation (12), one obtains:

$$F\dot{\Delta} + P\dot{V}_f + \sum_i^k \mu_i \dot{N}_i - \dot{\Psi} = T\Lambda_c \geq 0 \quad (71)$$

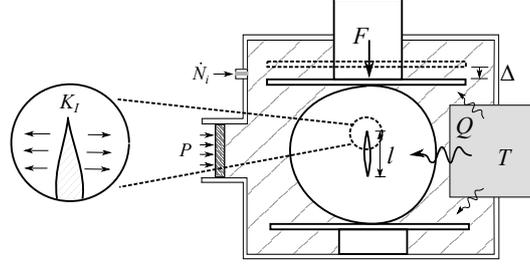


Figure 16: Schematic of a center-cracked particle immersed in adsorptive fluid mixture.

Substances in solid and fluid states, as well as those present in the form of surface excess, can carry both entropy and energy. Therefore, the total Helmholtz free energy can be decomposed into three contributions:

$$\Psi = \Psi_s + \Psi_f + 2\psi_{surf}l \quad (72)$$

where  $\Psi_s = \Psi_s(\Delta, l)$  is the elastic energy stored in the solid phase;  $\Psi_f = \Psi_f(V_f, N_i^f)$  is the fluid free energy;  $\psi_{surf} = \psi_{surf}(\Gamma_i)$  is the surface excess free energy per unit area. Since in this study the solid phase is not involved in mass exchanges, the composition of the solid is constant and the rate of change of its Helmholtz free energy is independent of the chemical potentials of the solid species. A similar argument can be used for the temperature,  $T$ , in that the restriction to isothermal processes allows the exclusion of  $T$  from the list of arguments of  $\Psi_s$ .

By substituting equations (70), (72) into (71), the following thermodynamic constraint can be derived:

$$\Phi_s + \Phi_f + 2l\dot{\varphi}_{surf} = T\Lambda_c \geq 0 \quad (73)$$

where

$$\Phi_s = \left( F - \frac{\partial \Psi_s}{\partial \Delta} \right) \dot{\Delta} + \left[ G_l - 2 \left( \psi_{surf} - \sum_i^k \mu_i \Gamma_i \right) \right] \dot{l} \quad (74)$$

is the rate of solid dissipation;

$$\Phi_f = P\dot{V}_f + \sum_i^k \mu_i \dot{N}_i^f - \dot{\Psi}_f \quad (75)$$

is the rate of fluid dissipation and

$$\varphi_{surf} = \sum_i^k \mu_i \dot{\Gamma}_i - \dot{\psi}_{surf} \quad (76)$$

is the rate of surface dissipation per unit area. Let us restrict the discussion to non-dissipative constitutive response of the fluid mixture and the surface (i.e.  $\Phi_f = \varphi_{surf} = 0$ ). In this case, equation (75) is the standard free energy balance for homogenous multicomponent fluid [Deh06] and equation (76) embodies a relation between the chemical potential and the number of adsorbed molecules of species  $i$ :

$$\mu_i = \frac{\partial \psi_{surf}}{\partial \Gamma_i} \quad (77)$$

To further interpret the energy balance for solid-fluid interface (76), it is convenient to define an energy potential from the Legendre transformation of  $\psi_{surf}$  as

$$\gamma(\mu_i) = \psi_{surf}(\Gamma_i) - \sum_{i=1}^k \Gamma_i \mu_i \quad (78)$$

The rate of change of this potential can be expressed as follows by substituting equation (78) into (76) noticing  $\varphi_{surf} = 0$ :

$$\dot{\gamma} = - \sum_{i=1}^k \Gamma_i \dot{\mu}_i \quad (79)$$

Equation (79) is the celebrated Gibbs adsorption isotherm for multicomponent adsorptive fluids. Indeed, the energy potential  $\gamma$  defined in (78) is the solid-fluid interfacial energy (or, more simply, the surface energy) that represents the energy consumption due to the creation of a unit area of such interface (assuming no other dissipation sources). The Gibbs adsorption isotherm states that the solid-fluid interfacial energy is a function of the chemical potential of each adsorptive species in the environment. For inert species that do not interact with the surface (i.e.  $\Gamma_i = 0$ ), their concentrations in the fluid phase do not alter the interfacial energy. Otherwise, the amount of adsorbed molecules  $\Gamma_i > 0$  is typically a monotonically increasing function of  $\mu_i$  [Sin85], which according to equation (79) implies a decrease of surface energy upon adsorption.

Substituting equation (78) and  $\Phi_f = \varphi_{surf} = 0$  into equation (73), one can derive again equation (13) that embodies the Rice's criterion (8). This concept is important to explain the results of experiments subjecting granular solids to changes in relative-humidity (e.g., [Old03]). In these tests, the controlled external variables are total stresses and strains, air relative humidity, and the total air pressure (typically kept at atmospheric value). As a result, the chemical potential of the water vapour  $\mu_w$  is the only varying state variable for the fluid phase, thus allowing the integration of equation (79) under constant  $\mu_i$  ( $i \neq w$ ):

$$\gamma(\mu_w) = \gamma_a - \int_{-\infty}^{\mu_w} \Gamma_w(\hat{\mu}_w) d\hat{\mu}_w \quad (80)$$

where  $\gamma_a$  is the solid surface energy in dry air (which can be treated as a constant) and the subscript  $w$  denotes water vapour. The chemical potential  $\mu_w$  can be further expressed in terms of partial vapour pressure  $P_w$  and thus the relative humidity [Deh06]:

$$\gamma = \gamma_a - RT \int_0^{P_w} \Gamma_w(\hat{P}_w) d \ln \hat{P}_w = \gamma_a - RT \int_0^{RH} \Gamma_w(\hat{RH}) d \ln \hat{RH} \quad (81)$$

where  $RH = P_w / P_{w,sat}$  is relative humidity; where  $P_{w,sat}$  is the saturation vapour partial pressure.

To further explore the implications of the Gibbs isotherm, the Langmuir isotherm is selected to relate  $\Gamma_w$  and  $\mu_w^*$  for the adsorption of ideal gases on solid surfaces [Cou10]:

$$\Gamma_w = \frac{\Gamma_w^{\max}}{1 + f_w / RH} = \frac{\Gamma_w^{\max}}{1 + f_w \exp(-\mu_w^* / RT)} \quad (82)$$

where  $f_w$  and  $\Gamma_w^{\max}$  are two model parameters that should be calibrated from adsorption tests. Here we consider a set of water-quartz and water-silica adsorption data (figure 17a) to locate a range of adsorption parameters  $f_w$  and  $\Gamma_w^{\max}$  for the surface of sand or rock particles. The curves A, B and C are generated by equation (82) using different values of  $f_w$  and  $\Gamma_w^{\max}$  (specified in the figure caption). Combining the equation (82) with the Gibbs isotherm (81), the surface energy can be expressed explicitly as a function of RH:

$$\gamma = \gamma_a - RT \Gamma_w^{\max} \ln(RH / f_w + 1) \quad (83)$$

The effect of varying surface energy on the fracture toughness of a brittle solid can be readily derived through the Irwin's relation (i.e.,  $2\gamma = G \sim K_{I0}^2$ ):

$$\frac{K_{I0}}{K_{I0,a}} = \sqrt{\frac{\gamma}{\gamma_a}} = \sqrt{1 - \frac{RT}{\gamma_a} \Gamma_w^{\max} \ln\left(\frac{RH}{f_w} + 1\right)} \quad (84)$$

where  $K_{I0,a}$  is the fracture toughness in dry air. The validity of equation (84) is examined through the experimental data reported by [Nar12] about the  $K_{I0}$  for different rocks subjected to changing relative humidity in figure 17b.

Combining equation (83) with the Maugis kinetics law (19) allows for an environment-dependent propagation of cracks. The performance of this model is compared

again with Nara’s data in figure 18. A good agreement is illustrated between model predictions and data for two sandstones at different values of RH.

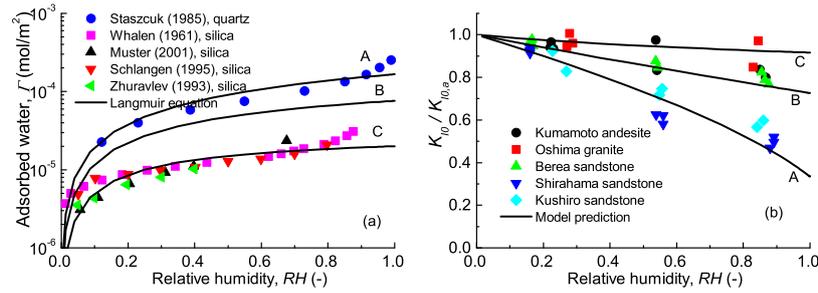


Figure 17: a) Calibration of the parameters for Langmuir adsorption model and b) predictions of fracture toughness subjected to various RH (data from [Nar12]). The adsorption parameters for curve A are  $f_w=5$  and  $\Gamma_w^{\max}=1\times 10^{-3}$  mol/m<sup>2</sup>, for curve B are  $f_w=1.5$  and  $\Gamma_w^{\max}=1.9\times 10^{-4}$  mol/m<sup>2</sup>, and for curve C are  $f_w=0.5$  and  $\Gamma_w^{\max}=3\times 10^{-5}$  mol/m<sup>2</sup>

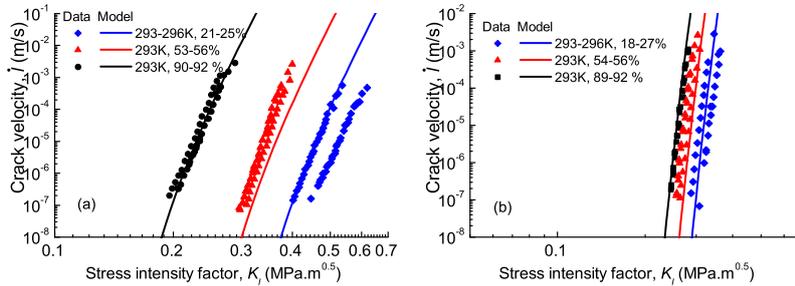


Figure 18: Predicted vs. observed subcritical crack growth curves under different values of relative humidity: a) Shirahama Sandstone ( $n=20$ ,  $\dot{l} = 10^{-8}$  m/s and  $K_{I0,a}=0.3$  MPa.m<sup>0.5</sup>); b) Berea Sandstone ( $n=40$ ,  $\dot{l} = 10^{-8}$  m/s and  $K_{I0,a}=0.22$  MPa.m<sup>0.5</sup>).

### 6.2 Breakage of a granular assembly in wet air

Now consider a three-dimensional (3D) REV of brittle particles in contact with wet air (figure 19). By assuming that all processes are sufficiently slow to ensure that heat transfer and water vapour flow can be neglected, the Clausius-Duhem inequality for the entire system can be expressed as follows [Cou10]

$$\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} + \sum_i^k \mu_i \dot{n}_i - \dot{\Psi} = \Phi \geq 0 \quad (85)$$

where  $n_i = N_i / V$  is the number of moles of species  $i$  per unit total volume of the REV, which as in equation (70) can be decomposed into two components:

$$n_i = n_i^a + A_s \Gamma_i \quad (86)$$

where  $n_i^a$  is the number of moles of free molecules in the air;  $A_s$  is the specific surface area (here defined as the solid surface area per unit total volume of the REV), and  $\Gamma_i$  is the number of moles of adsorbed molecules per unit solid surface area.

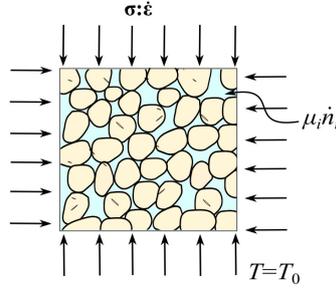


Figure 19: Schematic of a REV of granular skeleton immersed in ideal gas mixture.

Similar to equation (72), the total Helmholtz free energy consists of three components:

$$\Psi = \Psi_s + \Psi_a + A_s \psi_{surf} \quad (87)$$

where  $\Psi_s$ ,  $\Psi_a$  and  $A_s \psi_{surf}$  are the free energy stored in the solid skeleton, air, and solid-air interface respectively. The thermodynamic constraints for such system can be derived by substituting equations (86) and (87) into (85):

$$\Phi_s + \Phi_a + A_s \varphi_{surf} \geq 0 \quad (88)$$

where

$$\Phi_s = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} + P \dot{\phi} + \left( \sum_i^k \mu_i \Gamma_i - \psi_{surf} \right) \dot{A}_s - \dot{\Psi}_s \quad (89)$$

$$\Phi_a = -P \dot{\phi} + \sum_i^k \mu_i \dot{n}_i^a - \dot{\Psi}_a \quad (90)$$

$$\varphi_{surf} = \sum_i^k \mu_i \dot{\Gamma}_i - \dot{\psi}_{surf} \quad (91)$$

are energy dissipation terms due to changes of the state of solid skeleton, gas mixture and solid-gas interface, respectively. Equations (90) and (91) resemble the relations discussed in Section 6.1, in that they reflect the free energy balance for ideal gas mixtures and the adsorption isotherm. Hence, they can be treated as reversible non-dissipative processes, thus restricting the analysis to the constitutive relations of solid skeleton under the constraint  $\Phi_s \geq 0$ .

Equation (89) can be simplified by considering (78) and the relation  $\dot{\phi} = -\text{tr}(\dot{\boldsymbol{\epsilon}}) = -\mathbf{I} : \dot{\boldsymbol{\epsilon}}$  (which implies incompressible solid grains):

$$(\boldsymbol{\sigma} - P\mathbf{I}) : \dot{\boldsymbol{\epsilon}} - \gamma \dot{A}_s - \dot{\Psi}_s = \Phi_s \geq 0 \quad (92)$$

Substituting the Helmholtz free energy (43) and considering the hyperelastic relation, equation (92) reduces to

$$(E_B - \gamma A_{s,B}) \dot{B} + \boldsymbol{\sigma}' : \dot{\boldsymbol{\epsilon}}^p = \Phi_s \geq 0 \quad (93)$$

where  $\boldsymbol{\sigma}' = \boldsymbol{\sigma} - P\mathbf{I}$  is the effective stress (or the net stress if the pore fluid is air; [Nut08]);  $A_{s,B} = \partial A_s / \partial B$  is a function that reflects the intensity of surface area growth with respect to an infinitesimal change of GSD.

The rest of the development for a rate-dependent triaxial breakage model in surface reactive environment follows the same procedure as in Section 5.2. The same form of flow potential equation (67) is assumed:

$$w(\bar{E}_B, \bar{p}', \bar{q}) = \frac{a\gamma^b \dot{B}_0 \langle \bar{y}_{pB} \rangle^{n/2+1}}{2(1-B)^2 n/2+1} \quad (94)$$

with the yield surface in true stress space expressed as

$$y_{pB} = \frac{(1-B)^2}{a\gamma^b} (E_B - A_{s,B}\gamma) + \left( \frac{q}{Mp'} \right)^2 - 1 \quad (95)$$

Note that the critical breakage energy  $E_c$  in equation (67) is replaced by  $a\gamma^b$  based on grain-scale analysis [Zha16]. The evolution laws of the system can be derived in the same way

$$\begin{aligned} \dot{B} &= \frac{\partial w}{\partial \bar{E}_B} = \cos^2 \omega \dot{B}_0 \langle y_{pB} \rangle^{n/2} \\ \dot{\boldsymbol{\epsilon}}_v^p &= \frac{\partial w}{\partial \bar{p}'} = \frac{(E_B - A_{s,B}\gamma)}{p'} \sin^2 \omega \dot{B}_0 \langle y_{pB} \rangle^{n/2} \\ \dot{\boldsymbol{\epsilon}}_s^p &= \frac{\partial w}{\partial \bar{q}} = \frac{a\gamma^b}{(1-B)^2} \frac{q}{M^2 p'^2} \dot{B}_0 \langle y_{pB} \rangle^{n/2} \end{aligned} \quad (96)$$

The formulation is completed by a surface growth function which is here assumed to be given by

$$A_s = \frac{(1 + \mathcal{G}_A B) A_{s0}}{1 - \mathcal{G}_V B} \quad (97)$$

The derivation of this relation is discussed in detailed by [Zha18] and its agreement with reported data is shown in figure 20.

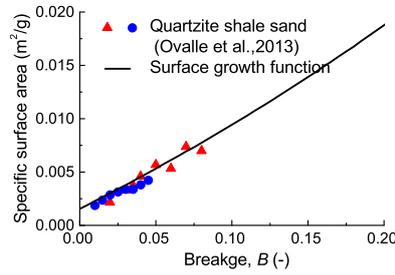


Figure 20: Performance of the surface growth function equation (97) against experimental data from oedometric compression test on quartzite shale sand (data after [Ova13]).

## 6.4 Model performance

Quantitative results can be obtained for the case of isotropic compression (i.e.  $q=0$ ). For linear elasticity, the comminution pressure can be derived by substituting equation (52) into (95)

$$p'_{CR} = \frac{1 - \mathcal{G}_M B}{1 - B} \sqrt{\frac{2K}{\mathcal{G}_M} \left[ a\gamma^b + (1 - B)^2 A_{s,B} \gamma \right]} \quad (98)$$

Similarly, the comminution pressure for pressure-dependent elasticity model can be expressed as follows:

$$p'_{CR} = p_r \frac{1 - \mathcal{G}_M B}{(1 - B)^{2/(2-m)}} \left[ \frac{(2-m)\bar{K}}{\mathcal{G}_M p_r} \left( a\gamma^b + (1 - B)^2 A_{s,B} \gamma \right) \right]^{\frac{1}{2-m}} \quad (99)$$

In both cases the incorporation of equation (83) into (98) or (99) provides  $p'_{CR}$  as a function of the water vapour potential,  $\mu_w$ . Such dependence can be explored by using curve A from figure 20. Figure 21 shows the variation of the comminution pressure  $p'_{CR}$  as a function of the chemical potential of water vapour,  $\mu_w^*$ , and of the relative humidity,  $RH$ . It can be noticed that the computed yield stress monotonically de-

increases upon increasing values of chemical potential (or, equivalently, upon decreasing values of  $RH$ ), thus reflecting a lower crushing resistance of a granular matrix embedded in a humid environment. Figure 22 plots elastic hardening and suction hardening characteristics of the yield surface in  $p'-q-B$  and  $p'-q-RH$  space, from which it is possible to note the growth of the yield stress upon comminution and drying. While the analysis is reported only with reference to a pressure-dependent elastic law, similar trends would have been obtained also in presence of a linear elastic granular matrix.

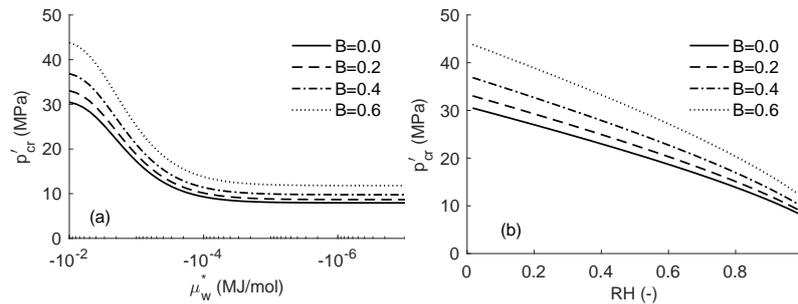


Figure 21: Isotropic yield pressure as a function of a) chemical potential of water vapour and b) relative humidity at various levels of  $B$ . Curves computed for a pressure-dependent elastic law (99).

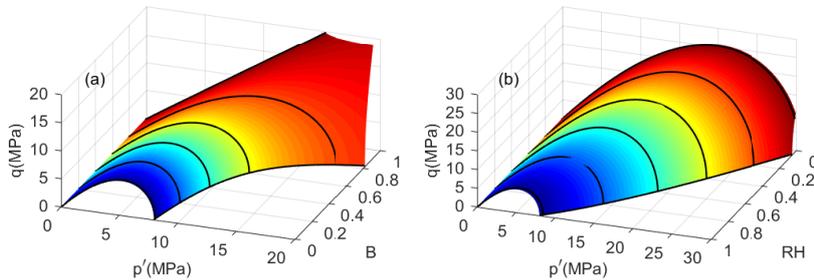


Figure 22: yield surface (95) in a)  $p'-q-B$  space plotted for  $RH=1$ , and b)  $p'-q-RH$  space plotted for  $B=0$ .

Now it is possible to simulate again Takei's oedometer test on quartz sand calibrated in figure 13 under alternating hydraulic conditions. The interruption of dry creep by flooding the specimens typically cause a substantial acceleration of the creep response, which can be explained as a signature of the feedback between water weakening and volumetric collapse. Figure 23a displays predicted two-stage creep responses at different stress levels for the same quartz sand is predicted to cause sharp increases in volumetric strain, as well as to activate creep effects even if they were not predicted under dry conditions (e.g., for  $\sigma_v=2.34$  MPa). The same phenomenon has been observed in numerous two-stage creep tests on sand [Che07, Brz14], for which

significant grain breakage is often observed after flooding and/or creep. These evidences are consistent with the model simulations depicted in figure 23b, where increase of specific surface area accompanied by flooding and creep is reported.

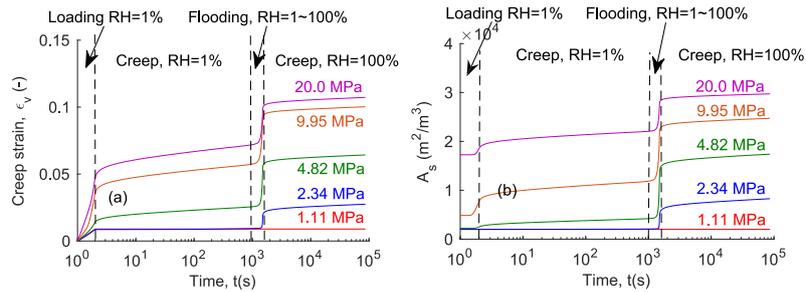


Figure 23: a) Simulated two-stage creep tests of quartz sand and b) corresponding surface growth curves.

The experimental data on Pancrudo rockfill reported by [Old03] provides an opportunity to evaluate the model under multiple values of relative humidity. Figure 24 refers to conventional oedometer tests involving a rapid vertical loading followed by a creep stage. The time interval  $\Delta t$  between two load increments is 1000 min, while the total time until completion of the test is approximately 9.5 days. These tests can be simulated more accurately by replacing the constant stress-rate loading with a series of sudden load increments followed by creep stages (i.e. a loading process which mimics directly how the data were obtained). In addition, the wetting stage of the test with RH=15.2% is also simulated by increasing RH in a similar incremental manner. Figure 24a shows that the computed compression curves agrees well with the experimental measurements at all levels of relative humidity. Fig. 24b plots the collapse strain vs. RH during the wetting stage of the tests with initial RH=15.2% and 95.9%. The model can provide a satisfactory quantitative agreement in terms of the relation

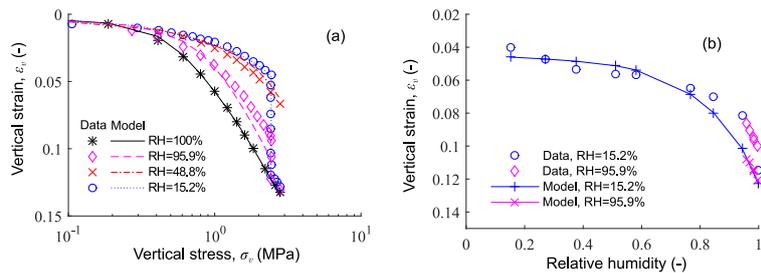


Figure 24: Oedometric compression with wetting stage: stars are data from Oldecop and Alonso (2003); open circles connected with lines are model performance

between collapse strains and relative humidity, RH. Such agreement is a direct consequence of the combined use of the Gibbs isotherm and Langmuir adsorption law, which corroborates the model hypothesis about the link between the weakening of granular materials in wet environments and the reduction of mineral surface energy due to water adsorption.

## 7 Conclusions

This paper has reviewed the application of energy principles in developing constitutive laws for fracture propagation in single particles and comminution in granular assemblies. The development of Breakage Mechanics theories for both rate-independent and rate-dependent scenarios has been revisited in analogy with the formalism of fracture mechanics by emphasizing the close similarity between fracture laws at the scale of individual particles and Breakage laws at the level of granular continua. Such analogy has led to fruitful applications and generalizations of the breakage mechanics theory, including the modelling of rate-dependent crushing starting from the kinetics of subcritical crack growth, as well as the incorporation of environmental effects due to stress corrosion at the crack tip. Similar lines of research developed over the last few years have shown that the Breakage Mechanics framework can be readily extended to a variety of materials and processes by exploiting such physical and conceptual duality between inelastic processes occurring at the grain scale and their implications at the macroscopic scale of granular solids. Examples of such recent developments include the interpretation of the grainsize-dependent properties of granular soils [Zha16], the formulation of critical state theories for crushable sand [Ten16], the mechanics of unsaturated granular soils [Bus12], cemented granular materials [Ten14, Bus15] and the effects of chemical deterioration in porous rocks [Ten14, Bus15].

Throughout the chapter, the benefits of TIV-based constitutive modelling have been emphasized by pointing out that: 1) TIV provides a well-defined hierarchical structure for modelling granular geomaterials that satisfies thermodynamic constraints; 2) this methodology is not limited to the mechanical response, but it can also model hydraulic, chemical, thermal and other multiphysical processes in a unified manner; 3) the degree of complexity and the number of features of the model can be flexibly adjusted by adding and/or removing ISVs; 4) TIV also provides a platform to integrate micro-mechanical considerations when combined with proper statistical homogenization schemes.

## 8 References

- [Atk80] Atkinson, B. K. (1980) Stress corrosion and the rate-dependent tensile failure of a fine-grained quartz rock. *Tectonophysics* **65(3)**:281-290.
- [Atk82] Atkinson, B. K. (1982) Subcritical crack propagation in rocks: theory, experimental results and applications. *Journal of Structural Geology* **4(1)**:41-56.
- [Atk84] Atkinson, B. K. (1984) Subcritical crack growth in geological materials. *Journal of Geophysical Research: Solid Earth (1978–2012)* **89(B6)**:4077-4114.
- [Atk87] Atkinson, B. K. & Meredith, P. G. (1987) The theory of subcritical crack growth with applications to minerals and rocks. *Fracture mechanics of rock* **2**:111-166.
- [Bau00] Baud, P., Zhu, W. & Wong, T.-F. (2000) Failure mode and weakening effect of water on sandstone. *Journal of Geophysical Research* **105(B7)**:16371-16389.
- [Ben10] Ben-Nun, O. & Einav, I. (2010) The role of self-organization during confined comminution of granular materials. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **368(1910)**:231-247.
- [Bou96] Boudet, J., Ciliberto, S. & Steinberg, V. (1996) Dynamics of crack propagation in brittle materials. *Journal de Physique II* **6(10)**:1493-1516.
- [Brz14] Brzesowsky, R., Hangx, S., Brantut, N. & Spiers, C. (2014) Compaction creep of sands due to time - dependent grain failure: Effects of chemical environment, applied stress, and grain size. *Journal of Geophysical Research: Solid Earth* **119(10)**:7521-7541.
- [Bus15] Buscarnera, G. & Das, A. (2015) Chemo - mechanics of cemented granular solids subjected to precipitation and dissolution of mineral species. *International Journal for Numerical and Analytical Methods in Geomechanics*.
- [Bus12] Buscarnera, G. & Einav, I. (2012) The yielding of brittle unsaturated granular soils. *Géotechnique* **62(2)**:147.
- [Cha58] Charles, R. (1958) Static fatigue of glass. I. *Journal of Applied Physics* **29(11)**:1549-1553.
- [Che07] Chester, F., Chester, J., Kronenberg, A. & Hajash, A. (2007) Subcritical creep compaction of quartz sand at diagenetic conditions: Effects of water and grain size. *Journal of Geophysical Research: Solid Earth* **112(B6)**.
- [Cil14] Cil, M. & Alshibli, K. (2014) 3D evolution of sand fracture under 1D compression. *Géotechnique* **64(5)**:351.
- [Col97] Collins, I. & Houlsby, G. (1997) Application of thermomechanical principles to the modelling of geotechnical materials. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **453(1964)**:1975-2001.
- [Col03] Collins, I. & Muhunthan, B. (2003) On the relationship between stress–dilatancy, anisotropy, and plastic dissipation for granular materials. *Géotechnique* **53(7)**:611-618.

- [Coo04] Coop, M., Sorensen, K., Freitas, T. B. & Georgoutsos, G. (2004) Particle breakage during shearing of a carbonate sand. *Géotechnique* **54**(3):157-164.
- [Cou10] Coussy, O. (2010) *Mechanics and physics of porous solids*. John Wiley & Sons.
- [Das13] Das, A., Nguyen, G. D. & Einav, I. (2013) The propagation of compaction bands in porous rocks based on breakage mechanics. *Journal of Geophysical Research: Solid Earth* **118**(5):2049-2066.
- [Deh06] Dehoff, R. (2006) *Thermodynamics in materials science*. CRC Press.
- [Dol84] Döll, W. (1984) Kinetics of crack tip craze zone before and during fracture. *Polymer Engineering & Science* **24**(10):798-808.
- [Ein07a] Einav, I. (2007) Breakage mechanics-part I: theory. *Journal of the Mechanics and Physics of Solids* **55**(6):1274-1297.
- [Ein07b] Einav, I. (2007) Breakage mechanics-Part II: Modelling granular materials. *Journal of the Mechanics and Physics of Solids* **55**(6):1298-1320.
- [Ein07c] Einav, I. (2007) Fracture propagation in brittle granular matter. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. The Royal Society, vol. 463, pp. 3021-3035.
- [Ein07d] Einav, I., Houlsby, G. & Nguyen, G. (2007) Coupled damage and plasticity models derived from energy and dissipation potentials. *International Journal of Solids and Structures* **44**(7):2487-2508.
- [Gri21] Griffith, A. A. (1921) The phenomena of rupture and flow in solids. *Philosophical transactions of the royal society of london. Series A* **221**:163-198.
- [Hac08] Hackl, K. & Fischer, F. D. (2008) On the relation between the principle of maximum dissipation and inelastic evolution given by dissipation potentials. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **464**(2089):117-132.
- [Har85] Hardin, B. O. (1985) Crushing of soil particles. *Journal of Geotechnical Engineering* **111**(10):1177-1192.
- [Hou02] Houlsby, G. & Puzrin, A. (2002) Rate-dependent plasticity models derived from potential functions. *Journal of Rheology (1978-present)* **46**(1):113-126.
- [Hou14] Houlsby, G. T. (2014) Dissipation rate functions, Pseudopotentials, Potentials and Yield Surfaces. In *Beyond the Second Law*. Springer, pp. 73-95.
- [Hou07] Houlsby, G. T. & Puzrin, A. M. (2007) *Principles of hyperplasticity: an approach to plasticity theory based on thermodynamic principles*. Springer Science & Business Media.
- [Kar10] Karimpour, H. & Lade, P. V. (2010) Time effects relate to crushing in sand. *Journal of Geotechnical and Geoenvironmental Engineering* **136**(9):1209-1219.
- [Kar13] Karimpour, H. & Lade, P. V. (2013) Creep behavior in Virginia Beach sand. *Canadian Geotechnical Journal* **50**(11):1159-1178.
- [Kik10] Kikumoto, M., Wood, D. M. & Russell, A. (2010) Particle crushing and deformation behaviour. *Soils and foundations* **50**(4):547-563.
- [Lad10] Lade, P. V. & Karimpour, H. (2010) Static fatigue controls particle crushing and time effects in granular materials. *Soils and foundations* **50**(5):573-583.

- [Lad14] Lade, P. V. & Karimpour, H. (2014) Stress relaxation behavior in Virginia Beach sand. *Canadian Geotechnical Journal* **52(7)**:813-835.
- [Lad96] Lade, P. V., Yamamuro, J. A. & Bopp, P. A. (1996) Significance of particle crushing in granular materials. *Journal of Geotechnical Engineering* **122(4)**:309-316.
- [Lub72] Lubliner, J. (1972) On the thermodynamic foundations of non-linear solid mechanics. *International Journal of Non-Linear Mechanics* **7(3)**:237-254.
- [Mau85] Maugis, D. (1985) Subcritical crack growth, surface energy, fracture toughness, stick-slip and embrittlement. *Journal of materials Science* **20(9)**:3041-3073.
- [McD98] Mcdowell, G. & Bolton, M. (1998) On the micromechanics of crushable aggregates. *Géotechnique* **48(5)**:667-679.
- [McD96] Mcdowell, G., Bolton, M. & Robertson, D. (1996) The fractal crushing of granular materials. *Journal of the Mechanics and Physics of Solids* **44(12)**:2079-2101.
- [McD02c] Mcdowell, G. & Humphreys, A. (2002) Yielding of granular materials. *Granular Matter* **4(1)**:1-8.
- [McD13] Mcdowell, G. R. & De Bono, J. P. (2013) On the micro mechanics of one-dimensional normal compression. *Géotechnique* **63(11)**:895.
- [Mic82] Michalske, T. A. & Freiman, S. W. (1982) A molecular interpretation of stress corrosion in silica. *Nature* **295**:511-512.
- [Moe99] Moës, N., Dolbow, J. & Belytschko, T. (1999) A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering* **46(1)**:131-150.
- [Nak99] Nakata, A., Hyde, M., Hyodo, H. & Murata (1999) A probabilistic approach to sand particle crushing in the triaxial test. *Géotechnique* **49(5)**:567-583.
- [Nak01b] Nakata, Y., Hyodo, M., Hyde, A. F., Kato, Y. & Murata, H. (2001) Microscopic particle crushing of sand subjected to high pressure one-dimensional compression. *Soils and foundations* **41(1)**:69-82.
- [Nak01a] Nakata, Y., Kato, Y., Hyodo, M., Hyde, A. F. & Murata, H. (2001) One-dimensional compression behaviour of uniformly graded sand related to single particle crushing strength. *Soils and foundations* **41(2)**:39-51.
- [Nar12] Nara, Y., Morimoto, K., Hiroyoshi, N., Yoneda, T., Kaneko, K. & Benson, P. M. (2012) Influence of relative humidity on fracture toughness of rock: implications for subcritical crack growth. *International Journal of Solids and Structures* **49(18)**:2471-2481.
- [Ngu09] Nguyen, G. D. & Einav, I. (2009) The energetics of cataclasis based on breakage mechanics. *Pure and Applied Geophysics* **166**:1693-1724.
- [Nut08] Nuth, M. & Laloui, L. (2008) Advances in modelling hysteretic water retention curve in deformable soils. *Computers and Geotechnics* **35(6)**:835-844.
- [Old01] Oldecop, L. & Alonso, E. (2001) A model for rockfill compressibility. *Géotechnique* **51(2)**:127-139.
- Oldecop, L. A. & Alonso, E. (2003) Suction effects on rockfill compressibility. *Geotechnique* **53(2)**:127-139.

- [Old07] Oldecop, L. A. & Alonso, E. (2007) Theoretical investigation of the time-dependent behaviour of rockfill. *Géotechnique* **57(3)**:289-301.
- [Old03] Oldecop, L. A. & Alonso Pérez De Agreda, E. (2003) Suction effects on rockfill compressibility. *Géotechnique* **53(2)**:127-139.
- [Ova13] Ovalle, C., Dano, C. & Hicher, P.-Y. (2013) Experimental data highlighting the role of surface fracture energy in quasi-static confined comminution. *International Journal of Fracture* **182(1)**:123-130.
- Perzyna, P. (1966) Fundamental Problems in Viscoplasticity. *Advances in applied mechanics* **9**:243.
- [Per66] Perzyna, P. (1966) Fundamental problems in viscoplasticity. *Advances in applied mechanics* **9**:243-377.
- [Pes95] Pestana, J. M. & Whittle, A. (1995) Compression model for cohesionless soils. *Géotechnique* **45(4)**:611-632.
- [Ric78] Rice, J. (1978) Thermodynamics of the quasi-static growth of Griffith cracks. *Journal of the Mechanics and Physics of Solids* **26(2)**:61-78.
- [Ric68] Rice, J. R. (1968) A path independent integral and the approximate analysis of strain concentration by notches and cracks. *Journal of Applied Mechanics* **35(2)**:379-386.
- [Ric71] Rice, J. R. (1971) Inelastic constitutive relations for solids: an internal-variable theory and its application to metal plasticity. *Journal of the Mechanics and Physics of Solids* **19(6)**:433-455.
- [Ros63] Roscoe, K., Schofield, A. & Thurairajah, A. (1963) Yielding of clays in states wetter than critical. *Géotechnique* **13(3)**:211-240.
- [Rus11] Russell, A. R. (2011) A compression line for soils with evolving particle and pore size distributions due to particle crushing. *Géotechnique Letters* **1(1)**:5-9.
- [Rus13] Russell, A. R. & Einav, I. (2013) Energy dissipation from particulate systems undergoing a single particle crushing event. *Granular Matter* **15(3)**:299-314.
- [Sam87] Sammis, C., King, G. & Biegel, R. (1987) The kinematics of gouge deformation. *Pure and Applied Geophysics* **125(5)**:777-812.
- [Sin85] Sing, K. S. (1985) Reporting physisorption data for gas/solid systems with special reference to the determination of surface area and porosity (Recommendations 1984). *Pure and applied chemistry* **57(4)**:603-619.
- [Sow65] Sowers, G., Williams, R. & Wallace, T. (1965) Compressibility of broken rock and the settlement of rockfills. In *Proc. 6th ICSMFE.*, vol. 2, pp. 561-565.
- [Sta03] Stazi, F., Budyn, E., Chessa, J. & Belytschko, T. (2003) An extended finite element method with higher-order elements for curved cracks. *Computational Mechanics* **31(1-2)**:38-48.
- [Tak01] Takei, M., Kusakabe, O. & Hayashi, T. (2001) Time-dependent behavior of crushable materials in one-dimensional compression tests. *Soils and foundations* **41(1)**:97-121.
- [Ten16] Tengattini, A., Das, A. & Einav, I. (2016) A constitutive modelling framework predicting critical state in sand undergoing crushing and dilation. *Géotechnique* **66(9)**:695-710.

- [Ten14] Tengattini, A., Das, A., Nguyen, G. D., Viggiani, G., Hall, S. A. & Einav, I. (2014) A thermomechanical constitutive model for cemented granular materials with quantifiable internal variables. Part I—Theory. *Journal of the Mechanics and Physics of Solids* **70**:281-296.
- [Ver17] Verberne, B. & Spiers, C. (2017) A quantitative microstructural investigation of depleted and undepleted reservoir sandstones. In *51st US Rock Mechanics/Geomechanics Symposium*. American Rock Mechanics Association.
- [Wan02] Wang, F., Sassa, K. & Wang, G. (2002) Mechanism of a long-runout landslide triggered by the August 1998 heavy rainfall in Fukushima Prefecture, Japan. *Engineering Geology* **63(1-2)**:169-185.
- [Wei74] Weidmann, G. & Holloway, D. (1974) Plastic flow-slow crack-propagation and static fatigue in glass. *Physics and Chemistry of Glasses* **15(3)**:68-75.
- [Wie67] Wiederhorn, S. (1967) Influence of Water Vapor on Crack Propagation in Soda - Lime Glass. *Journal of the American Ceramic Society* **50(8)**:407-414.
- [Wie70] Wiederhorn, S. & Bolz, L. (1970) Stress corrosion and static fatigue of glass. *Journal of the American Ceramic Society* **53(10)**:543-548.
- [Zha13] Zhang, C., Nguyen, G. & Einav, I. (2013) The end-bearing capacity of piles penetrating into crushable soils. *Géotechnique* **63(5)**:341.
- [Zha18] Zhang, Y. & Buscarnera, G. (2018) Breakage mechanics for granular materials in surface-reactive environments. *Journal of the Mechanics and Physics of Solids* **112**:89-108.
- [Zha16] Zhang, Y. D., Buscarnera, G. & Einav, I. (2016) Grainsize dependence of yielding in granular soils interpreted using fracture mechanics, breakage mechanics and Weibull statistics. *Géotechnique* **66(2)**:149-160.
- [Zhu97] Zhu, W. & Wong, T.-F. (1997) Shear-enhanced compaction in sandstone under nominally dry and water-saturated conditions. *International Journal of Rock Mechanics and Mining Sciences* **34(3-4)**:364. e1-364. e12.

©ALERT Geomaterials  
Laboratoire 3SR / Bâtiment Galilée  
CS 40700  
38 058 Grenoble cedex 9  
France

ISBN 978-2-9561359-2-0

Fon: +33 (0) 456 528 621  
Fax: +33 (0) 476 827 043  
president@alertgeomaterials.eu  
<http://alertgeomaterials.eu>

---

All rights reserved. No part of this book may be reproduced in any form or by any electronic or mechanical means, including information storage and retrieval systems, without written permission from the publisher or author, except in the case of a reviewer, who may quote brief passages embodied in critical articles or in a review.



**ALERT Doctoral School 2018**  
*Energetical methods in geomechanics*

---

Editors: I. Einav, E. Gerolymatou

M. Liu

Thermodynamics and Constitutive Modeling

I. Stefanou

Definition and Uses of the Principle of Virtual Power

F. Guillard

Energetics in Discrete Element Modelling

I. Einav

Hierarchical guide for constructing thermodynamically admissible constitutive models

E. Gerolymatou

Energetical background of common approaches in geomechanics

M. Sari, J. Lin, S. Alevizos, T. Poulet, M. Veveakis

Energy Based Constitutive framework for Multiphysics Geomechanics

Y. Zhang and G. Buscarera

Energetics of Crushable Granular Materials: from Particle Fracture to Breakage Mechanics

ISBN 978-2-9561359-2-0