Stochastic and Deterministic Models for Dense Granular Flow

by

Kenneth Norman Kamrin

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2008

Author .....
Department of Mathematics
May 5, 2008

Certified by .....
Martin Z. Bazant
Associate Professor, Thesis Supervisor

Accepted by .....
Alar Toomre
Chairman, Applied Mathematics Committee

Accepted by .....
David Jerison
Chairman, Department Committee on Graduate Students

Stochastic and Deterministic Models for Dense Granular Flow

by

Kenneth Norman Kamrin

Submitted to the Department of Mathematics
on May 5, 2008, in partial fulfillment of the
requirements for the degree of
DOCTOR OF PHILOSOPHY

Abstract

Granular materials such as sand or gravel surround us everyday and yet remain poorly understood. In this thesis, two models are developed for dense granular flow, each capable of predicting flows with accuracy in multiple environments. The models are based on differing perspectives of grain-level dynamics, with one deriving flow from a stochastic mechanism and the other from a deterministic deformation law.

The Stochastic Flow Rule (SFR): This work models granular flow as a sequence of localized collective grain displacements. As in the Spot Model for drainage (Bazant 2001), grain clusters move as dictated by "spots" which travel through the material as biased random-walkers. The SFR derives spot motion directly from the material stresses, thus generalizing and extending the Spot Model beyond drainage to any quasi-2D geometry with a computable stress field. Limit-State Mohr-Coulomb Plasticity is used to approximate the stress profile in a slow flowing granular assembly. The SFR then describes quantitatively how to convert the slip-line field and stresses into the necessary parameters to fully define a spot's random trajectory through the material and generate a steady flow profile. Results are compared to known flow data.

Nonlinear Granular Elasto-Plasticity: This work models granular deformation at the meso-scale as a deterministic consequence of the local stresses and state parameters. Recently proposed models for granular elasticity (Jiang and Liu 2003) and plastic flow (Jop *et al.* 2006) are combined into one universal granular continuum law, capable of predicting both flowing regions and stagnant zones simultaneously in any arbitrary 3D flow geometry. The unification is performed by first motivating physically, and then implementing a Kröner-Lee elasto-plastic decomposition. The model is then numerically solved in multiple geometries and results are compared to experiments and discrete simulations.

Thesis Supervisor: Martin Z. Bazant
Title: Associate Professor

Acknowledgments

My heart goes out to all those who provided academic and moral support for me over my years of PhD study at MIT. Academically, each member of my thesis committee deserves particular acknowledgement for shaping the way that I approach research. First, I must thank my adviser **Martin Bazant** for his steadfast guidance throughout my graduate career. Beyond his wealth of knowledge in several fields of physical science, Martin has taught me how to take raw ideas and compile them into something much bigger. His energetic, creative, and broad approach to problem-solving mirrors what I aspire to be, and so in Martin I have found an excellent model to follow. I cannot thank him enough for the support he has given me and his dedication to my future academic life.

I am also graciously indebted to **Lallit Anand**, who has had a strong influence on the way I think about material mechanics. His rigorous treatment of the subject matter suited my desire as a mathematician for a consistent and clear understanding of material response. Coupled with simulation methods and a basis in real experimental data, his approach has shown me the breadth and power of continuum mechanics and plasticity theory. I cannot thank him enough for granting me access to the ABAQUS software package, without which the latter half of this thesis would be unverified. And of course, I appreciate the hours of time he has spent meeting with me.

Throughout my graduate life, **Ruben Rosales** has always been there to provide me wisdom regarding the academic process and consultation on my research. In my many interactions with Ruben, he has been an excellent source of feedback on the various ideas I have had and has always kept the process positive through his infectious level of excitement and interest. I look forward to continuing this interaction in the upcoming years.

Other faculty who have impacted my graduate studies include **George Haller** and **Peter Shor**, who each participated in my qualifying examination and have enlightened me in the fields of dynamical systems and quantum information theory respectively. **Arshad Kudrolli** of Clark University has been a unifying factor be-

tween myself and the larger granular community, and I appreciate his years of support and firm knowledge of up-to-date results in granular physics. I must also thank **Igor Aranson** of Argonne National Laboratory, who has been an advocate for me and my research for over a year. Of course, I cannot forget those professors and teachers of my past who helped show me the wonders of applied mathematics and physical science years ago: **L. Craig Evans** and **Alexandre Chorin** of UC Berkeley, and **Tim Merrill** of Ygnacio Valley High School.

In terms of student collaborators, much of my work would not have been possible without the collaborative efforts of **Chris Rycroft**, whose discrete simulations have served well in enabling comparisons between theoretical predictions and actual results. Chris has been one of my closest partners in trying to decipher the granular puzzles of today, and for his contributions I am quite grateful. I also thank **Jeahyuk Choi**, another past member of the **MIT Dry Fluids Group**, for teaching me experimental methods in granular flow. The students in the **MIT Solid Mechanics Group** were instrumental to my learning of the ABAQUS software package. **Shawn Chester** deserves particular acknowledgement for his hours of administrative help, especially for setting up my account on the Truesdell cluster, without which many of my later results would have been almost impossible to achieve.

I am grateful to **Victor Chen** and **Joungkeun Lim** who co-organized the 2006-2007 Simple Person's Applied Math Seminar with me. This experience was also made more enjoyable thanks to the administrative aid of **Shirley Entzminger**. She and **Linda Okun** have both been consistently reliable throughout my graduate career and deserve many thanks. Others who have played a positive role in my research and classwork at MIT include applied math graduate students **Sabri Kilic**, **King Yeung Yick**, **Kevin Chu**, **Jeremy Levitan**, **Kevin Matulef**, and **Michael Baym**.

On the non-academic side, I am infinitely grateful to my family (**Diane**, **Terry**, and **Kacey Kamrin**) for always being there for me in times of good and bad. My upbringing within this family of fine-artists has instilled in me a deep appreciation for music and creativity, which I believe has helped me along the way as a scientist. My close extended family (including the **Harris** family and grandparents **Gert** and

Jerry Samuel) have been a wonderful and consistent influence throughout my life. The love, support, and companionship of my girlfriend **Jennifer French** has been invaluable to my well-being over the last year. I must also thank **Maeghan Weaver** for her positive involvement in my life during years past.

My MIT experience was also made more exciting thanks to my roommates **Nir Matalon**, **Denis Daly**, and **Alex McCauley**. I have made numerous story-esque memories with my friends from MIT and greater Boston: **Brian Selden**, **Mehdi Gazor**, **Korisha Ramdhanie**, **Raj Patel**, **Pedzi Makumbe**, **Tim Johnson**, **Morgan Jenkins**, **Alyssa Brown**, **Misha Leybovich**, **Ethan Butler**, and **Tim Nguyen**. Friends of mine from home, **Daron Carman**, **Robert Petersen**, and **Mark Schwartz** in particular, have continued to be there for me as they each have for well over a decade. I must also acknowledge my friends from UC Berkeley who continue to be a part of my life: **Priscilla Hernandez**, **Jennifer Kremen**, **Anand Uphadye**, **Jason Mabie**, and **Elson Lai** (who I promised years ago to one day mention in my dissertation!).

Finally, I would like to acknowledge the various groups I have had the pleasure of working with during my MIT career: the **MIT Musical Theater Guild**, the **Somerville Community Baptist Church Gospel Choir**, the **Boston Pops Gospel Choir**, and the **Swingin' Eight Balls Barbershop Quartet**.

Contents

1 General introduction **13**

2 The Stochastic Flow Rule **15**

2.1 Introduction 15

2.2 Concepts from continuum mechanics 18

2.2.1 Mohr-Coulomb plasticity: stresses 18

2.2.2 Mohr-Coulomb plasticity: flow rules 24

2.2.3 The rate-independence concept 26

2.3 Shortcomings of Mohr-Coulomb plasticity 28

2.3.1 Discontinuous "shocks" in stress and velocity 28

2.3.2 Physical difficulties with limit-state coaxiality 32

2.3.3 Incipient yield everywhere 34

2.3.4 Neglect of discreteness and randomness 35

2.4 The Stochastic Flow Rule 36

2.4.1 Diffusing "spots" of plastic deformation 36

2.4.2 General form of the flow rule 39

2.4.3 A simple model for steady flows 42

2.4.4 A mechanical theory of spot drift 43

2.4.5 Frame indifference 46

2.5 Applications to granular flow 47

2.5.1 Silos 48

2.5.2 Couette cells 52

2.5.3 Plate dragging 57

2.5.4 Slow heap flows . . . 59
2.6 Transition from the SFR to Bagnold rheology . . . 62
2.6.1 Breakdown of the SFR . . . 62
2.6.2 Bagnold rheology . . . 63
2.6.3 Slip-line admissibility . . . 64
2.6.4 Redistribution of excess shear stress . . . 66
2.6.5 A simple composite theory . . . 67
2.6.6 Some applications of the composite theory . . . 69
2.7 Conclusion . . . 73
2.7.1 Highlights of the present work . . . 73
2.7.2 Comparison with partial fluidization . . . 74
2.7.3 Future directions . . . 77

3 Nonlinear Granular Elasto-Plasticity **81**
3.1 Transition from the SFR . . . 81
3.2 Granular matter as a continuum . . . 83
3.3 Continuum statics . . . 87
3.3.1 Stress-only laws vs. elasticity . . . 87
3.3.2 Effective Medium Theory for bulk granular elasticity . . . 89
3.3.3 The Jiang–Liu granular elasticity law . . . 93
3.4 Continuum flow . . . 97
3.4.1 Bagnold scaling and relevant dimensionless quantities . . . 97
3.4.2 Flow regimes . . . 99
3.4.3 Review of past flow models . . . 103
3.4.4 The Jop–Pouliquen granular plasticity law . . . 105
3.5 Combining elasticity and plasticity: Physical considerations . . . 110
3.5.1 A simplified picture . . . 110
3.5.2 Microscopic evidence for elasticity during flow . . . 113
3.5.3 Modifying the Jiang–Liu elasticity law . . . 115
3.6 Combining elasticity and plasticity: Mathematical specifics . . . 119

3.6.1 Kröner–Lee decomposition . . . 120
3.6.2 Kinematic definitions . . . 122
3.6.3 Intermediate space variables . . . 125
3.6.4 Constitutive dependences . . . 128
3.6.5 Frame indifference . . . 130
3.6.6 The Coleman–Noll procedure . . . 133
3.6.7 Upholding symmetry: Reference body isotropy . . . 134
3.6.8 Upholding symmetry: Intermediate space isotropy . . . 137
3.6.9 Inserting the elastic and plastic models . . . 143
3.6.10 Summary of the equations . . . 144
3.7 Numerical implementation . . . 146
3.7.1 The material model . . . 146
3.7.2 Explicit procedure for global deformation and stress . . . 148
3.7.3 Artificial density reduction . . . 150
3.8 Results . . . 154
3.8.1 Rough-walled inclined chute . . . 155
3.8.2 Annular Couette cell . . . 159
3.8.3 Flat-bottomed silo . . . 164
3.9 Conclusion and future directions . . . 174
3.9.1 Quasi-static non-locality . . . 174
3.9.2 Dilation . . . 175
3.9.3 Flow condition . . . 177

4 General conclusion **179**

A Review and comparison of other granular plasticity models **183**
A.1 Critical state soil mechanics . . . 183
A.2 The Anand-Gu model . . . 185
A.2.1 Double-Shearing Flow Rule . . . 186
A.2.2 Hardening/Softening and Dilation . . . 191
A.3 Comparison to Critical State Theory . . . 194

B SFR in a hopper **197**

C Deciding the stochastic dilemma **201**

D Eulerian numerical method for fluid/solid interaction **207**

D.1 Past work on fluid/fluid interaction 207

D.2 Introductory large deformation elasticity 211

D.3 Eulerian solid mechanics 213

D.4 The solid/fluid jump relation 214

D.5 Process of the numerical solution 216

E Explicit VUMAT **223**

F Implicit VUMAT **233**

Chapter 1

General introduction

At first it may seem confounding how it is that our civilization, with all its modern technological and scientific advances, still has trouble understanding one of the most familiar and basic materials— sand. For years granular materials have resisted theoretical development, demonstrating non-trivial behavior that may resemble solid, liquid, or gas under different circumstances. The deceptively challenging mechanics and dynamics of granular materials combined with the ubiquity of such materials in day-to-day life have made this topic especially interesting to study.

The drive to understand granular physics is motivated not just by academic curiosity, but also by a real need for a predictive model. Advancements in granular materials science would have a direct effect on fields like geology and civil engineering, which could benefit from improved models for soil and gravel, as well as military research, where codes for bunker design have not been updated in recent memory and require enormous safety factors. In terms of energy production, better granular models could help optimize the handling and processing of coal, as well as streamline the engineering of pebble-bed nuclear reactors, in which billiard-size pebbles containing radioactive fuel are slowly drained through an apparatus. In industry, granular materials are second only to water as the most handled raw material, so it is expected that improved granular modeling could increase industrial efficiency dramatically.

This thesis develops two distinct models for granular media. It should be emphasized that these models attempt to describe different regimes of granular response and

thus are not contradictive. Both are primarily concerned with steady, dense flows. The Stochastic Flow Rule (SFR), developed in Chapter 2, is intended for slow flows. Chapter 3 develops a continuum elasto-plastic model that attempts to describe the liquid-like effects of faster flows while simultaneously describing statics. Each model offers sufficient agreement with known flow data.

Though the models cover a combined range that spans almost the entire range of dense flow behavior, individually, each model leaves out certain effects. At the end of Chapter 2 an attempt is made to extend the SFR into the faster flow regime, which appears to add some breadth. While the continuum model of Chapter 3 draws clear lines as to which flow behaviors should be accounted for, it remains future work to integrate a slow-flow theory. It would be ideal if the SFR could somehow be combined with the continuum model, though it remains unclear if this is possible given the different foundations of the models. Some speculation on a joint model is given in the general conclusion, chapter 4.

The following two chapters can be read independently. There has been intentionally very little cross-referencing between chapters because the two models are based on very different fundamental assumptions. The underlying theme among both models is unification. The SFR unifies an older theory for 2D granular stresses, Limit-State Mohr-Coulomb Plasticity [132], with more recent work on collective rearrangement during flow, the Spot Model [15]. The continuum elasto-plastic model unifies a recently proposed granular elasticity model, that of Jiang and Liu [68], with a newly developed plastic flow rule, that of Jop and Pouliquen [71].

Chapter 2

The Stochastic Flow Rule

2.1 Introduction

For centuries, engineers have described granular materials using continuum solid mechanics [101, 132, 60]. Dense granular materials behave like rigid solids at rest, and yet are easily set into liquid-like, quasi-steady motion by gravity or moving boundaries, so the classical theory is Mohr-Coulomb plasticity (MCP), which assumes a frictional yield criterion. The simplest model is the two-dimensional "Ideal Coulomb Material" at limit-state, where the maximum ratio of shear to normal stress is everywhere equal to a constant (the internal friction coefficient), whether or not flow is occurring. This model is believed to describe stresses well in static or flowing granular materials, but, as we explain below, it fails to predict flow profiles, when combined with the usual Coaxial Flow Rule of continuum plasticity. Indeed, it seems continuum mechanics has not yet produced a simple and robust model for granular flow.

In recent years, the sense that there is new physics to be discovered has attracted a growing community of physicists to the study of granular materials [64, 34, 72, 43, 58, 9]. Unlike the engineers, their interest is mostly at the discrete particle level, mo-

tivated by the breakdown of classical statistical mechanics and hydrodynamics due to strong dissipation and long-lasting, frictional contact networks. Dense granular materials exhibit many interesting collective phenomena, such as force chains, slow structural relaxation, and jamming. Similar non-equilibrium phenomena occur in glasses, foams, and emulsions, as in granular materials, so it is hoped that a general new statistical theory may emerge. Presumably from such a microscopic basis, continuum models of glassy relaxation and dense granular flow could be systematically derived, just as dissipative hydrodynamics for granular gases can be derived from kinetic theory with inelastic collisions [25].

This dream has not yet been achieved, but many empirical continuum models have been proposed [64, 9, 13]. The difficulty in describing dense granular flow is evidenced by the remarkable diversity of physical postulates, which include: coupled static and rolling phases [21, 22, 24, 23], Bagnold rheology [12] based on "granular eddies" [44], granular temperature-dependent viscosity [124], density-dependent viscosity [88, 19], non-local stress propagation along arches [95], self-activated shear events due to non-local stress fluctuations [111, 110], free-volume diffusion opposing gravity [85, 100, 102, 15, 120], "shear transformation zones" coupled to free-volume kinetics [83, 82], and partial fluidization governed by a Landau-like order parameter [7, 8]. Each of these theories can fit a subset of the experimental data [94], usually only for a specific geometry for which it was designed, such as a flowing surface layer [21, 22, 24, 23, 7], inclined plane [12, 44], Couette cell [88, 19], inclined chute [111, 110], or wide silo [85, 100, 102, 15, 120], and none seems to have very broad applicability. For example, we are not aware of a single model, from physics or engineering, which can predict velocity profiles in both draining silos and annular Couette cells, even qualitatively.

The theory of partial fluidization of Aranson and Tsimring has arguably had the most success in describing multiple flows within a single theoretical framework [7, 8]. Although setting boundary conditions for the order parameter usually requires additional *ad hoc* assertions, the model is nonetheless able to reproduce known flow behavior in inclined chutes, avalanches, rotating drums, and simple shear cells without many fitting parameters. It also describes some unsteady flows. However, the theory

lacks any clear microscopic foundation and is not directly coupled to a constitutive stress model for static materials. As such, it has only been applied to problems with very simple solid stress fields, limiting its current applicability to flows that depend on only one spatial variable.

In an attempt to describe arbitrary geometries, such as silos and Couette cells, we take the view that the engineers may already have a reasonable continuum description of the mean stresses, so we start with Mohr-Coulomb plasticity. However, discreteness and randomness clearly need to be taken into account in a granular material. For static stresses, quenched randomness in material properties is known to lead to statistical slip-line blurring in "stochastic plasticity" [104], but this says nothing about how plastic yielding actually occurs.

To describe yielding *dynamics*, we propose a "stochastic flow rule" (SFR) where local fluidization (stick-slip transition) propagates randomly along blurred slip-lines. We build on the recently proposed Spot Model for random-packing dynamics [15] by viewing "spots" of free volume as carriers of plasticity in granular materials, analogous to dislocations in crystals. Multiscale spot simulations can reproduce quite realistic flowing packings in silo drainage [120]; here, we introduce a mechanical basis for spot motion from MCP, which leads to a theory of considerable generality for bulk granular flows.

The paper is organized as follows. Since plasticity is unfamiliar to most physicists, we begin by reviewing key concepts from MCP in section 2.2, both for stresses and for dense flows. In section 2.3, we highlight various shortcomings of the classical theory, many of which we attribute to the Coaxial Flow Rule. We then introduce the general spot-based SFR and a specific simplification to be used for granular flow in section 2.4. Next we apply the theory to four prototypical examples: silo, Couette, heap, and plate-dragging flows in section 2.5. Then in section 2.6, we explain how the last two examples indicate a smooth transition from the SFR to Bagnold rheology, when slip-lines become admissible, and we present a simple composite theory, which extends the applicability of the model to various shear flows. In section 2.7, we conclude by further clarifying the range of applicability of the SFR and possible

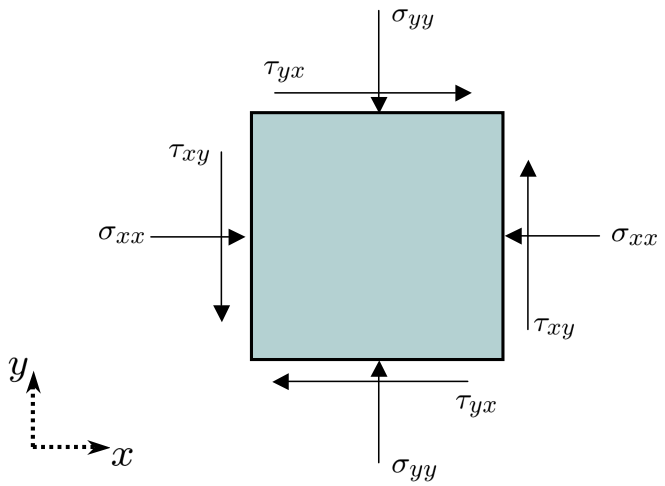


Figure 2-1: Stresses on a material element. All vectors are pointing in the positive direction as per our sign convention.

extensions to other granular flows and different materials.

2.2 Concepts from continuum mechanics

2.2.1 Mohr-Coulomb plasticity: stresses

In the eighteenth century, it was Coulomb, as a military engineer designing earthen fortresses, who introduced the classical model of a granular material, which persists to the present day: a continuous medium with a frictional yield criterion. His ideas were expressed in general continuum-mechanical terms by Mohr a century later, and a modern mathematical formulation of "Mohr-Coulomb plasticity" (MCP), which we also use below, is due to Sokolovskii[132]. Although other mechanical models exist, such as Drucker-Prager plasticity [112], MCP is perhaps the simplest and most widely used for granular materials in engineering [101]. As such, we choose to build our model of dense granular flow on the MCP description of stresses, as a reasonable and time-tested first approximation.

We begin in this section by reviewing relevant concepts from MCP, e.g. following Nedderman [101]. The fundamental assumption is that a granular material can be treated as an "Ideal Coulomb Material" (ICM), i.e. a rigid-plastic continuous media

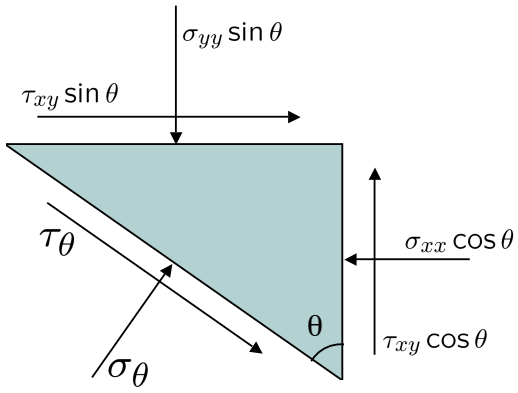


Figure 2-2: Force diagram for a wedge. Hypotenuse length assigned to unity.

which yields according to a Coulomb yield criterion

$$|\tau/\sigma| = \mu \equiv \tan\phi \tag{2.1}$$

where τ is the shear stress, σ is the normal stress, and ϕ is the internal friction angle, akin to a standard friction law with no cohesion. Throughout, we accept the common tensorial conventions for stresses with the key exception that normal stresses are deemed positive in compression. This is a standard modification in the study of non-cohesive granular materials since granular assemblies cannot support tension. We will also focus entirely on quasi-2D geometries.

Consider a small material element in static equilibrium and with no body forces present (see Figure 2-1). The normal stresses σ_{xx} and σ_{yy} can differ and the shear stresses τ_{xy} and τ_{yx} must be equal in order to balance moments. Likewise the variable τ_{yx} is redundant and will not be used again in this paper. To determine the stresses along any angle within this element, we place a new boundary within the material at some desired angle θ and observe force balance on the wedge that remains (see Figure 2-2). After algebraic simplification, this gives

$$\begin{aligned} \sigma_\theta &= \frac{1}{2}(\sigma_{xx} + \sigma_{yy}) + \frac{1}{2}(\sigma_{xx} - \sigma_{yy})\cos 2\theta - \tau_{xy}\sin 2\theta \\ \tau_\theta &= \frac{1}{2}(\sigma_{xx} - \sigma_{yy})\sin 2\theta + \tau_{xy}\cos 2\theta \end{aligned}$$

Now define

$$\begin{aligned} p &= \frac{1}{2}(\sigma_{xx} + \sigma_{yy}) \\ \tan 2\psi &= \frac{-2\tau_{xy}}{\sigma_{xx} - \sigma_{yy}} \\ R &= \sqrt{\left(\frac{\sigma_{xx} - \sigma_{yy}}{2}\right)^2 + \tau_{xy}^2} \end{aligned}$$

which allows us to write

$$\sigma_\theta = p + R\cos(2\theta - 2\psi) \tag{2.2}$$

$$\tau_\theta = R\sin(2\theta - 2\psi) \tag{2.3}$$

This implies that for all angles θ, the locus of traction stresses $(\sigma_\theta, \tau_\theta)$ is a circle centered at $(p, 0)$ with radius R. This circle is referred to as "Mohr's Circle".

We have just derived Mohr's Circle without accounting for the possible effects of body forces acting on the material element and gradients in the stress field. Adjusting for these effects, however, would change the results only negligibly as the element gets small in size. If we were to apply the same force-balancing analysis to a differentially small material element with a body force and stress gradients, we would find that the stress differences on the walls and the inclusion of the differentially small body force within only add differentially small terms to the equations for σ_θ and τ_θ. Thus we can always use Mohr's Circle to obtain traction stresses along a desired angle.

To ultimately define a full stress state for the material element, we need one more equation— we have 3 stress variables and only 2 force balance equations:

$$\frac{\partial \sigma_{xx}}{\partial x} - \frac{\partial \tau_{xy}}{\partial y} = F^x_{body} \tag{2.4}$$

$$\frac{\partial \sigma_{yy}}{\partial y} - \frac{\partial \tau_{xy}}{\partial x} = F^y_{body} \tag{2.5}$$

We say a material element is at *incipient failure* if the yield criterion is fulfilled along some direction and $|\tau/\sigma| \leq \mu$ along all others. A material in which incipient failure

occurs everywhere is said to be at a *limit-state*. In a limit-state, the Mohr's Circle at every point in the material must be tangent to the locus $|\tau/\sigma| = \mu$. As can be seen by applying trigonometry in Figure 2-3, this requirement means that $R = p\sin\phi$, enabling us to parameterize the stresses in terms of p and ψ only, thereby closing the equations. For this reason, we restrain our analysis to limit-state materials and refer to p and ψ as the stress parameters or Sokolovskii variables. (The limit-state stress treatment described here is also known as "Slip-Line Theory"; to avoid possible confusion, we specify this is not equivalent to Limit Analysis Plasticity concerned with upper and lower collapse limits.)

Solving for the original stress variables in terms of the stress parameters gives:

$$\sigma_{xx} = p(1 + \sin\phi\cos 2\psi) \tag{2.6}$$

$$\sigma_{yy} = p(1 - \sin\phi\cos 2\psi) \tag{2.7}$$

$$\tau_{xy} = -p\sin\phi\sin 2\psi \tag{2.8}$$

Using these expressions, we re-write equations (2.4) and (2.5):

$$\begin{aligned}(1+\sin\phi\cos 2\psi)p_x - 2p\sin\phi\sin 2\psi\ \psi_x + \sin\phi\sin 2\psi\ p_y \\ + 2p\sin\phi\cos 2\psi\ \psi_y = F^x_{body} \\ \sin\phi\sin 2\psi\ p_x + 2p\sin\phi\cos 2\psi\ \psi_x + (1-\sin\phi\cos 2\psi)p_y \\ + 2p\sin\phi\sin 2\psi\ \psi_y = F^y_{body}\end{aligned}$$

These will be referred to as the "stress balance equations". They form a hyperbolic system and thus can be solved using the method of characteristics. The system reduces to the following two characteristic equations:

$$\begin{aligned}dp \mp 2p\mu\ d\psi = F^y_{body}(dy \mp \mu\ dx) + F^x_{body}(dx \pm \mu\ dy) \\ \text{along curves fulfilling}\ \ \frac{dy}{dx} = \tan(\psi \mp \epsilon).\end{aligned} \tag{2.9}$$

To solve the stress balance equations, mesh the two families of characteristic curves

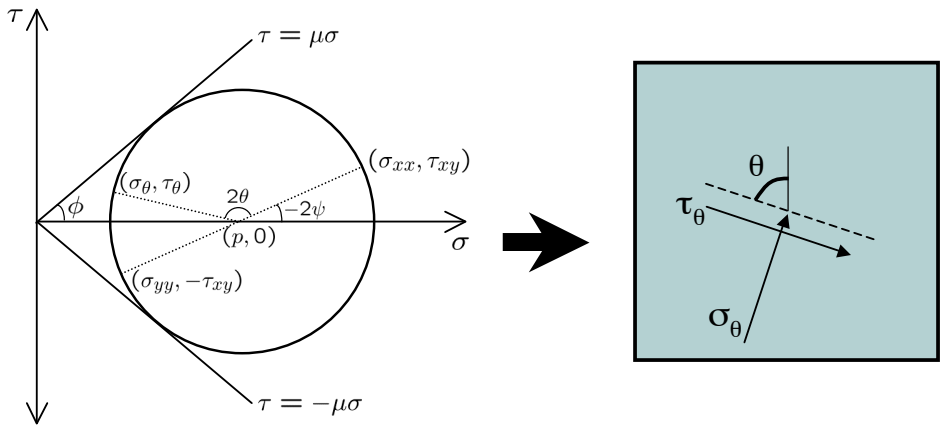


Figure 2-3: Using Mohr's circle jointly with the Coulomb internal yield locus ($\tau = \pm\mu\sigma$) to determine the traction stresses along any plane within a material element.

in the bulk, then march from the boundaries in, progressively applying the two differential relationships above to approximate the stress parameters at each intersection point in the mesh. More on this can be found in [62]. Other ways to solve the stress balance equations include the Two-Step Lax-Wendroff Method [105] and the Galerkin Method [55].

We return now to Mohr's Circle for a discussion of the stress properties within a differential material element. Equations (2.2) and (2.3) show that Mohr's Circle can be used as a slide-rule to determine the stresses along any angle θ: One arrives at the point $(\sigma_\theta, \tau_\theta)$ by starting at (σ_{xx}, τ_{xy}) and traveling anti-clockwise around Mohr's Circle for 2θ radians (see Figure 2-3). Also note on the diagram that the stresses along the x and y directions lie along a diameter of Mohr's Circle; any two material directions differing by an angle of $\pi/2$ lie along a diameter of the corresponding Mohr's Circle diagram. Utilizing this property in reverse is perhaps the easiest way to draw Mohr's Circle in the first place; draw the unique circle for which (σ_{xx}, τ_{xy}) and $(\sigma_{yy}, -\tau_{xy})$ are endpoints of a diameter.

Let $(\sigma_1, 0)$ and $(\sigma_3, 0)$ be the points of intersection between Mohr's Circle and the σ-axis, where $\sigma_1 > \sigma_3$. These points correspond to the two lines within a material element along which the shear stress vanishes and the normal stress is maximal or minimal. σ_1 (σ_3) is called the major (minor) principal stress and the line on which it

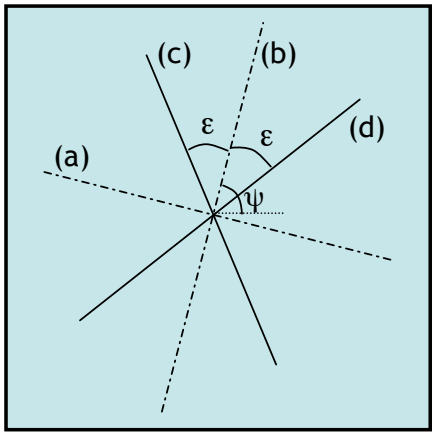


Figure 2-4: Important lines intersecting each material point: (a) Major principal plane / Minor principal stress direction; (b) Minor principal plane / Major principal stress direction; (c)-(d) Slip-lines.

acts is called the major (minor) principal plane.

Mohr's Circle shows that the major principal plane occurs at an angle ψ anti-clockwise from the vertical (see Figure 2-3). Thus the major principal stress points along an angle ψ anti-clockwise from the horizontal. This is the standard physical interpretation of ψ. One might think of ψ as the angle from the horizontal along which a force chain would be predicted to lie.

By right-triangle geometry, a line segment connecting the center of Mohr's Circle to a point of tangency with the internal yield locus would make an angle of $\pi/2 - \phi$ with the σ-axis. Each point of tangency represents a direction along which the yield criterion is met, i.e. a slip-line. Mohr's Circle indicates that the slip-lines are angled $(\pi/2 - \phi)/2$ up and down from the minor principal plane. But since the major and minor principal planes are orthogonal, the major principal stress points along the minor principal plane. Defining $\epsilon = \pi/4 - \phi/2$, we deduce that slip-lines occur along the angles $\psi \pm \epsilon$ measured anti-clockwise from the from the horizontal. Looking back at the characteristic equations, we see that the slip-lines and the characteristic curves coincide. This means that information from the boundary conditions propagates along the slip-lines to form a full solution to the stress balance equations.

It is worth noting that the stress balance equations are written for static materials and do not appear to account for dynamic behavior like dilatancy and convection

stresses. The theory of critical state soil mechanics [127] was the first to rigorously approach the issue of dilatancy (see appendix). It concludes that when material attains a flow state in which the density field stops changing in time, all points in the flow lie along a *critical state line* of the form $|\tau/\sigma| = \delta$ for δ constant. Since this exactly mirrors the Coulomb yield criterion, we can keep the stress balance equations and utilize $\delta = \mu$ (as in [66]). As for convection, adding the $\rho\mathbf{u}\cdot\nabla\mathbf{u}$ term into the stress equations couples the stresses to the velocity and makes the problem very difficult to solve. The practice of ignoring convection is justified by our slow-flow requirement and is commonly used and validated in basic solid mechanics literature [60, 132, 101]. So we conclude that dynamic effects in flowing materials do not preclude the use of the stress balance equations in slow, steady flows.

2.2.2 Mohr-Coulomb plasticity: flow rules

To calculate flow, we assert incompressibility and a flow rule— the flow rule is a constitutive law chosen to reflect the general behavior of the material at hand. The continuous nature of the ICM assumption suggests that symmetry should be kept with respect to the principal stress planes. Based on this, Jenike proposed adopting the Coaxial Flow Rule. The principle of coaxiality claims that material should flow by extending along the minor principal stress direction and contracting along the major principal stress direction; the principal planes of stress are aligned with the principal planes of strain-rate. The intuition for this constitutive rule is shown in Figure 2-5. Mathematically, this means that in a reference frame where the minor and major principal stress directions are the basis, the strain-rate tensor should have no off-diagonal components, i.e.

$$\mathbf{R}_\psi \dot{\mathbf{E}} {\mathbf{R}_\psi}^T \text{ is diagonal,} \tag{2.10}$$

where $\mathbf{R}_\psi$ rotates anti-clockwise by ψ and $\dot{\mathbf{E}}$ is the strain-rate tensor

$$\dot{\mathbf{E}} = \frac{1}{2}\left(\nabla\mathbf{u} + (\nabla\mathbf{u})^T\right). \tag{2.11}$$

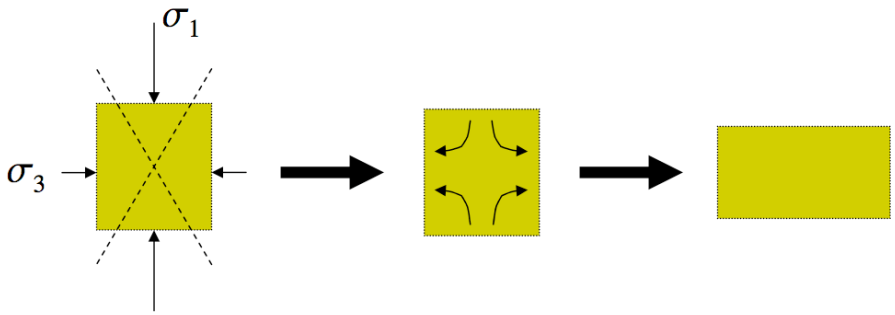


Figure 2-5: Sketch of the Coaxial Flow Rule.

where $\mathbf{u} = (u, v)$ is the velocity. Calculating the (1,2) component of the matrix in equation (2.10) and setting it to zero gives the equation of coaxiality,

$$\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) \tan 2\psi. \tag{2.12}$$

This flow rule has played a dominant role in the development of continuum plasticity theory and will be closely analyzed in this work.

Coaxiality with incompressibility comprises another hyperbolic system of equations enabling the velocity field to be solved via characteristics:

$$\begin{aligned} &du + \tan(\psi \mp \pi/4)\, dv = 0 \\ &\qquad \text{along curves fulfilling} \quad \frac{dy}{dx} = \tan(\psi \mp \pi/4). \end{aligned} \tag{2.13}$$

So, given $\psi(x, y)$ from the stress balance equations, information about the flow travels from the boundaries into the bulk along curves rotated $\pi/4$ off from the principal stress planes— using Mohr's Circle, we observe that these are the lines for which the shear stress is maximal (and the normal stresses are equal).

Other flow rules have been suggested instead of coaxiality. Of specific note, A.J.M. Spencer [133] has proposed the double-shearing flow rule. Unlike coaxiality which can be understood as a simultaneous equal shearing along both slip-line families, double-shearing allows the shearing motion to be unequally distributed between the two families in such a way that the flow remains isochoric. For steady flows, the double-

shearing flow rule is

$$\begin{aligned} & \sin 2\psi \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) - \cos 2\psi \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) \\ = \quad & \sin \phi \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} - 2\mathbf{u} \cdot \nabla \psi \right) \end{aligned} \tag{2.14}$$

It can be seen that when the material neighboring a particle rotates in sync with the rotation of the principal planes (i.e. as tracked by the rate of change of ψ), the right side goes to zero and the rule matches coaxiality. Under double-shearing, the characteristics of stress and velocity align, easing many aspects of the numerics. Some recent implementations of granular plasticity have utilized principles of double-shearing [3]. Though in this paper we deal primarily with the comparison of coaxiality to our new theory, this equation will be mentioned again in a later section.

2.2.3 The rate-independence concept

We now more fully address the conceptual basis for the flow theory just introduced. The theory is fundamentally different from traditional fluids where force-balance (including convection and viscous stresses in the case of Newtonian fluids) can be used alongside incompressibility to fully determine the fluid velocity and pressure fields. Unlike a fluid, granular materials can support a static shear stress and thus force-balance plus incompressibility alone is an underconstrained system. Rather, the stress-strain relationship for granular material is presumed to be *rate-independent* in the slow, quasi-static regime we study.

This concept is best understood tensorially. We can rewrite the equations of coaxiality and incompressibility equivalently as:

$$\dot{\mathbf{E}} = \lambda \mathbf{T_0} \tag{2.15}$$

where

$$\mathbf{T} = \text{Stress tensor} = \begin{pmatrix} -\sigma_{xx} & \tau_{xy} \\ \tau_{xy} & -\sigma_{yy} \end{pmatrix} \tag{2.16}$$

$$\mathbf{T_0} = \mathbf{T} - \frac{1}{2}(\text{tr}\mathbf{T})\mathbf{I} = \text{Deviatoric stress tensor}, \tag{2.17}$$

and λ is a multiplier that can vary in space. Equation (2.12) is merely the ratio of the (1,2) component and the difference of the (2,2) and (1,1) components of equation (2.15), thus canceling λ, and incompressibility is automatic since we relate to the deviatoric stress tensor. Equation (2.15) gives a simple and highly general form for plastic material deformation applicable to a broad range of deformable materials and so it is ideal for illustrating the role of rate-dependency. In MCP, we solve for $\mathbf{T_0}$ a priori from the stress balance equations. λ adds the extra degree of freedom necessary to make sure the strain-rate field is actually compatible with a real velocity field— λ is not any specific function of the stress or strain-rate variables and it adjusts to fit different velocity boundary conditions. Thus the stress alone does not imply the strain-rate and vice versa.

Supposing on the other hand that we were dealing with a *rate-dependent* (i.e. visco-plastic) material like Newtonian fluid, the above tensorial equation would still apply but we cannot claim to know $\mathbf{T_0}$ in advance since material motion changes the stresses. Instead we prescribe a functional form for λ, like $\lambda = \text{viscosity}^{-1} = \text{constant}$, and write the force balance equations in terms of $\dot{\mathbf{E}}$. Thus $\dot{\mathbf{E}}$ is computed very differently for the two cases: in the rate-independent case, (2.15) is solved using a known form for $\mathbf{T_0}$ and in the rate-dependent case, (2.15) is solved using a known form for λ.

The physical intuition for rate-independent *flow* can be easily understood with an example. Suppose we slide two frictional blocks against each other at two different non-zero sliding rates. In most rudimentary dry friction laws, the shear stress required to slide one block against another is proportional to the normal stress— there is no mention whatsoever of the rate of sliding. Thus the two sliding rates are modeled

to be attainable with the same shear stress and likewise the stress-strain relationship is deemed rate-independent. For slow granular flows with long-lasting interparticle contacts, comparisons with this example are especially instructive.

2.3 Shortcomings of Mohr-Coulomb plasticity

The use of the stress balance equations with incompressibility and the Coaxial Flow Rule will be referred hitherto as Mohr-Coulomb Plasticity (MCP). The theory has the benefit of being founded on mechanical principles, but does have some marked drawbacks. We point out a few:

- The theory frequently predicts highly discontinuous velocity fields.
- The Coaxial Flow Rule is conceptually troubling in some simple geometries.
- The assumption of limit-state stresses is overreaching.
- MCP is a continuum theory and thus cannot model discreteness and randomness.

We will now discuss these four points in detail.

2.3.1 Discontinuous "shocks" in stress and velocity

The two stress PDEs and two flow PDEs are each fully hyperbolic systems meaning that continuous solutions do not necessarily exist for arbitrary choices of the boundary conditions. Instead, discontinuous solutions are constructed utilizing intuitive jump conditions. For stresses, a jump in the stress parameters across a discontinuity line is only allowable if such a jump places no net forces on a small control volume surrounding the line thereby ensuring particle stability. This means the normal and shear stresses along the direction of the discontinuity must be the same on both sides of the discontinuity. However, the normal stress along the perpendicular direction can have a jump upon crossing the discontinuity as this places no net force on the control volume (see Figure 2-6).

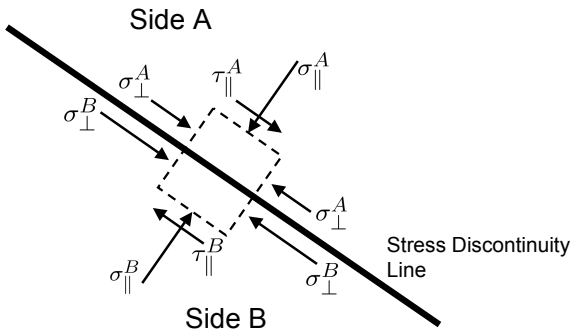


Figure 2-6: Stresses on a control volume intersected by a discontinuity. Note how a jump in $\sigma_\perp$ places no net force on the control volume.

In terms of the stress parameters, this means that p and ψ can jump as long as

$$\frac{(1+\sin\phi\cos(2\Theta-2\psi^B))}{(1+\sin\phi\cos(2\Theta-2\psi^A))}=\frac{\sin(2\Theta-2\psi^B)}{\sin(2\Theta-2\psi^A)}=\frac{p^A}{p^B} \tag{2.18}$$

where Θ is the angle from the vertical along which the stress discontinuity lies.

As for velocity, incompressibility forces us to impose a simpler jump rule in that the component tangent to the velocity discontinuity is the only one allowed to jump. We note that whenever a stress shock exists, the jump in the stress parameters will usually place a jump in the flow rule and may cause a velocity discontinuity to form coincident with the stress shock. A velocity discontinuity can form even when the stress field is smooth since the velocity PDEs are themselves hyperbolic. (It is worth noting that when shocks are allowed in the solution, multiple solutions sometimes arise to the same problem; introduction of the so-called "entropy condition" can be used to choose the best of the possible solutions [38, 101].) Overall, the MCP equations are mathematically very poorly behaved, and have also been shown to give violent instabilities and finite-time singularities [106, 126].

Aside from its mathematical difficulties, MCP theory also does not match experiments or our everyday experience of granular flows. In particular, MCP commonly predicts complicated patterns of velocity discontinuities in situations where experiments indicate smooth flow in steady-state. In Figure 2-7, the numerically determined stress field for a wedge hopper with only slightly non-radial boundary conditions on

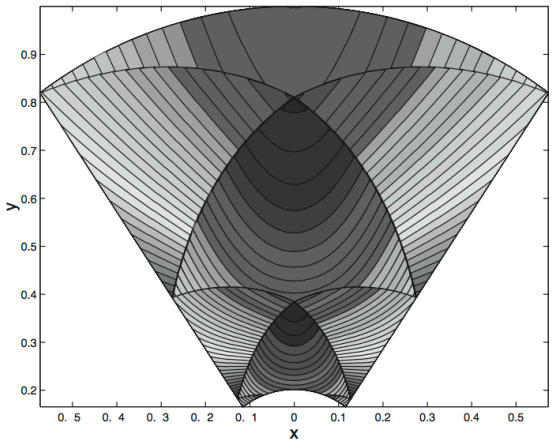


Figure 2-7: Numerical solution to MCP in a wedge hopper with non-radial stresses on the top boundary. Normal stress in the radial direction displayed.

the top surface exhibits a fan-like array of shocks [1]. The associated velocity field (not shown) will at best exhibit a similar pattern of discontinuities and at worse add even more discontinuities. Such a broken velocity field is clearly unphysical. As the grain size becomes very small (sands), discontinuous velocity fields with no relationship to the stress field have been observed, but these are only temporary; the shocks commonly blur away immediately after the onset of flow, which has been attributed to some instability mechanism [39]. Literature on the topic [101] is quick to concede that infinitesimally sharp velocity jumps are physically nonsensical and should be understood as being spread over at least a few particle widths. Below, we will see that our model naturally provides a mechanism for the blurring of velocity shocks even in the presence of a stress shock, with large velocity variations occurring only at the scale of several particle diameters.

Typically, to avoid the task of having to track/capture shocks in the stress/velocity field, approximations to MCP are invoked which give continuous solutions either by altering the boundary conditions or simplifying the PDE's. Smooth stress approximations are especially useful when attempting to solve for the velocity field— tracking flow shocks coming from both a discontinuous stress field *and* hyperbolicity in the

[1]Figure reprinted from J. Comput. Phys., **166**, P. A. Gremaud and J. V. Matthews, On the Computation of Steady Hopper Flows, 63 - 83, Copyright (2001), with permission from Elsevier.

velocity equations is an enormous job. To our knowledge, a full solution to MCP has never been obtained either numerically or analytically in cases where the underlying stress field has shocks. Instead, shock-free approximate solutions have mostly been pursued.

Arguably, the two most successful and commonly used results of MCP are actually approximations, not full solutions. The Jenike Radial Solution [66, 67] for wedge hopper flow solves the MCP equations exactly, and with no discontinuities, but does not allow for a traction free top surface. It is a similarity solution of the form

$$p = r f(\theta) \tag{2.19}$$

$$\psi = g(\theta) \tag{2.20}$$

$$\mathbf{v} = -\frac{h(\theta)}{r}\hat{r} \tag{2.21}$$

which reduces the entire system to 3 ODE's with (r, θ) the position; r is the distance from the hopper apex and θ is measured anticlockwise from the vertical. Though this solution enables the material to obey a wall yield criterion along the hopper walls, the stress parameters at the top surface have very little freedom. This is why most claim the Radial Solution to only hold near the orifice, considerably away from the actual top surface.

Another commonly used simplification is called the Method of Differential Slices, although it only applies to stresses and not flow (our focus here). Originally proposed by Janssen in 1895 and significantly enhanced since then, it is used to determine wall stresses in bins and containers. The method makes some very far-reaching assumptions about the internal stresses: p is presumed to only depend on height and the ψ field is assumed to be identically $\pi/2$ or 0. These assumptions reduce the stress balance equations to one ODE and ultimately give the famous result that wall stresses increase up to a certain depth and then saturate to a constant value. (This saturation behavior is not a byproduct of the approximation; the discontinuous, full solution to the stress balance equations in a bin also gives similar stress saturation behavior.) While this effect has been verified extensively in experiment, the underlying assump-

tions clearly cannot hold since, for example, the walls exert an upward shear stress on the material which contradicts the assumption about ψ. [101].

In summary, the equations of MCP theory have very limited applicability to granular flows. There are very few, if any, solutions available (either numerical or analytical) for many important geometries such as planar or annular Couette cells, vertical chutes, inclined places, etc. In the case of silos, MCP has been used extensively to describe stresses, although the equations are difficult to solve and poorly behaved from a mathematical point of view, as noted above. There have also been some attempts to use MCP to describe granular drainage from silos, in conjunction with the coaxial flow rule, but this approach has met with little success, as we now elaborate.

2.3.2 Physical difficulties with limit-state coaxiality

It is instructive to review the existing picture of silo drainage in MCP theory, to highlight what we will view as a major concern in the use of coaxiality for granular flows. Suppose we have a flat-bottomed quasi-2D silo with smooth side-walls. Under standard filling procedures, the walls provide only enough pressure to keep particles from sliding farther out. These wall conditions, known as the "active case", give the following solution to the stress balance equations as found by marching down characteristics starting from the flat, pressure free, top surface:

$$\psi(x,y) = \pi/2 \tag{2.22}$$

$$p(x,y) = \frac{f_g y}{1+\sin\phi} \tag{2.23}$$

where $f_g \equiv \rho g$ is the weight density of the material and y is positive downward. Since the ψ field is identically $\pi/2$ everywhere, the slip-lines are thus perfectly straight lines angled at $\pm\epsilon$ from the vertical.

Refer again to Figure 2-5. The material deforms based solely on principal plane alignment. For a slow, dense flow in the silo geometry, coaxiality is troubling. Since the major principal stress is everywhere vertical, coaxiality requires material to stretch horizontally, thus making it geometrically impossible for it to converge and exit

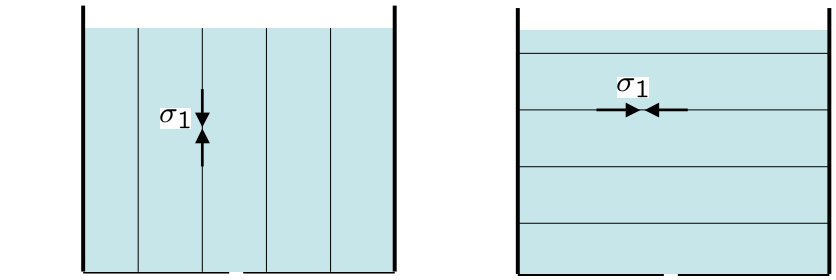


Figure 2-8: Major principal stress chains in a quasi-2D silo for the (a) active case, and (b) passive case.

through the silo orifice. This apparent paradox has traditionally been handled by asserting a sudden change in wall stresses that occurs once the orifice opens, such that the walls drive the flow, not gravity [101]. The silo is claimed to enter a "passive state" where the walls are squeezing the material through the orifice. Even with this questionable assumption, the solution predicted by equation (2.12) is unrealistic; it predicts the only non-stagnant regions in the silo are two narrow, straight channels which converge on the silo opening and are angled at $\pm 45^o$ from the vertical.

Coaxiality can also violate principles of thermodynamics. The equation itself only ensures there is no shear strain-rate in the principal stress reference frame and actually does not directly enforce that of the two principal strain-rate axes, the axis of maximal compression (i.e. the major principal strain-rate direction) must align with the major principal stress direction, as was the physical intuition shown in Figure 2-5. Coaxiality just as easily admits solutions for which the minor principal stresses align with the major principal strain-rate. When this happens, the plastic power dissipated per unit volume can be written

$$\begin{aligned} P &= \mathbf{T} : \dot{\mathbf{E}} \\ &= \begin{pmatrix} -\sigma_1 & 0 \\ 0 & -\sigma_3 \end{pmatrix} : \begin{pmatrix} |\dot{\gamma}| & 0 \\ 0 & -|\dot{\gamma}| \end{pmatrix} \\ &= |\dot{\gamma}|(\sigma_3 - \sigma_1) \\ &< 0 \end{aligned}$$

where $\mathbf{A} : \mathbf{B} = \Sigma A_{ij} B_{ij}$ and $\pm|\dot{\gamma}|$ are the principal strain-rate values. This type of behavior violates the second law of thermodynamics as it implies that material deformation *does* work on the system and likewise is non-dissipative. In more advanced plasticity theories, the thermodynamic inequality is upheld by requiring λ in equation (2.15) to be non-negative, but in the basic limit-state framework we discuss, this constraint cannot be directly enforced.

We should briefly mention that in constructing the limit-state stress field for the discharging silo, we have used as a boundary condition that flow ensues when the pressure p above the hole drops differentially from the value it takes when the hole is closed. This claim allows us to preserve the continuous stress field described in equations (2.22) and (2.23) for slow, quasi-static flow.

2.3.3 Incipient yield everywhere

The fundamental assumption of a limit-state stress field at incipient yield everywhere is also questionable. Granular flows can contain regions below the yield criterion within which the plastic strain-rate must drop to zero. For example in drainage from a wide silo with a small orifice, the lower regions far from the orifice are completely stagnant [123, 28], and thus can hardly be considered to be at incipient yield. In fact, discrete-element simulations show that grains in this region essentially do not move from their initial, static packing [120]. Simulations also reveal that high above the orifice, where the shear rate is reduced, the packing again becomes nearly rigid [121], suggesting that the yield criterion is not met there either. As the silo example illustrates, a more general description of stresses coming from an elasto-plasticity theory may be necessary to properly account for sub-yield material [60, 69].

Elasto-plasticity also alleviates another major concern with MCP which is that it is only well-defined in two dimensions. Three-dimensional stress tensors have six free variables, too many to be uniquely described by just force balance and incipient failure (altogether only 4 equations). We mention in passing that extensions of MCP to axisymmetric three-dimensional situations have been developed. For example, the Har Von Karman hypothesis, which assumes that the intermediate principal stress

σ_2, where $\sigma_1 \leq \sigma_2 \leq \sigma_3$, is identical to either σ_1 or σ_3, is frequently utilized in solving for conical hopper flow. However, elasto-plasticity can handle three dimensions without this hypothesis, while also allowing for stress states below the yield criterion in different regions.

2.3.4 Neglect of discreteness and randomness

Beyond its practical limitations and mathematical difficulties, the most basic shortcomings of MCP are in its assumptions. Above all, a granular material is not continuous. The microscopic grains composing it are usually visible to the naked eye, and significant variations in the velocity often occur across a distance of only several particle lengths, e.g. in shear bands and boundary layers. Of course, the general theory of deterministic continuum mechanics is only expected to apply accurately when the system can be broken into "Representative Volume Elements" (RVE's) of size L fulfilling $d \ll L \ll L_{macro}$ for d the microscale (particle size) and L_{macro} the size of the system [61], which is clearly violated in many granular flows. Therefore, the discrete, random nature of the particle packing must play an important role in the deformation process. To incorporate this notion coherently, it may be useful to seek out a dominant meso-scale object as a substitute for the RVE, upon which mechanical flow ideas apply, but in a non-deterministic, stochastic fashion (see Figure 2-9). This concept is somewhat comparable to the "Stochastic Volume Element" in the theory of plasticity of heterogeneous materials [104]. In that setting, it is known that (what physicists would call "quenched") randomness in material properties leads to blurring of the slip-lines, but, to our knowledge, this effect has not been considered in MCP for granular materials.

More importantly, however, since the meso-scale should only be a few grains in width, there must also be randomness in the *dynamics* of yielding to an applied stress or body force, since the theoretical concept of a continuous slip-line is incompatible with the reality of a discrete, random packing. A stochastic response also seems fundamentally more consistent with the assumption of incipient yield: If the material is just barely in equilibrium, it must be very sensitive to small, random fluctuations,

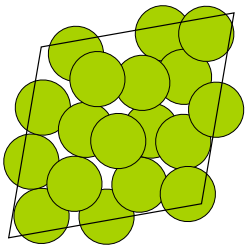


Figure 2-9: A meso-scale object containing a small number of randomly packed, discrete grains, which controls the dynamics of dense granular flow, analogous to the "representative volume element" in classical continuum mechanics.

causing localized yielding.

We conclude that the shortcomings of MCP may be related to the deterministic Coaxial Flow Rule, so we now proceed to replace it with a more physically appropriate Stochastic Flow Rule. The Mohr-Coulomb description of stresses is more clearly grounded in principles of solid mechanics and is widely used in silo engineering, so we will assume that it still holds, on average, in the presence of slow flows, as a first approximation.

2.4 The Stochastic Flow Rule

2.4.1 Diffusing "spots" of plastic deformation

It has been noted in a variety of experiments that dense granular flows can have velocity profiles which resemble solutions to a diffusion equation. By far the best example is drainage from a wide silo, which has a well-known Gaussian profile near the orifice, spreading vertically as the square root of the height (with parabolic streamlines) in a range of experiments [102, 140, 123, 92, 28]. Recently, experiments in the split-bottom Couette geometry have demonstrated precise error-function profiles of the velocity spreading upward from the shearing circle, albeit with more complicated scaling [47]. Shear bands, when they exist, tend to be exponentially localized near moving rough walls, but we note that these too can be viewed as solutions to a steady drift-diffusion equation with drift directed toward the wall.

It seems, therefore, that a successful flow rule for dense granular materials could be based on a stochastic process of deformation, consistent with our general arguments above based on discreteness and randomness. This begs the question: What is the diffusing carrier of plastic deformation? In crystals, plasticity is carried by dislocations, but it is not clear that any such defects might exist in an amorphous material. The Gaussian velocity profile of granular drainage was first explained independently by Litwiniszyn [85, 86] and Mullins [100] as being due to the diffusion of voids upward from the orifice, exchanging with particles to cause downward motion. However, this model cannot be taken literally, since granular flows have nearly uniform density with essentially no voids and with far less cage-breaking than the model would predict.

Instead, the starting point for our theory lies in the work of Bazant [15], who proposed a general model for the flow of amorphous materials (dense random packings) based on diffusing "spots" of cooperative relaxation, as illustrated in Figure 2-10. The basic idea is that each random spot displacement causes a small block-like displacement of particles in the opposite direction. This flow mechanism takes into account the tendency of each particle to move together with its cage of first neighbors, so we would expect the size of a spot to lie in the range three to five particle diameters, $L_s \approx 3 - 5d$, for simple cohesionless materials. This expectation has been confirmed in dense silo drainage as the length scale for spatial velocity correlations in the experiments of Choi et al. [29, 28] using glass beads (data shown in Figure 2-11) as well as the discrete-element simulations of Rycroft et al. [120] using a variety of force laws with monodisperse spheres. Although other material properties, such as grain shape and contact friction, could increase velocity correlations as suggested in Ref. [6], for purposes of illustration and comparison to a variety of experiments below, we will view $L_s = 3 - 5d$ as the typical range of spot sizes in this paper. In continuum mechanics terminology, we are proposing the spot as an appropriate meso-scale replacement for the RVE, which reflects velocity correlations resulting from cooperative displacements of the particles.

A major motivation for our work comes from the recent demonstration by Rycroft et al. that the Spot Model can be used as a basis for realistic multiscale simulations of

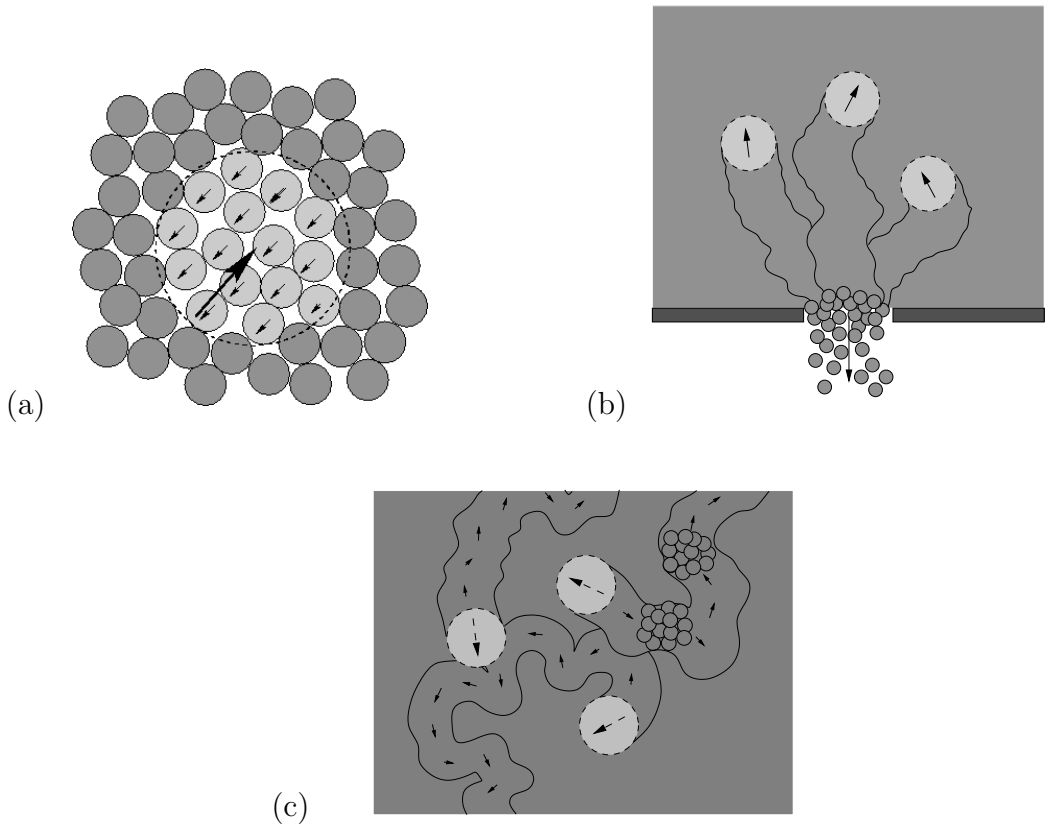


Figure 2-10: Spots as carriers of plastic deformation in amorphous materials. (a) Cartoon of basic spot motion. A spot of local fluidization, carrying some free volume, moves to the upper right causing a cooperative displacement of particles, on average to the lower left, opposing the spot displacement. (b) In silo drainage, spots are injected at the orifice and perform random walks biased upward by gravity, causing downward motion of particles. (c) In other situations, we conjecture that spots are created during initial shear dilation and perform random walks biased by local stress imbalances and body forces during steady flow.

dense granular drainage in a wide silo, assuming that spots perform upward random walks, biased uniformly by gravity [120]. Using only five fitting parameters (the size, volume, diffusivity, drift speed, and creation rate of spots), the spot simulations were able to accurately reproduce the statistical dynamics of several hundred thousand frictional, visco-elastic spheres in discrete-element simulations of drainage from a wide silo. This suggests that a general understanding of dense granular flows may come from a mechanical theory of spot dynamics.

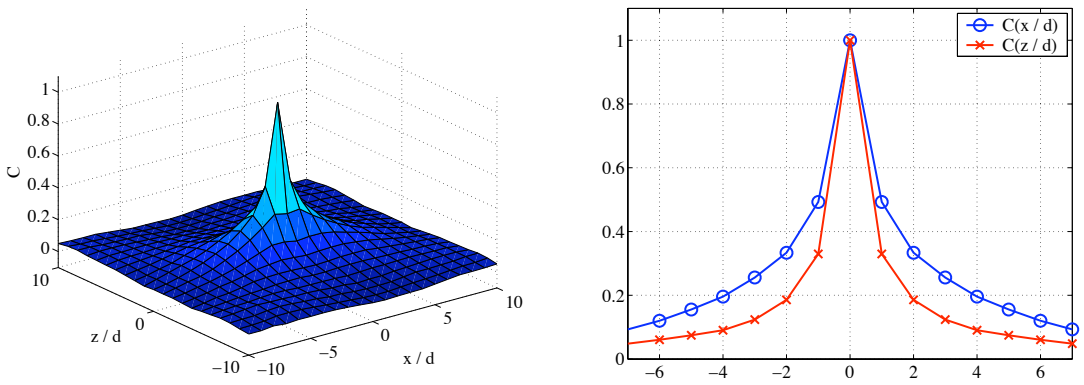


Figure 2-11: Spatial velocity correlations in silo drainage experiments as in Refs. [29, 28] with glass beads ($d = 3$mm) obtained high-speed digital-video particle tracking. The correlation coefficient of instantaneous displacements of a pair different particles is plotted as a function of their separation, averaged over all pairs and all times in the video. (Courtesy of the authors of Ref. [29].)

2.4.2 General form of the flow rule

In this work, we propose such a mechanical theory, based on the assumption that MCP provides a reasonable description of the mean stresses in slow dense granular flows. The key idea is to replace the Coaxial Flow Rule with a "Stochastic Flow Rule" based on mechanically biased spot diffusion. In the continuum approximation, the general form of the flow rule thus consists of two steps [15]: (i) a Fokker-Planck (drift-diffusion) equation is solved for the probability density (or concentration) of spots, $\rho_s(\mathbf{r}, t)$:

$$\frac{\partial \rho_s}{\partial t} + \frac{\partial}{\partial x^\alpha}\left(D_1^\alpha \rho_s\right) = \frac{\partial}{\partial x^\alpha}\frac{\partial}{\partial x^\beta}\left(D_2^{\alpha\beta}\rho_s\right) \tag{2.24}$$

where $\{D_1^\alpha\}$ is the drift vector and $\{D_2^{\alpha\beta}\}$ the diffusivity tensor of spots, determined by the stress field (below); and (ii) the mean drift velocity of particles $\mathbf{u} = \{u^\alpha\}$ is constructed to oppose the local flux of spots:

$$\begin{aligned} u^\alpha \;=\; -\int d\mathbf{r}'\, w(\mathbf{r},\mathbf{r}') \,\Big[& D_1^\alpha(\mathbf{r}',t)\rho_s(\mathbf{r}',t) \\ & - \frac{\partial}{\partial x^\beta}\left(D_2^{\alpha\beta}(\mathbf{r}',t)\rho_s(\mathbf{r}',t)\right)\Big] \end{aligned} \tag{2.25}$$

where $w(\mathbf{r}, \mathbf{r}')$ is the (dimensionless) spot "influence function" specifying how much a particle at $\mathbf{r}$ moves in response to a spot displacement at $\mathbf{r}'$. Without making

the continuum approximation, the same physical picture can also be the basis for a multiscale model, alternating between macroscopic continuum stress computation and discrete spot-driven random-packing dynamics [120].

The mean flow profile (2.25) is derived from a nonlocal stochastic partial differential equation for spot-driven particle dynamics, in the approximation that spots do not interact with each other [15]. Here, we have assumed the centered Stratonovich definition of stochastic differentials [114], which means physically that the spot influence is centered on its displacement. In contrast, Bazant used the one-sided reverse-Îto definition, where the spot influence is centered on the end of its displacement, which leads to an extra factor of two in the last term [15]. This choice is mathematically unrestricted (the "stochastic dilemma" [114]), but we find the centered definition to be a somewhat more reasonable physical hypothesis. Rycroft et al. have also found that the centered definition produces more realistic flowing random packings in multiscale spot simulations of granular drainage, when compared to full discrete-element simulations [120]. If we use the simple approximation $w \approx \delta(|\mathbf{r} - \mathbf{r}'|)$ in the integral for particle velocity, the Stratonovich interpretation has the benefit of automatically upholding volume conservation.

Without even specifying how local stresses determine spot dynamics, the general form of the flow rule (2.25) predicts continuous velocity fields, even when the mean stresses are discontinuous. For example, shocks in the MCP stress field may lead to discontinuities in the spot drift vector, $\mathbf{D}_1$, and thus the spot flux. However, due to the nonlocal nature of the spot model, the particle flux is a convolution of the spot flux with the spot influence function, thus preserving a continuous velocity field, which varies at scales larger than the spot size, L_s. This is a direct consequence of the geometry of dense random packings: The strong tendency for particles to move with their nearest neighbors smears velocity changes over at least one correlation length.

In the simplest approximation, the spot influence is translationally invariant, $w = w(\mathbf{r} - \mathbf{r}')$, so that spots everywhere in the system have the same size and shape. The spot influence decays off for $r > L_s$, as a Gaussian among other possibilities.

In (2.25) we allow for the likelihood that the spot influence may not be transla-

tionally and rotationally invariant, e.g. since the local stress state always breaks symmetry. This is actually clear in the experimental measurements of Figure 2-11, where velocity correlations are more short-ranged, without roughly half the decay length in the vertical direction (parallel to gravity) compared to the horizontal (transverse) direction. This suggests that spots are generally non-spherical and may be more elongated in the directions transverse to their drift (or the body force). If anisotropy in the spot influence were taken into account, it would also be natural to allow for an anisotropic spot diffusivity tensor, which depends on the local stress state. However, such effects seem to be small in the granular flows we consider below, which are well described by a much simpler model.

Another likely possibility is that spots come in a range of shapes and sizes. There could be a statistical distribution of regions of local fluidization or plastic yield related to the local packing and stress state. It is thus more reasonable to think of the spot influence function $w(\mathbf{r}, \mathbf{r}')$ as averaging over this distribution, just as a spatial velocity correlation measurement averages over a large number of collective relaxations. One advantage of taking the continuum limit of a Fokker-Planck equation (2.24) in applying the SFR is that such details are buried in the coefficients, which could in principle be derived systematically from any microscopic statistical model, or simply viewed as a starting point for further physical hypotheses (as we do below).

Finally, we mention that there are also surely some nontrivial interactions between spots, which would make the SFR nonlinear and could lead to interesting new phenomena, such as spontaneous pattern formation. For example, spots may have a medium range attraction, since it is more difficult to propagate particle rearrangements and plastic deformation into less dense, less mobile regions; there could also be a short range repulsion if the spot density gets too high, since grains will collapse into overly dilated regions. Such effects may be responsible for intermittent density waves in draining hoppers with narrow orifices [14, 107], and perhaps even shear banding in other amorphous materials, such as metallic glasses, with a different plastic yield criterion. However, we will see that the hypothesis of non-interacting spots already works rather well in cases of steady, dense granular flows.

2.4.3 A simple model for steady flows

Due to efficient dissipation by friction, granular materials subjected to steady forcing typically relax very quickly to a steady flowing state. For example, when a silo's orifice is opened, a wave of reduced density (spots) propagates upward, leaving in its wake a nearly steady particle flow, which we associate with a steady flow of spots. This initial density wave can be seen very clearly in discrete-element simulations of various hopper-silo geometries [121]. For a narrow orifice, we have noted that intermittent flows with density waves can be observed [14, 107], but typical drainage flows are rather steady in time [29, 28]. Similarly, when a rough inner cylinder is set into uniform rotation in a Couette cell, shear dilation propagates outward, raising the level of the packing, until a steady flow profile is reached. We interpret this initial dilation as signaling the creation of spots on the cylinder, which quickly reach a steady distribution in the bulk.

Hereafter, we focus on describing steady flow profiles, with equilibrium spot densities. For simplicity, let us assume isotropic spot diffusion, $D^{\alpha\beta} = D_2\delta_{\alpha\beta}$, since fluctuations are dominated by the (largely isotropic) geometry of dense random packings. Using the spot size L_s as the natural length scale, we can express the spot drift speed, $|\mathbf{D}_1| = L_s/\Delta t_1$, and diffusivity, $D_2 = L_s^2/2\Delta t_2$, in terms of the times, Δt_1 and Δt_2, for drift and diffusion to reach this length.

The flow profile of a draining silo, normalized by the outflow speed, is approximately constant over a wide range of flow speeds, as has recently been verified to great precision in the experiments of Choi et al. [29]. Not only is the mean velocity profile independent of flow rate (over an order of magnitude in mean velocity), but fluctuations about the mean, such as vertical and horizontal diffusion and measures of "cage breaking", also depend only on the distance dropped, and not explicitly on time (or some measure of "granular temperature"). In statistical terms, changing the flow rate is like watching the same movie at a different speed, so that the random packing goes through a similar sequence of geometrical configurations regardless of the velocity. Similar features have also been observed in shearing experiments in

Couette cells [99] and numerical simulations of planar shear [87].

The experimental and simulational evidence, therefore, prompts the crucial approximation that $\Delta t_1/\Delta t_2 =$ constant so as to uphold statistical invariance of the particle trajectories under changing the overall flow rate. This can be justified if spots perform random walks with displacements selected from a fixed distribution, set by the geometry of the random packing [15]. Here, we make the stronger assumption that the characteristic length of these random walks is the spot size L_s, so that $\Delta t_1 = \Delta t_2 \equiv \Delta t$. Our physical picture is that a spot represents a "cell" of localized fluidization (or plastic yield) of typical size L_s, which triggers further fluidization ahead of it and randomly propagates to a neighboring cell of similar extent. This picture is also consistent with the interpretation of L_s as a velocity correlation length above.

With these hypotheses, the Fokker-Planck equation (2.24) takes the simple time-independent form,

$$\nabla \cdot (\hat{\mathbf{d}}_s \, \rho_s) = \frac{L_s}{2} \nabla^2 \rho_s \tag{2.26}$$

where $\hat{\mathbf{d}}_s(\mathbf{r}) = \mathbf{D}_1/|\mathbf{D}_1|$ is the spot drift direction, determined by the mechanics of plastic yielding (below). The flow field is then

$$\mathbf{u} = -\frac{L_s}{\Delta t} \int d\mathbf{r}' \, w(\mathbf{r}, \mathbf{r}') \left(\hat{\mathbf{d}}_s(\mathbf{r}') \rho_s(\mathbf{r}') - \frac{L_s}{2} \nabla \rho_s(\mathbf{r}') \right) \tag{2.27}$$

Equations (2.26) and (2.27) define a simplified Stochastic Flow Rule, with only one parameter, L_s, which need not be fitted to any flow profile. Instead, it can be measured independently as the velocity correlation length, which may be viewed as a dynamical material property.

2.4.4 A mechanical theory of spot drift

The main contribution of this paper is a simple theory connecting the spot drift direction to the stresses in MCP. The basic idea is to view the displacement of a spot as being due to a local event of material failure or fluidization. To make a

quantitative prediction, we first define a cell of the material as the roughly diamond shaped region encompassed by two intersecting pairs of slip-lines, separated by L_s. When a spot passes through this cell, it fluidizes the material and thus locally changes the friction coefficient from the static value μ to the kinetic value μ_k. This upsets the force balance on the cell and may cause a perturbative net force to occur.

The force diagram for a material cell occupied by a spot can be broken into the sum of two diagrams (Figure 2-12), one which is the static diagram multiplied by μ_k/μ and one which contains only normal contact force contributions and a body force term. MCP requires the static diagram to be balanced, thus the latter is the only cause for a net force. A well-known corollary of the divergence theorem enables us to express the surface integral of normal contact stresses in terms of a gradient of p giving

$$\mathbf{F}_{net} = \left(1 - \frac{\mu_k}{\mu}\right)\left(\mathbf{F}_{body} - \cos^2\phi\nabla p\right) \tag{2.28}$$

as an effective force which pulls on a cell as it is fluidized by a passing spot, causing the spot to preferentially drift in the opposite direction.

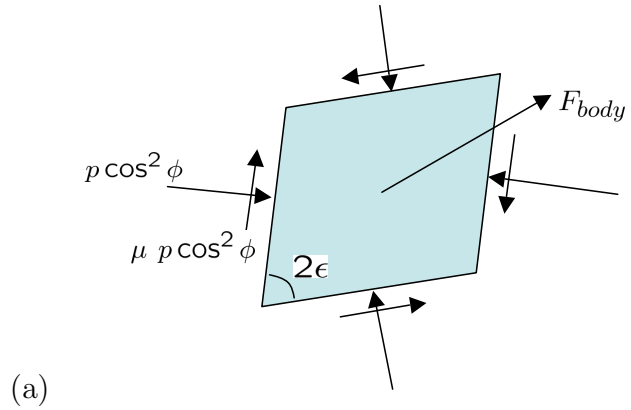
A spot cannot move in an arbitrary direction, however, since the material is at incipient yield only along the two slip lines. Therefore, the net force is constrained to pull the material cell along one of the slip-lines. The spot drift direction is then obtained by projecting (minus) the force, $-\mathbf{F}_{net}$, onto the slip-lines and averaging these two projection vectors with equal weight:

$$\xi^{(\pm)} = -(\mathbf{F}_{net}\cdot\hat{n}_{\psi\pm\epsilon})\hat{n}_{\psi\pm\epsilon} \tag{2.29}$$

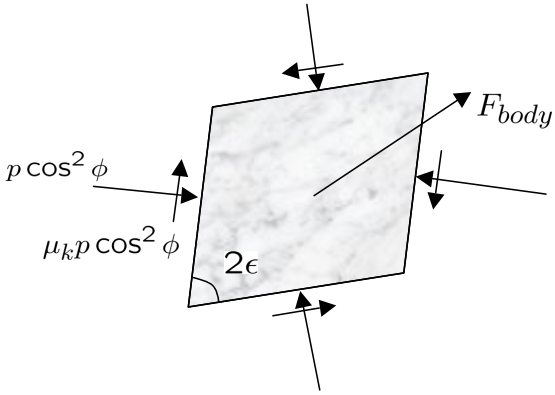
$$\hat{\mathbf{d}}_s = \frac{\xi^{(+)} + \xi^{(-)}}{|\xi^{(+)} + \xi^{(-)}|} \tag{2.30}$$

where $\hat{n}_\theta = (\cos\theta, \sin\theta)$. With a formula for $\hat{\mathbf{d}}_s$ now determined, the SFR as stated in equations (2.26) and (2.27) is now fully defined and ready for use.

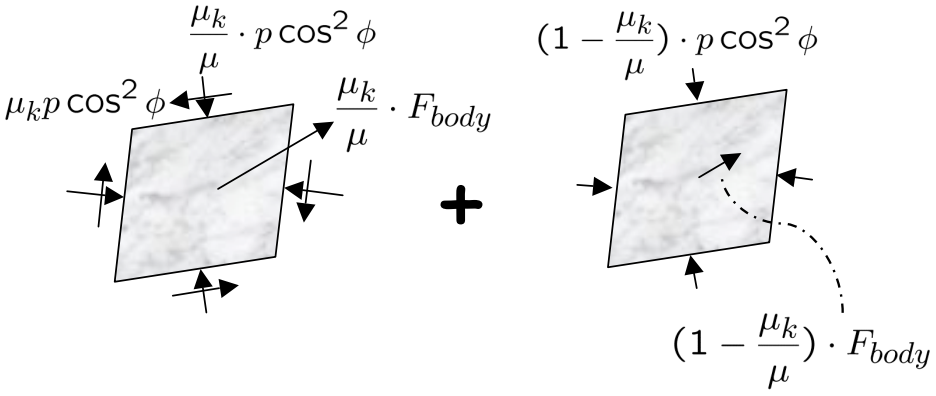
This continuum mechanical theory of spot drift also helps us understand the sources of spot diffusion. As noted above and sketched in Figure 2-9, a material cell is a small fragment of a random packing, which is unlikely to be able to accom-



(a)



(b)



(c)

Figure 2-12: (a) Material cell in static equilibrium.(b) A spot enters the cell fluidizing the material. In the force diagram, this means μ decreases to μ_k. (c) The force diagram for the fluidized material cell is best analyzed by breaking it into the sum of two diagrams.

modate shear strain precisely along the slip-lines of the *mean* continuum stress field. Instead, the instantaneous slip-lines are effectively blurred by the discrete random packing. Still, we preserve the picture of spots moving along slip-lines in constructing $\hat{\mathbf{d}}_s$, but represent the additional blurriness in the slip-line field by enforcing isotropic spot diffusivity.

2.4.5 Frame indifference

Finally, we must check that our flow rule satisfies frame indifference; solving for flow in different rigidly moving reference frames cannot give different answers in a fixed reference frame. Since the SFR is a 2D steady-state flow rule, the only flows we need to check for indifference are those with rotational/translational symmetry. In these cases, the particle velocity is a function of only one spatial variable and equation (2.26) for ρ_s becomes a second-order ODE. In solving the boundary value problem, we must ensure that grains along the walls have a velocity vector tangent to the walls. This constrains one of the two degrees of freedom in the set of possible solutions for ρ_s. Since (2.26) is homogeneous, the other degree of freedom must come out as a multiplicative undetermined constant. Thus the velocity profile is unique up to a multiplicative constant.

With only one constant, we cannot match boundary conditions for particle speed along more than one wall in general. So to solve for a flow between two walls, we must add rigid-body motions to the reference frame of the observer until we have the unique frame for which a solution exists matching both boundary conditions. This is an unexpected and welcome bonus of the SFR. Most flow rules in continuum mechanics enforce material frame indifference directly, i.e. the flow rule itself is derived to be automatically satisfied by any rigid-body motion, ensuring the same results independent of reference frame. Coaxiality achieves this by relating stress information only to strain-rate variables for instance. The SFR, however, upholds frame indifference indirectly in that the solution does not exist *unless* the problem is solved in exactly one "correct" frame of reference.

We have thus integrated the spot concept with the theory of plastic stresses treat-

ing spots as the "carriers of plasticity". We note that up to our granular-specific determination of the drift direction, the SFR principle can be applied to any amorphous isotropic material with a small characteristic length scale (dominant randomness) and a yield criterion.

2.5 Applications to granular flow

The Stochastic Flow Rule is quite general and in principle can be applied to any limit-state plasticity model of stresses, with different choices of the yield function to describe different materials. In this section, we apply the simplest SFR (2.26)–(2.27) to granular materials with MCP stresses and compare its predictions to a wide range of existing experimental data for steady dense flows. In calculating stresses, we assume a typical friction angle of $\phi = 30^o$. It is known that for spherical grains, the friction angle usually lies in a somewhat narrow range of about $20^o - 30^o$ and can be as large as 50^o for some anisotropic, highly angular materials [101]. In the examples we consider, however, varying the ϕ value in this range has very little macroscopic effect in our model.

The spot size L_s has a much larger effect, so we focus on its role in a variety of dense flows. We emphasize that we do not fit L_s to any flow profile below. Instead, we simply use the range $L_s = 3 - 5d$ for dense flowing sphere packings inferred independently from spatial velocity correlations in silo drainage experiments [29] and simulations [120] (see Fig. 2-11). This is consistent with our view of the correlation length, L_s, as a fundamental geometrical property of a flowing granular material.

Without any fitting parameters, we will apply the simple SFR to several prototypical flows. Each has different forcing and symmetries and, to our knowledge, they cannot be simultaneously described by any existing model. Our first example is granular drainage to a small orifice in a wide flat-bottomed silo, driven entirely by gravity. Our second example is shear flow in an annular Couette cell driven by a moving rough inner cylinder, where gravity plays no role. Our third example is the dragging of a loaded plate over a semi-infinite material at rest, which combines

gravitational forces and boundary forcing. Lastly we apply the SFR to a canonical free-surface flow, the continuous avalanching of a granular heap. The transition to a rapidly flowing surface shear layer on a heap will also lead us into a discussion of how rate-dependent effects, such as Bagnold rheology, may naturally extend into our model.

Throughout our treatment of the various examples, the first step will be to solve (2.26) to obtain the "unconvolved" velocity field

$$\mathbf{u}^* = -L_s \hat{\mathbf{d}}_s \rho_s + \frac{L_s^2}{2} \nabla \rho_s \tag{2.31}$$

which corresponds to the SFR velocity (2.27) for a point-like influence function $w = \delta(|\mathbf{r} - \mathbf{r}'|)$. For the most part, $\mathbf{u}^*$ is the "skeleton" for the full solution $\mathbf{u}$ because convolving $\mathbf{u}^*$ with a general spot influence merely blurs out the sharper features of $\mathbf{u}^*$. In some situations with wide shear zones, such as silo flow, the convolution has only a minor effect, but in others with narrow shear bands, at the scale of the spot size, the convolution step is essential for self-consistency and accuracy.

2.5.1 Silos

The flow profile in a flat-bottom silo geometry is well-known for its striking similarity to the fundamental solution of the diffusion equation. As noted above, early models of silo flow explained this based on the upward diffusion of voids from the orifice [85, 86, 100]. Without reference to a specific microscopic mechanism, Nedderman and Tüzun later derived the same equations based on a continuum constitutive law [102, 101]. They asserted that the horizontal velocity component u is proportional to the horizontal gradient of the downward velocity component v,

$$u = b \frac{\partial v}{\partial x} \tag{2.32}$$

since particles should drift from regions of slow, dense flow toward regions of faster, less dense (more dilated) flow. Assuming small density fluctuations, mass conserva-

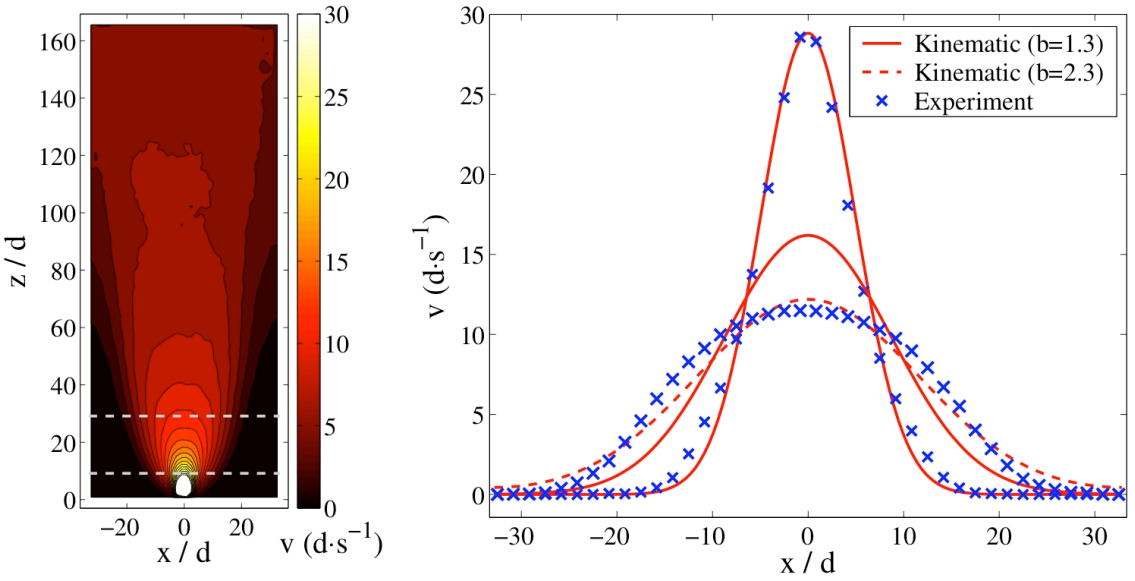


Figure 2-13: The mean velocity profile in a wide quasi-2d silo of 3mm glass beads from Ref. [28]. Horizontal slices of the downward velocity component near the orifice, indicated in the complete flow profile on the left, are shown on the right, and compared to the predictions of the Kinematic Model with two choices of the parameter b. The Stochastic Flow Rule for MCP for a wide silo (without side walls) gives a similar velocity field with $b \approx 1.5 - 2.5d$.

tion applied to the 2D velocity field, $\mathbf{u} = (u, -v)$ then yields the diffusion equation for the downward velocity,

$$\frac{\partial v}{\partial z} = b \frac{\partial^2 v}{\partial x^2} \tag{2.33}$$

where the vertical direction z acts like "time". The diffusivity b is thus really a "diffusion length", to be determined empirically. An advantage of the continuum formulation is that it avoids the paradox (resolved by the Spot Model [15]) that the classical picture of void random walks requires $b \ll d$, while experiments invariably show $b > d$.

Solving the Kinematic Model in the wide flat-bottomed silo geometry with a point orifice gives the familiar Green function for the diffusion equation,

$$v(x, z) = \frac{e^{-x^2/4bz}}{\sqrt{4\pi bz}}. \tag{2.34}$$

This gives an excellent match to experimental data close to the orifice (e.g. see

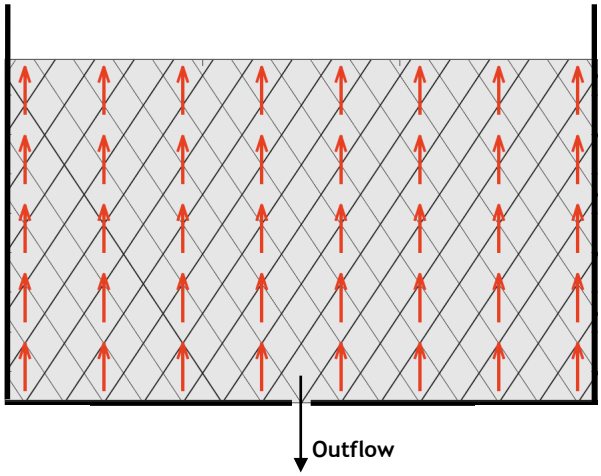


Figure 2-14: The flat-bottomed silo geometry. The intersecting black lines represent the slip-line field as determined from solving the stress balance equations of MCP, and the red vector field is the spot drift direction, as determined from the SFR. In this highly symmetric geometry, the spot drift precisely opposes the gravitational body force, $\hat{d} = -\hat{g}$.

Figure 2-13 [2]), although the fit gradually gets worse with increasing height, as the flow becomes somewhat more plug-like. Nevertheless, Gaussian fits of experimental data have provided similar estimates of $b = 1.3d$ [29], $1.3 - 2.3d$ [28], $2.3d$ [102], $3.5d$ [123], and $2d - 4d$ [92] for a variety of granular materials composed of monodisperse spheres.

We now apply our theory to this flow geometry and see how it connects to the Kinematic Model. Applying equation (2.30) using the stress field described by equations (2.22) and (2.23) gives uniform upward spot drift; $\mathbf{F}_{net}$ comes out as pointing uniformly downward and the slip-lines are symmetric about the vertical (see Figure 2-14). The SFR (2.26) then reduces to

$$\frac{\partial \rho_s}{\partial z} = \frac{L_s}{2}\left(\frac{\partial^2 \rho_s}{\partial x^2} + \frac{\partial^2 \rho_s}{\partial z^2}\right) \tag{2.35}$$

although we emphasize that this form applies only when the walls are smooth or

[2]Figure reprinted from J. Phys.: Condensed Matter, **17**, J. Choi, A. Kudrolli, and M. Z. Bazant, Velocity profile of granular flows inside silos and hoppers, S2533 - S2548, Copyright (2005), with permission from IOP.

equivalently when the silo width is large. The last term, which represents vertical diffusion of spots (relative to the mean upward drift), makes this equation for the spot density differ somewhat from the simple diffusion equation for the downward velocity of the Kinematic Model. Consistent with our model, vertical diffusion, with a similar (but not identical) diffusion length as horizontal diffusion, has been observed in recent silo-drainage experiments [29, 28].

The general solution of (2.35) can be expressed as a Fourier integral,

$$\rho_s = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} A(k) e^{\frac{z}{L_s}(1-\sqrt{1+L_s^2 k^2})} dk \tag{2.36}$$

where $A(k)$ is the Fourier transform of the spot density at the bottom ($z = 0$). The narrowest possible orifice allowing for flow is the case of a point source of spots, $\rho_s(x, 0) \propto \delta(x)$, $A(k) \propto 1$ (which is also the Green function). Convolution with a spot influence function of width L_s produces a downward velocity profile on the orifice of width L_s. Unlike the Kinematic Model (or any other continuum model which does not account for the finite grain size), our theory thus predicts that flow cannot occur unless the orifice is at least as wide as one spot, $L_s = 3 - 5d$.

The details of flow very close to the orifice, $z = O(L_s)$, are controlled by the choice of boundary condition, reflecting the dynamics of dilation, contact-breaking, and acceleration at the orifice, which are not described by our bulk dense-flow model. Rather than speculate on the form of this boundary condition, we focus on the bulk region slightly farther from the orifice. For $z \gg L_s$ (and $L_s k \ll 1$), the vertical diffusion term becomes unimportant, and the Green function tends to a Gaussian

$$v(x, z) \sim \frac{e^{-x^2/2\sigma_v^2(z)}}{\sqrt{2\pi\sigma_v^2(z)}} \tag{2.37}$$

where the variance is

$$\sigma_v^2(z) \sim L_s z + O(L_s^2). \tag{2.38}$$

(The second term is an offset from convolution with the spot influence function, which also depends on the choice of boundary conditions.)

There has been no prior theoretical prediction of the kinematic parameter b, which we interpret as the spot diffusion length [15]. Comparing (2.34) and (2.37), we obtain $b = L_s/2 = 1.5 - 2.5d$ without any fitting, beyond the independent determination of L_s from velocity correlations. This prediction is in excellent agreement with the experimental measurements listed above. However, the model does not predict the apparent increase of b with height, as the flow becomes more plug like. This may be due to the breakdown of the assumption of incipient yield higher in the silo, where the shear is greatly reduced, and it may require modeling stresses more generally with elasto-plasticity.

In any case, we are not aware of any other model of silo flow with a plausible basis in mechanics. It is noteworthy that we assume *active* silo stresses (driven by gravity), as typically assumed in a quasi-static silo. As a result, we do not require a sudden switch to passive stresses (driven by the side walls) upon flow initiation, as in existing plasticity theories based on the Coaxial Flow Rule [101]. Our use of the standard MCP model for stresses in quasi-static silos also suggests that the SFR may predict reasonable dependences on the geometry of the silo or hopper, wall roughness, and other mechanical parameters. In contrast, the Kinematic Model fails to incorporate any mechanics, and, not surprisingly, fails to describe flows in different silo/hopper geometries in experiments [28]. Testing our model in the same way would be an interesting direction for future work, since it has essentially no adjustable parameters.

2.5.2 Couette cells

The key benefit of our model is versatility; we will now take exactly the same model, which is able to describe wide silo flows driven by gravity, and apply it to Taylor-Couette shear flows in annular cells driven by a moving boundary. The granular material is confined between vertical rough-walled concentric cylinders and set into motion by rotating the inner cylinder. The flow field has been characterized extensively in experiments and simulations, and several theories have been proposed for this particular geometry [88, 19, 80, 94]. For example, a good fit of experimental data for Couette cells can be obtained by postulating a density and temperature dependent

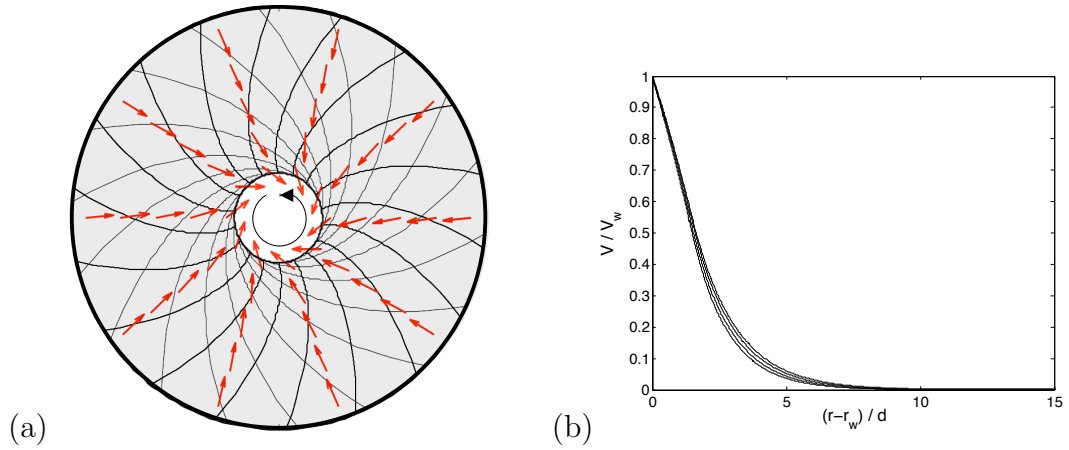


Figure 2-15: (a) A plane view of the annular Couette cell geometry, where the granular material is confined between concentric vertical cylinders. The rough wall is rotated anti-clockwise while the outer wall is held fixed. The crossing black lines within the material are the slip-lines as found from MCP, and the red vector field is the spot drift as determined by the SFR. (b) Normalized SFR velocity from as a function of distance from the inner wall with inner cylinder radius 15d, 25d, 50d, and 100d (from bottom to top curves, respectively). The friction angle is $\phi = 30°$, and the spot size is $L_s = 3d$.

viscosity in a fluid mechanical theory [19], but it is not clear that the same model can describe any other geometries, such as silos, hoppers, or other shear flows.

To solve for the MCP stresses in the annular Couette geometry, we first convert the stress balance equations to polar coordinates (r, θ) and require that p and ψ obey radial symmetry. This gives the following pair of ODEs:

$$\frac{\partial \psi^*}{\partial r} = -\frac{\sin 2\psi^*}{r(\cos 2\psi^* + \sin\phi)} \tag{2.39}$$

$$\frac{\partial \eta}{\partial r} = -\frac{2\sin\phi}{r(\cos 2\psi^* + \sin\phi)} \tag{2.40}$$

where $\eta = \log p$ and $\psi^* = \psi + \frac{\pi}{2} - \theta$. Although ψ^* has an implicit analytical solution, η does not, so we solve these equations numerically using fully rough inner wall boundary conditions $\psi^*(r_w) = \frac{\pi}{2} - \epsilon$ and any arbitrary value for $\eta(r_w)$. The resulting slip-lines are shown in Figure 2-15(a).

In the Couette geometry, the average normal stress, p, decreases with radial distance, which implies that the fluidization force on material, $\mathbf{F}_{net}$, is everywhere di-

rected outward. We then apply equation (2.30) to calculate the drift direction $\hat{\mathbf{d}}_s(r)$ by projecting the vector $\mathbf{F}_{net}$ onto slip-lines, and implement the SFR, exploiting symmetry which allows only a nonzero velocity in the $\hat{\theta}$ direction. This implies

$$\mathbf{u}^* \cdot \hat{r} = 0 = -L_s(\hat{\mathbf{d}}_s \cdot \hat{r})\rho_s + \frac{L_s^2}{2}\frac{\partial \rho_s}{\partial r} \tag{2.41}$$

which yields a solution for ρ_s up to a scalar factor. We then use ρ_s to calculate the θ component of the (unconvolved) velocity once again using the SFR equation,

$$\mathbf{u}^* \cdot \hat{\theta} = -L_s(\hat{\mathbf{d}}_s \cdot \hat{\theta})\rho_s. \tag{2.42}$$

It turns out, as we may have expected, that $\mathbf{u}^*$ has a shear band at the inner wall with nearly exact exponential decay. The length scale of this decay is the spot size, L_s, since this is the velocity correlation length, beyond which the inadmissible shear at the inner cylinder can be effectively dissipated by the material.

The thinness of the shear band requires that, for consistency, we must take into account the finite spot size in reconstructing the velocity field through the convolution integral (2.27). For simplicity we will use a uniform spot influence function, i.e.

$$w(\mathbf{r}) = \frac{4}{\pi L_s^2} H\left(\frac{L_s}{2} - |\mathbf{r}|\right) \tag{2.43}$$

where $H(x)$ is the Heaviside step function. To evaluate the integral (2.27), we also must make a hypothesis about how spots operate when they overlap one of the walls. Random packing dynamics near walls is different than in the bulk and sensitive to surface roughness, and further detailed analysis of experiments and simulations will be required to elucidate the collective mechanism(s). Here, the precise shape of spots near the wall has little effect, except to flatten out the spike in velocity that occurs near the wall in the unconvolved velocity $\mathbf{u}^*$. As a simple first approximation, used hereafter in this paper, we will view the space beyond each boundary as containing a bath of non-diffusive spots at uniform concentration whose flux is such that the particle velocity invoked "inside" the boundaries directly mimics the rigid motion of

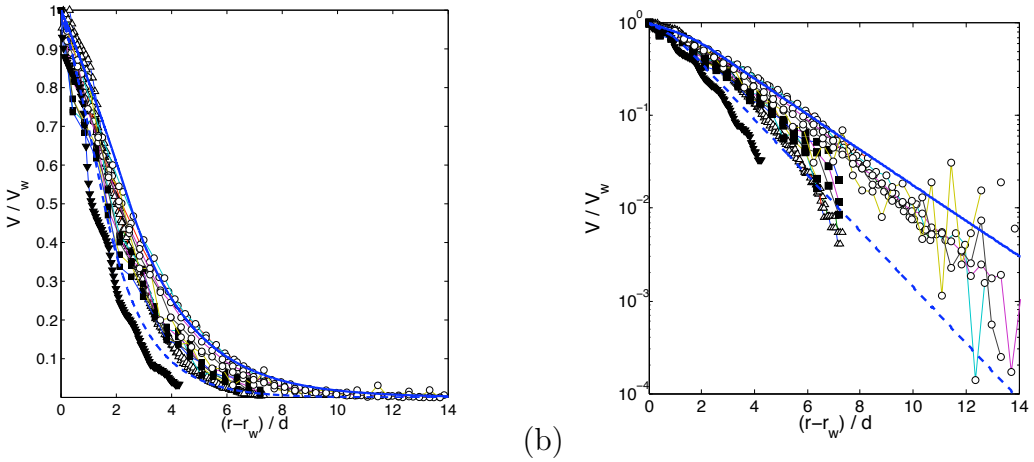


Figure 2-16: Theory versus experiment for the normalized velocity in annular Couette cells on (a) linear and (b) semilog plots. The dashed curve is the predicted SFR velocity field with $L_s = 3d$, while the solid line is for $L_s = 5d$; both curves are for an inner cylinder radius of $r_w = 80d$ and $\phi = 30°$. Experimental measurements (points) for a wide range of inner and outer cylinder radii are shown from the compilation of data shown in Figure 3c of [94]. (The experimental data is courtesy of GDR Midi and originates from the work of [32], [19], [98], and [27].)

the walls. This effectively "folds" part of the spot influence back into the granular material when it overlaps with the wall. The resulting velocity field is shown in Figure 2-15(b), where normalized velocity is shown versus distance from the inner cylinder wall for $L_s = 3d$ for a wide range of inner cylinder radii.

The predicted flow field – without any fitting – is in striking agreement with a large set of data from experimental and discrete-element simulations for different cylinder radii and grain sizes [88, 19, 80, 94]. As shown in Figure 2-16, the experimental data compiled by GDR Midi [94] falls almost entirely within the predicted SFR velocity profiles, by setting the spot size to the same typical range of correlation lengths, $L_s = 3 - 5d$, measured independently in a quasi-2D silo (Figure 2-11). Viewing the data on a semilog plot shows that the agreement extends all the way into the tail of the velocity distribution. We emphasize that the same simple theory, with the same range of L_s values, also accurately predicts silo flows above, as well as other situations below. Unifying all of this data in a single simple theory without any empirical fitting constitutes a stringent, quantitative test.

It is interesting to note the behavior close to the wall, especially in thin Couette cells. In experiments [94], annular flow profiles are known to have a Gaussian correction term when the thickness of the cell becomes non-negligible in terms of particle size. This slight flattening near the wall is apparent in our model as well and is a byproduct of convolving with the spot influence. We thus interpret this feature as another sign of the strongly correlated motion of particles, primarily with the "cage" of nearest neighbors, as approximately described by the spot mechanism. In this calculation, we used a uniform spot influence, but have noticed relatively little sensitivity of the predicted flow profile, for different strongly localized influence functions, such as a Gaussian, $w \propto e^{-2r^2/L_s^2}$. A detailed comparison of the model to experimental data may provide fundamental insights into the spot influence, and thus the collective dynamics of random packings, near a rough wall at the discrete particle level.

The experimental results shown in Figure 2-16 come from apparati with inner wall radii ranging from $14d - 100d$. The relatively small variations in the data sets over such a large range of inner radii clearly indicates that the inner wall radius is not a crucial length scale in the flow. The plotted theoretical prediction uses an inner radius of $\approx 80d$, but, as can be seen in Figure 2-15(b), our results depend only minimally on the inner cylinder radius. Indeed, the meso-scale correlation length of $L_s = 3 - 5d$ is the dominant length scale in our theory for this geometry.

To substantiate an earlier claim, we now consider how the friction angle ϕ affects the flow properties (holding L_s fixed) according to our model. We can see this most clearly by observing how the shear band half-width (i.e. the distance from the wall to the location where velocity is half-maximum) varies over the ϕ range for usual granular materials ($\approx 20^o - 50^o$). As shown in Figure 2-17 the half-width changes by $< 0.4d$ over the entire range and by $< 0.1d$ for the range of laboratory-style spherical grains. This very weak influence of internal friction agrees with simulations in the Couette geometry by Schöllmann [128].

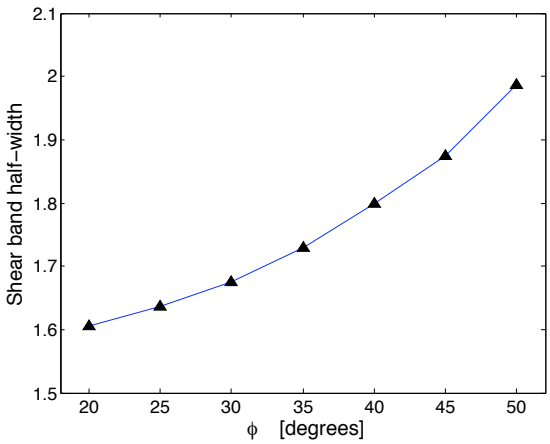


Figure 2-17: Predicted variation in the width of the shear band with SFR over the standard range of granular friction angles for the annular Couette geometry. ($L_s = 3d$, $r_w = 80d$)

2.5.3 Plate dragging

We now examine perhaps the simplest situation where gravity affects the shear band caused by a moving rough wall. Consider slowly dragging a rough plate horizontally across the upper surface of a deep (semi-infinite) granular material. The plate maintains full contact by pressing down on the surface with pressure $p_0 \cos^2 \phi$. The profile of the shear band that forms below the plate depends on the relative loading pressure, $q_0 = p_0/f_g$, where f_g is the weight (gravitational body force) density of the material.

The plate-dragging flow field can be found using a procedure analogous to the annular Couette cell, but enforcing horizontal instead of radial symmetry. With y measuring distance below the plate, the stress balance equations give

$$\begin{aligned} \psi_y &= \frac{-\sin 2\psi}{2q(\cos 2\psi - \sin \phi)} \\ q_y &= \frac{\cos 2\psi}{\cos 2\psi - \sin \phi} \end{aligned}$$

where $q(y) = p(y)/f_g$ is the average normal stress scaled to the weight density. The fluidization force will push material downward and spots upward resulting in a flow profile that decays close to exponentially near the moving wall.

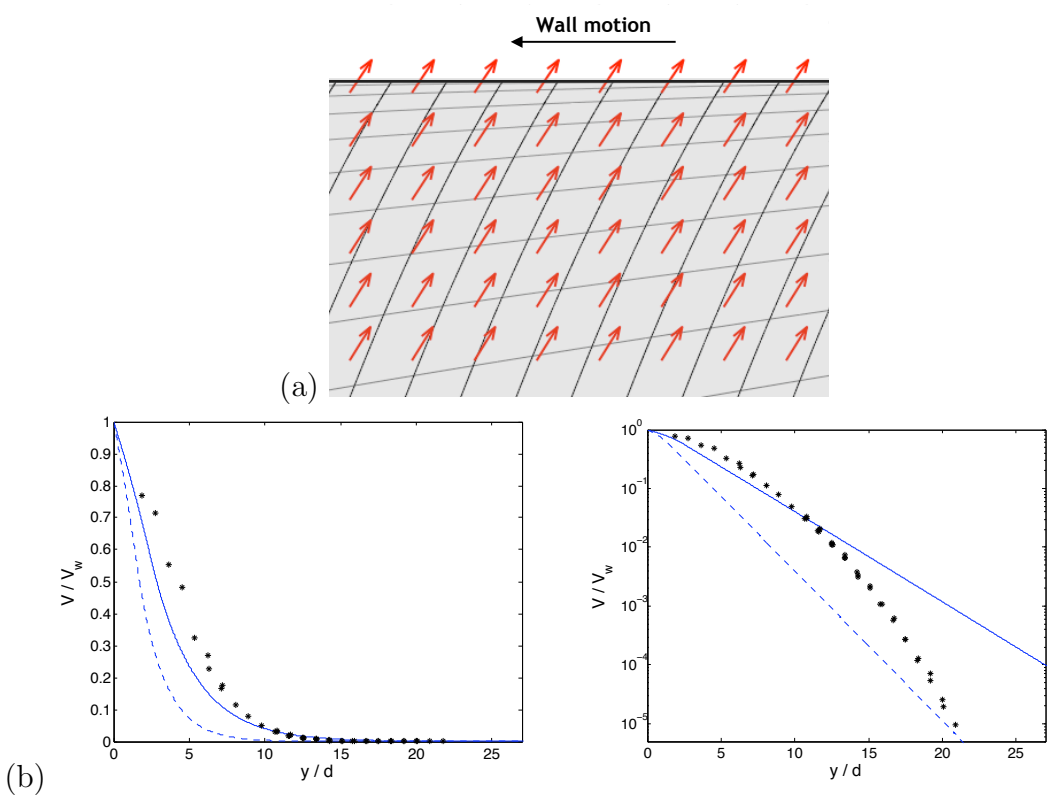


Figure 2-18: (a) The plate-dragging geometry. The top wall drags along the top of a bed of granular material. The crossing black lines within the material are the slip-lines as found from MCP, and the vector field is the spot drift as determined by the SFR. (b) Theory against experiment for the plate dragging geometry. Theory: L=3 (- - -), L=5 (—). Experiment (∗) courtesy of the authors of [137] .

Experiments [138, 129] and simulations [135, 141, 65] offer differing assessments of the details of the flow profile away from the shear band, but the dominant exponential decay behavior is clearly observed in all. The displayed SFR prediction (Figure 2-18) uses loading parameters from Tsai and Gollub [137] in order to appropriately compare with their results. Although the general properties of the flow appear to be represented well by the model, we do notice that the predicted range of typical flows is too small to fully encompass the experimental data. There could be a number of reasons for this discrepancy, but it is worth pointing out that the quasi-2D plate-dragging geometry is rather difficult to realize in experiments. For example, this experiment was executed by rotating a loaded washer-type object on top of an annular channel, and it was observed that the sidewalls pushing in the third dimension actually

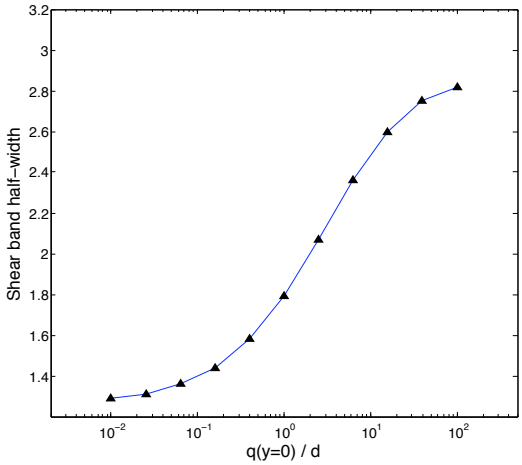


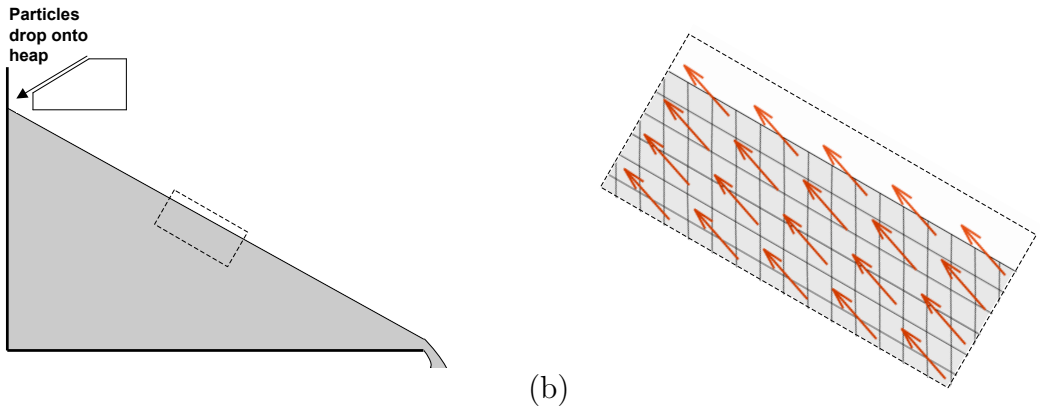
Figure 2-19: Plot of theoretical shear band half-width vs. relative loading pressure $q_0 = p(y = 0)/f_g$ of the top plate (where f_g is the weight density). The calculation assumes $L_s = 5d$

did play some role.

In Refs. [141] and [1], simulations of this environment indicate that the shear band width increases with increasing loading of the top plate. As can be seen in Figure 2-19, our theory captures this general trend of increasing loading causing increasing shear band width. However, the swing in band size predicted by our theory is not large enough to match the range of band sizes in simulations [141] and [1] in which the shear band half-width can be as large as several tens of particle diameters for large enough q_0 and diverges as $q_0 \to \infty$ (i.e. zero gravity). In cases such as these where the value of q_0 becomes very large, as we will discuss in more depth after the next section, we believe a new phenomenon begins to dominate our meso-scale argument and that this phenomenon may be attributed to a particular property of the slip-line field.

2.5.4 Slow heap flows

We now examine a prototypical free surface flow. Very close to the repose angle, a granular heap which is slowly but consistently re-fed grains undergoes a particular type of motion characterized by avalanching at the top surface and a slower "creep-



(a) (b)

Figure 2-20: (a) The heap flow setup. (b) The dashed rectangle in (a) is enlarged; the slip-line field from MCP is plotted along with the drift field from the SFR.

ing" motion beneath. This type of flow has been studied in experiments [84] and simulations [131]. Though heap flows with faster top shear layers have also been studied [76, 94] we will focus for now on the slower regime, which more closely resembles a quasi-static flow where the SFR might apply.

This kind of flow is stable, but indeed quite "delicate" in the sense that relatively small changes to the system parameters (i.e. flow rate, height of the flowing layer) can invoke large changes to the qualitative flow profile especially in the top layers [131]. We will describe and attempt to explain this effect more in the next section.

The heap geometry is depicted in Figure 2-20 along with the corresponding spot drift field and slip-line field. Any gravity driven free surface flow problem for which the stresses and flow are approximately invariant in the direction parallel to the surface will have limit-state stresses that obey the following relations:

$$\psi = -\epsilon \tag{2.44}$$

$$p = \frac{f_g y}{\cos\phi} \tag{2.45}$$

where y is the depth measured orthogonally from the free surface. Note that in limit-state theory, for self-consistency, the static angle of repose is identical to the internal friction angle ϕ, which is a reasonable assertion but still debated in the community. (By "static repose angle" we refer to the angle of inclination below which a flowing

system jams; in simulations of flow down a rough inclined plane, it has been shown that this angle does vary in a narrow range depending on the height of the flow [131, 130].) Applying equation (2.30) to these equations gives a simple expression for the spot drift vector:

$$\hat{\mathbf{d}}_s = \frac{(1+\sin^2\phi,\ -\sin\phi\cos\phi)}{|(1+\sin^2\phi,\ -\sin\phi\cos\phi)|} \tag{2.46}$$

We may then apply the SFR, which simplifies upon requiring that the flow run parallel to the free surface (i.e. $\mathbf{u}^* = (-u, 0)$).

Since the drift field is uniform, we obtain an analytical solution for the unconvolved velocity:

$$u = u(0)\exp\left(\frac{-y\sin 2\phi}{L_s}\sqrt{\frac{2}{5-3\cos 2\phi}}\right) \tag{2.47}$$

Thus our model predicts that the velocity decays exponentially off the free surface. In cases like these, where the boundary of the flow makes no contact with a rigid wall, it is less clear how the spots (and free volume) might behave near the flowing free surface. To avoid addressing this issue in detail, we neglect convolution with the spot influence function and simply assume $\mathbf{u} \approx \mathbf{u}^*$.

In their experiments on slow heap flow, Lemieux and Durian [84] have shown that the velocity profile in the flowing top layers is indeed well approximated by an exponential decay. Furthermore, they found the flow in this regime to be continuous and stable. The decay law they obtained is

$$u/u(0) \approx \exp(-\frac{y}{4.5d})$$

which is very close to our predicted solution for $L_s = 3d$:

$$u/u(0) = \exp(-\frac{y}{4.58d})$$

Silbert *et al.* [131] report finding a similar decay profile at low flow rates in simulations of flow down a rough inclined plane, although the avalanching at the surface was intermittent. Komatsu et al [76] have conducted faster heap flow experiments (discussed below in more detail) and found an exponentially decaying region of much

smaller decay length, $\approx 1.4d$, beneath the rapid flowing top layer (see Figure 2-23). We note that this flow is qualitatively different than the slower heap flows we attempt to describe here in which the decay behavior persists to the surface and is measured in the top layers. In conclusion, we have demonstrated a fourth, qualitatively different situation where the same simple MCP/SFR theory predicts the flow profile, without adjusting any parameters.

2.6 Transition from the SFR to Bagnold rheology

2.6.1 Breakdown of the SFR

In the last two examples, plate dragging and slow heap flow, there are limits where the SFR fails to predict the experimental flow profiles. In this section, we will explain why the breakdown of the SFR is to be expected in these cases and others, whenever slip-lines approach "admissibility" and coincide with shear planes. In this singular limit of the SFR, we postulate a transition to Bagnold rheology. The stochastic spot-based mechanism for plastic yielding is thus replaced by a different physical mechanism, the free sliding along shear planes.

For example, consider the case of plate dragging above. At large relative loading, the flow field resembles that of a zero-gravity horizontal shear cell (between shearing flat plates), and it appears that the SFR breaks down: With body forces and $\boldsymbol{\nabla} p$ both going to 0, equation (2.30) gives $\mathbf{F}_{net} = \mathbf{0}$ implying that spots have no drift and consequently the only SFR solution is $\mathbf{u} = \mathbf{0}$.

Problems also occur with flow down a rough inclined plane: Slightly increasing the flow rate (and consequently the flow height) or inclination angle causes the velocity vs. depth relationship to exit the exponential decay regime detailed above and undergo significant changes, passing first through a regime of linear dependence [5, 17] to a regime resembling a 3/2 power-law of depth [108, 130, 113, 11] opposite in concavity to the exponential decay regime.

Why does the velocity profile for inclined plane flow undergo many different qual-

itative regimes depending delicately on system parameters, while others (e.g. silos, annular cells) appear to be only weakly affected and almost always exhibit the same (normalized) velocity profile? Tall inclined plane flows and zero-gravity planar shear flows have been successfully described in multiple experiments and simulations [33, 87, 108, 113, 130] by the empirical scaling law of Bagnold [12]. In this section, we suggest a means to reconcile and perhaps eventually combine these theories into a coherent whole.

2.6.2 Bagnold rheology

Let us briefly review Bagnold's classical theory of granular shear flow. In its original form, "Bagnold scaling" expresses a particular rate-dependency for granular flow whenever the solid fraction is uniform throughout:

$$\tau \propto \dot{\gamma}^2. \tag{2.48}$$

where $\dot{\gamma}$ is the rate of simple shear. To account as well for static stresses arising from the internal friction, a related variant of this scaling law is commonly used [41]:

$$\tau - \mu\sigma \propto \dot{\gamma}^2. \tag{2.49}$$

It is in some sense a law for how the yield criterion can be exceeded when non-negligible strain-rates can absorb the extra shear stress. This constitutive law alone is an incomplete flow theory since it provides no way of predicting whether or not the solid fraction will be uniform during flow or how a non-uniform solid fraction affects the above rheology. Bagnold originally explained the quadratic relationship between stress and strain-rate in terms of binary collisions as the joint effect of both the particle collision rate and the momentum loss per collision being directly proportional to the strain-rate [12]. Despite this collision-based argument, however, Bagnold scaling has been observed to hold well into the dense regime, whenever the solid fraction is approximately constant throughout the system. This seemingly contradictory obser-

vation can be justified in the hard-sphere limit (without body forces) by a Newtonian invariance argument [87], although it calls into question the underlying physical mechanism.

Zero-gravity planar shear flow and thick inclined plane flow both exhibit nearly uniform density and thus have been employed as test cases for Bagnold scaling. In the planar-shear environment, the shear and normal stresses acting on the shear planes are spatially constant throughout the flow. Equation (3.4.1) therefore implies that the strain-rate is uniform; as a result, the velocity varies linearly from one wall to the other. This result is known as Uniform Shear Flow (USF) and is easily verified in simulations of Lees-Edwards boundary conditions. For example, the rheology (3.4.1) has been demonstrated in the simulations of [33].

Applying Bagnold scaling to the inclined plane geometry, slightly above static repose, gives a shear stress excess which grows linearly with depth and thus a shearing rate that grows as the square root of depth. This implies a velocity profile of the form

$$u \propto h^{3/2} - y^{3/2} \tag{2.50}$$

for y the depth variable and h the height of the flowing material (with no-slip bottom boundary condition). In this way, Bagnold scaling successfully explains the 3/2 power law dependence noted above.

2.6.3 Slip-line admissibility

The seemingly disparate flow mechanisms of the SFR and Bagnold rheology can be reconciled very naturally by considering the geometry of the slip-lines. In plasticity theory, all flows can be classified based on "slip-line admissibility". For admissible slip-lines, boundary conditions are such that the flow can, and presumably does, take place by continuous shearing along only one family of slip-lines. In mathematical terms, the slip-lines are admissible for a given flow, whenever the double-shearing continuum flow-rule (2.14) allows multiple solutions to the boundary value problem.

Slip-line admissibility is the exception, not the rule, since it is highly unlikely that

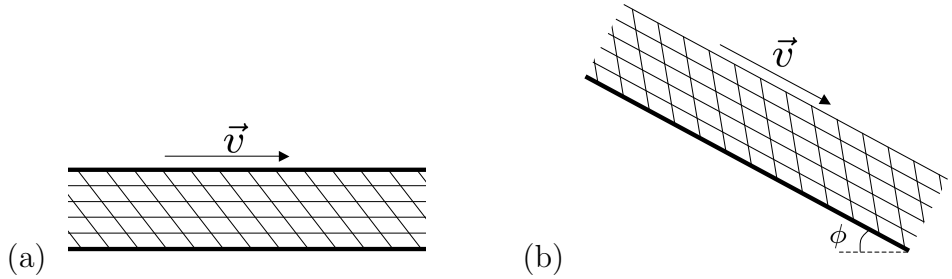


Figure 2-21: (a) Slip-line field for gravity-free planar shear flow, (b) Slip-line field for inclined plane flow. Note that in both cases the shear planes will be aligned completely with one slip-line family.

the prescribed velocity boundary conditions are fulfilled by a continuous shear on either slip-line family. Be that as it may, it so happens that planar shear flow and inclined plane flow are both slip-line admissible. This special property is shared by no other flow geometry studied in this paper, or, to our knowledge, elsewhere in the granular materials community. (Contrast the slip lines in Figure 2-21 with those in Figures 2-14 and 2-15.)

There is also an interesting difference in the density distributions. For admissible flows, the volume fraction is nearly uniform, and Bagnold rheology has a reasonable physical justification. For the more common case of inadmissible flows, as in silos and Couette cells, the volume fraction is typically highly nonuniform. In such cases, the SFR seems to provide an excellent description of the flow, and Bagnold rheology clearly does not apply.

These observations motivate us to think of admissibility as a criterion for two very different microscopic mechanisms for granular flow: In admissible flows, material motion is a viscous dragging of material "slabs" along one slip-line family (Bagnold dominated), whereas in inadmissible flows there is no clear choice as to which slip-line family should control the motion and thus material randomly chooses between both slip-line families (SFR dominated). Perhaps admissibility in the slip-line field *causes* the solid fraction to remain roughly uniform because the material in no sense has to collide head-on into neighboring material for it to move. In flows where non-uniform

dilation does occur, experiments have shown the motion is nearly independent of any local Bagnold rheology, encouraging our strong distinction between the dynamics of flow problems of differing admissibility status [134, 19].

These considerations all lead us to the fundamental conjecture:

> Slip-line admissibility is a geometrical and mechanical indicator as to the relative importance of rate-dependency (Bagnold rheology) over rate-independency (SFR) in a dense granular flow.

This means that a flow which has an admissible limit-state stress field will be dominated by rate-dependent effects when the yield criterion is only slightly exceeded.

2.6.4 Redistribution of excess shear stress

A more rigorous physical justification of our conjecture can be made utilizing limit-state stresses and observing the effect of pushing the system above yield. Bagnold rheology is a statement connecting the shear stress excess (i.e. amount by which τ exceeds $\mu\sigma$) along a shear plane to the rate of simple shear along the plane. We must emphasize that shear planes and slip-lines are *not* equivalent terms; slip-lines are defined by the quasi-static stresses as lines along which $\tau - \mu\sigma = 0$, whereas shear planes are defined entirely by the velocity profile. In an admissible system, the shear planes coincide with one slip-line family. In inadmissible systems, the shear planes almost everywhere do not coincide with slip-lines.

With admissible slip-lines, excess shear stress tends to be uniformly distributed throughout the system, resulting in global Bagnold rheology. For example, consider a zero-gravity planar shear cell. If we were to apply additional shear stress to the body in a manner aligned with the admissible slip-line family, e.g. by increasing the wall shear stress above yield by some amount $\Delta\tau$, that additional shear stress would distribute itself within the material precisely along the slip-lines. Every horizontal slip-line within the bulk would thus receive a boost in shear stress of size $\Delta\tau$. In limit-state theory the slip-lines have the highest possible $\tau - \mu\sigma$ value a quasi-static material element can take – zero. Adding $\Delta\tau$ additional shear stress to the slip-lines means

that τ will maximally exceed $\mu\sigma$ precisely along the slip-lines, and, by admissibility, every shearing plane. As a result, there will be a Bagnold contribution everywhere. Similarly, if we took a limit-state inclined plane geometry and increased the tilt angle some amount, an analogous boost in shear stress along the admissible slip-line family would occur causing τ to exceed $\mu\sigma$ precisely along all the shear planes.

In contrast, with inadmissible slip-lines, excess shear stress tends to remain localized where it is applied, and the SFR dominates the rest of the flow. For example, consider annular Couette flow. As can be seen in Figure 2-15, the slip-lines only coincide with the shear planes (which in this case are concentric circles) along the inner wall of the cell. If the inner wall were given an increase $\Delta\tau$ in applied shear stress, torque balancing requires the shear stress along any concentric circle within the material to receive a boost of $\Delta\tau \cdot r_w/r$. Suppose the shear cell has inner wall radius $40d$ and the boost in wall shear is significant, say $\Delta\tau = \tau_w/10$. Solving for $\tau - \mu\sigma$ along the shear planes in this situation gives a very different result than in the previous case— here τ will only exceed $\mu\sigma$ along the shear planes that are less than $1.4d$ off the inner wall. So, regardless of whether the density is or is not uniform, Bagnold scaling would at best only apply in an almost negligibly thin region near the wall. If the wall friction were less than fully rough, this region would further decrease.

2.6.5 A simple composite theory

The preceding discussion indicates that, in general, one can use the admissibility status of the system to choose whether or not the flow should obey the SFR or Bagnold rheology, or perhaps some combination of the two. Indeed, it seems reasonable that when slip-lines are *approaching* admissibility (e.g. plate-dragging with high q_0) or when an admissible system is only slightly pushed above yield (e.g. inclined plane flow near static repose) we must account for contributions from both effects simultaneously. It is beyond the scope of this paper to postulate the precise microscopic dynamics (and derive corresponding continuum equations) for this regime, but we can at least give a sense of how the more general theory might look.

In general, we envision a smooth transition from rate-independent SFR dynamics

to rate-dependent Bagnold dynamics controlled by the distribution of shear stress excess. This implies the coexistence of (at least) two different microscopic mechanisms: SFR and admissible shear. The SFR contribution would derive from the usual spot-based, quasi-static stochastic dynamics; the Bagnold contribution would come from a rate-dependent shearing motion along the appropriate slip-line family whenever there is a small excess shear stress (beyond the limit state) applied on a boundary which causes shear stress excess along the shear planes within.

The two mechanisms should have different statistical signatures. For shear deformation along admissible slip-lines, we would expect anisotropic velocity correlations. In the direction perpendicular to the shear plane, the correlation length should be somewhat shorter than the typical spot size, since slip-line admissibility allows flow to occur with less drastic local rearrangements, farther from jamming. In the directions parallel to the shear plane, however the correlation length could be longer, since material slabs sliding along shear planes may develop more rigid, planar regions. It would certainly be interesting to study velocity correlations in heap flows at different inclinations and plate-dragging experiments under different loads to shed more light on the microscopic mechanisms involved in the SFR to Bagnold transition.

For the remainder of this section, we make a first attempt at a composite model, simply a linear superposition of SFR and Bagnold velocity fields:

$$\mathbf{u} = \alpha \mathbf{u}_{\text{SFR}} + \beta \mathbf{u}_{\text{Bag}}. \tag{2.51}$$

which could have its microscopic basis in a random competition between the two mechanisms, when slip-lines are near admissibility. Here, $\mathbf{u}_{\text{SFR}}$ is an SFR solution for the flow, using the limit-state stress field everywhere, and $\mathbf{u}_{\text{Bag}}$ is obtained from the excess shear stress on a boundary by integrating the Bagnold strain rate $\dot{\gamma} = \sqrt{\tau - \mu\sigma}$ over those shear planes for which $\tau - \mu\sigma$ has been boosted above zero. (Note that we ignore the condition of uniform density for Bagnold rheology since we conjecture that uniform density is actually a geometric consequence of slip-line admissibility and will arise naturally whenever Bagnold rheology dominates the flow.)

A reasonable first approximation is that the SFR and Bagnold solutions individually fulfill the necessary boundary conditions for the velocity profile since, under the right circumstances, either can be made to dominate the other. The constant β is the Bagnold proportionality constant which may depend on the density of the flow among other parameters [33]. Since the SFR is a rate-independent flow model, $\mathbf{u}_{\text{SFR}}$ can always be multiplied by a positive constant (observe that equation (2.26) is homogeneous in ρ_s), and thus we allow the scalar multiple α. Given some determinable form for β, α is chosen such that $\mathbf{u}$ fits the velocity boundary conditions. This seems reasonable for moving walls (as in plate dragging), but not for free surfaces, whose boundary velocity should also be predicted by the theory. In such cases, where α is not clearly defined in this simple model, one could use other empirical relations, such as the Pouliquen Flow Rule for inclined plane flows [108], to deduce the free boundary velocity, and thus α.

2.6.6 Some applications of the composite theory

Using our very simple composite theory, we will now revisit a few geometries that were troublesome for the SFR alone. Extending the theory with a smooth transition to Bagnold scaling controlled by slip-line admissibility seems to resolve the experimental puzzles and capture the basic physics of granular shear flows. In the cases we consider below, we do not change the value of α as we increase the shear stress excess; this way the relative importance of Bagnold effects are easier to isolate.

Planar shear cell

In a zero-gravity planar shear cell, $\mathbf{u}_{\text{SFR}} = 0$, *but* the Bagnold solution for any amount of shear stress excess is of the form $\mathbf{u}_{\text{Bag}} = ky$, and thus the composite solution, regardless of the values α and β, is a homogeneous flow between the two rough plates. The lack of a "background" SFR solution in this case may explain why Bagnold rheology is almost exactly observed in simulations of this geometry over a wide range of strain rates [87].

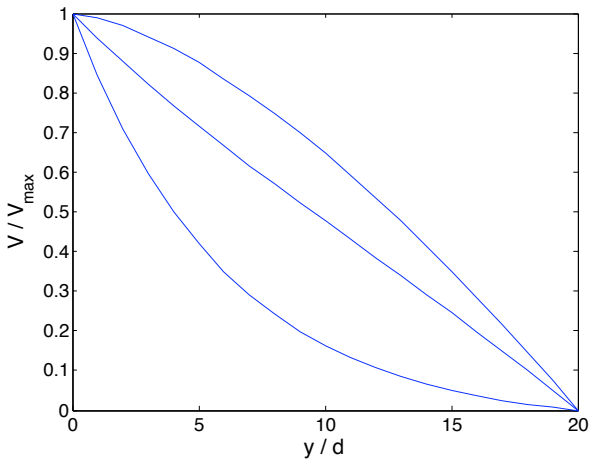


Figure 2-22: Some predicted velocity profiles for flow down a rough inclined plane as a function of depth. Note that all three known flow behaviors; exp decay, linear, 3/2 power law; appear in proper relationship to the inclination (see Ref. [131]). (Bottom) Incline near static repose, fully SFR dominated ($L_s = 4d$); (Middle) Increased inclination angle, Bagnold to SFR ratio of 3:1; (Top) Further increase to inclination; fully Bagnold dominated.

Rough inclined plane

For flow down a rough inclined plane, the SFR solution is an exponential decay (2.47), and the Bagnold solution is a 3/2 power law (2.50). When the material is only slightly above static repose, a shear stress excess along the shear planes will exist but will be very small; it goes as $\sqrt{\Delta\theta}$ for an incline $\Delta\theta$ above static repose [41]. As a result, $\mathbf{u}_{\text{Bag}}$ will be small in magnitude, and the SFR solution will show through as the "creeping flow" with exponential decay. As $\Delta\theta$ increases, the increased shear stress excess will cause the Bagnold contribution to increase, and the flow will eventually morph into the 3/2 power law dependence of pure Bagnold scaling. In between, where both contributions are of similar magnitude, the superposition of the two flow fields gives a predicted profile that appears approximately linear, since the SFR and Bagnold solutions are of opposite concavity. Thus, the composite SFR-Bagnold formulation appears to be able to explain the various flow regimes in inclined plane flow, which have been observed in experiments and simulations (see Figure 2-22).

Recent experimental work of Pouliquen [109] seems to support this analysis; it is

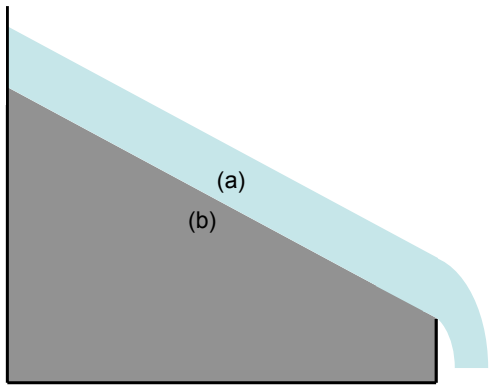


Figure 2-23: Standard heap flows enable one to see both the SFR and Bagnold contributions separately in one flow geometry. (a) The surface region is dominated by Bagnold scaling. (b) The creep-region beneath adheres to the SFR.

found that inclined plane flow occurring at lower inclination angles exhibits spatial velocity correlations near the typical spot size (as the SFR would imply), but as inclination increases, the correlation length appears to decrease, an effect we might attribute to an increased dominance of Bagnold scaling (a phenomenon not governed by a correlation length) over the SFR.

Rapid heap flows

The composite theory also seems consistent with rapid heap flows. When the flow rate down the heap increases, the region near the surface resembles inclined plane flow in any one of its various flow regimes, whereas the region beneath the surface flow undergoes creep motion which decays close to exponentially [76, 94] (see Figure 2-23). We can justify this in terms of slip-line admissibility: The slip-lines throughout the system (see Figure 2-21) have the same form in both regions. In the surface region, the slip-lines are admissible because there is nothing blocking the motion from being a simple shearing along the slip-lines. In the creep region, however, the gate (or the ground) prevents global shearing along the slip-lines and thus the slip-lines are inadmissible and the SFR dominates.

We can equivalently explain heap flow in terms of shear stress excess. The excess incurred by increasing the heap angle will distribute itself differently in the two re-

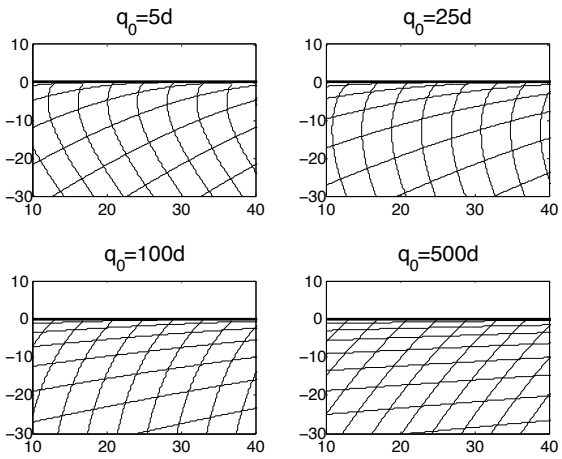


Figure 2-24: Plate-dragging slip-lines approaching admissibility as q_0 increases.

gions. In the surface region, the excess can only be absorbed along the shear planes by inducing a strong Bagnold dependence. However, the gate at the bottom of the heap will support any shear stress excess on the creep region. (Note that the slip-lines in the creep region all hit the gate, or the ground.) Thus the full flow will be the sum of an exponentially decaying SFR solution superposed with a significant Bagnold-type solution which starts at the surface and cuts off at the interface with the creep zone.

Plate dragging under a heavy load

We will now explain the comment made at the end of the plate-dragging section. The slip-lines of a plate-dragging geometry can be pushed drastically close to full admissibility by simply increasing the relative loading of the top plate, q_0, above a certain non-excessive amount (see Figure 2-24). To see the effects of approaching admissibility more carefully, say we take a limit-state plate-dragging setup and pull the plate slightly harder, inducing a super-yield shear stress boost of $\Delta\tau$ under the plate. Bagnold effects should appear wherever, as a consequence of stress balancing, a shear stress excess results along a shear plane. The shear planes are horizontal lines in this case, and at limit-state, the stresses along any horizontal obey

$$\tau - \mu\sigma = -\mu f_g y = -\mu \frac{p_0}{q_0} y. \tag{2.52}$$

When the extra shear stress is applied, a shear stress excess of $\Delta\tau - \mu p_0 y/q_0$ will form for $0 \leq y \leq \frac{\Delta\tau}{\mu p_0} q_0$. Accordingly the Bagnold flow contribution will have a shear zone whose depth extends into the granular bed as an increasing linear function of q_0 for a fixed shear stress boost $\Delta\tau$ and fixed downward plate-pressure (i.e. we decrease the material weight density to increase q_0). Integrating the Bagnold shear rate gives that the relative size of the Bagnold contribution should also increase with increasing q_0. With q_0 large enough, therefore, the SFR contribution will be dwarfed by a Bagnold term with a larger shear band. As $q_0 \to \infty$ the slip-lines become completely admissible and the shear band width diverges as we would expect.

The mismatch in shear band size between the data of Tsai and Gollub and the SFR could be due, among other possible reasons, to the fact that $q_0 \approx 380d$ was large enough to make the Bagnold contribution sizeable. This could also explain the large shear bands found in [141]. A detailed comparison of experiments and simulations with different versions of a composite SFR/Bagnold theory would be an interested direction for future work.

2.7 Conclusion

2.7.1 Highlights of the present work

We have proposed a stochastic flow rule (SFR) for granular materials, assuming limit-state stresses from Mohr-Coulomb plasticity (MCP). In the usual case where slip-lines are inadmissible (inconsistent with boundary conditions), we postulate that flow occurs in response to diffusing "spots" of local fluidization, which perform random walks along slip-lines, biased by stress imbalances. The spot-based SFR corrects many shortcomings of classical MCP and allows some of the first reasonable flow profiles to be derived from limit-state stresses, which engineers have used for centuries to described the statics of granular materials.

Our theory notably differs from all prior continuum theories (cited in the introduction) in that it is derived systematically from a microscopic statistical model [15].

The Spot Model is already known to produce realistic flowing random packings [120], and what we have done is to provide a general mechanical theory of spot dynamics. Evidence for spots has been consistently found in spatial velocity correlations in simulations [120] and experiments [29] (Fig. 2-11) on silo drainage.

Beyond its fundamental physical appeal, the SFR seems to have unprecedented versatility in describing different granular flows. It has only two parameters, the friction angle ϕ and correlation length (spot size) L_s, which are not fitted; they are considered properties of the material which can be measured independently from flow profiles. For monodisperse frictional spheres, the SFR can predict a variety of different flows using the same spatial velocity correlation length, $L_s \approx 3 - 5d$, measured in experiments and simulations. This is perhaps the most compelling evidence in favor of the spot mechanism which underlies the SFR.

We have shown that the SFR can describe a rather diverse set of experimental data on granular flows. Some flows are driven by body forces (silo and heap flows); others have body forces, but are driven by applied shear (plate-dragging); still others are driven by applied shear without body forces (annular shear flow). Some geometries have straight boundaries (silos, heaps, plate-dragging), and yet the theory works equally well for highly curved boundaries (annular shear flow). Some of the flows exhibit shear localization (annular shear flow, plate-dragging, heap flow), and yet the theory correctly predicts wide shear zones in silo flow. It is noteworthy that the same, simple model, correctly predicts and places shear bands in geometries where they arise for very different reasons— gravity causes the shear band in plate-dragging, and yet the geometry (through the $\boldsymbol{\nabla} p$ term in the drift) causes the shear band in annular Couette flow. We are not aware of any other theory (including classical MCP) which can quantitatively describe more than one of these flows, let alone without empirically fitting the velocity profiles.

2.7.2 Comparison with partial fluidization

It is interesting to compare our approach to the continuum theory of partial fluidization of Aranson and Tsimring [7, 8]. Although it lacks any microscopic basis, their

theory also introduces a diffusing scalar field to control the dynamics, as opposed to a classical stress/strain-rate relation. The analog of our spot density is the "order parameter" ρ, which measures the degree of "fluidization" of the continuum by mixing two different types of stresses, corresponding to distinct "liquid" ($\rho = 0$) and "solid" ($\rho = 1$) phases. Given the stress tensor for the material in a static solid state, σ^0_{ij}, the stresses in a flowing granular material are modeled by adding some degree of viscous stresses, as in a Newtonian liquid:

$$\sigma_{ij} = (\rho + (1-\rho)\delta_{ij})\sigma^0_{ij} + \eta\dot{E}_{ij} \tag{2.53}$$

where η is the viscosity. The order parameter controlling the balance of these two stress tensors is postulated to obey a reaction-diffusion equation,

$$(\Delta t)\frac{\partial \rho}{\partial t} = l^2\boldsymbol{\nabla}^2\rho + \rho(1-\rho)(\rho - \delta) \tag{2.54}$$

for collision time Δt, grain length scale l, and a function δ of the stress state, which is greater than 1 where the material is above the static yield criterion, less than 0 where below the dynamic yield criterion, and between 0 and 1 otherwise. One benefit of this model is that it can be used for unsteady flows. In principle, the SFR may also describe time-dependence through the spot Fokker-Planck equation (2.24), but we have only developed and tested the theory so far for steady flows, starting from (2.26).

For the sake of comparison, consider a steady flow modeled by partial fluidization and the SFR. The difference is that the spot equation (2.26) couples diffusion to a drift depending on frictional yielding, whereas the order parameter equation (2.54) balances diffusion with a nonlinear source term, resembling a chemical reaction rate, which indirectly mimics the effect of a Coulomb yield criterion. Interestingly, if the SFR could be extended to an elasto-plastic model without making the incipient failure assumption (see below), a similar non-linear source term may have to be added to the spot equation to account for the need to destroy (and create) spots when they enter zones below (or closer to) yield. More generally, a reaction term could describe the

creation and destruction of spots, e.g. in response to plastic work as in the theory of shear transformation zones [83].

As a higher-order yield effect, this term could also be chosen to reflect bistability in granular materials, where solid-like and fluid-like local phases can both be dynamically stable and coexist depending on the stress state. Equation (2.54) describes this through the dependence of δ on both μ and μ_k and implies that stable flow can occur beneath the static yield criterion as long as the nearby kinetic yield criterion is still exceeded, as should be expected. Our current model describes the case where static incipient failure and the transition $\mu \to \mu_k$ are both valid assertions, so it is effectively a *flow* theory more than a complete dynamic theory of both flow and solidification. To increase the prominence of μ_k would not just modify our flow rule, but also necessitate a more complicated stress model, given our physical hypotheses. Unsteady flows, such as a traveling avalanche front, may require accounting for bistability, but for fully developed steady flows as described above there is often some degree of motion, even in nearly stagnant regions. In any case, solid-like regions do not seem to greatly affect the flows we consider above. We should also point out that the SFR does allow a *global* no-flow solution since $\rho_s = 0$ always solves the spot equation.

It is also notable that our argument for why a spot drifts, i.e. a localized stick-slip type of shear stress decrease along the spot boundary, is reminiscent of equation (2.53) wherein the shear stress goes down in the presence of fluidization. In this sense, a higher spot concentration in our model is similar to a higher degree of partial fluidization.

One difficulty with the partial fluidization approach is that it cannot easily describe rate-independent effects since the motion stems from a viscous form in the stress tensor. Also in sharp contrast to our approach based on plasticity, partial fluidization does not provide a clear theory of the static solid stresses in the limit of no flow, opting instead to deal with environments for which the open components of this tensor are not needed (simple shear flows). This could perhaps be modified. These considerations as well as selecting boundary conditions on the order parameter, seem to be the primary limitations in testing partial fluidization in more general situations.

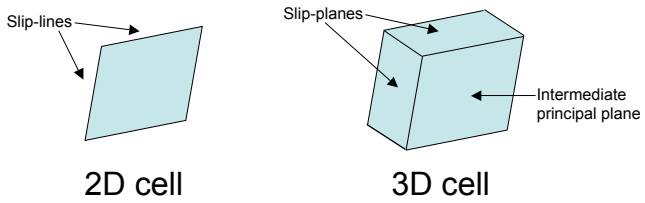


Figure 2-25: The difference between 2D and 3D continuum cells which could be used to construct the SFR.

2.7.3 Future directions

In spite of some successes, we still do not have a complete theory of dense granular flow. There are at least three basic limitations of the SFR, about which we can only offer some preliminary ideas to guide future work.

Slip-line inadmissibility

Although most slip-line fields are inadmissible, the SFR breaks down as slip-lines approach admissibility. We have already begun to extend the model into this regime by conjecturing that slip-line admissibility is associated with Bagnold rheology, as excess shear stress (above the limit-state) drives a local shear rate along the most admissible slip-lines. We have shown that a simple linear superposition of Bagnold and SFR flow fields with appropriate boundary conditions can describe a variety of composite flows, exhibiting both Bagnold and SFR behavior in different limits or segregated into different regions. These include planar zero-gravity shear, various inclined-plane and heap flows, and plate-dragging at large loading. However, more work is needed to develop and test a composite SFR/Bagnold theory, both at the continuum level and in terms of the two microscopic mechanisms.

2D symmetry

Through MCP, the SFR is currently used only in quasi-2D geometries. In efforts to extend the theory to 3D, a good test case would be the split-bottom Couette cell, which displays a wide, diffusive shear band [47], reminiscent of a draining silo. The 2D limitation may not be so difficult to overcome, although any plasticity theory

is more complicated in three dimensions, than in two. As usual, constructing a 3D limit-state stress field requires an additional hypothesis to close the stress equations. In 3D, a general material point at incipient failure with distinct principal stresses $\sigma_1 > \sigma_2 > \sigma_3$ is intersected by a pair of slip planes angled 2ϵ apart. We cannot therefore encase a 3D cell of material within slip planes as we are able to do to a 2D cell with slip-lines. However, the principal plane on which σ_2 acts, the intermediate principal plane, can be used along with the slip-planes to encase a 3D material cell. This is legitimate because, if such a cell underwent slip-plane fluidization, the net material force would be guaranteed to point parallel to the intermediate principal plane; since the intermediate principal plane offers no shear resistance, the material can slide along this plane, while simultaneously sliding along a slip-plane.

To apply the SFR then, the drift vector should still be calculated from equations (2.28) and (2.30), but all vectors must be projected first into the $\sigma_1\sigma_3$-plane since the σ_2 direction is not involved in slip-plane fluidization. The shape of a spot and its diffusivity would likely be anisotropic, with different values in the intermediate direction, since the main source of diffusion is slip-plane fluidization.

If ever the intermediate principal stress equals either the major or minor principal stress, as in the Har Von Karman hypothesis, incipient failure is upheld on a cone instead of intersecting slip-planes. When this degenerate case occurs, the material cell can be encased several different ways depending on the surrounding stress states. This must be determined before we can rigorously define how to apply fluidization and the SFR.

Incipient yield everywhere

The SFR assumes stresses near a limit-state. While incipient yield is believed to be a good hypothesis in many situations, even during dense flow (which should be checked further in DEM simulations), it clearly breaks down in some cases, at least in certain regions. This is perhaps the most difficult limitation to overcome, since the limit-state assumption is needed to fully determine the stress tensor. Without it, the material effectively enters a different state most likely governed by some non-linear

elastic stress strain law which is far more difficult to apply.

We have already argued that such a transition away from incipient yield must exist in some granular flows, due to the strong tendency of granular materials to compactify into a rigid solid state when shaken (e.g. by nearby flowing regions) but not sufficiently sheared. A good example is a tall narrow silo with smooth side walls, where the SFR holds near the orifice, but breaks down in the upper region, resembling a vertical chute. The broad shear band localizes on the side walls, as a rigid central plug develops, which likely falls below incipient yield.

A more robust elasto-plastic theory for the stress state would relax our limit-state constraints and allow for material to fall below the yield criterion where it is described by elasticity. The SFR could then be applied only where the material is at yield and everywhere else the material does not deform plastically. Elasto-plasticity theory operates just as well in 3D as in 2D which is a key benefit over limit-state plasticity. However, our model as we have already presented it is *far* simpler than elasto-plasticity and yet still manages accurate results when applied to limit-state stress fields.

As the SFR matures as a theory of granular flow, it would also be interesting to apply it to other amorphous materials, such as metallic glasses, and to develop new simulation methods. The basic idea is very general and applies to any material with a yield criterion. It has already been suggested that the Spot Model could have relevance for glassy relaxation [15], and the SFR provides a general means to drive spot dynamics, based on solid mechanical principles. The Spot Model also provides a multiscale algorithm for random-packing dynamics, which works well for silo drainage [120], so the SFR could enable a general framework for multiscale modeling of amorphous materials. The idea would be to cycle between continuum stress calculations, meso-scale spot random walks, and microscopic particle dynamics.

Chapter 3

Nonlinear Granular Elasto-Plasticity

3.1 Transition from the SFR

Many new results have come out since the SFR was developed. Some offer a firm challenge to physical hypotheses on which the SFR was built. We briefly summarize the major results now, and delve into details shortly.

Discrete element simulations of Chris Rycroft [122] point out two important facts about mesoscopic deformation within a bulk quasi-2D flow. For one, coaxiality *is* approximately upheld in regions of significant plastic flow. Others have reached this conclusion as well: Depken and coworkers [36, 35] find significant coaxiality in 3D simulations of split-bottom shear flow, and 2D polydisperse disk simulations of [33, 87, 32, 130, 113, 136] show an adherence to coaxiality in steady flows within annular shear, inclined plane, and planar shear geometries. Secondly, Rycroft's results find the limit-state assumption is markedly *incorrect*— the local friction state can take on a range of values in both static and flowing regions. The assumption of a constant state of incipient failure may have some validity in the very slow limit of flow, as observerd in [33, 115]. However, Rycroft's data suggests that the internal friction increases with the flow-rate in faster flows, which is in broad agreement with the findings of other researchers [33, 135, 97, 125, 26, 108]. Rycroft's data also suggests that in solid-

like zones, the stress state can take on values below incipient failure. This has been observed as well in other discrete simulation studies both in 2D [135, 141] and 3D [35]. These results contradict the underlying assumptions that were used in developing the SFR— the SFR inherits the limit-state assumption from Mohr-Coulomb Plasticity and non-coaxiality from the Spot Model. While the predicted flows obtained from the SFR look convincing, the fact that they come from a model based on incorrect physical assumptions is reason for alarm.

One should be especially careful not to misjudge the merits of the SFR or the notion of spots from this information alone. The SFR is intended only for very slow flows where the stress state is independent of the flow rate. The discrete simulations were unable to show whether or not coaxiality is upheld in slower flowing regions, only that more rapid, liquid-like zones have the coaxial property. As shall be discussed in depth during this chapter, it is still reasonable to believe that the SFR is a valid principle in truly quasi-static flows where the stresses happen to be near the static yield criterion. However, the presumption of a universal limit-state stress field is most likely false.

In any case, it seems a different approach may be necessary capture these newest observations. Since the local friction state is not necessarily constant in a static material, elasticity may be a better model for static stresses. Since the flow appears to be locally connected to the stress state, a formal plasticity law may be in order so that the stresses determine the deformation rate. Combining these two notions into one material model would imply an elasto-plastic treatment, where liquid-like flow occurs whenever a yield criterion is met, but below yield the material acts as an elastic solid. This type of model is in striking contrast to the *stochastic* dynamics of the SFR. Basic elasto-plasticity is *deterministic*; given the local state parameters and stresses, an element of an elasto-plastic material should undergo a predictable deformation increment.

Elasto-plasticity is a well-established field of material mechanics. Unlike the SFR, which attempts to describe a very new mechanism for material flow, by utilizing elasto-plasticity one gains the benefits of years of theoretical development. Some

issues that were problematic for the SFR are non-existent in elasto-plasticity. For one, the SFR is frame-indifferent in steady-state but it is not clear if any way exists to extend the SFR to transient behavior and still maintain frame-indifference. As will be shown, elasto-plasticity applies as well to transient flows and is derived so as to always ensure frame-indifference. There are also no issues with adding the third dimension. And without hyperbolicity in the stress equations, elasto-plasticity should not invoke stress shocks like those encountered in limit-state Mohr-Coulomb plasticity.

This chapter shall construct and test a new elasto-plastic model for granular materials. More than just a mathematical model, the foundational principles are emphasized, showing piece-by-piece how and why such a model should be physically appropriate. Starting out with arguments for a continuum treatment (section 3.2), past continuum models for granular statics and flow are then discussed (section 3.3, 3.4). From these options, the Jiang–Liu elasticity relation [68] and the Jop–Pouliquen flow rule [71] are selected. The precise method for joining these two behaviors will then be physically motivated (section 3.5). The physical picture is then translated into a complete set of mathematical relationships (section 3.6). The procedure for solving these equations is then described in depth, including particular techniques that are specific to our interests (section 3.7). Direct numerical results are then compared against known data (section 3.8).

3.2 Granular matter as a continuum

Before all else, it is essential to lay out an argument for why it should be acceptable to model granular matter as a continuum. At the outset, there appear to be a few reasons why such a treatment may be objectionable.

Most obvious is the fact that the microconstituents of granular matter, the individual grains, may not be small enough to warrant a continuum description. Typical continuum laws are only expected to apply when there is a strong *separation of scales* [59] between the micro-scale and the macro-scale size of the flow geometry. When

the scales separate, the material response should be dictated by the dynamics of a meso-length Representative Volume Element (RVE). A well-homogenized RVE undergoes deterministic homogeneous deformation when loaded under a uniform boundary stress state. While the separation of scales clearly holds for, say, a Newtonian fluid, the situation is not so clear for granular matter.

Discrete Element Method (DEM) simulations have shed light recently on exactly how "small" a meso-element of grains must be to capture the flow properties of an RVE [122]. The DEM algorithm simulates 3D granular flow on a particle-by-particle basis, tracking the individual motions of several hundred thousand grains on a parallel computer. The particles are modeled as visco-elastic spheres (among other possible shapes) allowing incremental grain motion to follow from contact dynamics. Since interparticle contact forces are known throughout, DEM also enables one to construct the space-averaged stress tensor over any localized collection of grains at any instant.

For d the particle diameter, monodisperse flow simulations were performed in three "conventional" geometries, where the smallest macroscopic length was always $\geq 6d$ and flow was driven by gravity or some slow to moderate wall motion. Results appear to indicate that a granular RVE need only be $5d$ wide. At this length, the average stresses and deformation rate appear to vary smoothly between neighboring volume elements. Within flowing elements, the eigenvectors of the Cauchy stress and deformation rate approximately align, suggesting the possibility of a deterministic connection between the two tensor fields. Other deterministic relationships can be observed on this size scale such as the dependence of the packing fraction on the pressure and shearing rate. It seems that a $5d$ wide element could be large enough to represent a predictive rheology.

It should be emphasized that what matters most in these findings is not the precise measurement $5d$, but rather the more general observation that the size scale for representative behavior is in the range of several particle widths. How small this is may come as a surprise. We normally think of determinacy at the element level as stemming from the combined effects of an approximately infinite number of microconstituents. It is indeed a welcome and unexpected result that only a couple

hundred grains replicates this phenomenon. Beyond monodisperse spheres, we would speculate that other granular media possess a similar size scale for RVE behavior, though certain features may affect the specifics such as bidispersity, anisotropy of the grain shape, and grain contact roughness.

The $\sim 5d$ element width for dense flow is not entirely outlandish in light of the past observations in Glasser and Goldhirsch (2001) [49] and Goldenberg *et al.* (2006) [52], where in depth studies were performed to quantify the effects of spatial averaging on granular stresses. They utilized a more ellaborate general form for space-averaging, involving a spatial coarse-graining function $\phi(\mathbf{R})$, which is a positive semidefinite normalized function containing a single maximum at $\mathbf{R} = \mathbf{0}$. A complete, continuum mechanically compatible expression for the Cauchy stress tensor field can be derived in terms of ϕ. For a static assembly,

$$T_{ij}(\mathbf{r}) = \frac{1}{2} \sum_{a \neq b} f_{abi} r_{abj} \int_0^1 \phi(\mathbf{r} - \mathbf{r}_a + s\mathbf{r}_{ab}) ds \tag{3.1}$$

where i and j are cartesian coordinates, a and b are particle labels, $\mathbf{f}_{ab}$ is the contact force of grain a on grain b, and $\mathbf{r}_{ab} = \mathbf{r}_a - \mathbf{r}_b$. The spatial averaging conducted by Rycroft [122] could be thought of as a simple case of the above, in which $\phi = H(|x - 5d/2|)H(|y - 8d/2|)H(|x - 5d/2|)/V_{box}$ for H the Heaviside function.

Using the above form for spatial stress averaging, Glasser and Goldhirsch found that rapidly sheared granular gases *do not* undergo scale separation. That is, the computed stress field has an irreconcilable order one dependence on the width of the spatial coarse-graining function. On the other extreme, Goldenberg *et al.* finds that scales *do* separate for static 2D material under a localized force. If in addition the computed stress profiles are averaged over 40-50 realizations, the width of the coarse-graining function necessary for an "objective", or resolution-independent definition of Cauchy stress can be less than one particle diameter. Given the shear-rates present in Rycroft's work, the fact that (qualitatively) resolution-invariant behavior appears for coarse-graining widths $\sim 5d$ could fit appropriately into this trend of behaviors. This speculation is bolstered in Glasser and Goldhirsch, where it was observed that

by decreasing shear-rate, a granular gas can become resolution-independent, though with a large coarse-graing width ($\gtrsim 15d$).

Previous arguments against a granular continuum, particularly in the static phase but also during flow, have focused on the presence of force chains in granular matter. In 2D experiments and simulations of disk assemblies [16, 48], concentrated chains of interparticle force have been shown to exist over many particle widths. It was argued that if forces are not homogeneous at a meso-scale, continuum relations at this scale cannot exist. However, in 3D simulations of flowing monodisperse sphere packings, we observe that the force chains have a dramatically shorter length [122]. A possible geometric explanation for this phenomenon is that a 3D granular assembly has a much higher average coordination number, reducing the likelihood that only two contacts maintain the majority of the force on one grain. These simulations also include interparticle contact friction, which may contribute to the smoothing out of force chains as has been previously shown in 2D static disk assemblies [54]. Typically speaking, whether flowing or static, we observe that a $5d$ granular volume element contains a diffuse network of contact forces, enabling a sufficient degree homogenization of the stresses at this scale.

The limits of any continuum approach can be drawn from basic statistics. In a first-order constitutive model as we shall propose, where stress relates directly to the homogeneous part of deformation, gradients in the fields must be sufficiently small over the width of the RVE. Our previous use of the word "conventional" could in this sense be thought of as a common limit on the size of such gradients. In the presence of large gradients, the mean behavior within an element would vary significantly in space, reducing the multiplicity of the state within and consequently reducing the determinacy of the continuum relation. Other averaging schemes (ensemble average, steady-state time average) may still hold relevance in the presence of large spatial gradients, but our goal of accurately predicting a single realization of granular flow would be lost.

To appropriately model dense granular media, the model must be able to encompass both static and flowing granular behavior. Frequently, a granular material that

has been set into motion will express both types of behavior, possessing regions of fluid-like flow adjacent to essentially solid-like regions. Examples of such flows include wide draining silos [28, 123] and hoppers (so-called "core flow") where a broad column of material extending upward from the orifice flows like a fluid, while regions closer to the side walls remain almost completely static. As will be discussed later, perfectly clear solid/fluid interfaces are rarely observed in granular flow, which has led some to believe that the solid-like zone is actually just a 'highly viscous' fluid region [124, 88, 19]. Solid-like material does undergo intermittent rearrangement events when close to a zone of moderate flow-rate, but we find that the stresses in these regions have essentially zero rate-dependence. For example, when a DEM simulation of silo flow is suddenly arrested, say by shutting the orifice, the stresses in the solid-like regions remain virtually unchanged, supporting static shear stress like a solid. It appears that a mechanism disconnected from the flow-rate or any notion of viscosity is responsible for maintaining the stress tensor in solid-like granular matter, even if occasional failure events are occurring within.

3.3 Continuum statics

3.3.1 Stress-only laws vs. elasticity

This last argument suggests that a continuum description for granular material could be formed by augmenting a statics law with a plastic flow law describing yield behavior. First, let us review past work on continuum granular statics. Static granular matter can support shear stresses so force balance in 2D or 3D never provides enough equations to determine, respectively, the 3 or 6 unique components of the stress tensor. To close the equations, two basic approaches have been applied: stress-only laws and elasticity.

Stress-only relationships constrain the stresses directly, by asserting that the stress tensor must satisfy some a priori relationship. Examples include: Janssen's law of differential slices (originally proposed by Janssen in 1895) where vertical and hori-

zontal stresses are set to be proportional, limit-state Mohr-Coulomb plasticity or the "Incipient Failure Everywhere" (IFE) hypothesis [132] where a Mohr-Coulomb failure line is required to exist at all locations, and "Oriented Stress Linearity" [20] where stresses propagate in directions aligned with the presumptive microstructure of the packing.

While stress-only relationships are convenient and have had some success, their physical assumptions can be questionable. For example, static granular matter is rarely in a limit-state of incipient failure [122] and wall shear is not compatible with a Janssen-style stress tensor [101]. Most stress-only laws are defined only for 2D media, which brings out issues of generality. They often predict a "hyperbolic" character to the stress profile, where stress quantities propagate in certain directions from the boundaries. The notion of force propagation was backed chiefly by the observation of a double-peak pressure distribution beneath a bed of grains undergoing a point force from the top. Work by Goldenberg and Goldhirsch [54] has shown, however, that in the presence of interparticle friction and a large system size to particle size ratio (as commonly found in engineering applications) the pressure distribution is indeed a broad single peak beneath the point force, as one would expect for an elastic bulk media.

This brings us to elasticity, which shall be our preferred method for granular statics. It is a sensible approach seeing as the grains themselves are elastic bodies presumably enabling generalization to an Effective Medium Theory (EMT) where grain-level elasticity extends statistically to a continuum mean-field theory [40, 37, 142, 103]. Reversible (elastic) deformations have been observed in granular matter for strains less than 10^{-4} [79]. This is negligibly small compared to the size of typical plastic deformation. However, grains are commonly composed of stiff material (i.e. glass has elastic moduli on the order of 10 GPa) indicating the important role that small elastic strains may play in the development of the stress profile. One may argue that the barely noticeable elastic strain of a static assembly is what impelled scientists years ago to seek stress-only laws, believing that whatever determines the stresses should do so under rigid-body assumptions. Rather, it seems there is no

generally applicable stress-only constraint to accurately define a 3D stress tensor. The elastic strains are small but non-ignorable, and bear essential importance to the stress description.

Bulk elasticity for cohesionless grains is not likely to have a simple form since, for example, the material is unable to support tension and thus the small strain response cannot be linear. Nonlinear EMT has been derived from Hertz-Mindlin interparticle contact mechanics [40, 96] and modified by others [45], offering reasonable albeit not completely satisfactory agreement with experiments [79, 51]. A recently proposed elasticity model encompassing many of the same features as EMT was proposed by Jiang and Liu [68] in 2003. This state-of-the-art formulation has had multiple successes and is well-suited to our end goal of combining with a plasticity model.

3.3.2 Effective Medium Theory for bulk granular elasticity

To start, let us review the basics of Hertzian contact laws. In the classical work of Hertz, two perfectly elastic spheres that are pressed into contact will repel each other with a contact force that depends on the radii of the spheres and the apparent overlap. That is, we pretend for ease of computation that the spheres penetrate each other. For two spheres of identical radius R located at $\mathbf{x}_1$ and $\mathbf{x}_2$, the normal force contact law is

$$F_n(\delta) = \frac{2}{3} k_n R^{1/2} \delta^{3/2} \tag{3.2}$$

where δ is the normal overlap $(1/2)\,(2R - ||\mathbf{x}_2 - \mathbf{x}_1||)$. The parameter k_n is an effective stiffness equal to $4G_g/(1-\nu_g)$ where G_g and ν_g are the shear modulus and Poisson ration of the sphere material.

There is a rather basic way to extend the Hertzian normal force law into an EMT for the bulk modulus of an isotropic collection of grains. A full derivation is difficult to find in the literature, so we go about constructing our own. It is a quintessential result that is worth understanding.

The Cauchy stress tensor $\mathbf{T}$ over a volume V that contains a static collection of

N objects in contact can be expressed according to [30] as

$$\mathbf{T} = \frac{1}{V} \sum_{i<j}^{N} \mathbf{r}^{(ij)} \otimes \mathbf{f}^{(ij)} \tag{3.3}$$

where $\mathbf{r}^{(ij)}$ is the vector that connects the centroid of particle j to the centroid of particle i, and $\mathbf{f}^{(ij)}$ is the force of particle i acting on particle j, which will be $\mathbf{0}$ unless particle i and j are in contact. In the case of elastic spheres in Hertzian contact, $\mathbf{f}^{(ij)} \cdot (\mathbf{r}^{(ij)}/2R) = -F_n(\delta)$ for the function F_n as defined above.

In a gravity free environment, consider an unstressed container of N spheres that are barely in contact. By this we mean the geometry of the contact network is established, but no force is being exerted through any contact. Now suppose we isotropically compress the container— that is, a particle that started at $\mathbf{X}$ is now located at $\mathbf{x} = a\mathbf{X}$, for some $a < 1$. This is equivalent to a strain of

$$\mathbf{E} = (1/2)\left(\frac{\partial \mathbf{x}}{\partial \mathbf{X}} + \frac{\partial \mathbf{x}}{\partial \mathbf{X}}^{T}\right) - \mathbf{1} = (a-1)\mathbf{1}$$

in the small-strain limit. The idea that isotropically compressing the boundaries results in all particles displacing in an affine fashion comes with our mean field approximation, where an isotropic distribution of contacts is presumed. We concede that any realization of a true random packing is likely to have some internal non-affine shifting occur during compression, as micro-level equilibrium is not a guarantee. This particular feature of granular relaxation has been studied in depth in [91, 53, 90].

Each particle contact is now accompanied by a normal force. Any two particles in contact were originally separated by $2R$ and are now separated by $2Ra$ implying $\delta = R(1-a)$. Thus,

$$\mathbf{T} = \frac{1}{V} \sum_{\text{Contact pairs } i<j}^{N} \mathbf{r}^{(ij)} \otimes \left(-F_n\left(R(1-a)\right)\mathbf{r}^{(ij)}/2R\right)$$

and consequently the macroscopic pressure is

$$p = -(1/3)\mathrm{tr}\mathbf{T} = \frac{1}{3V}\frac{F_n\left(R(1-a)\right)}{2R}\sum_{\text{Contact pairs } i<j}^{N}\mathrm{tr}(\mathbf{r}^{(ij)}\otimes\mathbf{r}^{(ij)})$$

$$\rightarrow\ p = 2RF_n\left(R(1-a)\right)\frac{1}{3V}\times(\text{number of contacts})$$

If the mean packing fraction in the container is Φ, then the container contains roughly $N = \Phi V/V_g$ spheres each of volume V_g. If the average coordination number is Z, then the mean number of contacts within the container should be $NZ/2$. In full, we have the following mean field result:

$$p = \frac{1}{6\pi}\Phi Z k_n (1-a)^{3/2}$$

This implies, in the absence of any shear deformation, that the mean field bulk modulus κ is a nonlinear function of the compressive strain:

$$\kappa = \frac{1}{18\pi}\Phi Z k_n (-\frac{1}{3}\mathrm{tr}\mathbf{E})^{1/2}$$

where the nonlinearity can be seen as arising directly from the nonlinearity of Hertz's contact law. The major point is that the bulk modulus scales with (compressive strain)$^{1/2}$, or equivalently as $p^{1/3}$ under isotropic compression. This has been verified directly in large-scale DEM simulations of compressed sphere packings both with and without interparticle contact friction [90].

Supposing frictionless spheres, a mean-field shear modulus can also be derived in a similar fashion. Instead of an isotropic compression, one could analyze the case of an arbitrary affine deformation. The mathematics is a bit lengthier because the average over orientations does not immediately collapse. In the end, one finds that the bulk and shear moduli scale similarly to the above form for the bulk modulus in pure compression:

$$\kappa \propto G \propto \Phi Z(-\mathrm{tr}\mathbf{E})^{1/2} \propto (\Phi Z)^{2/3}p^{1/3} \tag{3.4}$$

It so happens that the inclusion of shear strain renders EMT less accurate. The moduli agree with the above scalings but only under low pressures (<10 MPa) [90]. Experiments verify that the dependence on p is characterized by a larger exponent at higher pressures [51]. This could be because shearing under higher pressures tends to make the affine displacement assumption less valid. Time-dependent relaxation occurs, which significantly complicates a determination of the shear modulus.

Attempts to improve the theoretical elastic moduli during shear are not vastly helped by accounting for tangential forces between particles. It was not until the 1950's that Mindlin [96], determined a form for the tangential force between two spheres that are not just pressed into contact, but also sheared with respect to each other. During the displacement, the friction is such that rolling and slippage do not occur. Then the resulting incremental tangential force is

$$\Delta F_t = k_t (R\delta)^{1/2} \Delta s \tag{3.5}$$

where the effective tangential stiffness k_t is equal to $8G_g/(2-\nu_g)$. The quantity Δs measures the lateral displacement increment of the contact region's center with respect to the center of one of the spheres. Note that the tangential force is path dependent under this description:

$$F_t = \int_{\text{Path in } (\delta,s) \text{ space}} k_t (R\delta)^{1/2} ds$$

In other words, compressing by δ and then shearing by s results in a different tangential force than if one were to compress by $\delta/2$, shear by s, and then compress by $\delta/2$. This effect is due to the fact that the tangential force stems from microscopic tractions within the circular contact region between spheres. The shear tractions within an annular region of the contact zone depend on how far that region was sheared from when it originally became part of the contact area.

EMT that utilizes Mindlin tangential forces (and some prescribed loading history) produces a modified shear modulus and an unchanged bulk modulus. The new shear modulus scales the same way as in the frictionless case, differing only by a constant

prefactor in terms of k_t. This is not enough to rescue EMT, as a more dramatic difference would be necessary that somehow incorporates a greater importance of the shear strain on the moduli.

3.3.3 The Jiang–Liu granular elasticity law

With the inability of EMT to provide a reliable theory of granular elasticity in the presence of shear, Jiang and Liu proposed an alternate approach in 2003 [68]. Rather than continue laboring on a microscopically derivable mean field theory, augment the results from EMT that *do* work, with presumptive forms that capture known macroscopic behavior.

Consider an elastic free energy density of the following form:

$$\psi(\mathbf{E}) = B\sqrt{\Delta}\left(\frac{2}{5}\Delta^2 + \gamma^2/\xi\right) \tag{3.6}$$

B is a relative stiffness that can vary with packing fraction. The compressive strain and shear strain are measured respectively by

$$\Delta = -\mathrm{tr}\mathbf{E}\ , \quad \gamma = |\mathbf{E}_0|$$

where $\mathbf{A}_0 \equiv \mathbf{A} - (1/3)\mathrm{tr}\mathbf{A}$ and $|\mathbf{A}| \equiv \sqrt{\sum_{i,j} A_{ij}^2}$. The dimensionless parameter ξ will be discussed shortly. In classical small-displacement elasticity theory, the stress tensor arises from the free energy function according to

$$\mathbf{T} = \frac{\partial \psi}{\partial \mathbf{E}}. \tag{3.7}$$

Thus, we convert the elastic free energy directly into a form for the stress tensor:

$$\mathbf{T} = 2\underbrace{\frac{B\sqrt{\Delta}}{\xi}}_{G(\Delta)}\mathbf{E}_0 + \underbrace{B\sqrt{\Delta}\left(1 + \frac{\gamma^2}{2\xi\Delta^2}\right)}_{\kappa(\Delta,\gamma)}(\mathrm{tr}\mathbf{E})\mathbf{1} \tag{3.8}$$

Under isotropic compression, the pressure is proportional to $\Delta^{3/2}$, in agreement with

the most successful result of EMT. The shear modulus scales with $\Delta^{1/2}$ as in EMT, but the added nonlinearity in the full form of the bulk modulus allows for some important properties not represented well by EMT. Consider an applied strain of (Δ, γ). Equation 3.8 implies the pressure obeys

$$p = (-1/3)\mathrm{tr}\mathbf{T} = B\Delta^{3/2}\left(1 + \frac{\gamma^2}{2\xi\Delta^2}\right).$$

The equivalent shear stress τ, obeys

$$\tau = |\mathbf{T}_0|/\sqrt{2} = \frac{B\gamma\sqrt{2\Delta}}{\xi}.$$

The ratio of equivalent shear stress to pressure is called the Drucker-Prager μ and represents a 3D friction state. From the above, we obtain the result

$$\mu = \frac{\frac{B\gamma\sqrt{2\Delta}}{\xi}}{B\Delta^{3/2}\left(1 + \frac{\gamma^2}{2\xi\Delta^2}\right)} = \frac{\sqrt{2}}{\xi\left(\frac{\gamma}{\Delta}\right)^{-1} + \frac{1}{2}\left(\frac{\gamma}{\Delta}\right)}.$$

Let $r = \gamma/\Delta$. Note that μ has a maximal value if the denominator of the above has a minimum. To check for this, we set the derivative of the denominator to 0 and solve for r:

$$-\xi r^{-2} + \frac{1}{2} = 0 \quad \rightarrow \quad r = \sqrt{2\xi}$$

Consequently, μ has a maximal value determined by ξ:

$$\mu_{\max} = \frac{\sqrt{2}}{\frac{\xi}{\sqrt{2\xi}} + \frac{\sqrt{2\xi}}{2}} = \frac{1}{\sqrt{\xi}}.$$

Thus, the Jiang–Liu elasticity formulation comes with a powerful macroscopic property: no elastic strain state can produce a stress state where μ exceeds $1/\sqrt{\xi}$. Thus, by selecting ξ accordingly, we can prevent certain states of friction from ever arising.

In the work of Jiang and Liu, a static yield criterion μ_s was declared and ξ was set to μ_s^{-2} so as to require that no purely elastic state can exist above μ_s. Their interest was in the study of statics, so cutting off elasticity precisely at static yield is a sensible

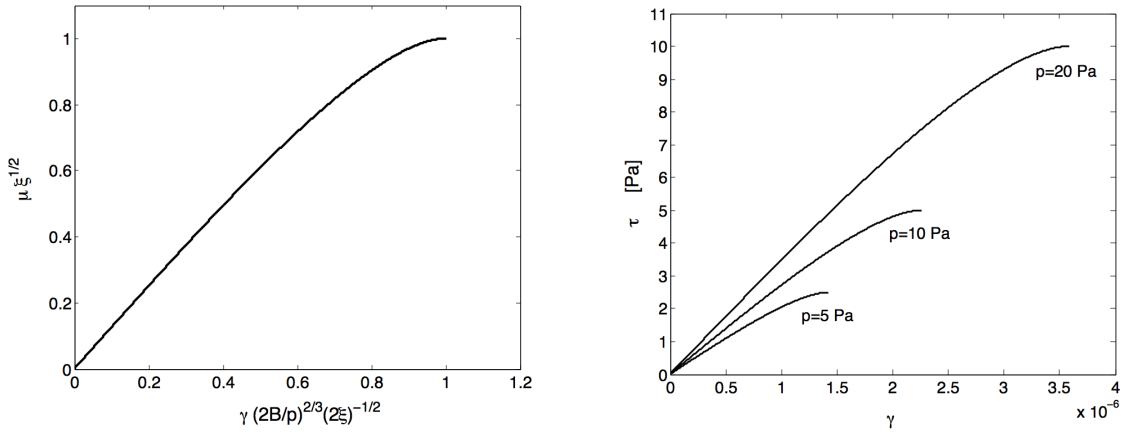


Figure 3-1: The shear stress vs. shear strain relation under Jiang–Liu elasticity represented as a single dimensionless plot (left), and plotted equivalently as a family of shear stress vs. shear strain curves in SI units using $\xi = 4$, $B = 7 \times 10^9$ Pa, each curve determined by the applied compressive pressure (right).

assumption. Our work shall attempt to integrate elasticity within a complete elasto-plastic framework, so the basis for selecting ξ will be slightly different.

At the point where the elastic stress state reaches μ_{max}, the relation loses convexity and ceases to represent a valid elasticity law. More specifically, the Jiang–Liu model is only physical for shear states fulfilling $|\mathbf{E}_0| \leq \left(\frac{p}{2B}\right)^{2/3} \sqrt{2\xi}$. Beyond this point, any additional strain cannot be purely elastic. Figure 3-1 elucidates the fashion by which μ increases to its maximum under increasing shear strain. The plots are all discontinued in regions where the relation is not convex.

More than just a model that connects with yield, the Jiang–Liu elasticity model has had some convincing successes in representing granular statics:

1. The nonlinear form of $\mathbf{T}$ produces a stiffness tensor $\frac{\partial T_{ij}}{\partial U_{kl}}$ that agrees to a satisfactory extent with the form of the stiffness tensor extracted from experimental data [79].

2. The model predicts Janssen-type saturation of the wall stresses in a tall silo. The ratio of vertical to horizontal compressive stresses in the silo is found to be approximately constant when not close to the walls. This verifies the commonly used notion of a "coefficient of redirection", which has been verified in DEM

simulations [122].

3. The model predicts that a granular material under simple shear stress responds anisotropically to the addition of a point-load at a surface. Such anisotropy under pre-stress is a well-known granular phenomenon that is captured appropriately in the nonlinearity of the Jiang–Liu model.

4. It has been observed that preparation history is largely responsible for the stress dip that one often observes under the peak of a sand pile. This fact is reproduced by the elasticity model when solved assuming an initial packing fraction distribution that one might expect for a conical pile constructed from grains flowing out of a nozzle.

The Jiang–Liu nonlinear elasticity model indeed comes with many perks. To the theorist who wishes a micro-to-macro understanding of granular elasticity, the model encompasses many successful results from EMT, like the $p^{1/3}$ scaling of the bulk modulus and the $p^{1/6}$ dependence of the sound speed. The preparation history, which is known to affect stress dependence, is included in the theory via the dependence of B on the packing fraction. This feature is also borrowed from EMT, though the Jiang–Liu model does not account for the effect of the average coordination number like EMT. We note, however, that in disordered systems the packing fraction and coordination number tend to be closely related.

At the same time, the Jiang–Liu model predicts several known macroscopic phenomena as listed above, many of which cannot be explained by EMT. The model also makes some account of γ in the definition of the elastic moduli, a welcome addition whose lacking within EMT was problematic. It should be emphasized that while the Jiang–Liu model is versatile, its qualitative capabilities still outweigh its quantitative exactness at this stage. It appears to have the necessary *form* to describe many static granular phenomena.

3.4 Continuum flow

3.4.1 Bagnold scaling and relevant dimensionless quantities

Bagnold's seminal work on granular flow followed from shear experiments on granular/fluid suspensions. In the "grain-inertia" regime where the effects of the interstitial fluid are small, Bagnold found that both shear and normal stresses on the shearing wall depend quadratically on the shear rate [12], a phenomenon that came to be known as "Bagnold scaling". Bagnold scaling has been verified for dry grains in both the collisional and dense flow regimes [87, 130, 108, 113, 33]. An explicit form in the case of simple shear of a dense configuration of dry grains can be expressed as the following pair of dimensionless relations:

$$\Phi = f(I) \quad \text{for } I = \frac{\dot{\gamma} d}{\sqrt{P/\rho_s}} \tag{3.9}$$

$$\mu = g(I) \quad \text{for } \mu = \frac{\tau}{P} \tag{3.10}$$

In the above, Φ is the packing fraction, P is the pressure on the shearing plate, τ is the shear stress, and the steady shear rate is $\dot{\gamma}$. The dimensionless number I is commonly referred to as the inertial number or normalized flow rate, and μ is the effective friction.

The simplest way to understand these equations is through dimensional analysis. The major physical quantities involved in a gravity-free simple shearing between long rough plates are the material parameters d and ρ_s, and the variable quantities Φ, τ, P, and $\dot{\gamma}$. These are the only relevant quantities if we tacitly ignore the possibility of any other length-scales playing a role and presume that collisions are fully dissipative (pressure high enough to damp out restitution), two assumptions whose consequences are important and will be discussed shortly. The particle-on-particle contact friction μ_p is also ignored. Granted, μ_p does affect μ, but it has been found to merely translate the μ vs. I relationship vertically [33].

Now, suppose the plates of the shear cell are pushed together with pressure P and

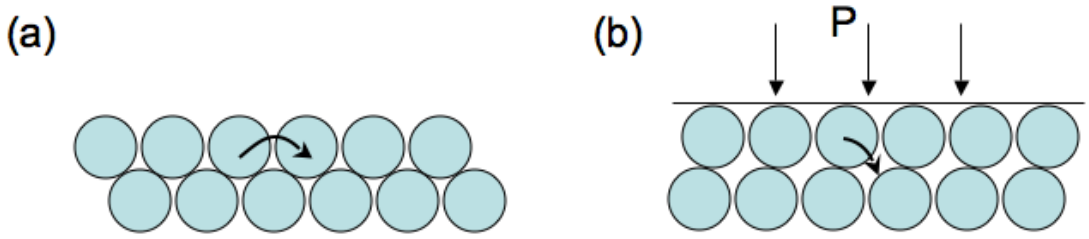


Figure 3-2: (a) Illustration of the macroscopic time scale $1/\dot{\gamma}$ representing the time for a particle to move one particle's width under the action of the applied shear deformation rate, and (b) an illustration of the inertial time scale $(1/d)\sqrt{P/\rho_s}$ representing the time for a particle to move one particle's width under the unbalanced action of the applied pressure.

then one of the plates is moved with respect to the other to produce a shear rate of $\dot{\gamma}$. Then the effective friction and the packing fraction should arise uniquely from P and $\dot{\gamma}$ as a matter of cause and effect. Nondimensionalizing this argument tells us that μ and Φ should depend only on I. Unlike a Newtonian fluid where a temperature time-scale exists, the quadratic dependence of stress on shear rate can be seen to arise from the fact that to scale $\dot{\gamma}$, the only choice is to use the square root of one of the stress quantities.

But dimensional arguments, while effective, sometimes leave the physics of the situation unclear. A simple picture to understand the basic dynamics of a granular material is difficult to come by since so many phenomena can happen at once (e.g. frictional sliding, rolling, elastic compression, impact dissipation, packing evolution, collisional restitution). We shall try our best, however, to gain a stronger physical grasp with the help of a 2D toy example that simplifies planar shear:

> **Toy Model**: Suppose two horizontal layers of disks are compressed against each other and the top layer is sheared across the bottom layer at some fixed shear rate $\dot{\gamma}$ and some compressive pressure P. Within each layer, the particle positions are fixed and equally spaced, and particles cannot rotate, so each layer is essentially a "bumpy surface". (See figure 3-2)

This toy model, which ellaborates upon a similar one introduced in [33], will

be used as a tool for understanding throughout this work. With grain positions fixed within each layer, dilation is constricted to the vertical direction. Collective rearrangement phenomena is absent from the toy as well, as the two layers have constrained horizontal degrees of freedom. The blocking out of rotations may be less of a concern, since many interesting granular phenomena carry over for frictionless particles. What the toy does provide, is a simple framework within which to attempt to explain the rheological properties of granular media. Geometrically, it maintains a basic picture of the dynamics of grains passing through a field of other grains.

I is a quantity that compares a macroscopic time-scale $1/\dot{\gamma}$ to an inertial time $(1/d)\sqrt{P/\rho_s}$. To understand the meanings of these scales using the toy, observe the process by which a particle in the upper layer traverses over a bottom-layer particle. The duration of the entire process is represented by the macro-time. Generally speaking, once the upper particle has climbed past the crest of the lower particle, it "snaps downward" under the applied pressure P (see figure 3-2b). The inertial-time represents the duration of the snap-down, or more specifically, it is the time required for the upper particle to be pushed downward a distance proportional to d under the unbalanced action of the confining pressure.

3.4.2 Flow regimes

It is important to clarify the various conditions that underlie the validity of equations 3.9 and 3.10 for use in a general dense flow law. In our dimensional argument, we presumed zero collisional restitution. The assumption is valid if the energy of a collision is always dissipated on impact, presumably in the form of heat and sound. Particles are in fact *visco*-elastic bodies, and while viscous dissipation is inconsequential for granular statics, it can be responsible for a significant amount of energy loss during flow. Here again, the toy picture is instructive. Consider what happens at the end of the inertial phase, as the upper particle is about to land on the next lower particle (see figure 3-3a). If I is small enough, the applied pressure will be sufficiently large compared to the shear rate to ensure the upper particle does not bounce to any noticeable extent on impact. We could think of this as immediate collapse upon

impact— The competition between applied compression and inter-particle repulsion is "won" by the compression, resulting in a short-lived, sound-inducing oscillation that immediately damps out the momentum in the normal direction.

Rate-sensitivity is a signature of this type of flow, as faster flows have both a higher rate of contact formation and dissipate more energy per impact. Or in terms of our dimensionless variables, I becomes one-to-one with μ, and $\dot{\gamma}$ is immediately determined by τ and P as in a non-Newtonian fluid. Therefore, as a general classification scheme, we say that a flow rate is "moderate" when I is large enough for rate-dependence, but small enough for the flow to remain dense as per the collisional collapse argument. Data of da Cruz *et al.* [33], would suggest this regime lies within the band $10^{-3} < I < 10^{-1}$.

Moderate flows have the property of *shearing dilation*, where increasing the normalized flow rate causes the steady-state packing fraction to decrease (i.e. f in equation 3.9 becomes a decreasing function). This should not to be confused with *shear dilation*, which refers to the drop in packing fraction as a function of total shear that occurs to a dense assembly at the beginning stages of a shear deformation. Flows too slow to be deemed moderate may still undergo shear dilation due to geoemtric packing constraints, but rate effects like shearing dilation only set in for faster flows.

The toy offers a cute illustration of shearing dilation in moderate flows. Let us observe the dynamics of the inertial phase by studying the trajectory of one upper particle. Presume a basic situation where the upper particle loses contact upon passing the crest, and is flung against the next particle beneath it (as pictured in 3-3a) resulting in impact dissipation. We may use kinematics to understand this ballistic trajectory. With pressure and shear rate fixed, this is akin to a projectile accelerating downward by a force proportional to P and moving laterally at a speed proportional to $\dot{\gamma}$. The parabolic trajectory and the height of the lift-off point is determined by $\dot{\gamma}/\sqrt{P} \propto I$, with wider trajectories associated to higher values of I. If I increases, the particle lands higher, implying that the mean separation between layers increases. Consequently, we have a picture for why the time-average packing fraction decreases in response to increasing I. In the limit as applied pressure goes to

zero, this effect becomes even more obvious— shearing without any pressure induces no compaction.

Of course, the toy is heavily oversimplified. It is exceedingly rare for a grain in a 3D flow to ever lose contact with all its neighbors, so ballistic trajectories are overreaching. But the general idea is that as the macro-time shrinks with respect to the inertial time, particles have less and less ability to compactify under the confinement stresses before being "tugged out" by macroscopic agencies.

Dilute or "collisional" flows occur in general for $I > 10^{-1}$ and correspond to the breakdown of the zero restitution assumption. When I becomes this large, particle collisions are accompanied by some additional "bounce-back" akin to a gas (see 3-3c). The collisions are chiefly binary, and particles rarely maintain long lasting contacts. These flows require a temperature quantity to store fluctuational energy and are well-described by dissipative Boltzmann kinetics.

On the other side of the spectrum, where $I < 10^{-3}$, we enter the "quasi-static" regime where intermittent motion is prevalent. The packing fraction in simple shear does not vary noticeably with I in the (time-averaged) steady limit— the inertial time is always small enough for the particles to find tight compaction (see 3-3b). Without a significant contribution from collisional dissipation, rate-dependence subsides, and more complicated dissipation mechanisms dominate like frictional sliding and stick-slip dynamics. The stress/strain-rate relationship becomes singular; driving the system with a range of quasi-static normalized shear rates all give the same time-average value for μ. This is definitive rate-independence, in the sense that the size of the strain rate variables cannot be deduced from the size of the stress variables. It is akin to the rudimentary picture of a block sliding on a table, where as long as the applied force equals $\mu_k N$ for normal force N, a moving block can be made to slide at any speed.

Our discussion thus far has focused almost entirely on describing simple shear, where stresses and time-average flow are spatially uniform. While the other regimes display a strong local rheology, it has been observed that quasi-static flows are sensitive to gradients in the fields as expressed through some non-local term with a new

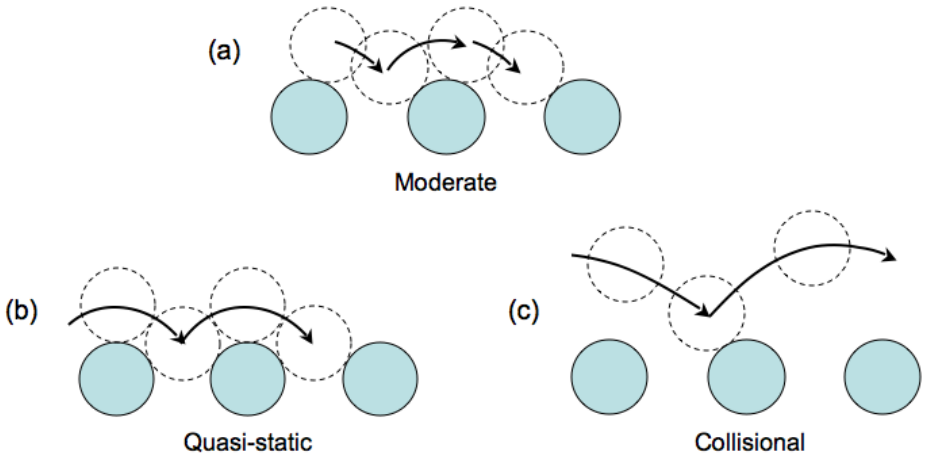


Figure 3-3: Extending the cartoon in Figure 3-2 to understand the various flow regimes. (a) Moderate ($\sim\ 10^{-3} < I < 10^{-1}$): Dissipation primarily rate-sensitive due to energy loss during contact formation, yet packing remains dense. (b) Quasi-static ($\sim\ I < 10^{-3}$): Dissipation primarily frictional and rate-independent. Packing fraction appears independent of I, and grain-level specifics are more important to flow dynamics. (c) Collisional ($\sim\ I > 10^{-1}$): Flow becomes dilute and gas-like. Dynamics modeled best by dissipative Boltzmann kinetics.

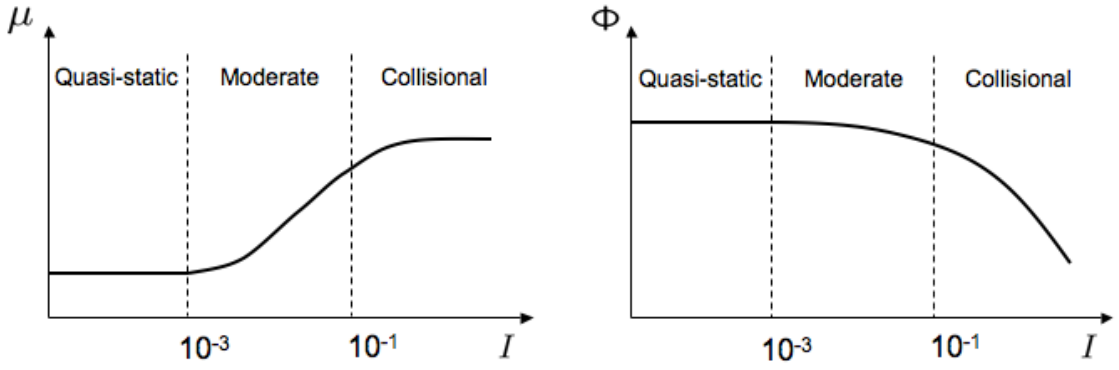


Figure 3-4: Qualitative diagrams (primarily due to [33]) showing the variation of dimensionless parameters through the various flow regimes under simple shearing. In other geometries, the quasi-static regime of μ is not as clear to define in terms of I due to meso-scale effects. The coefficient of restitution affects Φ vs. I in the collisional regime. The moderate regime is relatively well-determined as a sole function of I.

length-scale [94, 15, 73, 8]. For example, consider steady flow in an annular Couette cell. Slow moving material far from the inner wall is observed to constantly creep [94], even though the stress state should be below yield by all common local measures. Indeed, the motion appears to be caused by a non-local effect where faster flow near the inner wall has effectively "bled out" into neighboring material. Grain-level specifics such as roughness, grain shape, and configurational statistics (including wall effects) should affect the non-local flow behavior through the new length-scale. The size, dynamics, and general interpretation of the length-scale are object of debate, though most agree its size should be on the order of several particle widths.

Observe figure 3-4 for a schematic view of how μ and Φ vary throughout the flow regimes (in simple shear). The major take home points of this discussion are: moderate flow is much simpler than quasi-static, and collisional flow is outside our interest as it is not dense. Moderate flows should be characterized by a local rheology relating I to both μ and Φ. The regime is definitively rate-dependent, so equation 3.10 inverts into a fluid-like law wherein the flow rate can be determined uniquely from the stress state. As I decreases to the quasi-static limit, quantities we were previously able to ignore become important. Grain motion is locally correlated at some length-scale ushering in a greater role for grain-level properties. The flow does not permit a fluid-like treatment as before, since dissipation is largely due to rate-independent mechanisms. The motion is indeed hardly fluid-like, as it is intermittent, involving small, rapid failure events separated by long periods of effectively no movement.

3.4.3 Review of past flow models

It is worthwhile to discuss past work on dense flow in terms of quasi-static and moderate flow behavior. Quasi-static models are commonly applied in situations of small-strain such as soil deformation and high pressure geophysical phenomena. To obtain stresses during flow, a stress law must be included that determines stress invariants based on behavior unrelated to the plastic flow rate. Elasticity is normally employed, where the total deformation has a small elastic component that determines the size of the stress. Since the elastic response is not the focus of these models,

simple linear elasticity tends to be used, though as previously noted, this may be an oversimplification that carries more consequences at lower stresses.

As a brief review, Critical State Theory [127], Rudnicki-Rice-type modeling [119], and the recent model of Anand and Gu [3] each can be used for granular deformation in this fashion. These models propose various forms for how the packing fraction evolves with stress, how the yield parameters evolve with packing fraction (and what the yield surface looks like in stress space), and how the directionality of the stresses at yield determines the directionality of the plastic deformation. These approahces have had some recent direct validation in small-strain discrete simulations [118]. Other rate-independent models couple to the IFE stress formulation in 2D, such as pure coaxiality [101] and the Stochastic Flow Rule [73].

While rate-independent models have had success, moderate flow rates require rate-sensitivity. A non-hardening flow model that leaves out rate effects will predict that stresses exceeding static yield induce a shear-rate that will continually increase without bound. An analogy in the block-table friction picture would be to pull a moving block with a force above $\mu_k N$; the block always feels a net force and thus consistently accelerates unless the force is removed. Rate-independent models commonly use displacement control to circumvent "runaway". But in many flows of interest to us, tractions or body forces (gravity) drive the motion, not kinematic boundary conditions, and yet steady flows are still observed. Bagnold's results among others previously cited verify the existence of the moderate flow regime, where non-accelerative, dense, steady flow occurs for stress states higher than those which initiate yield, consistent with our picture of the moderate regime.

Beyond rate-sensitivity, it seems the bigger issue in representing quasi-static and moderate behaviors in one model is the fact that local rheology dominates moderate behavior (that is, equations 3.9 and 3.10) whereas a more complicated non-local rheology emerges when the flow slows down. Some theories that may connect the moderate and quasi-static regimes are based on new definitions of temperature. It is plausible that non-locality could follow from the "heat equation" in which a Laplacian is scaled by an imposed length. Flow models of this type include Shear Transforma-

tion Zone (STZ) Theory [83, 46], which relies on an effective temperature governing STZ creation, the dense flow theory of S. B. Savage [124], which defines a granular temperature to measure strain-rate fluctuations, and Edwards statistics [42], which utilizes a temperature-like 'compactivity' derived from an entropy per free volume. These models provide interesting physical insight, but do not appear to be at the point of development that simulating arbitrary flows would be possible— some are restricted to 2D, the boundary conditions for the new temperature are rarely obvious, and the equations may not be closed except in a few symmetric test cases. For our purposes, complete closure under arbitrary 3D flow situations is desired.

The quasi-static flow regime, though important, appears at the moment to be too difficult to account for appropriately within a simple continuum framework. Thus, we shall neglect any further attempt at representing quasi-static behavior, and opt instead to construct a model combining a statics law with moderate rheology. This concession is not as harsh when viewed in light of our initial goal: We seek a formulation to understand the first-order dependence of stress on the deformation of an RVE. The effects of an imposed length-scale would be of second-order dependence, most likely scaling a second spatial derivative of some sort. Besides temperature approaches, this could occur through the inclusion of a diffusing state parameter [73, 8] or through a more general strain-gradient plasticity theory [56, 57]. It is hoped that the model we are about to construct may serve as the backbone for a fuller model that also incorporates the dependence of a length-scale on the slow dynamics. This possibility shall be considered in more depth when we compare predictions of the model directly to experimental/DEM data. But for now, we accept inaccuracy in describing quasi-static motion in exchange for a closed, general model capable of giving worthwhile predictions over the full range of dense material behavior, accounting for both statics and flow.

3.4.4 The Jop–Pouliquen granular plasticity law

A closed form law to predict moderate flow must now be selected. We start from equation 3.10 for it states that there should be a function g that relates I directly to

μ without knowing the packing fraction. Since the moderate flow regime is (monotonically) rate-dependent, we have the bonus that g should be invertible. Recall that increasing μ corresponds to increasing I. This may seem counterintuitive, since it contradicts the notion that μ should *decrease* to a kinetic value as the rate of sliding picks up. But recall, for moderate flow rates, the impact dissipation is the more significant effect. In slower flows, shear weakening is indeed observed for "overconsolidated" material, and accounted for in various models via hardening parameters [3, 127].

Following general results from numerical simulations of planar shear [33, 63], and successful extensions to plane-strain inclined chute flows [94, 130], the experiments of [70] were conducted to quantify g^{-1} for glass beads:

$$I = g^{-1}(\mu) = I_0 \frac{\mu - \mu_s}{\mu_2 - \mu} \quad \text{for } \mu > \mu_s. \tag{3.11}$$

The values of the parameters were measured at $I_0 = 0.279$, $\mu_s = \tan 20.9°$, and $\mu_2 = \tan 32.76°$. The relation states that the normalized shear rate I increases as the material is sheared with higher μ. But μ must exceed some static yield value μ_s before any plastic flow ensues. There is also some maximal μ value called μ_2, and all steady shear rates should be tenable for a value of μ less than μ_2. If the material is stressed with an applied μ that exceeds μ_2, the shearing motion will be accelerative. In other words, when $\mu > \mu_2$, all typical dissipation mechanisms become overloaded by the applied shear stress, causing a continual increase in the shear rate.

The inclined plane geometry in 2D is quite instructive here. The slope of the incline gives a good approximation for μ. For slopes less than μ_s, no flow occurs. Between μ_2 and μ_s the material flows steadily down the incline. Above μ_2, the material accelerates down the incline under gravity's pull.

After a long debate concerning the importance of side walls in experiments of inclined plane flow, the rheology of equation 3.11 was accepted and attempts at converting this simple shear law into a general 3D flow law were made. The first attempt was met with high success. In Jop *et al.* [71], codirectionality was attempted,

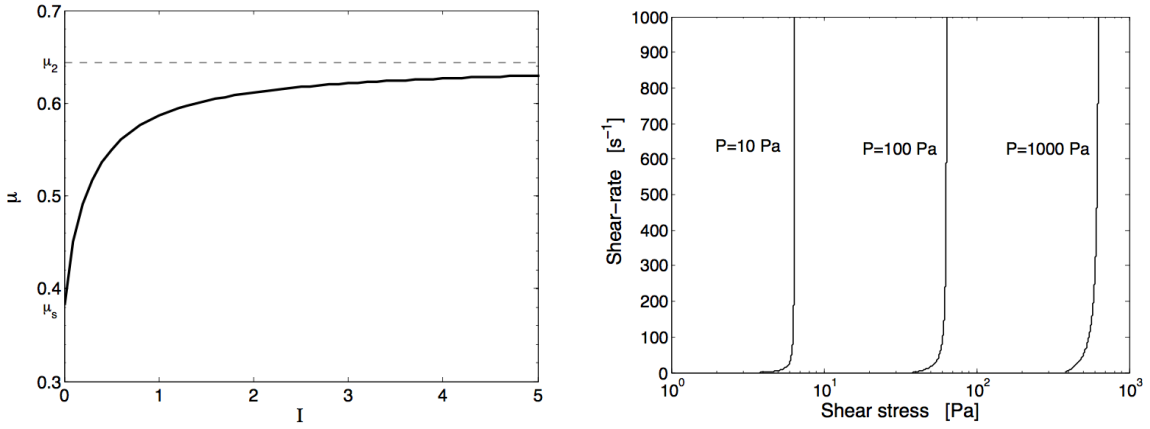


Figure 3-5: The plastic flow rheology in simple shear plotted as one dimensionless relationship (left), and plotted equivalently as a family of shear-rate vs. shear stress curves in SI units, each curve determined by the applied compressive pressure (right).

which presumes the deformation rate tensor

$$\mathbf{D} = (1/2)\left(\boldsymbol{\nabla}\mathbf{v} + (\boldsymbol{\nabla}\mathbf{v})^T\right)$$

is proportional to the deviatoric stress tensor

$$\mathbf{T}_0 = \mathbf{T} - (1/3)(\mathrm{tr}\mathbf{T})\mathbf{1}.$$

Written in full, Jop proposed the following generalization of equation 3.11:

$$\mathbf{D}\frac{d}{\sqrt{P/\rho_s}} = I_0\frac{\mu - \mu_s}{\mu_2 - \mu}\frac{\mathbf{T}_0}{\tau} \tag{3.12}$$

where now $P = -(1/3)\mathrm{tr}\mathbf{T}$ and $\mu = \tau/P$ where $\tau = |\mathbf{T}_0|/\sqrt{2}$ is the equivalent shear stress. When $\mu < \mu_s$, we take $\mathbf{D} = \mathbf{0}$ establishing a yield criterion. In particular, this 3D extension of frictional yielding is the Drucker-Prager criterion. Since the flow condition being used is codirectionality, the law is non-associative in the sence that $\mathbf{D}$ is not directed normal to the yield criterion. The codirectionality hypothesis implies that when flow occurs, its does so in a manner analogous to an incompressible fluid.

Since $\mathbf{D}$ is proportional to a deviatoric tensor, the flow rule asserts plastic incom-

pressibility. We know that dilation in dense flow does occur, but it is typically on the order of only a few percent and quickly reaches a steady value over large deformations. We would still need to keep track of Φ as a state parameter were it not for the fact that the (steady) shear rate can be uniquely expressed in terms of the stress. This implies that the assertion of plastic incompressibility should have little effect on the velocity field of a dense steady flow. However, it should be pointed out that the evolution of the packing fraction to its steady value has not yet been quantified in this context. Some quasi-static flow models attempt this [3, 127, 119] but rather than try to modify one of these, we go along with the presumption of the Jop–Pouliquen model and ignore plastic dilation. The bigger impact of this assertion is not on the flow, but rather on the stresses in the static regions— If plastic flow transiently passes through a region that becomes static in the steady-state, how much dilation occurred there is important since the elastic moduli depend on the packing fraction.

Codirectionality is one of many possible ways of directing the stress with respect to the deformation rate in 3D. Others have proposed double-shearing [3] based on A.J.M. Spencer's original formulation [133]. Unlike codirectionality which permits deformation in all 3 dimensions when the Drucker-Prager failure criterion is met, double-shearing constructs $\mathbf{D}$ when Mohr-Coulomb failure is met, by only allowing plastic sliding (and possibly dilation) to occur on the two internal slip-systems that satisfy Mohr-Coulomb yield (see appendix A for more details). The DEM results of Depken *et al.* [35] show a strong agreement with the codirectionality hypothesis, but less favorable agreement was observed in Rycroft *et al.* [122].

In non-Newtonian fluid mechanics, a fluid with stress-sensitive viscosity that, in particular, remains rigid when below a yield criterion is known as a *Bingham fluid* [18]. Accordingly, equation 3.12 represents granular matter as a Bingham fluid obeying a Drucker-Prager yield criterion. Inside rigid zones, Bingham fluids have undefined stresses. However, even without knowing the rigid stresses, the velocity profile everywhere can be uniquely determined from equation 3.12 as the limitting solution obtained by letting viscosity go to infinity in sub-yield regions. Algorithms that solve for Bingham fluid flow are more complicated than Newtonian fluid algorithms since

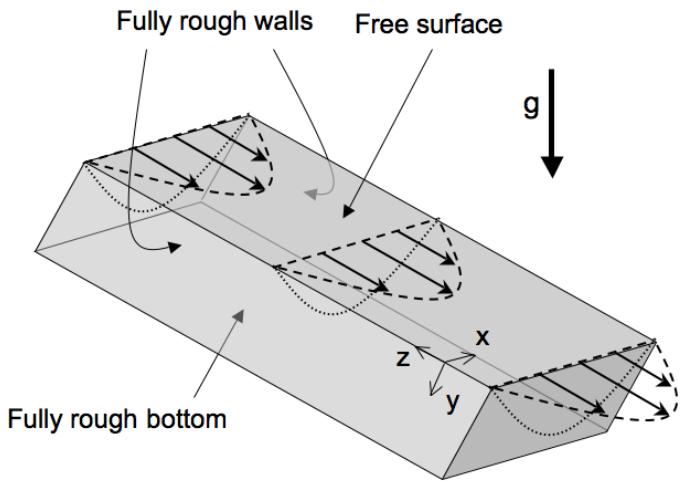


Figure 3-6: The geometry utilized by Jop *et al.* [71] to determine if codirectionality appropriately extends Bagnold scaling to 3D.

the Navier-Stokes-type equations cannot be directly implemented in solid zones, where infinite viscosity and a zero flow rate cause $\infty \times 0$ to appear. Instead, the solution is numerically approached using higher-level iteration and/or variational methods such as those presented in [116, 50]. Some materials commonly described with Bingham fluid models include mud [93, 144] and paste [89].

To test the law, Jop *et al.* solved for steady flow in a non-standard geometry: a long inclined chute with rough bottom and rough sidewalls (see figure 3-6). This geometry enables a flow that is symmetric in the down-incline direction, but spatially non-trivial in the other two. The results were shown to match experiments to a high degree (always within 15%) even while varying several parameters (e.g. inclination angle, flow height and width). The model was heralded as a major breakthrough.

While the results are promising, there are several reasons why this Bingham fluid treatment of granular matter is not enough:

1. There is no stress computation in the static regions. The Bingham approach only gives the stresses in the flowing regions.

2. It is not compatible with arbitrary traction boundary conditions. If an applied traction happens to be adjacent to a solid-like region of grains, the traction

cannot be accounted for.

3. While codirectionality does well, we may not be able to test other flow conditions under Bingham fluid assumptions. In the general case of double-shearing, for instance, the intermediate principal stress is disconnected from the flow rate. A Bingham fluid approach would be unable to determine this stress causing ambiguities when balancing momentum.

4. Fluid approaches miss some distinctive granular behaviors such as stick-slip and bistability.

5. Bingham fluid algorithms commonly have numerical issues near the fluid/solid interface, where the viscosity undergoes a discontinuity.

Many if not all of these drawbacks will be resolved when the Bingham fluid approach is replaced with elasto-plasticity. But properly mixing an elasticity formulation with a plastic flow law takes some theoretical development. We henceforth proceed to determine the proper way to unite the Jop–Pouliquen plasticity law (equation 3.12) with the Jiang–Liu elasticity law (equation 3.8).

3.5 Combining elasticity and plasticity: Physical considerations

3.5.1 A simplified picture

At the outset, this problem would appear to be quite confusing. Over the course of a complicated deformation path, how is one supposed to disentangle the elastic and plastic pieces of the end deformation? We shall attempt to solve this problem using a popular decomposition originally proposed by Kröner and Lee in the 1960's [78, 81]. Let us motivate the decomposition with a cartoon. We model a granular element in an admittedly over-idealized fashion as two elastic rectangles stacked one on top of the other, with μ representing the surface friction between rectangles (see figure 3-7).

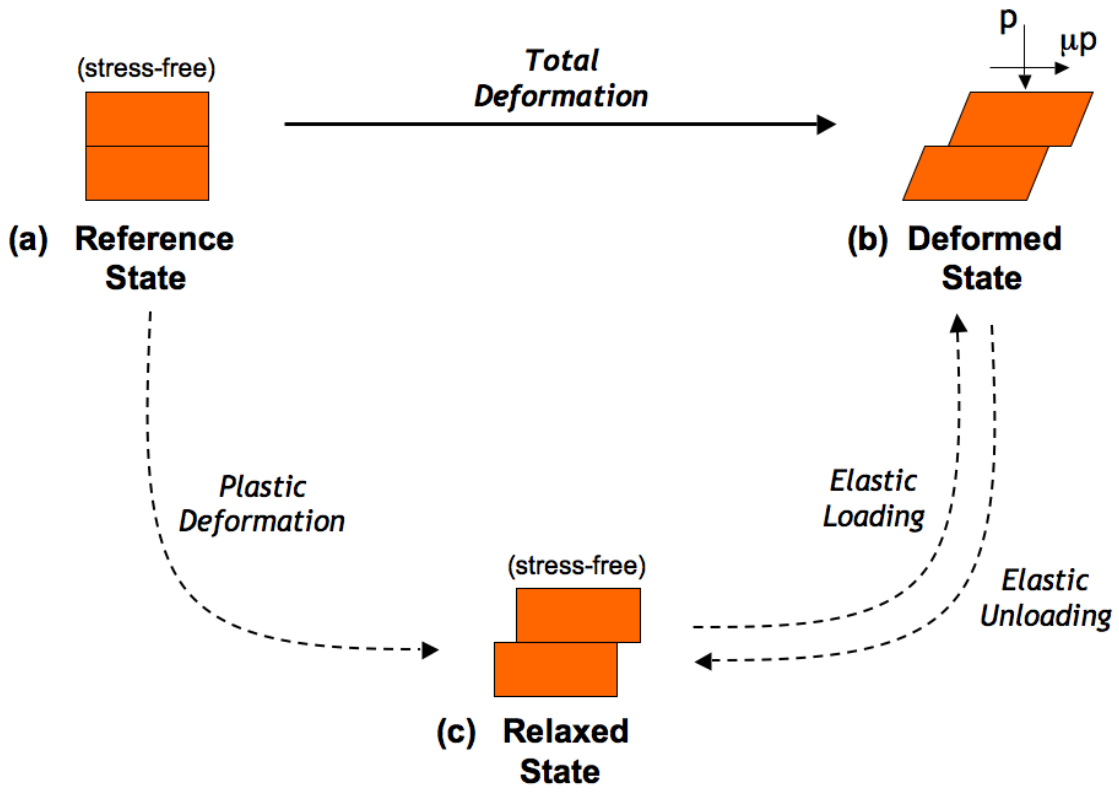


Figure 3-7: (a) The blocks start in the "reference state", which we imagine is the original stress-free configuration. (b) The application of a yield stress makes one block slide with respect to the other at some constant sliding rate. A snapshot of this at some time is called the "deformed state". (c) By removing the stresses that act on the deformed configuration, the elastic part of the total deformation unloads, revealing an underlying "relaxed state".

As a boundary condition, we constrain the bottom surface of the bottom block vertically, and pin the lower left corner of the bottom block. Let us consider a starting state, the so-called "reference state", where the blocks are unstressed and have not undergone any sliding or elastic deformation. Stresses are then applied to the top surface of the top block inciting motion. Any arbitrary motion may occur: the top block may slide side-to-side, it may even lift off the surface of the lower block at some point. Then, at some point in time later, we take a snapshot of the situation and call it the "deformed state". It represents some total deformation from the reference state. As pictured in figure 3-7b, let us consider a case where the top block has already slid some distance and is in the process of sliding when the snapshot is taken.

Note, as indicated in the diagram, that while sliding, the blocks themselves maintain an elastic deformation. The reason for this is simple; observing the stresses on each block during sliding, it is clear that each block is still in a state of uniform stress and elastic solids must deform under stress.

With everything that could have happened inbetween, how can one decompose the deformed configuration into a unique elastic and plastic contribution? A good way to do this is to simply release all the stresses from the deformed state. What remains is a "relaxed state" in which no additional sliding has occurred, but the elastic mechanism for supporting stress is relieved. Directly connecting the reference state to the relaxed state defines a unique plastic deformation. Given the relaxed state, one can bring the system back to the deformed state by reloading it exactly opposite to how it was unloaded. Thus, however the deformed state may have actually been achieved, the total deformation can be reconstructed in two successive steps: first let the reference state be mapped to the relaxed state by a purely plastic deformation, and then elastically load the relaxed state to the deformed.

This picture still applies, but in a less obvious sense, if the coefficient of friction is rate sensitive. Suppose μ increases with the sliding rate and the deformed snapshot has the top block in steady motion above static yield. The elastic deformation of the blocks should still reflect the total stress, but the idea of a relaxed state is less clear—One cannot go from a stress-free intermediate state to the deformed state without invoking additional plastic sliding. That is, suppose the deformed state stresses are released. Any loading path that brings the blocks back to the deformed state stresses over a finite time period cannot be purely elastic, as the friction state must pass through static yield during the process. We can understand the aforementioned elastic/plastic decomposition in this situation under the added pretense that we "fix" the plastic deformation while loading from the relaxed to the deformed configuration. Or, an equivalent patch is to say that the elastic loading step happens very fast, so that additional plasticity has no time to occur during loading.

This basic picture provides intuition for how we shall eventually decompose any arbitrary 3D deformation of a granular element into elastic and plastic parts. Other

approaches exist to handle elasto-plasticity, such as hypo-plasticity, which deals explicitly with Eulerian stress rates and strain rates instead of separating the elastic and plastic parts of a total deformation. Hypo-plastic approaches have lost favor in the continuum mechanics community over the last decade as thermodynamic effects are difficult to model and elastic strains must be negligibly small. While hypo-plasticity could conceivably be applied to a granular material, we opt instead for the more rigorous hyper-plastic approach sketched above. As shall be seen, the clear separation of elastic and plastic parts makes the process of specifying connections to state parameters and stresses much clearer and less arbitrary.

Several pieces of the puzzle are still missing though. Granular materials are definitely more complicated than the cartoon. Can the microscopic level of flow be reconciled with the notion that elastic stresses always reflect the total stress? And then there are mathematical details: Are rotations accounted for in the elastic or the plastic part? In what sense is the relaxed state a globally obtainable configuration and to what extent is it a useful albeit fictitious aid? We also must quantify precisely what a deformation means mathematically before rigorously referring to related quantities such as strain and deformation rate.

We answer these questions successively. First let us give some physical evidence for representing the stresses in a flowing granular material within the elasticity of the grain medium. Doing so shall require a small modification to the Jiang–Liu law. We shall then proceed through a mathematical theory of 3D, incompressible, athermal, finite-deformation elasto-plasticity.

3.5.2 Microscopic evidence for elasticity during flow

According to the basic picture shown in figure 3-7, the stress tensor can always be well-described, even in flowing regions, in terms of how much the grains elastically deform. Of course, in a static assembly it is obvious that no other possible stress mechanisms exist. But during flow, a couple other agencies are available to reflect additional stress. If we wish to use the elasto-plastic framework above, we need to provide some evidence for why these other stress agencies are negligible.

First of all, let us write out the complete form for the stress tensor within a granular element in terms of the behavior occurring during a snapshot. As can be obtained from a basic internal/external power argument, the stress formula is:

$$\mathbf{T} = \frac{1}{V}\left(\sum_{i<j}^{N} \mathbf{r}^{(ij)} \otimes \mathbf{f}^{(ij)} \; + \; \sum_{i=1}^{N} m^{(i)}(\mathbf{v}^{(i)} - \bar{\mathbf{v}}) \otimes (\mathbf{v}^{(i)} - \bar{\mathbf{v}})\right) \tag{3.13}$$

The average velocity within an element is $\bar{\mathbf{v}}$ and $\mathbf{v}^{(i)}$ is the velocity of the i^{th} particle. The right-most term on the right-hand side is new (compare to the static-only case, equation 3.3), and we shall refer to it as the "gas-like" contribution. We call it gas-like because the stresses in an ideal gas are entirely of this form, where velocity fluctuations about the mean represent a time-average momentum transfer rate. Physically speaking, the gas-like term accounts for any acceleration of the individual grains caused by boundary tractions as measured in the element's center-of-mass reference frame. The gas-like stress term is an important agency to consider when assessing flow-related extra stress.

According to the data of Chris Rycroft, it so happens that the gas contribution is remarkably small in DEM simulations of moderately flowing sphere packings. So small, that it can be neglected from the stress computation with essentially no consequences. This evidence verifies our picture of the moderate flow regime as one in which the kinetic energy from new contact formation is quickly damped out by viscous processes within the particles. Together with the denseness of the flow, particles never have the chance to accelerate to any appreciable velocity with respect to the element center. Thus, if the grains were at all reminiscent of a gas, the "temperature" of the grains would be close to absolute zero.

What remains is a formula for the Cauchy stress in terms of the contact forces $\mathbf{f}^{(ij)}$ and their corresponding contact separations. Since the grains are visco-elastic, each contact force can be decomposed additively into elastic and viscous parts, as in a damper/spring system in parallel:

$$\mathbf{f}^{(ij)} = \mathbf{f}_e^{(ij)} + \mathbf{f}_v^{(ij)}$$

Perhaps it is not surprising given the previous zero temperature observation, but it is also observed in DEM simulations that the viscous force contribution is incredibly small compared to the elastic, implying

$$\mathbf{T} \cong \frac{1}{V} \sum_{i<j}^{N} \mathbf{r}^{(ij)} \otimes \mathbf{f}_e^{(ij)}$$

The DEM evidence indeed suggests that elastic stress contributions are all that is necessary to represent the full Cauchy stress in a granular element. In other words, just given the spring-like compressions and shears in every grain (that is, δ and s as defined in section 3.3.2), Hertz-Mindlin elasticity is responsible for the full stress tensor even during flow.

It is understandable if this point is somewhat non-intuitive. Unlike the picture in figure 3-7, rate-dependent effective friction occurs in moderate flowing grains even though grain-on-grain sliding friction is modeled as rate-independent. We shall attempt to explain this particular detail with the help of the section 3.4.1 toy model. But no toy model alone suffices to justify the rather important and unobvious fact that the elastic agencies within the material provide all the stresses, even during flow. For this, we must rely on the evidence from DEM simulations. We are left to believe that the simple block-on-block cartoon for decomposing and handling elasticity with plasticity is a plausible starting point for an elasto-plastic model for granular matter.

3.5.3 Modifying the Jiang–Liu elasticity law

As motivated in the last subsection, we need to modify the Jiang–Liu elasticity law to permit elastic behavior during flow. Since flowing materials can have μ increase above μ_s, the elasticity must reflect this. Hence, we propose the following important modification to Jiang–Liu elasticity: Set $\xi = \mu_2^{-2}$ instead of $\xi = \mu_s^{-2}$.

The ξ parameter was engendered solely to capture the macroscopic repose angle of a static granular assembly. Its value was not determined from any quantitative microscopic requirements. So we have few qualms about this adjustment, as it only improves the law's ability to represent macroscopic phenomena.

With $\xi = \mu_2^{-2}$, there can never be an elastic stress state that has $\mu > \mu_2$. Looking back to equation 3.12, notice that the flow rule would completely break down if given a stress state that exceeds μ_2. Since the Jiang–Liu law admits a cap on the value of μ, one may use this property to prevent elastic stress states from entering the forbidden regime of the Jop–Pouliquen flow law. This feature is important when performing explicit time integration, where stresses outputted from the elasticity relation are used to compute the plastic flow rate the next time step. This is not an issue, however, for implicit time integration methods. Analytically speaking, capping the elastic stresses at μ_2 is not necessary for a solution; an elasticity law permitting stresses above μ_2 would also be acceptable, though the plastic response would always preclude such states from arising. In this regard, capping the elasticity at μ_2 is not a mathematical necessity, but a choice. A larger cap could be used, but for the best carry-over from Jiang and Liu's original formulation, we cap at μ_2 since it is closest to μ_s.

We already have DEM evidence indicating that elasticity describes the stresses during flow. While this point is difficult to prove from basic physical arguments, we may be able to explain a particular byproduct of this general fact. Let us try to understand the elastic basis for why μ continues to increases as I increases in the moderate flow regime. What complicates this question is the fact that the interparticle contact friction actually plays only a small role in the development of the effective friction μ. Most notably, grains flow at a steady rate with a non-zero effective friction even when the coefficient of interparticle contact friction is zero.

Since contact forces dominate the calculation of the stresses, let us focus on how the statistics of contact forces imply different effective frictions given different normalized flow rates. In [33], simple shear simulations were performed on a 2D collection of bidisperse disks with zero contact friction. The distribution of contact forces as a function of angle from the horizontal was computed for many I values. With zero contact friction, all forces on a particle are normal to the contact surface. The effective friction is the ratio of equivalent shear stress to pressure. The more anisotropic the contact force distribution is, the higher the effective friction. For example, a completely isotropic force distribution acting on a grain, as pictured in figure 3-8a,

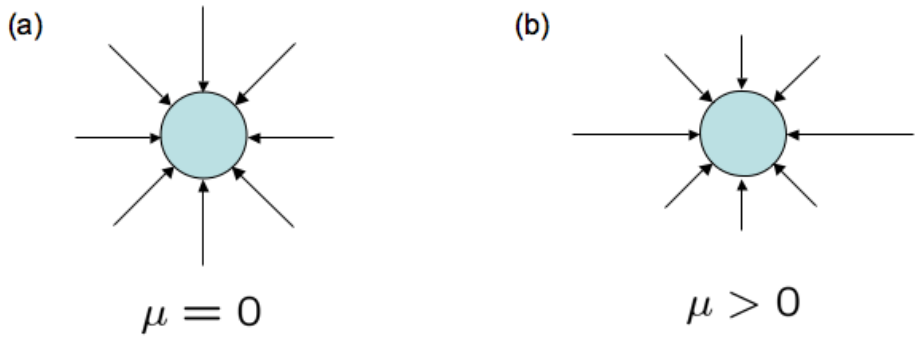


Figure 3-8: (a) Isotropic force distributions on the surface of a disk make for a stress tensor with zero shear stress and likewise $\mu = 0$. (b) As the distribution becomes more anisotropic, the μ value associated to the corresponding stress tensor grows.

sums to a stress tensor that is purely compressive with zero equivalent shear stress, and consequently $\mu = 0$. If the contact force distribution skews in one direction, as in figure 3-8b, the stresses gain a deviatoric part that is larger for larger amounts of skew.

The data of [33] show unambiguously that flows with higher normalized shear rates are accompanied by a contact force distribution that is more anisotropic (see figure 3-9 [1]). As I increases, contact forces appear to concentrate more and more heavily in the range $90^o(270^o) < \theta < 180^o(360^o)$ and less so in the other two quadrants. Thus, even in the case of zero contact friction, the observed presence and variation of contact force anisotropy provides a successful explanation for effective friction's dependence on I.

Once again, for intuition as to why this happens, we return to the toy model (see figure 3-3a). Again, suppose a constant downward pressure P and shear rate $\dot{\gamma}$, which is equivalent to holding I constant during the flow. Consider the case of zero interparticle friction. The contact force distribution on a lower particle as an upper particle traverses it, gives a time-average stress state of

$$\mathbf{T} = \frac{1}{V} \langle \mathbf{f} \otimes (2R\mathbf{n}) \rangle = \frac{1}{V t_{\text{period}}} \int_0^{t_{\text{period}}} 2RF_n(\delta)\mathbf{n} \otimes \mathbf{n} \, \text{dt}. \tag{3.14}$$

[1]Figure reprinted from Phys. Rev. E, F. da Cruz *et al.*, Rheophysics of dense granular materials: Discrete simulation of plane shear flows, **72**:021309, Copyright (2005) with permission from APS.

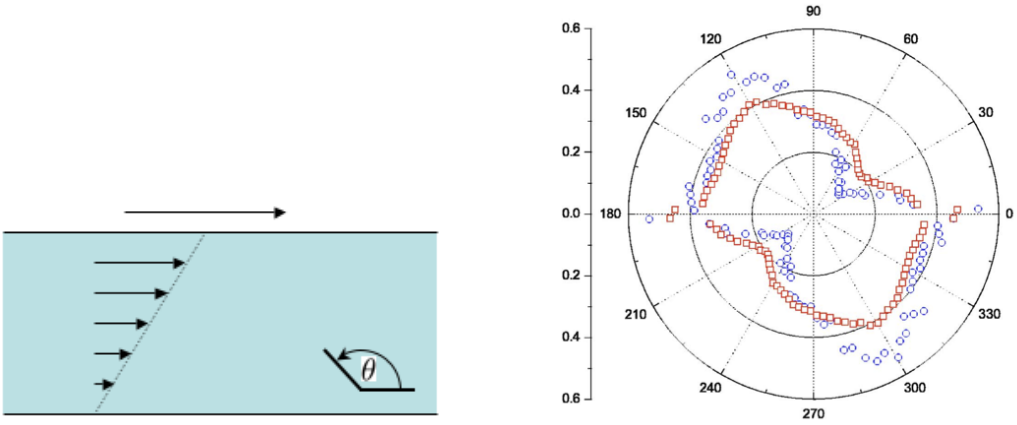


Figure 3-9: Reproduction of a major result from da Cruz *et al.* [33]. (Left) Consider the simple shearing of a collections of frictionless polydisperse disks and define θ oriented as displayed. (Right) The corresponding contact force distribution for $\mu_p = 0$ as a polar plot in terms of the angle of contact θ. $I = 0.005$ ($\square$) and $I = 0.13$ ($\circ$).

The volume V in the above is difficult to define for a two particle sample but is inconsequential in our current analysis.

As previously discussed in reference to packing fraction, increasing I causes the upper particle to leave the surface of the lower particle at a higher vertical position. It also lands higher up on the face of the next particle. Compare this to the low I extreme, where the upper particle practically never leaves the surface of the lower during one period. It is evident that the former case must correspond to a higher μ value, as the contact distribution on the lower grain extends over a smaller total angle with contacts occurring primarily on the upper left quadrant of the disk surface, and practically no contact with the back face.

The contact forces at the heart of this argument are all stemming from Hertzian elasticity. Hence, this argument gives physical reasoning for why the grain-level elasticity must be enabled to produce stress states that respect the I vs. μ interdependence during flow.

In fact, the toy model also gives physical reasoning for other bulk friction phenomena. Let us bring back the "old" definition of effective friction μ, as the ratio of shear stress to pressure on the shearing plate of a simple shear cell. This definition has μ

increase for higher degrees of anisotropy with respect to the direction normal to the shearing plane. Experience tells us that the μ required to initiate flow is higher than the μ necessary for flow to continue. This can be seen in terms of contact anisotropy in the toy example. Right before flow, static equilibrium implies the contacts are all oriented 45° from the vertical (for $\mu_p = 0$). The distribution during slow flow spreads more uniformly over the surface of the disks thereby decreasing μ. But, in concert with the previous argument, once flow has started, increasing the normalized flow rate concentrates the contact distribution in such a fashion that μ increases. The effect of contact mobilization on the force distribution is seen in this sense as a possible explanation for the bistability phenomena observed in inclined plane and heap flows at the onset of failure.

3.6 Combining elasticity and plasticity: Mathematical specifics

With the physical motivation provided, we now go about providing a mathematically rigorous framework for the model. The following derivation is based on previous theories of amorphous continuum mechanics as found in [4, 2].

At the outset, we know that Newton's equations of motion need to be upheld. In the continuum parlance, force balance takes the form

$$\nabla \cdot \mathbf{T} + \rho \mathbf{g} = \rho \dot{\mathbf{v}}. \tag{3.15}$$

where $\dot{}$ represents the material time derivative. Torque balance is equivalent to the requirement that

$$\mathbf{T} = \mathbf{T}^T. \tag{3.16}$$

These relationships comprise an underdetermined system. The equations are closed via a *constitutive process* whereby we implement our assertions about the material's constitution. In our case, a joint elastic and plastic response must be inserted.

3.6.1 Kröner–Lee decomposition

We now attempt to provide an appropriate extension of our basic picture, figure 3-7, to arbitrary 3D continuum deformations.

At $t = 0$, let $\mathscr{B}$ define the body of material in its unstressed reference state. A deformation process occurs to the body and, at the current time t, the body is in a deformed configuration $\mathscr{B}_t$. We, throughout, shall represent positions in the reference body by $\mathbf{X}$ and positions in the deformed by $\mathbf{x}$. Let us define the *motion function* $\boldsymbol{\chi}_t$ as a function that assigns every material point in the reference space to where it currently resides in the deformed space:

$$\mathbf{x} = \boldsymbol{\chi}_t(\mathbf{X}).$$

The gradient of the motion function readily admits a way to describe local deformation. Define the *deformation gradient* $\mathbf{F}$ as:

$$\mathbf{F}(\mathbf{X}, t) = \frac{\partial \boldsymbol{\chi}_t(\mathbf{X})}{\partial \mathbf{X}}.$$

$\mathbf{F}$ is a tensor with the property that when it operates on a small oriented material filament $d\mathbf{X}$ in the reference body, it outputs what that filament now looks like, $d\mathbf{x}$, in the deformed body. Hence $d\mathbf{x} = \mathbf{F}\ d\mathbf{X}$.

The idea illustrated with the two blocks can now be stated clearly in this framework. We want the deformation gradient to be represented as the composition of two deformation gradients, one elastic and one plastic. Given the idea that elastic loading is the final step in such a decomposition, we can write:

$$\mathbf{F} = \mathbf{F}^e\mathbf{F}^p \tag{3.17}$$

In this manner, a local neighborhood in the reference body deforms plastically according to $\mathbf{F}^p$, bringing it into some "intermediate" or "relaxed" space free of stresses. Then the material elastically loads via $\mathbf{F}^e$ to the appropriate deformed state configuration at time t.

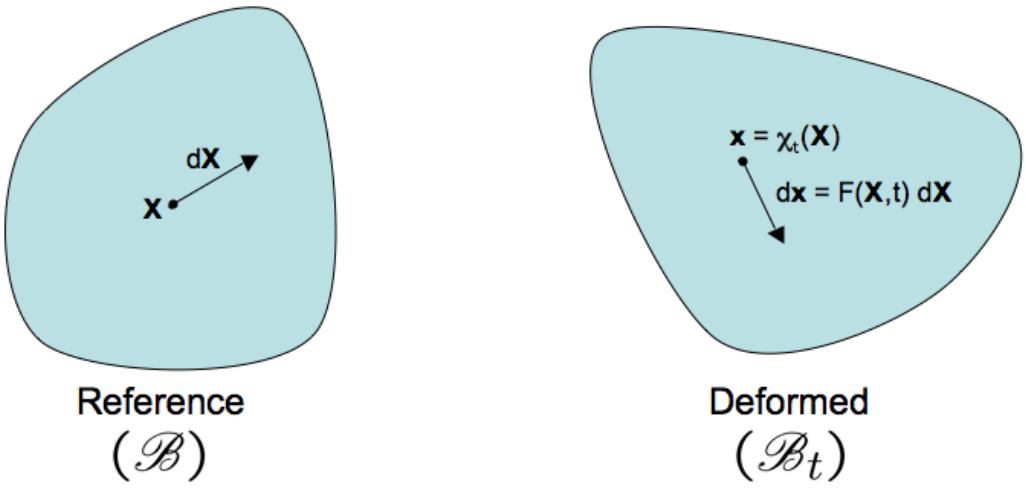


Figure 3-10: An illustration of the various continuum terms. A reference body is deformed into a deformed body. The material point $\mathbf{X}$ is located at $\mathbf{x}$ at time t and the small material filament $d\mathbf{X}$ (not to scale) that extends from $\mathbf{X}$ is stretched and rotated to $d\mathbf{x}$.

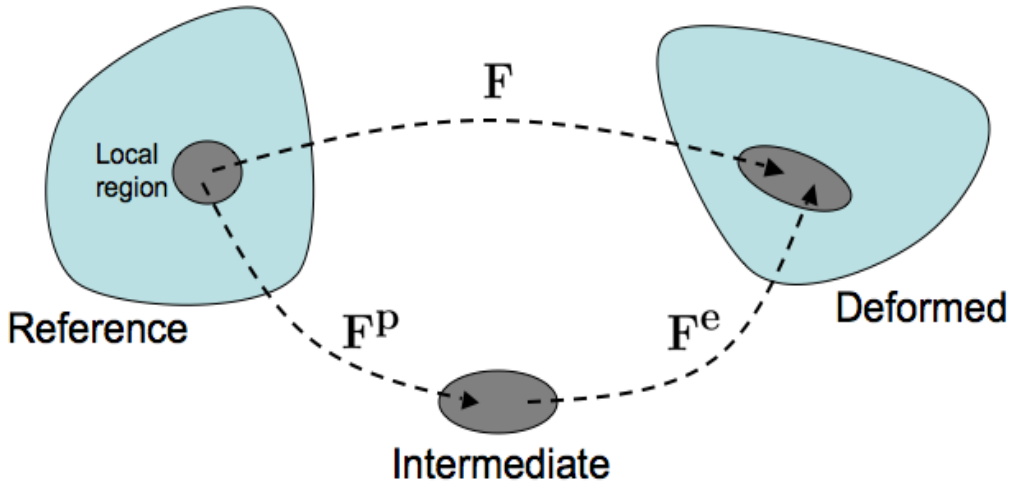


Figure 3-11: The Kröner–Lee decomposition of the deformation gradient. Locally, a material element is modeled as passing through an intermediate space where all plastic deformation is accounted for before elastically loading to the deformed space.

We do not expect that either $\mathbf{F}^p$ nor $\mathbf{F}^e$ correspond to the gradients of some global elastic and plastic motion functions. There are simple cases, like the block-on-block deformation, where elastic and plastic motions do exist and one completely relaxed intermediate body is attainable. But for inhomogeneous $\mathbf{F}$, one should expect that the decomposition we are suggesting only applies within a local neighborhood of each material point. That is, an arbitrarily deformed elasto-plastic body is unlikely to be capable of completely relaxing everywhere, but any small material element within should be able to if it were hypothetically removed from the surrounding material. We reiterate though, that the *product* of $\mathbf{F}^e$ and $\mathbf{F}^p$ must be the gradient of the motion $\boldsymbol{\chi}_t$.

In this regard, we accept that the intermediate state is somewhat fictitious. It represents the idea that every local region should have the capability of unloading elastically, and for this to happen, the deformation should be sequentially decomposable. But in the larger sense that just letting go of a deformed body should reveal a relaxed body, it is a slightly false pretense. Regardless, we accept the Kröner–Lee decomposition as a useful modeling tool.

3.6.2 Kinematic definitions

With our decomposition now stated, the next step is to define various kinematical quantities in terms of this decomposition. To wit, any fluid-like law must relate to a consistent notion of deformation rate and any elasticity law must have an equally consistent definition of stretch and rotation.

First, let us start with the more basic question of how to determine these quantities as they relate to a deformation gradient $\mathbf{F}$, before attempting to specialize their definitions to the elastic and plastic parts. Let us define *the spatial velocity field* $\mathbf{v}(\mathbf{x}, t)$, in terms of $\boldsymbol{\chi}_t(\mathbf{X}) \equiv \boldsymbol{\chi}(\mathbf{X}, t)$:

$$\mathbf{v}(\mathbf{x}, t) = \left.\frac{\partial \boldsymbol{\chi}(\mathbf{X}, t)}{\partial t}\right|_{\mathbf{X}=\boldsymbol{\chi}_t^{-1}(\mathbf{x})} = \dot{\boldsymbol{\chi}}(\boldsymbol{\chi}_t^{-1}(\mathbf{x}), t).$$

The material time derivative represented by the dot, is indeed the partial derivative

of time holding the reference location $\mathbf{X}$ constant.

In fluid mechanics, the primitive quantity from which stresses are calculated is the spatial velocity gradient $\mathbf{L}(\mathbf{x},t) \equiv \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$. Mathematically it follows that,

$$\begin{aligned}\dot{\mathbf{F}}(\mathbf{X},t) &= \frac{\partial}{\partial t}\left(\frac{\partial \boldsymbol{\chi}(\mathbf{X},t)}{\partial \mathbf{x}}\right) = \frac{\partial}{\partial \mathbf{X}}\left(\dot{\boldsymbol{\chi}}(\mathbf{X},t)\right) = \frac{\partial}{\partial \mathbf{X}}\left(\mathbf{v}(\boldsymbol{\chi}_t(\mathbf{X}),t)\right) \\ &= \mathbf{L}(\mathbf{x},t)|_{\mathbf{x}=\boldsymbol{\chi}_t(\mathbf{X})}\,\frac{\partial \boldsymbol{\chi}_t(\mathbf{X})}{\partial \mathbf{X}} = \mathbf{L}(\mathbf{x},t)|_{\mathbf{x}=\boldsymbol{\chi}_t(\mathbf{X})}\,\mathbf{F}(\mathbf{X},t).\end{aligned}$$

where the transition from the first to the second line invokes the chain rule. Omitting arguments, we may write out the important relationship that defines the spatial velocity gradient in terms of the deformation gradient:

$$\mathbf{L} = \dot{\mathbf{F}}\mathbf{F}^{-1} \tag{3.18}$$

It is known from Eulerian fluid mechanics, and basic mathematics, that the velocity gradient has the property that when it operates on a small material filament, it outputs the rate of change of that filament. The above expression clearly captures this fact: A filament in the deformed body is taken to its pre-image in the reference body by $\mathbf{F}^{-1}$, and then the result is acted upon by $\dot{\mathbf{F}}$, which outputs the rate of change of the filament in the deformed body.

We shall now break up $\mathbf{L}$ into its symmetric and skew parts:

$$\begin{aligned}\mathbf{D} &\equiv \mathrm{sym}(\mathbf{L}) = \frac{1}{2}(\mathbf{L}+\mathbf{L}^T) \\ \mathbf{W} &\equiv \mathrm{skw}(\mathbf{L}) = \frac{1}{2}(\mathbf{L}-\mathbf{L}^T)\end{aligned}$$

The deformation rate $\mathbf{D}$ is the part of the velocity gradient that corresponds to the rate of stretching of the material element. The spin $\mathbf{W}$ indicates the rigid-body type rotation rate of the element about some axial vector.

Now that we have discussed rates, let us delve into the meanings of total stretch and total rotation in terms of $\mathbf{F}$. The determinant of $\mathbf{F}$, which we shall denote J,

represents the ratio of an element's volume in the deformed space to the volume of its preimage in the reference space. Thus for physical reasons, we can assert that $J > 0$ so as to ensure that an element can never "turn inside out". Since J is positive, we invoke a useful tensorial decomposition, the *polar decomposition*, which allows one to take any tensor of positive determinant and write it as the product of a rotation tensor and a symmetric positive definite tensor:

$$\mathbf{F} = \mathbf{R}\mathbf{U}.$$

We refer to $\mathbf{R}$ as the rotation and $\mathbf{U}$ as the right stretch, as it appears to the right of the rotation. Furthermore, since $\mathbf{U}$ is symmetric, we may express it in term of an orthonormal principal basis:

$$\mathbf{U} = \sum_{i=1}^{3} \lambda_i \hat{\mathbf{r}}_i \otimes \hat{\mathbf{r}}_i$$

The λ_i are the right principal stretches and the eigenvectors $\hat{\mathbf{r}}_i$ are the right principal stretch directions.

To determine $\mathbf{U}$, simply note that $\mathbf{F}^T\mathbf{F} = \mathbf{U}^T\mathbf{R}^T\mathbf{R}\mathbf{U} = \mathbf{U}^T\mathbf{U}$. Since $\mathbf{U}$ is symmetric, we see that given $\mathbf{F}$, we may obtain $\mathbf{U}$ by

$$\mathbf{U} = \sqrt{\mathbf{F}^T\mathbf{F}}.$$

The square root is uniquely defined by constraining that the λ_i must all be positive as per the symmetric positive definiteness of $\mathbf{U}$. Once $\mathbf{U}$ is obtained, the rotation $\mathbf{R}$ follows immediately by $\mathbf{R} = \mathbf{F}\mathbf{U}^{-1}$.

With these kinematical quantities all rigorously defined, we now generalize their usages in terms of the elastic and plastic deformation gradients. The elastic and plastic velocity gradients are defined as

$$\mathbf{L}^e = \dot{\mathbf{F}}^e\mathbf{F}^{e-1}$$

$$\mathbf{L}^p = \dot{\mathbf{F}}^p\mathbf{F}^{p-1}$$

which in turn enables definitions of the elastic and plastic deformation rate and the elastic and plastic spin:

$$\mathbf{D}^e = \text{sym}(\mathbf{L}^e) \qquad \mathbf{D}^p = \text{sym}(\mathbf{L}^p)$$
$$\mathbf{W}^e = \text{skw}(\mathbf{L}^e) \qquad \mathbf{W}^p = \text{skw}(\mathbf{L}^p).$$

Using these definitions, we can write an expression for the full velocity gradient in terms of the elastic and plastic pieces:

$$\mathbf{L} = \dot{\mathbf{F}}\mathbf{F}^{-1} = \left(\frac{\partial}{\partial t}\left(\mathbf{F}^e\mathbf{F}^p\right)\right)\mathbf{F}^{p-1}\mathbf{F}^{e-1} = \mathbf{L}^e + \mathbf{F}^e\mathbf{L}^p\mathbf{F}^{e-1} \tag{3.19}$$

In the same manner as before, performing the polar decomposition on $\mathbf{F}^e$ yields

$$\mathbf{F}^e = \mathbf{R}^e\mathbf{U}^e$$

where $\mathbf{U}^e$ is the elastic right stretch and $\mathbf{R}^e$ is the elastic rotation. The plasticity model we wish to use is incompressible and as such we have $J^p = \det\mathbf{F}^p = 1$. Consequently, $J^e = \det\mathbf{F}^e = J$.

3.6.3 Intermediate space variables

Since we are dealing with three different spaces— reference, intermediate, and deformed —it behooves us when relating quantities that reside in different spaces to create surrogate quantities that all reside in the same space. To be clear, we say a tensor field $\mathbf{A}$ defined over $\mathbf{X}$ or $\mathbf{x}$ "resides" in a vector space $\mathscr{S}$, when the tensor is only meaningfully defined operating on vectors from $\mathscr{S}$, and only outputs vectors in $\mathscr{S}$. Case in point, the Cauchy stress $\mathbf{T}$ resides in the deformed space. This is because it gives the force per unit area at the current moment t given the orientation of the plane of interest at time t. If the stress causes the plastic deformation rate, for example, how is one to relate Cauchy stress to $\mathbf{L}^p$ when $\mathbf{L}^p$ is not defined in terms of deformed space variables? To resolve this issue, we now rewrite the necessary phys-

ical quantities in the intermediate space, as this space directly connects to both the deformed and reference spaces.

As will be a common theme throughout this process, we now use the second law of thermodynamics to our aid. For isothermal motions, the law is equivalent to the following *dissipation inequality*

$$\rho\dot{\psi} - \mathbf{T} : \mathbf{L} \leq 0 \tag{3.20}$$

The quantity ψ is the Helmholtz free energy per unit mass and, for clarity, ρ is the density in the deformed space. The leftmost term indicates the rate at which energy being expended on the element by outside agencies is being converted into stored internal energy. The other term represents the total power expenditure on the element in the form of stress-work. Altogether, equation 3.20 is stating that no more than the total energy being expended on an element can be stored.

Mass conservation implies the density of an element in the intermediate space is $\rho_I = \rho J^e$. And likewise, the free energy per unit intermediate volume can be expressed by $\psi_I = \rho_I \psi$. Note that $\dot{\psi}_I = \rho_I \dot{\psi}$ since plastic incompressibility ensures ρ_I constant. Multiplying both sides of equation 3.20 by J^e and writing $\mathbf{L}$ as per equation 3.19 gives

$$\dot{\psi}_I - J^e\mathbf{T} : \mathbf{L}^e - J^e\mathbf{T} : \left(\mathbf{F}^e\mathbf{L}^p\mathbf{F}^{e-1}\right) \leq 0 \tag{3.21}$$

Let us review some identities relating to the tensorial dot product:

1. For $\mathbf{A}$ symmetric, $\mathbf{A} : \mathbf{B} = \mathbf{A} : \mathrm{sym}(\mathbf{B})$

2. For $\mathbf{A}$ deviatoric (i.e. traceless), $\mathbf{A} : \mathbf{B} = \mathbf{A} : \mathbf{B}_0$

3. $\mathbf{A} : (\mathbf{BC}) = (\mathbf{B}^T\mathbf{A}) : \mathbf{C} = (\mathbf{AC}^T) : \mathbf{B}$

In observing the center term on the left side of equation 3.21, identity (1) implies the reduction

$$J^e\mathbf{T} : \mathbf{L}^e = J^e\mathbf{T} : \mathbf{D}^e$$

But we also know

$$\begin{aligned}\mathbf{D}^e &= \frac{1}{2}\left(\dot{\mathbf{F}}^e\mathbf{F}^{e-1} + (\dot{\mathbf{F}}^e\mathbf{F}^{e-1})^T\right) \\ &= \frac{1}{2}\left(\dot{\mathbf{F}}^e\mathbf{F}^{e-1} + \mathbf{F}^{e-T}\dot{\mathbf{F}}^{eT}\right) \\ &= \frac{1}{2}\mathbf{F}^{e-T}\left(\mathbf{F}^{eT}\dot{\mathbf{F}}^e + \dot{\mathbf{F}}^{eT}\mathbf{F}^e\right)\mathbf{F}^{e-1}\end{aligned}$$

Defining $\mathbf{C}^e \equiv \mathbf{F}^{eT}\mathbf{F}^e = \mathbf{U}^{e2}$, note that the term in parenthesis in the last step is $\dot{\mathbf{C}}^e$. Hence we have

$$J^e\mathbf{T} : \mathbf{D}^e = J^e\mathbf{T} : \left(\frac{1}{2}\mathbf{F}^{e-T}\dot{\mathbf{C}}^e\mathbf{F}^{e-1}\right) = \frac{1}{2}\left(J^e\mathbf{F}^{e-1}\mathbf{T}\mathbf{F}^{e-T}\right) : \dot{\mathbf{C}}^e$$

where identity (3) was invoked in the last equivalence. Let us provide a special denotation for the term in parenthesis in the last expression:

$$\mathbf{T}^e \equiv \mathbf{F}^{e-1}\mathbf{T}\mathbf{F}^{e-T}J^e$$

The stress quantity $\mathbf{T}^e$ resides in the intermediate space. To make this point more clear, note the following. The two last terms in the above, $\mathbf{F}^{e-T}J^e$, are an expression known mathematically as the *cofactor* of $\mathbf{F}^e$. Suppose a vector $\mathbf{a}_I$ is normal to a bounded plane of area $|\mathbf{a}_I|$ in the intermediate space. The image of the bounded plane in the deformed space is defined by the cofactor in the sense that the new area vector is $\mathbf{a}_D = (\mathbf{F}^{e-T}J^e)\mathbf{a}_I$. Therefore, $\mathbf{T}^e\mathbf{a}_I = \mathbf{F}^{e-1}\mathbf{T}\mathbf{a}_D$. By the definition of Cauchy stress, $\mathbf{T}\mathbf{a}_D$ is equal to the force $\mathbf{f}_D$ that acts on the bounded plane corresponding to the area vector $\mathbf{a}_D$. We may thus understand $\mathbf{T}^e$ as a stress measure that inputs an area vector $\mathbf{a}_I$ in the intermediate space and outputs a force vector $\mathbf{f}_I$ in the intermediate space; the force that is exerted on the plane represented by the image of $\mathbf{a}_I$ in the deformed space, is the image of $\mathbf{f}_I$ in the deformed space. The expression for $\mathbf{T}^e$ is closely related to a common stress measure known as *the second Piola stress*.

Moving on to the last expression on the left side of equation 3.21, identity (3) and

some algebraic manipulation implies

$$J^e\mathbf{T} : (\mathbf{F}^e\mathbf{L}^p\mathbf{F}^{e-1}) = J^e(\mathbf{F}^{eT}\mathbf{T}\mathbf{F}^{e-T}) : \mathbf{L}^p = (\underbrace{\mathbf{F}^{eT}\mathbf{F}^e}_{=\mathbf{C}^e}\underbrace{\mathbf{F}^{e-1}\mathbf{T}\mathbf{F}^{e-T}J^e}_{=\mathbf{T}^e}) : \mathbf{L}^p$$

The term in parenthesis deserves its own name. We shall refer to $\mathbf{M} = \mathbf{C}^e\mathbf{T}^e$ as the "Mandel stress". It can be shown in a manner similar to the above that $\mathbf{M}$ also resides in the intermediate space, but has a slightly different physical interpretation. Writing $\mathbf{T}$ in terms of the Mandel stress gives

$$\mathbf{T} = J^{e-1}\mathbf{F}^e\mathbf{M}\mathbf{F}^{eT} \tag{3.22}$$

The Cauchy stress is symmetric, but at this point we cannot say the same of the Mandel stress. Since the plastic deformation is incompressible, $\mathbf{L}^p$ is traceless, and likewise by identity (2) we see that $\mathbf{M} : \mathbf{L}^p = \mathbf{M}_0 : \mathbf{L}^p$.

The dissipation inequality can now be stated completely in terms of the free energy per unit intermediate space ψ_I and tensors that reside in the intermediate space:

$$\dot{\psi}_I - \frac{1}{2}\mathbf{T}^e : \dot{\mathbf{C}}^e - \mathbf{M}_0 : \mathbf{L}^p \leq 0 \tag{3.23}$$

The two last terms reflect the power expended per unit intermediate space. Hence this version of the dissipation inequality states that the rate at which free energy is stored per volume of intermediate space can be no greater than the power expended per volume of intermediate space.

3.6.4 Constitutive dependences

Now, we seek to implement mathematically the physical properties we believe to be true of the material at hand. We shall do so in the form of constitutive dependences, where particular quantities in equation 3.23 are asserted to be functions of other quantities. We choose from experience, but without loss of generality, that ψ_I, $\mathbf{T}^e$, and $\mathbf{L}^p$ compose the set of dependent variables, and that $\mathbf{M}$ and $\mathbf{C}^e$ are independent. Thus,

every quantity that arises in equation 3.23 is an independent variable, dependent variable, or the time derivative of one of these variables.

We believe, backed by our DEM simulations, that the stresses in a granular material, flowing or static, are deducible solely from the elastic deformation through some elasticity relationship. Since our goal is not to impose material behavior, but rather suggest dependences and let the physics determine the specifics, we shall assert our belief in the weakest way possible:

$$\psi_I = \tilde{\psi}_I(\mathbf{C}^e) \ , \quad \tilde{\psi}_I(\mathbf{1}) = 0 \tag{3.24}$$

$$\mathbf{T}^e = \tilde{\mathbf{T}}^e(\mathbf{C}^e) \ , \quad \tilde{\mathbf{T}}^e(\mathbf{1}) = \mathbf{0} \tag{3.25}$$

The $\tilde{}$ is used to represent constitutive functions, so as not to confuse the arguments with the spaces over which the quantities are defined. Thus, the free energy in the system is deemed to arise solely from the elastic deformation of the grains (no defect energy). When $\mathbf{F}^e = \mathbf{1} = \mathbf{C}^e$, no deformation has occurred, and consequentially we require both the stress and the free energy to be zero.

As for the plastic motion, we assert our belief that that the stress determines the plastic flow rate, in line with the results of Jop:

$$\mathbf{L}^p = \tilde{\mathbf{L}}^p(\mathbf{M}) \ , \quad \tilde{\mathbf{L}}^p(\mathbf{0}) = \mathbf{0} \ , \quad \mathrm{tr}\left(\tilde{\mathbf{L}}^p(\mathbf{M})\right) = 0 \tag{3.26}$$

The last two relations reflect the fact that an unstressed material cannot flow plastically, and that the material is being modeled as plastically incompressible.

One might ask: Why do we treat $\mathbf{M}$ as an independent variable when it is just the product of $\mathbf{C}^e$ and $\mathbf{T}^e$? This is simply a matter of convenience. In fact, by equation 3.25, we could express the Mandel stress as a sole function of $\mathbf{C}^e$. But as we shall see in the subsections to come, it helps in terms of algebraic cleanliness to keep these variables separate, if only by name.

3.6.5 Frame indifference

What makes this continuum approach so appealing is that we claim very little about the material behavior. As we shall see, basic physical requirements will narrow down the appearance of the relationships significantly.

The first requirement we shall impose is that our relationships be *frame indifferent* or *objective*. Suppose two observers are viewing the same deformation process. Suppose that one observer is rotating and translating with respect to the other. Since they are observing the same deformation, the only differences perceived by the two observers should stem from the fact that they observe from different rigid-body reference frames. Thus, the moving observer should report the same results except that some of his fields appear to rotate and/or translate appropriately. Consistent with our previous definition, we can imagine that at $t = 0$ the two reference frames align so that the reference body is the same for both observers.

First, we clarify how each of the fields *should* appear to the moving observer, in terms of what the stationary observer would measure. The moving observations will be labeled by *. To formalize, we can define the motion as observed in the * frame by

$$\boldsymbol{\chi}_t^*(\mathbf{X}) = \mathbf{Q}(t)(\boldsymbol{\chi}_t(\mathbf{X}) - \mathbf{o}) + \mathbf{y}(t)$$

where at any time t, $\mathbf{Q}(t)$ is a rotation tensor, $\mathbf{y}(t)$ is a translation vector, and $\mathbf{o}$ represents some fixed origin. Consequently,

$$\mathbf{F}^* = \frac{\partial \boldsymbol{\chi}_t^*(\mathbf{X})}{\partial \mathbf{X}} = \mathbf{Q}(t)\frac{\partial \boldsymbol{\chi}_t(\mathbf{X})}{\partial \mathbf{X}} = \mathbf{Q}(t)\mathbf{F}$$

Suppose both observers were monitoring one oriented material plane at time t. The traction acting on the plane and the orientation of the plane itself would both appear in the * frame to be rotated by $\mathbf{Q}(t)$ from the observations in the stationary frame. Since both the input and output of the Cauchy stress tensor rotate by $\mathbf{Q}(t)$ in the * frame, we can state

$$\mathbf{T}^* = \mathbf{Q}(t)\mathbf{T}\mathbf{Q}(t)^T.$$

The free energy ψ is a scalar field defined on a material element. Thus, regardless of frame, two observers viewing the same material element should agree on the value of ψ for all t. Hence

$$\psi^* = \psi.$$

We have now determined, based on physical necessity, how the fields $\mathbf{F}$, $\mathbf{T}$, and ψ should transform under frame-change. We shall now mathematically deduce the consequences on the intermediate state variables. Since ρ_I is a constant for plastically incompressible materials, we deduce

$$\psi_I^* = \rho_I^* \psi^* = \rho_I \psi = \psi_I.$$

Since the deformation gradient has elastic and plastic parts, we can write

$$\mathbf{F}^{e*}\mathbf{F}^{p*} = \mathbf{Q}(t)\mathbf{F}^e\mathbf{F}^p.$$

We presume that the frame change only modifies the deformed space, leaving the intermediate configuration the same for the two observers. This allows the simplification

$$\mathbf{F}^{e*} = \mathbf{Q}(t)\mathbf{F}^e \ , \quad \mathbf{F}^{p*} = \mathbf{F}^p.$$

and indicates that extra rotation applied after deforming an object will be recorded in the elastic part.

Therefore, we have

$$\mathbf{C}^{e*} = \mathbf{F}^{eT*}\mathbf{F}^{e*} = \mathbf{F}^{eT}\mathbf{Q}(t)^T\mathbf{Q}(t)\mathbf{F}^e = \mathbf{F}^{eT}\mathbf{F}^e = \mathbf{C}^e$$

Using what has been deduced thus far and the fact that $J = J^e$ is unaffected by frame change, the stress $\mathbf{T}^e$ must transform as

$$\mathbf{T}^{e*} = J^{e*}\mathbf{F}^{e*-1}\mathbf{T}^*\mathbf{F}^{e*} = J^e\left(\mathbf{F}^{e-1}\mathbf{Q}(t)^T\right)\left(\mathbf{Q}(t)\mathbf{T}\mathbf{Q}(t)^T\right)\left(\mathbf{Q}(t)\mathbf{F}^e\right) = J^e\mathbf{F}^{e-1}\mathbf{T}\mathbf{F}^e = \mathbf{T}^e$$

Since $\mathbf{T}^e$ and $\mathbf{C}^e$ are frame-invariant, consequently the Mandel stress $\mathbf{M} = \mathbf{C}^e\mathbf{T}^e$ is frame-invariant: $\mathbf{M}^* = \mathbf{M}$. And since $\mathbf{F}^p$ is frame invariant, the plastic velocity gradient $\mathbf{L}^p = \dot{\mathbf{F}}^p\mathbf{F}^{p-1}$ must also be: $\mathbf{L}^{p*} = \mathbf{L}^p$.

Now, we go about checking that our constitutive dependences, equations 3.24, 3.25, and 3.26, uphold these rules. That is, both observers should be able to use the constitutive dependences to relate the fields as they see them, and their results should compare as determined above. First of all, utilizing equation 3.24, we can write

$$\psi_I^* = \tilde{\psi}_I(\mathbf{C}^{e*}) \quad \text{and} \quad \psi_I = \tilde{\psi}_I(\mathbf{C}^e)$$

But since $\mathbf{C}^{e*} = \mathbf{C}^e$ the above guarantees $\psi_I^* = \psi_I$ as we desire. Thus, equation 3.24 respects frame indifference.

By equation 3.25, we can write

$$\mathbf{T}^{e*} = \tilde{\mathbf{T}}^e(\mathbf{C}^{e*}) \quad \text{and} \quad \mathbf{T}^e = \tilde{\mathbf{T}}^e(\mathbf{C}^e)$$

And since $\mathbf{C}^{e*} = \mathbf{C}^e$, we are assured $\mathbf{T}^{e*} = \mathbf{T}^e$, implying that equation 3.25 obeys frame indifference.

In a completely analogous manner, one can quickly show that equation 3.26 is also frame indifferent, as it too relates two frame invariant fields. So, as written, we have verified that all three constitutive dependences proposed thus far pass the frame indifference test and need not be altered as yet.

This reflects our wisdom in choosing the dependences that we did. Commonly, frame indifference tests reveal that changes need to be made. As a simple example, suppose we believed that ψ_I should relate more generally to $\mathbf{F}^e$ than to $\mathbf{C}^e$. Since $\psi_I^* = \psi_I$, then we would end up with the statement that a function operating on $\mathbf{Q}(t)\mathbf{F}^e$ gives the same answer for all possible $\mathbf{Q}(t)$. This cannot hold unless that function always ignores the rotational part of its argument, $\mathbf{Q}(t)\mathbf{R}^e$. Thus we would deduce that ψ_I can only relate to the stretch $\mathbf{U}^e$, or for simplicity to $\mathbf{C}^e = \mathbf{U}^{e2}$. So we would have been led to equation 3.24 regardless of the specifics of our first guess.

3.6.6 The Coleman–Noll procedure

The next physical requirement that our relationships must obey is that there must *never* be any situation that allows a violation of the dissipation inequality 3.23. The method by which we ensure this point is known as the "Coleman–Noll Procedure" after its originators [31].

Observe the time derivative of ψ_I:

$$\dot{\psi}_I = \frac{\partial \tilde{\psi}_I(\mathbf{C}^e)}{\partial t} = \frac{\partial \tilde{\psi}_I(\mathbf{C}^e)}{\partial \mathbf{C}^e} : \dot{\mathbf{C}}^e$$

Accordingly, we can rewrite inequality 3.23 as

$$\underbrace{\left(\frac{\partial \tilde{\psi}_I(\mathbf{C}^e)}{\partial \mathbf{C}^e} - \frac{1}{2}\tilde{\mathbf{T}}^e(\mathbf{C}^e)\right)}_{\equiv \tilde{\mathbf{A}}(\mathbf{C}^e)} : \dot{\mathbf{C}}^e - \underbrace{\mathbf{M}_0 : \tilde{\mathbf{L}}^p(\mathbf{M})}_{\equiv \tilde{B}(\mathbf{M})} \leq 0 \tag{3.27}$$

At some fixed time t_0 and material point $\mathbf{X}_0$, observe that $\dot{\mathbf{C}}^e(\mathbf{X}_0, t_0)$ can be chosen independent of $\mathbf{C}^e(\mathbf{X}_0, t_0)$. To make this point more clear, note that $\mathbf{C}^e$ is deemed independent in terms of the constitutive relations, and thus we should consider it to be a variable that can be changed at will. Thus, the value of $\mathbf{C}^e$ at some place and time does not constrain its value one time-step later; we may *choose* that next value as we please. Hence $\mathbf{C}^e$ and $\dot{\mathbf{C}}^e$ are both independent at any instant.

Next, take note of the following (basic) mathematical fact. Since the author is not aware of a direct proof of this fact, we provide our own.

Theorem: *For some vector-valued functions* $\{\mathbf{f}_k\}$ *and scalar-valued functions* $\{g_j\}$*, suppose*

$$\left(\sum_k \mathbf{f}_k(\mathbf{a}_k) \cdot \mathbf{b}_k\right) + \left(\sum_j g_j(\mathbf{c}_j)\right) \leq 0$$

for any sequences $\{\mathbf{a}_k\}$*,* $\{\mathbf{b}_k\}$*,* $\{\mathbf{c}_j\}$ *of real valued vectors. Then* $\mathbf{f}_k$ *is the zero function for all* k*.*

Proof: We prove this quickly by contradiction. Suppose $\mathbf{f}_K$ is not the zero function for

some K in the index set. Then we may choose a sequence $\{\mathbf{a}_k\}$ for which $\mathbf{f}_K(\mathbf{a}_{k'}) \neq \mathbf{0}$ for some k' in the index set. Let us also choose the sequence $\{\mathbf{b}_k\}$ so that the k' entry is

$$\mathbf{b}_{k'} = \frac{\mathbf{f}_K(\mathbf{a}_{k'})}{|\mathbf{f}_K(\mathbf{a}_{k'})|^2} \left(1 - \left(\sum_{k \neq k'} \mathbf{f}_k(\mathbf{a}_k) \cdot \mathbf{b}_k \right) - \left(\sum_j g_j(\mathbf{c}_j) \right) \right)$$

The left-hand side of the inequality in the theorem statement is now necessarily 1, and consequently greater than 0. □

In inequality 3.27, $\mathbf{C}^e$, $\dot{\mathbf{C}}^e$, and $\mathbf{M}$ can each be independently selected and, in light of the theorem, can each be deemed sequences with only one term. Hence, the theorem implies

$$\tilde{\mathbf{A}}(\mathbf{C}^e) = \mathbf{0}.$$

Consequently, we can write the important law

$$\tilde{\mathbf{T}}^e(\mathbf{C}^e) = 2\frac{\partial \tilde{\psi}_I(\mathbf{C}^e)}{\partial \mathbf{C}^e}. \tag{3.28}$$

What remains is the reduced inequality

$$\tilde{B}(\mathbf{M}) = -\mathbf{M}_0 : \tilde{\mathbf{L}}^p(\mathbf{M}) \leq 0. \tag{3.29}$$

The Coleman–Noll procedure is quite powerful, as it has just enabled a *derivation* of elasticity theory (equation 3.28) and greatly simplified the dissipation inequality.

3.6.7 Upholding symmetry: Reference body isotropy

For our current purposes, let us delete the meanings of $\tilde{\mathbf{A}}$ and $\tilde{B}$ from the last subsection. The Jiang–Liu elasticity law (equation 3.8) and the Jop–Pouliquen flow law (equation 3.12) both take the general form

$$\mathbf{A} = \alpha_1(\mathscr{I}_\mathbf{B})\mathbf{B}_0 + \alpha_2(\mathscr{I}_\mathbf{B})(\mathrm{tr}\mathbf{B})\mathbf{1}$$

where $\mathbf{A}$ and $\mathbf{B}$ are the stress and strain respectively in elasticity, or the deformation rate and stress in the flow rule. The term $\mathscr{I}_{\mathbf{B}}$ represents a set of invariants for $\mathbf{B}$. Constitutive laws of this form have the feature of *isotropy*. A more general definition of an isotropic tensorial function of a tensor is one in which

$$\mathbf{Q}^T\tilde{\mathbf{A}}(\mathbf{B})\mathbf{Q} = \tilde{\mathbf{A}}(\mathbf{Q}^T\mathbf{B}\mathbf{Q}) \quad \text{for all rotations } \mathbf{Q}. \tag{3.30}$$

Hence, a rotation of the input tensor will simply cause a rotation of the output. An isotropic material, physically, has no underlying lattice or ordering scheme that would make the material response depend on the local orientation.

A difficult question arises when trying to piece together a granular elasto-plasticity law: Are granular materials isotropic? If so, we should attempt to properly insert isotropy into the framework. If not, we have the more difficult task of determining what the preferred directions are, how do they evolve during a deformation process, and what is their symmetry group.

There are two ways to handle this issue. The simplest response, which is also a valid one, is to reiterate that our work is primarily concerned with gluing together two preexisting models for grains, one for statics and one for flow. Both models of concern assert isotropy, and consequently, we should employ isotropy in constructing the joint model. We could take this standpoint and immediately move on from symmetry considerations.

But on the other hand, isotropy is a physically nontrivial assumption for a flowing granular collection. Monodisperse spheres, for example, have several viable crystal packing arrangements. The correctness of the isotropy assumption and how to appropriately enforce it within the $\mathbf{F}^e\mathbf{F}^p$ framework deserves further commentary.

Let us start with the reference space. Granular materials in general begin as a completely random packing— they are formed via a thermalized process such as pouring, and then rapidly quench to a static configuration. In the static state (below yield, $\mathbf{F} = \mathbf{F}^e$), only elasticity applies. Since the random structure never changes during elastic deformation it is indeed a sensible physical assertion to claim that

purely elastic motions are isotropic.

Now suppose plastic flow occurs. Here, the reference and intermediate spaces become disparate. The reference space should always have no preferred directions as per the previous argument. Thus, our laws should always uphold invariance to reference space rotations. In other words, given a spherical reference body and some total deformation $\mathbf{F}$, the resulting stress and free energy fields in the deformed body should look no different than they would if someone were to have first rotated the body by some $\mathbf{Q}$ and *then* applied $\mathbf{F}$.

Letting * represent the fields corresponding to a pre-rotated reference body, we can write

$$\mathbf{F}^* = \mathbf{F}\mathbf{Q}$$

which motivates

$$\mathbf{F}^{p*} = \mathbf{F}^p\mathbf{Q} \quad \text{and} \quad \mathbf{F}^{e*} = \mathbf{F}^e.$$

By the previous physical argument, we wish to ensure

$$\mathbf{T}^* = \mathbf{T} \quad \text{and} \quad \psi^* = \psi$$

at all spatial points in the deformed body.

Just as in the frame indifference validation, we may construct intermediate space quantities in the pre-rotated * case from these above equivalences. One finds

$$\mathbf{C}^{e*} = \mathbf{C}^e \ , \ \mathbf{L}^{p*} = \mathbf{L}^p \ , \ \mathbf{T}^{e*} = \mathbf{T}^e \ , \ \mathbf{M}^* = \mathbf{M} \ , \ \psi_I^* = \psi_I$$

Since all the intermediate space fields are invariant to pre-rotation, we are guaranteed that our constitutive dependences (equations 3.26 and 3.28) uphold these transformation rules. What we have just shown is that the constitutive dependences are written in a form that permits an isotropic elasticity law for purely elastic deformations, and in general, that our dependences enforce the necessary consequences of an isotropic reference body.

3.6.8 Upholding symmetry: Intermediate space isotropy

Now comes the more physically challenging question: Is the intermediate space isotropic? That is, if the stresses that act on a flowing granular element are suddenly removed, will what remains appear to have no internal ordering? The presumption in the Jop–Pouliquen flow law is that the answer is yes. Their law as written is isotropic; no outside anisotropy measure appears. But is this a physically justifiable fact?

Polydisperse granular media have relative difficulty achieving any sort of internal ordering in a static or flowing state. This fact is even more resounding in 3D than in 2D. But with monodisperse materials, crystalline ordering may emerge in flow environments that geometrically support a crystal packing. For example, in 2D simulations of monodisperse disks undergoing planar shear, a hexagonal lattice commonly emerges after a few hundred percent strain. In 3D however, measurements by Tsai *et al.* [139] of monodisperse spheres under shear, reveal that ordering properties stay close to random packing for several orders of magnitude longer, commonly up to $10^5 - 10^6$ percent strain. Experiments were performed in a periodic shear cell $800d$ long, $30d$ wide, and $24d$ high, and it was observed that more than 3 hours of consistent steady shearing by a rough top plate ($12d/s$) is necessary before any crystalline ordering can be measured for fluid-submerged grains, and the duration increases ten-fold for dry grains.

The flows of interest to us occur over a time window on the order of several minutes, not tens of hours. Periodic flow environments are useful to the scientist, but uncommon in everyday life. Typically, the flow environment is not as amenable to crystallization, locally or globally, as a shear cell. But even then, monodisperse flows of spheres arrange into a crystal very, very slowly. It appears that the day-to-day flows we wish to model do not pose an appreciable risk of flow-induced local ordering.

DEM simulations of Rycroft [122], seem to uphold this point as well. But before these results can be interpreted, a representation theorem for isotropic tensor functions needs to be established.

Theorem: *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be* 3×3 *tensors and let* $\mathbf{B}$ *be symmetric. Suppose*

$\mathbf{A} = \tilde{\mathbf{A}}(\mathbf{B})$ *is an isotropic tensor function (i.e. fulfills 3.30). Then the eigenvectors of* $\mathbf{A}$ *align with those of* $\mathbf{B}$*. This property is known as* coaxiality*.*

In [143], it is proven that an isotropic function acting on a symmetric tensor must always output a symmetric tensor. This fact can be seen as an immediate consequence of the above theorem, since a tensor is symmetric if and only if it has orthogonal eigenvectors. The author is unaware if the full theorem above has been previously proven, and so a proof shall now be supplied.

Proof: Let $\{B_i\}$ and $\{\hat{\mathbf{e}}_i^B\}$ be the eigenvalues and corresponding orthonormal eigenvectors of $\mathbf{B}$. For an integer j such that $1 \leq j \leq 3$, define

$$\mathbf{Q}_j = \sum_{i=1}^{3} (2\delta_{ij} - 1)\hat{\mathbf{e}}_i^B \otimes \hat{\mathbf{e}}_i^B$$

so that $\mathbf{Q}_j$ is a rotation by π radians about the j^{th} eigenvector of $\mathbf{B}$. Observe that

$$\begin{aligned}
\mathbf{Q}_j \mathbf{B} \mathbf{Q}_j^T &= \mathbf{Q}_j \left(\sum_{i=1}^{3} B_i \hat{\mathbf{e}}_i^B \otimes \hat{\mathbf{e}}_i^B \right) \mathbf{Q}_j^T \\
&= \sum_{i=1}^{3} B_i (\mathbf{Q}_j \hat{\mathbf{e}}_i^B) \otimes (\mathbf{Q}_j \hat{\mathbf{e}}_i^B) \\
&= \sum_{i=1}^{3} B_i ((2\delta_{ij} - 1)\hat{\mathbf{e}}_i^B) \otimes ((2\delta_{ij} - 1)\hat{\mathbf{e}}_i^B) \\
&= \sum_{i=1}^{3} \underbrace{(2\delta_{ij} - 1)^2}_{=1 \text{ for all } i,j} B_i (\hat{\mathbf{e}}_i^B \otimes \hat{\mathbf{e}}_i^B) \\
&= \mathbf{B}
\end{aligned}$$

By isotropy of $\tilde{\mathbf{A}}$, this implies

$$\mathbf{Q}_j \mathbf{A} \mathbf{Q}_j^T = \mathbf{Q}_j \tilde{\mathbf{A}}(\mathbf{B}) \mathbf{Q}_j^T = \tilde{\mathbf{A}}(\mathbf{Q}_j \mathbf{B} \mathbf{Q}_j^T) = \tilde{\mathbf{A}}(\mathbf{B}) = \mathbf{A}.$$

Consequently,

$$\mathbf{A}\hat{\mathbf{e}}_j^B = (\mathbf{Q}_j \mathbf{A} \mathbf{Q}_j^T)\hat{\mathbf{e}}_j^B = \mathbf{Q}_j (\mathbf{A}\hat{\mathbf{e}}_j^B)$$

implying that $\mathbf{A}\hat{\mathbf{e}}_j^B$ is an eigenvector of $\mathbf{Q}_j$ with eigenvalue 1. But, referring back to the definition of $\mathbf{Q}_j$, the only eigenvector of $\mathbf{Q}_j$ with this eigenvalue is $\hat{\mathbf{e}}_j^B$. Thus $\mathbf{A}\hat{\mathbf{e}}_j^B$ must align with $\hat{\mathbf{e}}_j^B$. Likewise $\hat{\mathbf{e}}_j^B$ is an eigenvector of $\mathbf{A}$. This must hold for all j, and thus $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvectors. □

The tensor functions we typically care about act on symmetric tensors like the Cauchy stress, deformation rate, strain, etc. Hence, coaxiality is a major signature of an isotropic material response. Recent DEM simulations of Rycroft [122] have shown that *instantaneous* coaxiality between Cauchy stress and deformation rate is strongly upheld in several different flow environments, using volume-averaged quantities over $5d \times 8d \times 5d$ elements (see figure 3-12). Average deviation from perfect alignment is roughly 12^o, and that number shrinks by half when averaged over a time window of 20 frames. Others have found high levels of coaxiality in 3D flows over longer time windows, such as Depken *et al.* [35] who used the notion of Shear Free Sheets to verify principal vector alignment in split-bottom troughs.

There is evidence in 2D polydisperse disk flows that coaxiality takes some time to develop [136]. The misalignment is transient and disappears by about 20% strain. This notion resonates with some theories of deformation, particularly STZ theory [46], in which the orientation of shear zones gradually align with the stresses and flow as influenced by some effective disorder temperature. While interesting, it is questionable how large this effect would be for 3D flows. A 3D sphere cluster can rearrange under loading in far more ways than a 2D disk cluster, and the extra freedom in contact geometry reduces the likelihood that a few directions are markedly weaker than others. In general, the energy landscape is a smoother function of shear direction in 3D than in 2D. While features like rupture zones and banding over a microscopic width may occur in 2D, it is not clear if any of these effects appear in 3D [39]. It could be that 3D coaxiality happens almost instantaneously. Indeed, it would be interesting future work to directly test whether a certain amount of shearing is needed for coaxiality to emerge in 3D assemblies.

Taking all these points into account— the strong lack of ordering in 3D shear

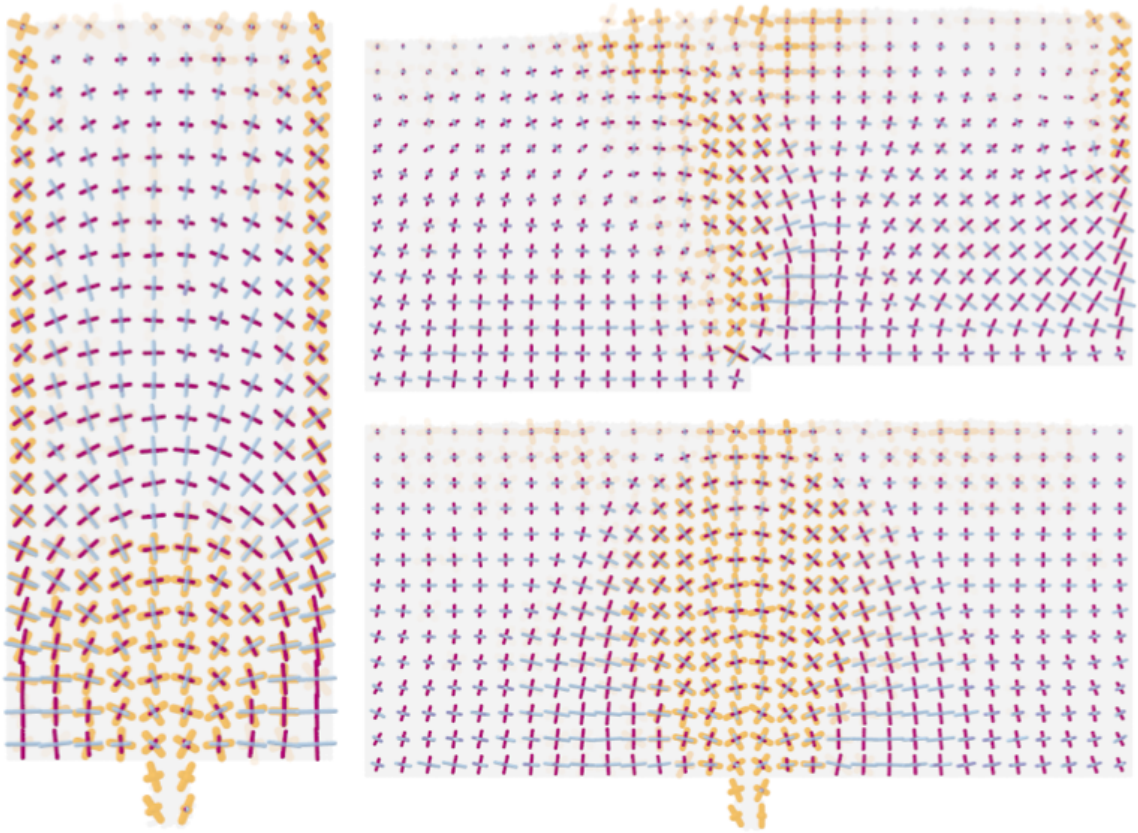


Figure 3-12: (Image c/o Chris Rycroft) Three flow environments, (left) gravity-driven drainage from a narrow quasi-2D silo, (top) steady upward pushing of the left half of a quasi-2D container under gravitational compression, (bottom) gravity-driven drainage from a wide quasi-2D silo. In all three environments, the eigenvectors of Cauchy stress and deformation rate are plotted with lengths corresponding to the square root of the magnitude of the corresponding eigenvalues. The maximum principal stress is colored purple with other principal stresses in blue. Where plastic deformation is significant, the deformation rate eigenvectors are colored in orange. Note the globally valid observation of alignment between stress and deformation rate eigenvectors.

flows, the coaxial agreement in developed flows, and the argument against applying 2D results to 3D —it appears that isotropy would be a decent presumption for flowing 3D granular media. As a consequence in our constitutive development, intermediate space isotropy ought to be enforced. That is, whatever structure exists in the relaxed space should be invariant to rotations.

We can think of the intermediate space as mapping to both the reference space (via $\mathbf{F}^{p-1}$) and the deformed space (via $\mathbf{F}^e$). Rotating the intermediate space by $\mathbf{Q}$ and then applying these maps gives the following:

$$\mathbf{F}^{e*} = \mathbf{F}^e\mathbf{Q} \ , \quad \mathbf{F}^{p*-1} = \mathbf{F}^{p-1}\mathbf{Q}$$

By isotropy, the mechanical variables we observe must remain unchanged by this rotation. Thus,

$$\mathbf{T}^* = \mathbf{T} \ , \quad \psi^* = \psi.$$

Computing the other variables from these transformations gives

$$\mathbf{C}^{e*} = \mathbf{Q}^T\mathbf{C}^e\mathbf{Q} \ , \quad \mathbf{L}^{p*} = \mathbf{Q}^T\mathbf{L}^p\mathbf{Q} \ , \quad \mathbf{T}^{e*} = \mathbf{Q}^T\mathbf{T}^e\mathbf{Q} \ , \quad \mathbf{M}^* = \mathbf{Q}^T\mathbf{M}\mathbf{Q} \ , \quad \psi_I^* = \psi_I$$

As before, we apply the constitutive dependences under these transformation rules. The results are

$$\tilde{\psi}_I(\mathbf{C}^e) = \tilde{\psi}_I(\mathbf{Q}^T\mathbf{C}^e\mathbf{Q}) \tag{3.31}$$

$$\mathbf{Q}^T\tilde{\mathbf{T}}^e(\mathbf{C}^e)\mathbf{Q} = \tilde{\mathbf{T}}^e(\mathbf{Q}^T\mathbf{T}^e\mathbf{Q}) \tag{3.32}$$

$$\mathbf{Q}^T\tilde{\mathbf{L}}^p(\mathbf{M})\mathbf{Q} = \tilde{\mathbf{L}}^p(\mathbf{Q}^T\mathbf{M}\mathbf{Q}) \tag{3.33}$$

Equation 3.31 implies that ψ_I can only depend on a set of invariants of $\mathbf{C}^e$. Let the eigenvalues of $\mathbf{U}^e = \sqrt{\mathbf{C}^e}$ be $\{\lambda_i^e\}$. We may now rewrite the functional dependence of $\tilde{\psi}_I$ as

$$\psi_I = \tilde{\psi}_I(\lambda_1^e, \lambda_2^e, \lambda_3^e).$$

Implementing equation 3.28 yields

$$\mathbf{T}^e = 2\sum_{i=1}^{3} \frac{\partial \tilde{\psi}}{\partial \lambda_i^e} \frac{\partial \lambda_i^e}{\partial \mathbf{C}^e}$$

The derivative $\frac{\partial \lambda_i^e}{\partial \mathbf{C}^e}$ can be shown from tensor calculus to equal $\frac{1}{2\lambda_i^e}\hat{\mathbf{r}}_i^e \otimes \hat{\mathbf{r}}_i^e$. As in the previous notation, $\{\hat{\mathbf{r}}_i^e\}$ are the eigenvectors of $\mathbf{C}^e$, which are also the right principal elastic stretch directions. Hence

$$\mathbf{T}^e = \sum_{i=1}^{3} \frac{1}{\lambda_i^e} \frac{\partial \tilde{\psi}_I}{\partial \lambda_i^e}\ \hat{\mathbf{r}}_i^e \otimes \hat{\mathbf{r}}_i^e \tag{3.34}$$

and consequently

$$\mathbf{M} = \mathbf{C}^e\mathbf{T}^e = \sum_{i=1}^{3} \lambda_i^e \frac{\partial \tilde{\psi}_I}{\partial \lambda_i^e}\ \hat{\mathbf{r}}_i^e \otimes \hat{\mathbf{r}}_i^e \tag{3.35}$$

Define $E_i^e = \log(\lambda_i^e)$ and define the *Hencky elastic strain tensor* as

$$\mathbf{E}^e = \sum_{i=1}^{3} E_i^e\ \hat{\mathbf{r}}_i^e \otimes \hat{\mathbf{r}}_i^e = \log(\mathbf{U}^e)$$

Then observe that,

$$\frac{\partial \tilde{\psi}_I}{\partial \lambda_i^e} = \frac{\partial \tilde{\psi}_I}{\partial E_i^e}\frac{\partial E_i^e}{\partial \lambda_i} = \frac{\partial \tilde{\psi}_I}{\partial E_i^e}\frac{1}{\lambda_i}$$

so that

$$\mathbf{M} = \sum_{i=1}^{3} \frac{\partial \tilde{\psi}_I}{\partial E_i^e}\ \hat{\mathbf{r}}_i^e \otimes \hat{\mathbf{r}}_i^e = \frac{\partial \tilde{\psi}_I}{\partial \mathbf{E}^e} \tag{3.36}$$

We take away from this development the following important points:

1. The tensors $\mathbf{T}^e$ and $\mathbf{C}^e$ are no longer necessary, as we have a simpler expression of the elasticity written in terms of $\mathbf{M}$ and $\mathbf{E}^e$ (equation 3.36).

2. In analogy with infinitesimal elasticity (equation 3.7), the Mandel stress takes the role of the Cauchy stress, and the Hencky strain takes the role of the small strain.

3. Since $\mathbf{E}^e$ and $\mathbf{M}$ are coaxial, then $\mathbf{U}^e$ and $\mathbf{M}$ are coaxial. Thus, the expression

for the Cauchy stress in terms of $\mathbf{M}$ reduces to

$$\mathbf{T} = J^{e-1}\mathbf{R^e}\mathbf{M}\mathbf{R}^{eT} \tag{3.37}$$

4. The Mandel stress has orthogonal eigenvectors and hence is symmetric.

Now, we determine how the plastic deformation is affected by these results. We have just established that $\mathbf{M}$ is symmetric. By one of the aforementioned dot product identities, the reduced dissipation inequality now reads

$$\mathbf{M}_0 : \tilde{\mathbf{D}}^p(\mathbf{M}) \geq 0 \tag{3.38}$$

But the symmetry of $\mathbf{M}$ also means coaxiality applies in equation 3.33. With $\mathbf{L}^p$ coaxial to $\mathbf{M}$, the eigenvectors of $\mathbf{L}^p$ must resultantly be orthogonal and hence $\mathbf{L}^p$ is symmetric. Let us restate this major result:

$$\mathbf{D}^p = \mathbf{L}^p \ , \quad \mathbf{W}^p = \mathbf{0}.$$

3.6.9 Inserting the elastic and plastic models

We are now in a position to close the model mathematically. We are left with only two functions that need to be filled in: $\tilde{\psi}_I(\mathbf{E}^e)$ and $\tilde{\mathbf{D}}^p(\mathbf{M})$. These functions are lifted directly from the elasticity and plasticity relations.

In the proper analogy with small strain elasto-statics, we write

$$\psi_I(\mathbf{E}^e) = B\sqrt{\Delta}\left(\frac{2}{5}\Delta^2 + \gamma^2/\xi\right) \quad \text{for} \quad \Delta = -\text{tr}\mathbf{E}^e \ , \ \gamma = \sqrt{\mathbf{E}_0 : \mathbf{E}_0} \tag{3.39}$$

when $\Delta > 0$. Otherwise $\psi_I = 0$. This implies the stress/strain law via equation 3.36:

$$\mathbf{M} = 2\frac{B\sqrt{\Delta}}{\xi}\mathbf{E}_0^e + B\sqrt{\Delta}\left(1 + \frac{\gamma^2}{2\xi\Delta^2}\right)(\text{tr}\mathbf{E}^e)\mathbf{1} \tag{3.40}$$

Now for the flow. The Jop-Pouliquen flow law is written in Eulerian as a relation between the Cauchy stress and the total deformation rate. We need to rewrite this

as per the elasto-plastic framework. The elastic deformation in a steady flowing assembly is *extremely* small compared to the plastic deformation. At steady flow, the size of the elastic stretches at a spatial point should stop changing altogether, and should be approximately constant in the Lagrangian frame. In light of equation 3.19, it is reasonable to presume that $\mathbf{D} \cong \mathbf{F}^e\mathbf{D}^p\mathbf{F}^{e-1} \cong \mathbf{R}^e\mathbf{D}^p\mathbf{R}^{eT}$ for use in equation 3.12.

Since the grain medium is elastically very stiff, $\mathbf{F}^e \cong \mathbf{1}$ and $J^e \cong 1$. Applying this in equation 3.37 implies $\mathbf{T} \cong \mathbf{R}^e\mathbf{M}\mathbf{R}^{eT}$. With these correspondences, we may rewrite equation 3.12 in terms of $\mathbf{D}^p$ and $\mathbf{M}$ as

$$\mathbf{D}^p = \frac{I_0}{d}\sqrt{\frac{P}{\rho_s}}\frac{\mu-\mu_s}{\mu_2-\mu}\frac{\mathbf{M}_0}{\tau} \tag{3.41}$$

where now $P = -(1/3)\mathrm{tr}\mathbf{M}$ and $\mu = \tau/P$ where $\tau = |\mathbf{M}_0|/\sqrt{2}$. Still, no plastic flow occurs for $\mu < \mu_s$. In light of the reduced inequality 3.38, observe that when $\mathbf{D}^p \neq \mathbf{0}$, or equivalently $\mu_2 > \mu > \mu_s$,

$$\mathbf{M}_0 : \mathbf{D}^p = \mathbf{M}_0 : \left(\frac{I_0}{d}\sqrt{\frac{P}{\rho_s}}\frac{\mu-\mu_s}{\mu_2-\mu}\frac{\mathbf{M}_0}{\tau}\right) = \frac{I_0}{d}\sqrt{\frac{P}{\rho_s}}\frac{\mu-\mu_s}{\mu_2-\mu}2\tau$$

which cannot be negative. Therefore, inequality 3.38 is guaranteeably upheld under all circumstances.

3.6.10 Summary of the equations

1. Balance of linear momentum:

$$\frac{\partial}{\partial \mathbf{x}} \cdot \mathbf{T} + \rho\mathbf{g} = \rho\dot{\mathbf{v}}$$

2. Balance of angular momentum:

$$\mathbf{T} = \mathbf{T}^T$$

3. Kröner-Lee decomposition:

$$\mathbf{F} = \mathbf{F}^e\mathbf{F}^p$$

4. Polar decomposition of elastic deformation gradient:

$$\mathbf{F}^e = \mathbf{R}^e\mathbf{U}^e$$

5. Definition of the elastic determinant, Mandel stress, Hencky elastic strain, and plastic velocity gradient:

$$J^e = \det\mathbf{F}^e$$

$$\mathbf{M} = J^e\mathbf{F}^{e-1}\mathbf{T}\mathbf{F}^e = J^e\mathbf{R}^{eT}\mathbf{T}\mathbf{R}^e$$

$$\mathbf{E}^e = \log(\mathbf{U}^e)$$

$$\mathbf{L}^p = \dot{\mathbf{F}}^p\mathbf{F}^{p-1}$$

6. Elasticity relation:

$$\mathbf{M} = 2G\mathbf{E}^{\mathbf{e}}_{\mathbf{0}} + \kappa\,(\mathrm{tr}\mathbf{E}^{\mathbf{e}})\,\mathbf{1}$$

where $\kappa = B\sqrt{\Delta}[1 + \gamma^2/(2\Delta^2\xi)]$ and $G = \sqrt{\Delta}B/\xi$, for $\Delta = -\mathrm{tr}\mathbf{E}^e$ and $\gamma = \sqrt{\mathbf{E}^e_0 : \mathbf{E}^e_0}$. If $\mathrm{tr}\mathbf{E}^{\mathbf{e}} > 0$, both κ and G are 0.

7. Flow rule:

$$\mathbf{L}^p = \mathbf{D}^p = \frac{I_0}{d}\sqrt{\frac{P}{\rho_s}}\,\frac{\mu - \mu_s}{\mu_2 - \mu}\frac{\mathbf{M}_0}{\tau}$$

where $P = -\mathrm{tr}(\mathbf{M})/3$, $\tau = \sqrt{\mathbf{M}_0 : \mathbf{M}_0/2}$, and $\mu = \tau/P$. If $\mu < \mu_s$, then $\mathbf{D}^p = \mathbf{0}$.

8. Initial conditions:

$$\mathbf{F}^p(t=0) = \mathbf{1}$$

$$\mathbf{L}^p(t=0) = \mathbf{0}$$

3.7 Numerical implementation

Solving the above equations in an arbitrary flow environment is itself highly nontrivial. The implementation method we are about to describe is not only very robust, but may also help the reader understand how the equations compose a closed system.

We shall employ ABAQUS/Explicit from the finite element software package ABAQUS 6.5. The body of material is decomposed into a mesh of elements that deform according to the above constitutive model. The algorithm propagates stresses through the material over many time-steps in a fashion we outline below.

3.7.1 The material model

A key component in this process is to describe how an arbitrary element develops stresses after an increment of deformation is applied to it. This is done in a Vectorized User Material (VUMAT) subroutine. For an element at time t and material point $\mathbf{X}$, the algorithm must compute the following:

$$\text{Given input: } \mathbf{F}(t,\mathbf{X}), \mathbf{F}^p(t,\mathbf{X}), \mathbf{T}(t,\mathbf{X}), \text{ and } \mathbf{F}(t+\Delta t,\mathbf{X})$$

$$\text{Compute output: } \mathbf{F}^p(t+\Delta t,\mathbf{X}) \text{ and } \mathbf{T}(t+\Delta t,\mathbf{X})$$

For our constitutive equations, here is an example of a pseudocode algorithm that leads from the input to the output:

1. Using the Kröner–Lee decomposition, calculate $\mathbf{F}^e(t,\mathbf{X}) = \mathbf{F}(t,\mathbf{X})\mathbf{F}^{p-1}(t,\mathbf{X})$.
2. Using the definition of Mandel stress, calculate $\mathbf{M}(t,\mathbf{X})$ from $\mathbf{F}^e(t,\mathbf{X})$ and $\mathbf{T}(t,\mathbf{X})$.
3. Using the flow rule, calculate $\mathbf{L}^p(t,\mathbf{X})$ from $\mathbf{M}(t,\mathbf{X})$.
4. Numerically integrate $\dot{\mathbf{F}}^p(t,\mathbf{X}) = \mathbf{L}^p(t,\mathbf{X})\mathbf{F}^p(t,\mathbf{X})$ one time-step to obtain $\mathbf{F}^p(t+\Delta x,\mathbf{X})$.
5. Using the Kröner–Lee decomposition, calculate $\mathbf{F}^e(t+\Delta t,\mathbf{X}) = \mathbf{F}(t+\Delta t,\mathbf{X})\mathbf{F}^{p-1}(t+\Delta t,\mathbf{X})$.

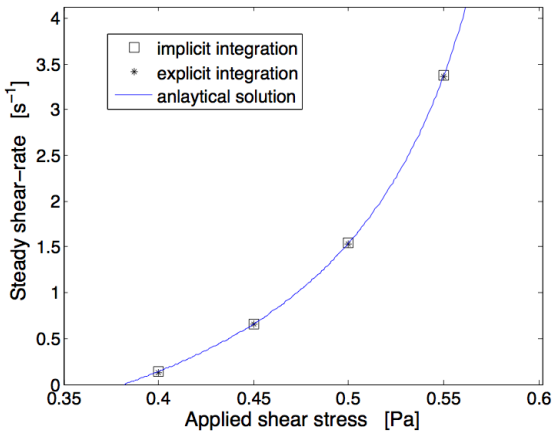


Figure 3-13: One element simple shearing tests using both the explicit and implicit VUMAT files.

6. Perform the polar decomposition on $\mathbf{F}^e(t+\Delta t, \mathbf{X})$ and compute $\mathbf{E}^e(t+\Delta t, \mathbf{X}) = \log \mathbf{U}^e(t+\Delta t, \mathbf{X})$.

7. Using the elasticity law, compute $\mathbf{M}(t+\Delta t, \mathbf{X})$ from $\mathbf{E}^e(t+\Delta t, \mathbf{X})$.

8. Inverting the definition of the Mandel stress, compute $\mathbf{T}(t+\Delta t, \mathbf{X})$ from $\mathbf{M}(t+\Delta t, \mathbf{X})$ and $\mathbf{F}^e(t+\Delta t, \mathbf{X})$.

The VUMAT file encodes this list of instructions and is provided in an appendix.

To improve the numerical stability of an ABAQUS/Explicit model, the integration step is commonly carried out *implicitly* where the time derivatives during the step, in particular the value of $\mathbf{L}^p$, correspond to the values of the quantities at the end of the step. This is opposed to the *explicit* technique outlined above. Implicit integration requires a nonlinear equation solver, typically an iterative Newton-Raphson routine, and is in general a more complicated algorithm than the rudimentary explicit integration procedure. An implicit VUMAT file for the material model is also included as an appendix. To clear up any confusion, the name "ABAQUS/Explicit" refers to the global solution process and indeed both implicit and explicit integration steps can be encoded and solved using ABAQUS/Explicit.

To check the VUMAT files for accuracy, one element tests were performed. For a list of material constants, see section 3.8. A cube-shaped element of width $5d$

(type C3D8R in ABAQUS), has its bottom four nodes fixed and its top four nodes enslaved to displace together so that the top surface moves like a rigid body. First a downward pressure of $P = 1$ Pa is applied to the top surface, and then various shear tractions are applied to the top in order to induce steady simple shear flow. The results are displayed in figure 3-13. It is clear that both VUMAT files perform the correct integration. It should be noted however that the explicit routine is only conditionally stable, and thus the time-step had to be significantly decreased in order for numerical stability at the higher shear stress values. The implicit routine is always stable numerically, but for higher shear stresses, the time step had to be reduced in order to ensure convergence of the Newton-Raphson solver.

3.7.2 Explicit procedure for global deformation and stress

While the VUMAT provides the necessary information to compute the stresses within an element given its deformation and state, how can we implement this globally to obtain the stresses and nodal motion for a mesh of elements given some traction or kinematic boundary conditions? The explicit dynamics method is well-illustrated using a 1D example that can be found in the ABAQUS 6.5 Documentation. We should quickly sketch this example as it admits a clear generalization to 3D deformations under arbitrary constitutive laws.

Consider the deformation induced when a bar fixed at its right end is compressed by some load P applied at its left end. Throughout, figure 3-14 should be used to visualize the development of the solution. We model the bar with 3 elements numbered 1 through 3, whose endpoints compose 4 nodes numbered 1 through 4. The mass of the bar is distributed onto the nodes.

At the beginning, $t = 0$, the force P is introduced to the system. The explicit algorithm is dynamic in nature, so the system as a whole does not immediately respond to this force— there is some time increment t_{inc} necessary for forces to propagate through an element. So only node 1 feels P at $t = 0$, and all other nodes feel no force during the time increment. Newton's second law is then applied to all the nodes. Since node 1 has a net force it undergoes an acceleration, but all other nodes stay

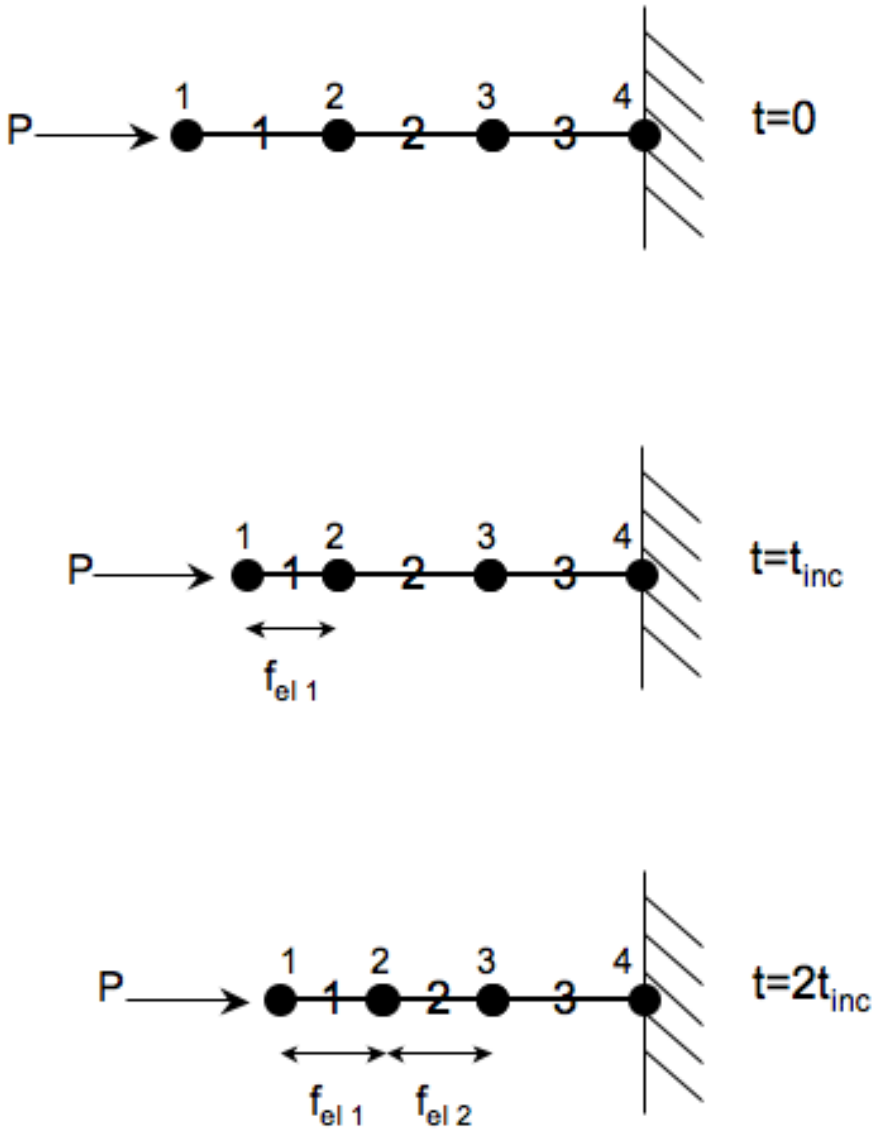


Figure 3-14: Three time steps of a bar compression problem as would be solved using the explicit algorithm.

fixed.

At the beginning of the next time step $t = t_{inc}$, element 1 has deformed. The deformation of element 1 causes a stress within element 1 as per some constitutive law. Hence, to compute this stress, the material model (VUMAT in our case) is called and applied to element 1. The output stress is then converted to reaction forces at the endpoint nodes, nodes 1 and 2. So now, nodes 1 and 2 are both experiencing force: node 1 is still being pushed by the applied load P but now is also experiencing a force $f_{el\ 1}$ from the reaction of element 1, and node 2 feels the equal and opposite reaction. Under these forces, Newton's second law is applied again at all nodes for one time increment.

At the beginning of the next time step $t = 2t_{inc}$, elements 1 and 2 are deformed. Once again, the material model is called to compute the appropriate reaction forces that these elements exert on their endnodes. Newton's second law then moves the nodes yet again under one time increment. This process of moving nodes, computing reaction forces, and then repeating the process continues until the desired total time is reached.

ABAQUS/Explicit, unless instructed otherwise, automatically computes the time increment t_{inc}. Before the solution phase of the algorithm initiates, a data collection routine is called. Several test strains are passed through the material and a numerical estimate of the highest frequency eigenmode is calculated. A characteristic element size L_e, which is commonly the width of the smallest element, is divided by the computed wave speed to yield t_{inc} as an approximation to the true stable time increment. Then the solution phase begins, as we previously described, and ABAQUS adjusts t_{inc} when deformations cause a change to the computed wave speed.

3.7.3 Artificial density reduction

By the time steady granular flow develops, a large total deformation has normally already occurred. In a Lagrangian numerical solver such as ABAQUS, large mesh distortions are accompanied by larger error. Since our interest is in steady flows primarily, we need to make a compromise so that a steady solution can emerge without

much mesh distortion.

ABAQUS 6.5 does have an "Arbitrary Lagrange-Eulerian" (ALE) procedure which can be used in problems of large strain so as to ease the inaccuracies of mesh distortion. Eulerian boundary conditions can be defined and nodal motion becomes independent of material motion under a mesh-sweeping algorithm. Unfortunately, it was this author's experience that the numerical inaccuracy brought on by convecting variables during a mesh sweep is unacceptably large, even under quadratic order convection. Error was most noticeable near boundaries, perhaps because of the reduced sample space for computation of a spatial gradient. This issue may have been improved in the most recent version of ABAQUS.

We instead adopt a different remedy based on artificially decreasing the material density in the simulation. To understand what effect this has, consider a system that closely mirrors the $\mathbf{F}^e\mathbf{F}^p$ elasto-plastic treatment— a spring and damper connected in *series* with one end attached to a wall and the other to a point mass m. The granular "spring" is very stiff so that the velocity of the point mass is roughly the extension rate of the damper. Suppose a constant force F is applied to the mass. The mass will ultimately reach the steady velocity F/c for damper viscosity c. Observe what happens if m is decreased— less mass implies less impediment to acceleration means the steady velocity can be more rapidly approached. Specifically, starting from rest, the ODE solution admits $v = f\left(\frac{(x-x_0)c}{m}, \frac{F}{c}\right)$, which indicates that steady motion occurs at smaller total displacement when m is decreased. This reasoning applies to the granular material model in that decreasing nodal masses, or equivalently the density, enables the system to find steady-state behavior before noticeable distortion has occurred.

Note in our final set of equations (section 3.6.10) that the mass density only appears in the equation of linear momentum balance. Except in the case where system curvature is large relative to the flow speed, the acceleration $\dot{\mathbf{v}}$ should always be $\approx \mathbf{0}$ in a steady flow. Thus, in general, the convection $\rho\dot{\mathbf{v}}$ should go to zero at steady-state regardless of ρ. The gravitational body force, however, is still dependent on the value of ρ even in steady state.

From all this we make the following analytical deduction: Suppose the body force is defined in terms of the true mass density but the inertial density is artificially decreased. When solving equations 3.6.10, this change will have no effect on the steady-state velocity and Cauchy stress fields under Eulerian boundary conditions. However, the time it takes for the system to reach steady-state can be made to decrease below the time necessary for large (Lagrangian) distortions to occur to the body.

Throughout the numerical implementation, this technique shall be used to obtain what appear to be Eulerian steady-state flows without having to use ALE. In fact, the elements *are* deforming, but their deformation is hardly noticeable by the time steady-state is reached. This comes with extra benefits but also with some compromises.

The universal stability of implicit integration usually makes it the preferred integration technique for user material models. However, the time increment goes as $\sqrt{\rho}$. By intensely shrinking ρ in the fashion just described, the time step becomes so small that explicit integration is rarely unstable. This notion was tested by running simulations of rough inclined chute flow and annular shear flow under both integration procedures. Nearly identical solutions for steady flow were found in both cases, though the implicit routine took slightly longer due to the iterative solver. The explicit procedure may even be the better for our purposes, since it cannot become entrapped in an infinite iteration loop and also saves a small amount of computational time.

The one element tests previously discussed all utilized some degree of artificial density reduction. When the true density is used, even with the implicit VUMAT, the integration loses accuracy giving an absurd end result. Density reduction factors on the order of 10^2 may be necessary to regain accuracy, but for the nodes to remain unnoticeably displaced as in an Eulerian problem, a reduction factor on the order of $10^4 - 10^5$ can be required.

This topic of accuracy and its relationship to the density requires more discussion. The elasticity model being used is stiff and nonlinear with elastic moduli going to zero as $\mathbf{E}^e \to \mathbf{0}$. Note the sensitive dependence of $\mathbf{M}$ on $\mathbf{E}^e$ places an equivalent sensitivity

of $\mathbf{L}^p$ on $\mathbf{E}^e$ through $\tilde{\mathbf{L}}^p(\tilde{\mathbf{M}}(\mathbf{E}^e)) \equiv \hat{\mathbf{L}}^p(\mathbf{E}^e)$, which is only enhanced by the fact that $\mathbf{L}^p$ also diverges as μ approaches μ_2. The plastic integration step in the explicit VUMAT is

$$\mathbf{F}^p_\tau = \exp(\mathbf{L}^p_t \; t_{inc})\mathbf{F}^p_t$$

where $\tau = t + t_{inc}$. Suppose for simplicity, that a single-element system has an applied deformation gradient corresponding to a constant, pure stretching: $\mathbf{F}(t) = \exp(\mathbf{A}t)$ for symmetric $\mathbf{A}$. Then the above reduces algebraically to the following integration of $\mathbf{E}^e$ to first-order in t_{inc}:

$$\mathbf{E}^e_\tau = \mathbf{E}^e_t + \left(\mathbf{A} - \hat{\mathbf{L}}^p(\mathbf{E}^e_t)\right) t_{inc}.$$

The integration of $\mathbf{E}^e$, and consequently the stress, will be inaccurate unless $\hat{\mathbf{L}}^p(\mathbf{E}^e)$ stays roughly constant over the time increment relative to the size of $\mathbf{A}$. Given the sensitivity of $\mathbf{L}^p$ on $\mathbf{E}^e$, this can be a strong restriction. And though explicit integration was used here, the point also applies to implicit integration since it too presumes $\mathbf{L}^p$ roughly constant over an increment.

ABAQUS calculates t_{inc} as the time necessary for the fastest wave to travel a characteristic element width L_{el}. Since the fastest wave is commonly a dilation wave, the time increment is thus delicately connected to the elastic bulk modulus $\kappa(\mathbf{E}^e)$, especially in a plastically incompressible model. One can approximate the requirement for numerical accuracy as:

$$|\hat{\mathbf{L}}^p(\mathbf{E}^e_\tau) - \hat{\mathbf{L}}^p(\mathbf{E}^e_t)| \approx \left|\frac{\partial \mathbf{L}^p}{\partial t}\right| t_{inc} \ll |\mathbf{A}|$$

Performing the chain rule and expanding the form for the time increment gives

$$\left|\frac{\partial \hat{\mathbf{L}}^p(\mathbf{E}^e)}{\partial \mathbf{E}^e}\left(\mathbf{A} - \hat{\mathbf{L}}^p(\mathbf{E}^e)\right)\right| L_{el}\sqrt{\frac{\rho}{\kappa(\mathbf{E}^e)}} \ll |\mathbf{A}|$$

Since $\hat{\mathbf{L}}^p$ is stiff and nonlinear, the inequality is primarily threatened by the term in absolute values. But, alas, this issue can be counteracted if ρ is selected to be small

enough. Consequently the accuracy of the integration procedure is enhanced by artificially reducing the density. Over N time increments, this refines the numerical integration over a smaller total time period Nt_{inc}. Since the steady-state emerges proportionally quicker, the fact that the total time represented by N time steps decreases is not a concern.

Since decreasing ρ speeds up the motion of waves, some particular physical features would be skewed. For instance, if one were interested in acoustic properties of the grains, this technique should not be used. How long a flowing system spends going through transients would also be skewed as this property is too closely tied to the wave-speed. But in general, for our purposes of obtaining the stresses and flow primarily in steady-state, these physical issues are not our concern.

3.8 Results

The model is tested in three different 3D flow environments— rough-walled inclined chute, annular Couette cell under gravity, and wide draining silo. Throughout, the model's six parameters are assigned the following values:

$B = 7 \times 10^9$ Pa	$I_0 = 0.279$
$\rho_s = 2450$ kg/m^3	$d = 0.003$ m
$\mu_s = \tan(20.9°)$	$\mu_2 = \tan(32.76°)$

Recall that ξ is also a parameter, but its value is tied directly to the value μ_2. Except for d, these values were all lifted from Jiang and Liu [68] and Jop *et al.* [71]. As both groups considered spherical glass beads in their work, it is assumed that the numbers should be representative of the same material. The particle diameter is set to 3mm, as this is the common value used in the experiments and DEM simulations of the MIT Dry Fluids Group, whose data this model will be compared against particularly in the case of the wide silo. Though we shall also compare against data

from other groups, the particle diameter is almost always scaled out in the non-dimensionalization. The elements used, unless otherwise stated, are hexahedral of type C3D8R in ABAQUS.

As a caveat, the following results are all rather new and may undergo some additional refinement in the near future. This pertains chiefly to the description of boundary conditions. The true interaction laws between a solid wall and a flowing granular media are very complicated. For now, we settle for simplified wall conditions, frequently utilizing fully-rough or kinematic constraints where acceptable.

ABAQUS/Explicit does not invoke a formal notion of "convergence", so the results are deemed to be correct based on evidence. When a solution appears valid, the mesh is then refined by a factor of two. If this only causes negligible changes to the velocity field, pressure field, and total kinetic energy, the refined mesh solution is accepted as correct.

The simulations are run long enough for transient behavior to disappear. We may refer to this as "steady" flow, but a more general meaning is implied, since environments like silo drainage do not correspond to an Eulerian boundary value problem (that is, the top surface always descends). The disappearance of transients is most apparent in the total kinetic energy, which becomes flat (compared to the kinetic energy transients) when steady-state has been reached. Basic inspection of the solution also tends to make it clear when steady flow has been established.

3.8.1 Rough-walled inclined chute

Figure 3-15(a) reviews the geometry and boundary conditions of the rough inclined chute problem originally solved by Jop *et al.* with a Bingham fluid treatment. The elasto-plastic formulation has added many new relationships and a different kinematical perspective on deformation. However, the steady flow profile it predicts should be inconceivably different from the Bingham fluid result of Jop *et al.*, since elastic stretching is negligible in a steady flow. Moreover, as finite elasto-plasticity is less prevalent in fluids communities, the ability to duplicate this result is intended to give credence to the method.

The chute is long such that the flow can be deemed invariant in the z direction. The bottom of the chute is at $y = 60d$ and the chute is $142d$ wide. The angle of incline of the chute with respect to the horizontal is 22.6°.

The chute model uses 2600 box-shaped elements arranged in a grid 65 elements wide in the x direction by 40 in the y. The assembly is only 1 element wide in the z direction since that direction is constrained to be symmetric. That is, the nodes on the back face of the xy slab are constrained to move with the exact displacements as their front face counterparts. The simulation must start in a reference state with no gravity. Gravity is then ramped up from 0 to its full value. Since the material must be compressed before any shear stress can be supported, the components of gravity are applied one at a time. The component in the y direction is first smoothly ramped up over 1×10^{-5}s to its final value of $f_g \cos(22.6°)$, where f_g is the correct body force per unit volume $\approx 0.6 g \rho_s$. After a 2×10^{-6}s pause for any needed relaxation, the component of gravity pointing down the chute, the one that initiates flow, is smoothly ramped up to its final value of $f_g \sin(22.6°)$ over 2×10^{-5}s. The assembly is then left to flow until a total time of 5×10^{-4}s has elapsed.

The inertial density was artificially decreased by a factor of several hundred thousand to a value of $\rho = 2 \times 10^{-3}$. At the free surface, theoretically, the stresses go to zero causing the elastic moduli to vanish. This is dangerous for an explicit solution procedure that is consistently sending waves through the material— any small perturbation could cause a node to accelerate out of control. Not to mention, the flow rule is undefined for a zero stress state. Thus, to keep the free surface in tact, a few Pascals of overpressure are applied.

While the flow is ultimately very steady looking to the eye, a more quantitative measure may be desired. From $t = 0 - 4 \times 10^{-5}$s, the motion is markedly transient due to the fact that gravity is being ramped up. The system then relaxes toward the steady state, first rapidly but then quite slowly as steady state is approached. Comparing the system's rate of relative kinetic energy change in the fast relaxation period to its value at the end of the run gives a rough criterion for how steady a flow is. At the end of the run, the rate of decrease of the relative kinetic energy is over 500

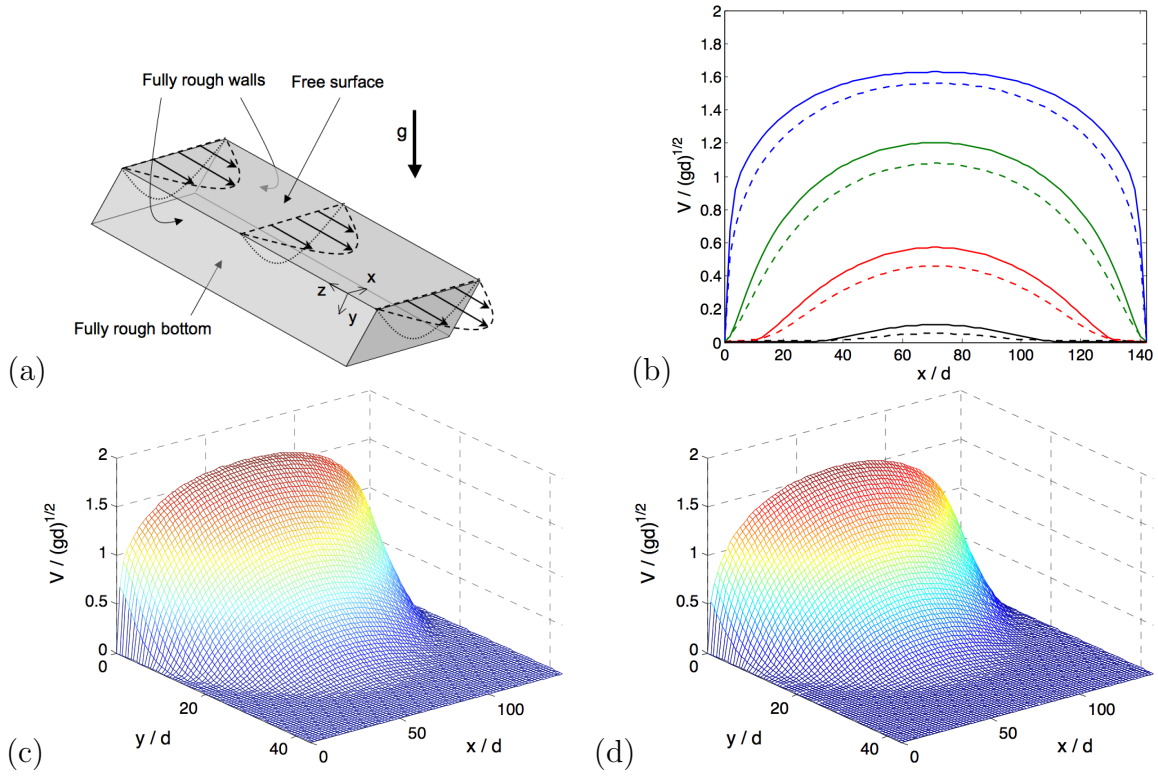
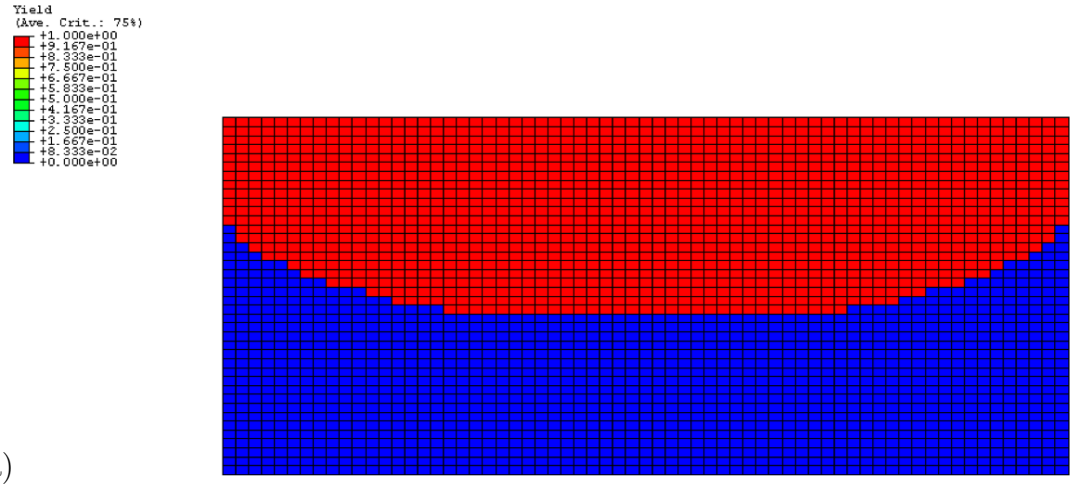


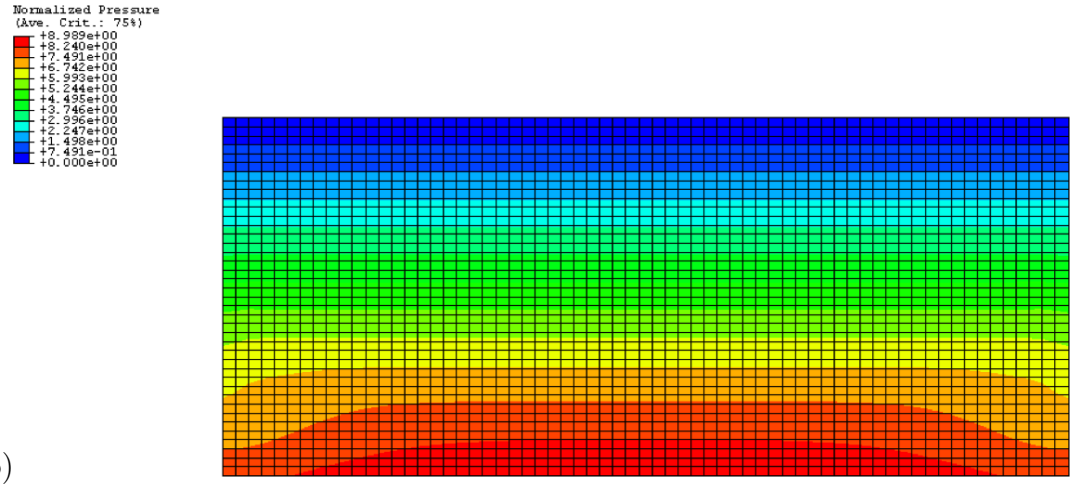
Figure 3-15: (a) The rough inclined chute setup. (b) The velocity as a function of x at depths $y/d = 0, 9.2, 18.4, 27.6$. The dashed curves are the numerical results of Jop *et al.* [71] via finite differences with a Stokes-type solver. The solid curves are from the steady solution of the elasto-plastic model implemented on ABAQUS/Explicit. (c) The full velocity field as solved on ABAQUS/Explicit and (d) the velocity field from Jop *et al.*.

times smaller than the value it takes during the initial relaxation, which indicates a strong steady-state behavior.

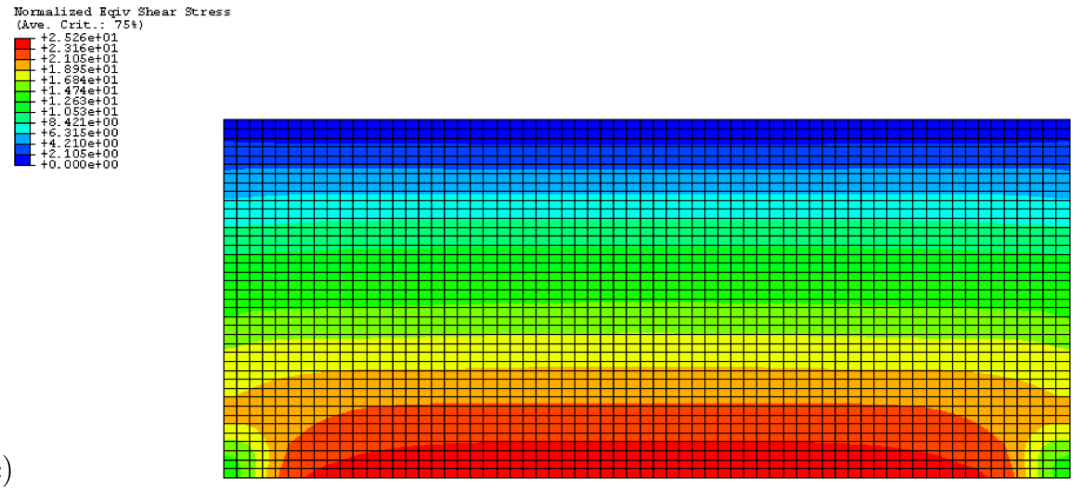
Comparisons to the numerical results of Jop *et al.* are displayed in figure 3-15. The agreement is quite good considering how disparate the solving methods are. Jop utilized a fixed grid finite-difference scheme solving a non-Newtonian Navier-Stokes-type problem, while the elasto-plastic results were obtained by a Lagrangian explicit procedure on nodes that are constantly moving during steady-state. While differences of numerical procedure can cause different sources of gain and loss, it seems more likely that the discrepancies are stemming from the free surface condition. The true deformation rate at the top corners is actually infinite. As this is numerically impossible for either scheme, large but finite gradients occur there numerically as



(a)



(b)



(c)

Figure 3-16: View of the xy plane with y downward. (a) Red is yielding material, blue is static. The normalized pressure is $P/\rho gd$ in (b) and normalized shear stress is $\tau/\rho gd$ in (c).

determined by the choice of the free surface treatment. This has a clear trickle-down effect on the global velocity field and, in lieu of the particular fashion in which the solutions differ, we suspect this effect causes much of the difference between the two numerical data sets.

What distinguishes the elasto-plastic model, though, is its ability to calculate both flow *and* stress *everywhere*. Figure 3-16 displays a plethora of different stress-based quantities. Note that the stresses always vary smoothly through the transition from yielding to static, as opposed to the interfacial stress issues that commonly occur with Bingham fluid models. In particular, observe that the pressure field goes from decreasing linearly in the flowing zone, to decreasing somewhat nonlinearly in the static zone. In this geometry, compressive stresses in the x, y and z directions of the steady flowing zone must all be identical under codirectionality. A hydrostatic pressure profile is induced as a result. But upon descending below the flowing layer, codirectionality no longer has this influence on the stresses and a somewhat more complicated elasto-static form for the pressure field becomes apparent. Discrete simulation data for this environment would be helpful for checking the validity of the stress profile in the static zone. Though, as previously admitted, the primitive fully-rough boundary conditions used may not be the most accurate reflection of the true conditions on the bottom and side walls.

3.8.2 Annular Couette cell

Since the elasto-plastic model is fully general, any well-posed boundary value problem can be solved. The equations and solving algorithm are geometry independent. No major changes are necessary to simulate flow in a non-chute type environment, thus we easily move on to the annular Couette cell, in which no previously known solution to the Jop–Pouliquen flow law exists.

The results shall be compared directly against the myriad of data on this environment compiled by G. D. R. Midi [94] and thus the geometric specifics and boundary conditions were chosen so as to give a good representation of the conditions used in these studies. Referring to figure 3-17a for general details on the environment, the

following values were chosen: $\omega_{\text{wall}} = 0.3\text{rad/s} \approx 0.1\text{rev/s}$, the distance from inner to outer wall is $x_{\text{out}} = 30d$, the height is $z_{\text{bottom}} = 10d$, and the inner wall radius is $40d$. At the walls, the material motion must match the wall motion in the θ and x directions, but material can slide without resistance up and down the walls.

Since flow and stress should be symmetric in the θ direction, the behavior as seen in a downward cut through the annular trough should represent the global behavior. A narrow sector of the annulus (total angle 0.1°) was likewise simulated using periodic boundary conditions on the front and back faces— nodal displacements on the front face are constrained to be identical to those on the back face except rotated appropriately by 0.1°. The sector is modeled using 40 elements in the x direction, 15 in the z, and a thickness of one element in the θ direction, for a grand total of 600 elements. Almost all the motion is known to occur near the inner wall in this environment. After preliminary tests of the elasto-plastic model produced the same conclusion, a bias was utilized to crowd nodes closer to the inner wall and improve precision. The bias resulted in half of the elements occurring within a distance of $6d$ from the inner wall.

The inertial density was reduced by a factor of over one hundred thousand to a value of $\rho = 0.01$. First, gravity is smoothly ramped up to its full value over $t = 0 - 1 \times 10^{-5}$s. During this time period, as before, a slight overpressure is also applied to the free surface for numerical ease. The rotation of the inner wall then commences, smoothly ramping up from a rotation rate of 0 to a final value of $\omega_{wall} = 0.3$rad/s over $t = 1 - 2 \times 10^{-5}$s. The simulation is then left to flow until a total time of $t = 1 \times 10^{-3}$s elapses.

After the wall has finished ramping up to ω_{wall}, the flow soon after begins to relax toward the steady state. At the beginning of the relaxation, the total kinetic energy is changing by a multiplicative factor of roughly 6×10^4 per second. By the end of the run, that number has dropped to 8.5×10^{-1} per second. Since the relative rate of change of total kinetic energy at the end of the run is dramatically smaller than during the material's natural transient period post ramp-up, by a factor of almost a hundred thousand, we deem the final flow state to be sufficiently steady.

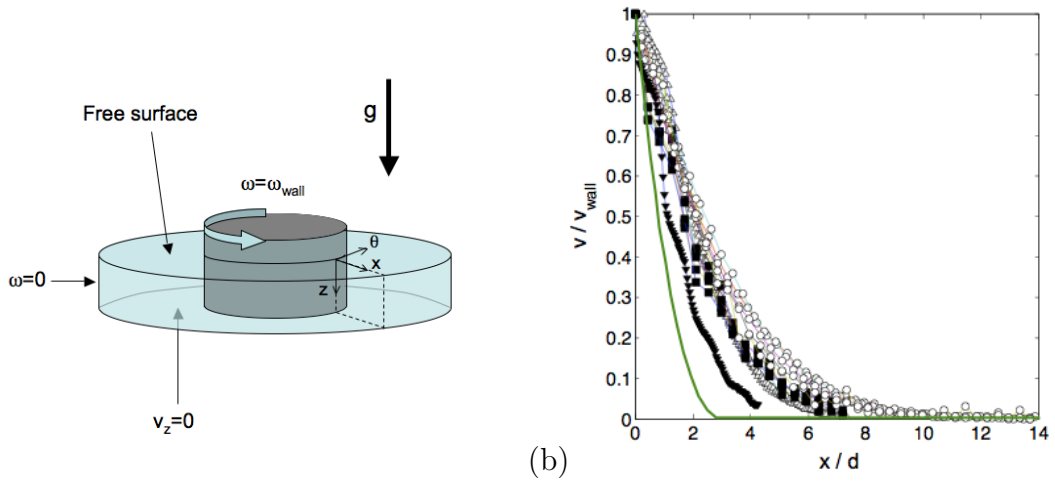
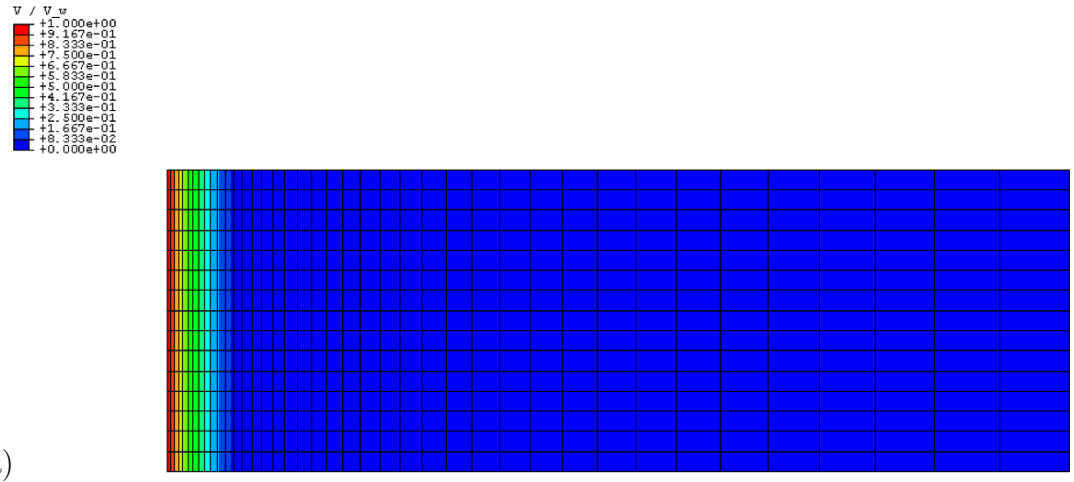


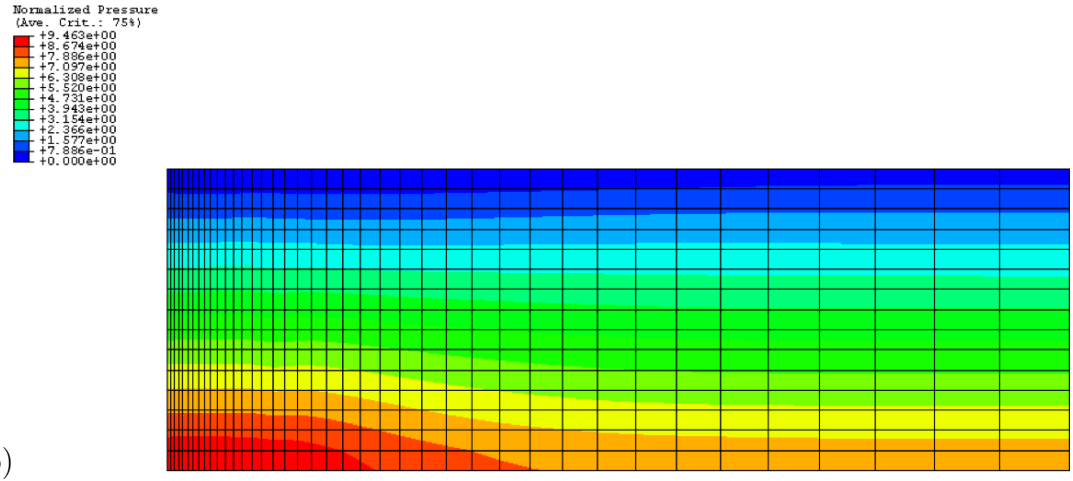
Figure 3-17: (a) The annular Couette setup. (b) The velocity profile normalized by the wall speed as predicted by the elasto-plastic model at half-height (–) compared against fifteen experimental and discrete element simulation data sets for this type of flow as compiled in [94].

The simulation was continued longer than necessary for steady flow conditions in order to check for an effect that will now be described. In this environment, the element thickness being small is instrumental to the flow reaching a numerical steady-state. Note that while a true annular flow can achieve steady-state, where local volume changes stop, the elements in our simulation have straight edges making this geometrically impossible. The cylindrical walls of the simulation are in truth not cylinders but 3600-gon prisms. As with any pair of concentric n-gons, rotating the inner one while holding the outer fixed is not an isochoric motion. Consequently, both the pressure field and flow field will not appear steady during the time window unless n is quite large, so that the walls very closely mimic cylinders. This was not an issue for the rough inclined chute where geometrically speaking, box-shaped elements can steadily flow without volume change. But for the annular Couette, even a 1° sector is too thick— the velocity field never actually becomes steady, rather a very slight rate of increase in the velocity can be observed until the end of the run. In general, this phenomena may occur whenever the exact isochoric steady flow, if imposed on the mesh, would cause inhomogeneous deformation within the elements.

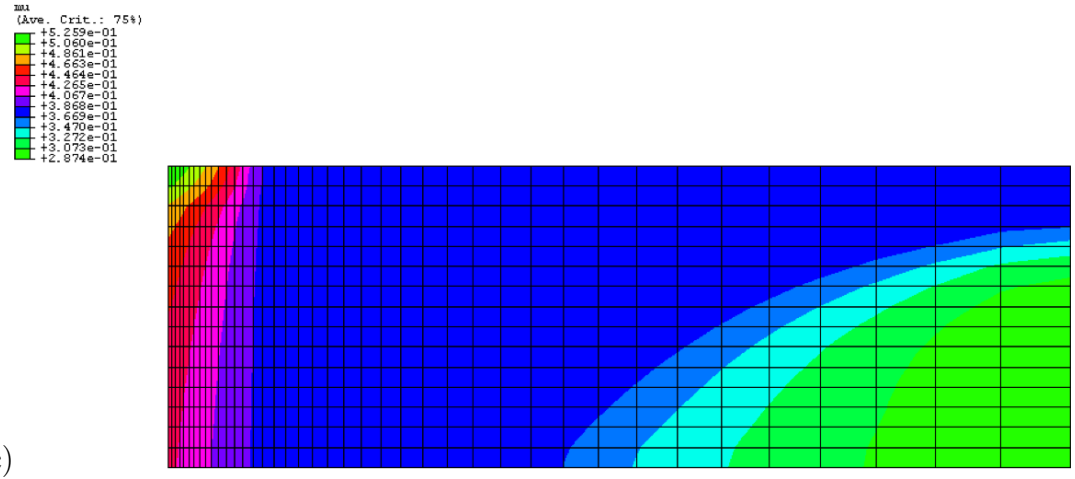
Observing figures 3-17b and 3-18a, we notice a few major qualitative points. For



(a)



(b)



(c)

Figure 3-18: In the above, the xz plane is shown with z downward. (a) The velocity of the material normalized by the wall speed. (b) The normalized pressure is $P/\rho gd$. (c) The stress ratio $\mu = \tau/P$.

one, the flow forms a clear band around the inner cylinder. The width of the band is on the order of several particles diameters. The experiments and DEM simulations of [94] verify the existence of a shear band on the order of this thickness surrounding the inner wall. While the predicted thickness is smaller than the experimental, the fact that it is on the same order is a major success for a model with no fitting parameters. The authors of [71] expressed their belief that the flow rule would be incapable of describing narrow shear bands. However, the above results show quite clearly that a complete elasto-plastic model incorporating the same flow rule has no problem resolving such features.

Another observation is that the velocity profile does not vary to any observable extent in the z direction. This result has been verified in DEM simulations of this environment [74], where it was found that almost no vertical fluctuation occurs in the fast zone near the inner wall. Of course, the elasto-plastic velocity field dies off much differently than experiment. Where the experimental data is shown to be well-fit with an exponential decay that extends throughout the Couette cell, the elasto-plastic solution predicts a sharp cutoff around $3d$ from the inner wall. This result should come as no surprise, since the model does not account for quasi-static, non-local behavior, of which slow exponential decay is a textbook case. The fact that the predicted velocity profile changes sharply over a width of less than one RVE is a strong indication that the model, which is only a first-order law for moderate flow rates and statics, is not equipped to describe all the details of this environment. While granular matter does not exhibit a sharp flow cutoff under annular shear, some related materials do such as muds and suspensions, so the prediction is not outlandish.

Viewing figure 3-18c, it is clear that while the velocity appears invariant in the z direction, other important stress-based quantities are not. Though μ has a non-trivial z dependence, note that the pressure field in the flowing zone displays a similar hydrostatic appearance as was observed in the rough inclined chute. Far from the inner wall, the pressure again appears constant in x, though the rate of growth in the z direction is slower. While the vertical compressive stresses should, and are observed to be the same at both walls, the disparity is caused by the different conditions imposed

on the other compressive stress components. In the flowing zone, codirectionality requires that the compressive stress in each of the coordinate directions be identical since the shear planes are closely aligned with the $z\theta$ surfaces. On the other hand, near the outer wall, a situation closer to an "active state" elasto-static solution forms, where the weight of the material causes compressive stress in the z direction to be larger than in the other coordinate directions.

Experimental data on this flow environment indicate that the normalized velocity profile does not change with flow rate over a common range of rates. However, it has been shown that in a regime of larger shear rates, the shear band width does get wider as flow rate increases [77]. It is likewise worthwhile to test how the elasto-plastic model responds to changes in wall speed. One might imagine that increasing the wall speed without changing gravity should cause the onset of yield to move farther from the inner wall. To check this, the model was simulated again but with the wall rotation rate reduced by a factor of 5. The resulting steady velocity field had its flow cutoff three times close to the inner wall. Indeed, decreasing the flow rate decreases the shear band width under the elasto-plastic model. In some sense it is not surprising that the model agrees with the rapid shear flow result, as higher shear rates bring more material into the moderate flow regime thereby increasing the dominance of the local rheology over any non-local behavior.

3.8.3 Flat-bottomed silo

While the past two flow environments show the versatility of this model in predicting flow, the stresses have yet to be directly checked. Experimentally, the stress tensor is a difficult quantity to measure. The stresses in 2D disk assemblies can be approximated using photoelastic grain material, however, there is not currently an experimental method available to measure the stress tensor within an arbitrarily flowing 3D granular material. For this measurement, the best option as yet is to utilize DEM and compute the local stress tensor from the contact forces as per equation 3.13.

Chris Rycroft has performed DEM simulations of wide silo drainage [122]. His

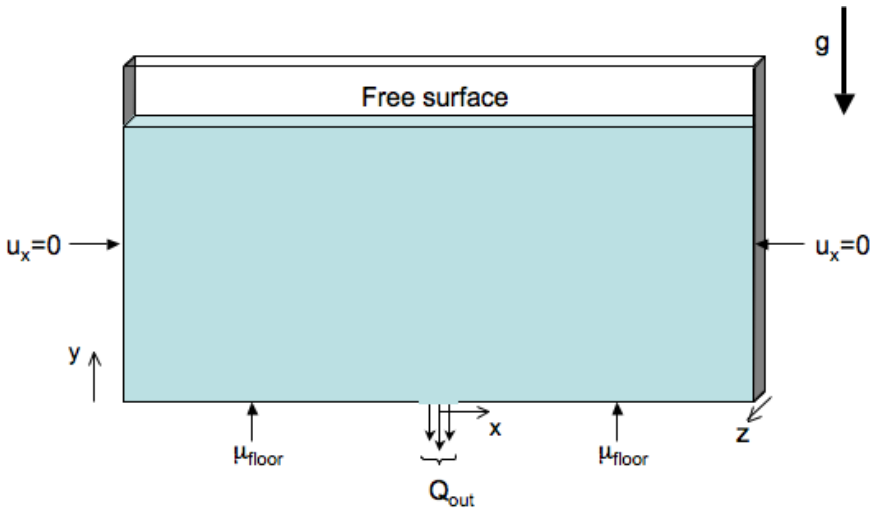


Figure 3-19: The flat-bottomed, quasi-2D silo setup.

computed stress and flow fields shall now be compared directly against the predictions of the elasto-plasticity model. A schematic diagram of the flow geometry is pictured in figure 3-19. In accordance with Rycroft's simulation, we model the silo as having an opening width of $6d$, a height of $70d$, and a full width of $150d$. In Rycroft's simulations, the z direction has periodic boundary conditions so that the silo has an apparent z thickness of $8d$ but has no wall-ordering effects. Since the flow should not fluctuate in the z direction, we simply model the silo as a plane-strain environment. Furthermore, the silo has left-right symmetry about the vertical center-line, which we take advantage of by modeling only the right half of the silo.

The floor of the silo is modeled as having a frictional interaction with the material characterized by a coefficient of friction $\mu_{\text{floor}} = 0.2$. This number was estimated loosely from Rycroft's simulation. While the surface of his particles had a friction coefficient of 0.5 with the floor, this is *not* the number we should use. Rather, we must input how an element containing many particles drags along the floor. To account for the weaker frictional bonds induced by the effects of particle rolling and rearrangement, the element/floor interaction was modeled with a 60% smaller friction coefficient. Future study would be necessary to determine the complete and precise form for such sliding interactions. While Rycroft's simulation utilized side walls made

of the same frictional material as the floor, for simplicity, the elasto-plastic simulation employs the simpler condition of no x displacement at the side-walls. This could be enhanced in the future, but in wide silos, the details of the side-walls have only a small effect on the dominating behavior.

Ideally speaking, the boundary condition at the silo orifice should be zero stress tractions. However, this is highly problematic numerically. In reality, silo flows develop a "free-fall arch" directly above the orifice [101]. The arch is typically only a few particles high and connects the edges of the opening. Once a particle passes through this hypothetical arch, it enters free-fall and becomes gas-like. A granular material element within the free-fall arch would realistically have a smaller packing fraction, but still support compressive stresses through internal random particle collisions. The elasto-plastic model, however, does not include gas-like effects. Such dilation would be assigned to the elastic part, causing the elastic moduli to vanish.

Our interest is not in the details within the immediate vicinity of the orifice, rather the bulk material behavior within the greater silo apparatus. But with a zero-traction condition on the opening, the situation described above dooms the simulation prematurely— once elements near the orifice undergo net dilation, they quickly destabilize. While free surfaces are usually taken care of by adding a slight amount of artificial compression, this remedy will not suffice at the opening because the applied pressure has too much of an effect on the evolution of the outflow rate. We are left with the alternative of using kinematic boundary conditions at the orifice. It would be too overreaching to assign any particular velocity profile at the orifice. Instead, we fix the total flux out the orifice and let the material response determine the shape of the flow profile. To match the outflux in Rycroft's simulation,

$$Q_{\text{out}}/2 = \int_{\text{Right half-opening}} v_y(x, y = 0)dx = 2.19 \times 10^{-3} \ \frac{\text{m}^2}{\text{s}}$$

was instituted at the orifice, encoded as an equation constraint in ABAQUS.

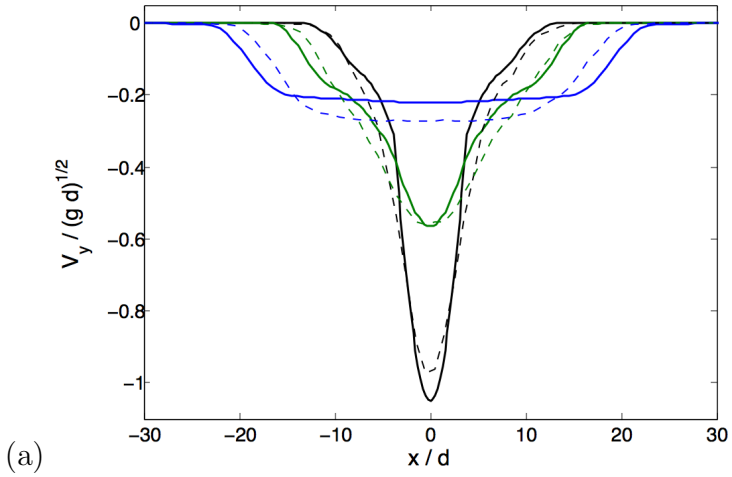
Not far from the opening, large inhomogeneous deformation occurs at small length-scales, necessitating many small elements to maintain accuracy. A grand total of

9750 elements were used in modeling the half-silo. To minimize discretization error, the orifice was modeled with a half-width of 15 elements. The adjacent silo floor was modeled with a width of 60 elements. The silo height was modeled with 130 elements. The element width was constant within the orifice, but bias was used along the other boundaries to maintain smooth changes in element sizes throughout. Elements shrink vertically as a sole function of distance from the silo bottom. The element width is uniform at the top surface, but moving downwards, elements crowd the center more and more so that the floor region has a smooth transition from wider elements near the wall to narrower elements adjacent to the orifice.

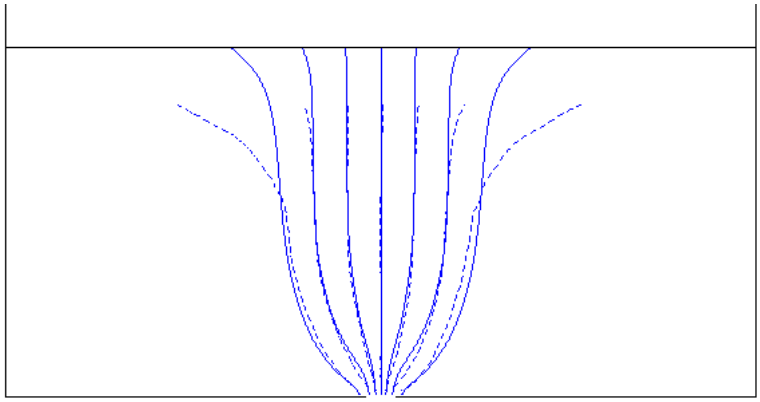
Due to the high number of elements and small minimal element size to system size ratio, this flow would take too long to compute on one processor. Instead, the 12-node Truesdell cluster of the MIT Solid Mechanics Group was employed to solve the problem in parallel. Using domain-level parallelization, the cluster splits the half-silo into 12 spatial regions. Each region is handled by one node and node-to-node messaging is used to communicate between regions. The problem solved was as follows: From $t = 0 - 5 \times 10^{-5}$s, gravity is gradually turned on while constraining the nodes along the opening from moving in the y direction in order to model a closed orifice. Then, from $t = 0.55 - 5.00 \times 10^{-4}$s, the orifice is gradually opened by smoothly ramping up the outflow rate from zero to $Q_{\text{out}}/2$. The simulation is then left to flow until a total time of $t = 10^{-3}$s has been reached.

This environment does not possess a steady-state since it is not an Eulerian boundary-value problem at the free surface. Instead, patternistic behavior eventually occurs, which signifies that transients have finished passing— starting at approximately $t = 5 \times 10^{-4}$s, the velocity and d^p fields appear to fluctuate regularly.

The first direct comparison that should be made is between the elasto-plastic and DEM flow profiles. To represent fully-developed mean behavior, Rycroft's flow data was averaged over 100 frames, during a period of what appears to be transient-free flow. Similarly, the elasto-plastic flow was also averaged over many frames of fully-developed motion. To improve the validity of this average, the model was run an extra 5×10^{-4}s longer and the time average was performed over the range $t =$



(a)



(b)

Figure 3-20: Comparison of elasto-plastic results (—) to DEM (- -). (a) The y velocity component as a function of x at heights $y = 5d, 10d, 30d$. (b) Trajectories predicted by the elasto-plastic model alongside the DEM trajectories. Container outline provided for ease of viewing.

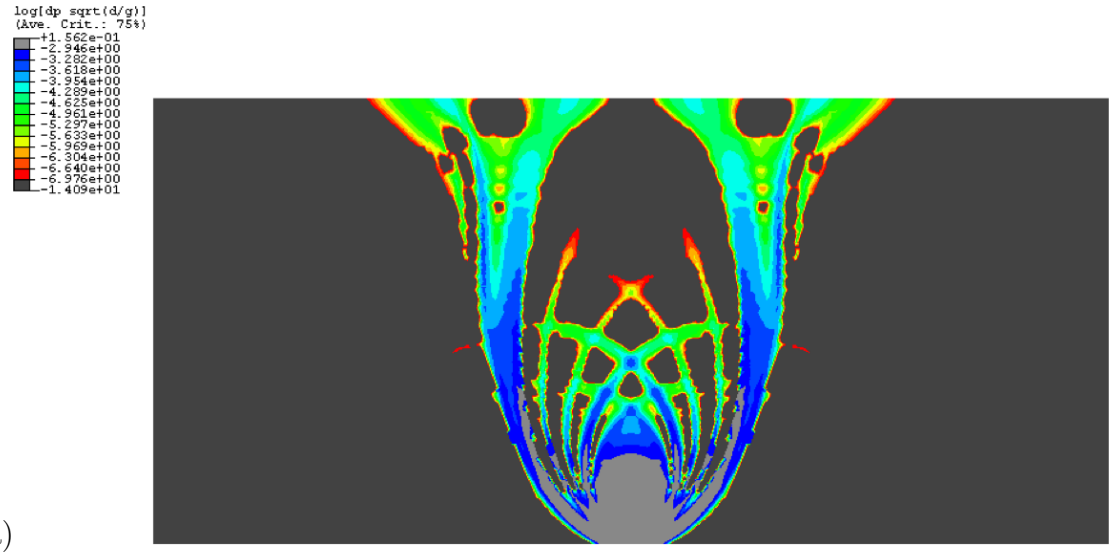
$0.5-1.5\times10^{-4}$s, comprising 127 frames. Figure 3-20 displays the comparison. Overall the agreement is sufficient. The particular way in which the peak in the downward velocity component broadens as height increases is well captured by the model. Once again, as is now a common theme, the elasto-plastic downward velocity appears to change more rapidly in space. Non-local effects such as diffusion could smooth out these sharper variations and possibly improve the agreement. Observing figure 3-20b, the DEM and elasto-plastic trajectories agree quite well below $y \approx 40d$. The elasto-plastic model and DEM both have the trajectories pealing outward higher up in the silo, though the onset of this behavior occurs at different heights for the two

models. The difference in the upper silo trajectories could stem partially from the fact that the free surfaces no longer match— the DEM free surface has already undergone a non-negligible downward displacement and is no longer flat during the pictured flow calculation whereas, thanks to artificial density reduction, the free surface of the elasto-plastic model still looks as it did in the beginning. It is also possible that the model is less accurate in the upper part of the silo because the flow is much slower there.

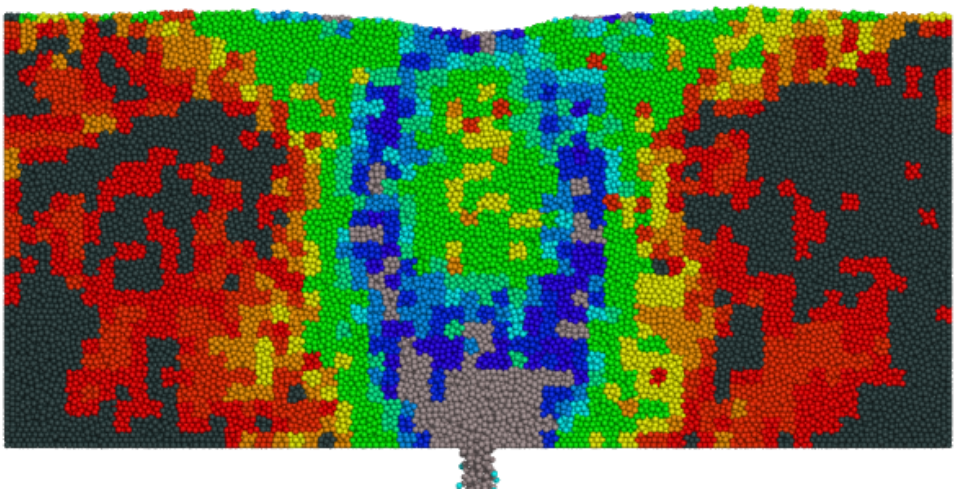
To compare instantaneous behavior, let us now observe snapshots of the shearing rate profile. As pictured in figure 3-21, the fully-developed elasto-plastic behavior involves two long "arms" of shearing that extend from the edges of the orifice to the top surface. From those large arms, smaller shear bands drop down one-by-one forming, at any instant, a criss-crossing field of shear bands within the two arms. The plot is logarithmic, so the shear bands are actually quite small, but the log-scale is helpful for accentuating small features that undergo patternistic events. The DEM solution does look similar, with two dominant arms of shearing. Note as well the similarities near the top of silo, with both plots showing the shear arms spreading out approaching the free surface. The most obvious difference, as has been previously noted, is that the elasto-plastic solution has sharp flow cutoffs whereas the shear rate always gradually tails off in the DEM. To reiterate, the model lacks non-local terms describing quasi-static behavior so flow cutoffs are to be expected. Even if shear bands did exist in the DEM, the box-average being performed would most assuredly obscure them.

Moving on to the stresses, we first check the principal stresses and directions during fully-developed flow. As is evident in figure 3-22, the principal stress orientations predicted by the elasto-plastic model closely match those of the DEM. Both show the major principal stresses forming "arches" about the orifice and, moving away from the opening, the principal stress chains transition to becoming more vertical.

In the DEM, it is apparent that the major principal stresses adjacent to the side walls have an orientation that slightly tilts away from the walls, whereas those of the elasto-plastic solution remain almost perfectly vertical. This is entirely due to



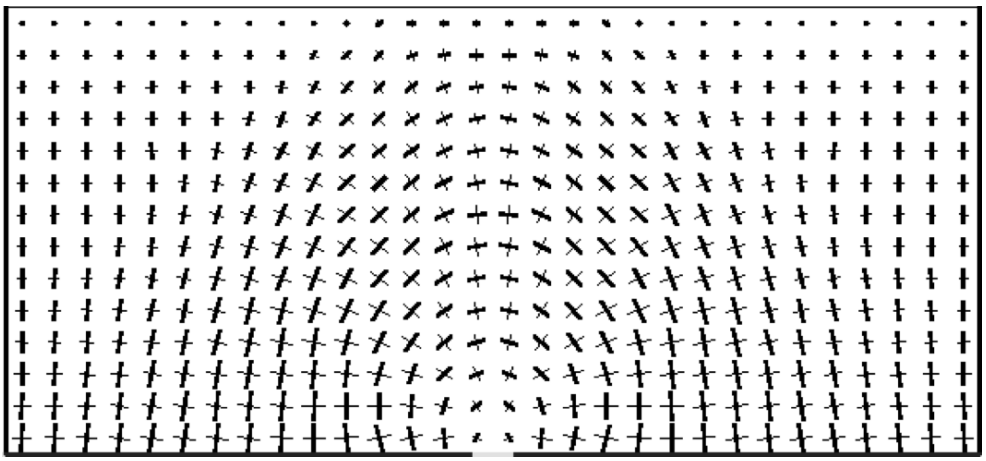
(a)



(b)

Figure 3-21: The plastic shearing rate d^p, expressed in units of $\sqrt{g/d}$, plotted in logarithm form. (a) The elasto-plastic solution: Note the intricate pattern of shear bands that fill the region between the two long shearing arms. The long-time behavior has the bands fall down one-by-one from the larger shearing arms. (b) The DEM solution: Similar to the elasto-plastic except blurred out by the box-averaging and the effects of non-local "diffusion" that are ignored by first-order elasto-plasticity. Both plots use the same color scale.

(a)



(b)

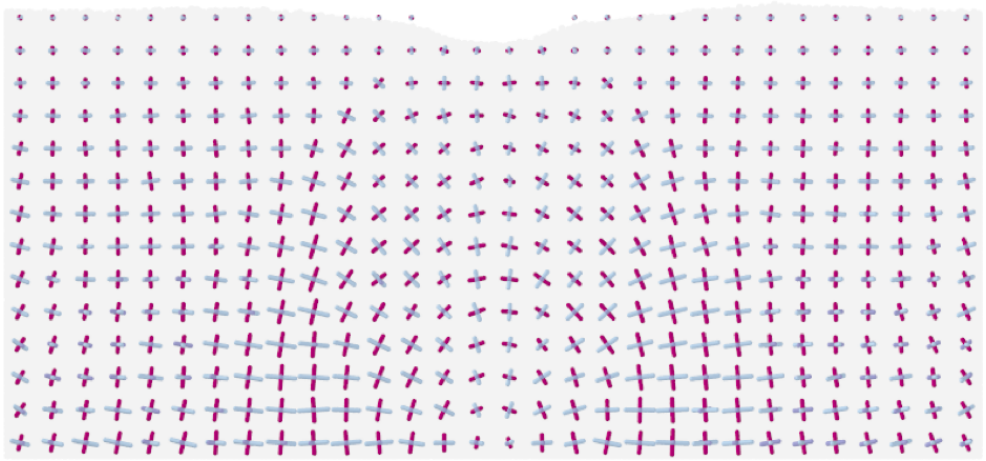


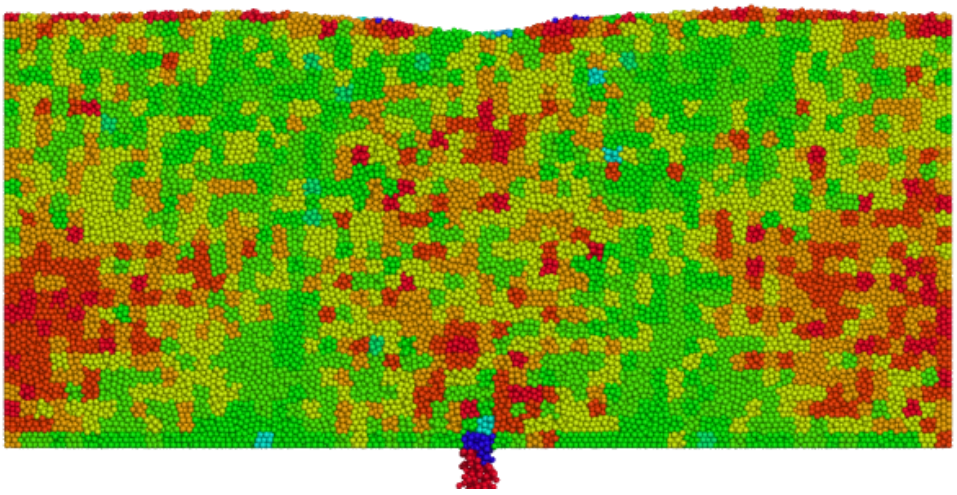
Figure 3-22: The instantaneous deviatoric principal stress directions plotted as crosshairs with lengths corresponding to the associated deviatoric stress eigenvalues. (a) The elasto-plastic solution: The thicker of the lines corresponds to the major principal stress direction. (b) The DEM solution: Major principal stress in purple, minor (and intermediate where visible) in blue.

the chosen boundary conditions. The DEM utilized frictional sidewalls and the tilt indicates the walls are exerting an upward shear on the material. This wall shear can be traced back to the filling process, where pouring causes an "active wall state" where walls respond to resist the downward motion of the grains. The side walls in the elasto-plastic model had no friction and thus cannot exert shear on the material. Consequently, vertical major principal stresses near the side walls come as no surprise. In the future, a side wall friction may be included to better model this effect, but it is truly a minor point.

The stress orientations flanking the orifice deserve some attention. In the DEM solution, the minor principal stresses on the floor immediately adjacent to the orifice are approximately parallel to the floor. This would seem to imply that the floor applies little to no shear resistance on the material that is about to leave the silo. This result is incompatible with the choice of boundary conditions used in the elasto-plastic model. By selecting a friction coefficient μ_{floor} to describe the floor/material interaction, any motion along the floor must be resisted with a shear stress proportional to the normal stress. This is evident in the tilted crosshairs next to the opening in the elasto-plastic solution. What this points out is that the actual boundary condition may not be as simple as a frictional interaction. Quantifying the interaction between flowing grains and a rough surface would be important future work.

To better observe the relative size of the stress components, figure 3-23 displays the stress ratio. Qualitatively, the two solutions show similar spatial changes. Both solutions show a region of lower μ that swoops up from the lower side walls together with scattered minima of μ in the upper-middle region of the silo. Quantitatively, it appears the μ values in the flowing zone are lower for the elasto-plastic solution. One possible explanation for this could be that grains in the DEM simulation had a rougher surface than the glass beads of [70] from which our μ_s and μ_2 were extracted. Perhaps now is a good time to reemphasize that the material parameters used in the elasto-plastic model are not necessarily those describing the beads in the DEM. Rather, our parameters came from two *different* papers with the only commonality being that both studies claimed to be using glass beads. Regardless, in [33] it was

(a)



(b)

Figure 3-23: The instantaneous stress ratio μ during fully developed flow. (a) The elasto-plastic solution. (b) The DEM solution. Both plots use the same color scale.

demonstrated that increasing the surface roughness of grains causes μ_s to increase. If the elasto-plastic model were implemented with a larger μ_s, the μ profile would indeed increase in the flowing regions, as larger μ would be needed to invoke the same plastic shear rate.

3.9 Conclusion and future directions

This work proposes a highly general, 3D granular continuum model that unifies the most recent results in granular elasticity and plasticity. The unification follows a rigorous $\mathbf{F}^e\mathbf{F}^p$ decomposition and is guided by the most basic physical principles. The model can be used to predict flows uniquely in any environment with mechanically well-posed boundary conditions and/or body forces. The model was implemented as a user-material in ABAQUS/Explicit and tested in three unrelated geometries. With no fitting, it appears to give qualitative, and in some case quantitative, predictions for both the stress and velocity field in arbitrary granular flow geometries. Even so, there are a few clear avenues of future work with regard to improving the current model.

3.9.1 Quasi-static non-locality

The most glaring effect absent from the model is that it cannot account for "blurring" in the flow fields. As described in depth in sections 3.4.2 and 3.4.3, when the normalized shear rate decreases low enough, the rheology is no longer determinable from a simple relation of the form

$$I = g^{-1}(\mu).$$

The plasticity model being used is of this form, but in truth, one would speculate a more complete form has higher order gradients in stress, flow rate, and/or state parameters balanced by some additional grain-level length-scale. In regions of moderate flow rate, these spatially second order effects should be dominated by the local

rheology. This is clear from the above tests, which show the model does indeed perform better where the flow is faster. But equally clear is the fact that these terms cannot be ignored when near static yield. Section 3.4.3 lists several possible theories to describe non-local behavior, but how each candidate would fit theoretically within the current model remains to be studied.

By skipping over quasi-static non-locality, the model is tacitly assuming that the material is capable of forming clean solid-like/fluid-like interfaces. This is almost always an idealization that is unrealized, as particles in a random packing rarely have the geometric ability to assemble in a fashion that avoids overlap with the predicted interface. A particle on the interface, being unable to both shear with the fluid and remain static with the solid, instead transmits some of the shearing behavior from the fluid-like zone into the solid zone, thereby explaining the gradual tails we see in actual granular flow profiles. Under particular circumstances, however, granular flow can be made to segregate clearly into flowing and completely static zones. Thompson and Grest [135] have shown that a monodisperse 2D disk assembly undergoing horizontal planar shear with downward gravity does indeed have a flow cutoff with zero flow occurring beneath a shear band at the top. This behavior is a rarity brought on by the fact that solid-like material in this geometry can and does form a hexagonal crystal and the horizontal fluid/solid interface happens to align perfectly with a crystal plane.

3.9.2 Dilation

The present model avoids all plastic dilation. Though an argument for why this is acceptable for our current purposes is presented in section 3.4.4, certain benefits would come with properly accounting for the small amount of dilation that occurs in dense flow.

To be clear, there are really two sources of dilation that can occur: plastic and gas-like. Plastic dilation reflects the opening up of space that occurs whenever material is sheared above yield. But even below yield, in particularly energetic surroundings, an element of grains can become gas-like inducing a packing fraction well below random close packing— for example, shaking a box of rubber balls produces dilute material

within, though the box is not deforming plastically. Granted, gas-like dilation is outside the dense flow regime we currently study. But even under circumstances of predominantly dense flow, it is possible that a small region becomes gas-like and consequently one should have some way of dealing with this type of behavior. Recall this issue was crucial in the silo geometry, which is almost entirely dense except within the small free-fall arch that encompasses the opening. Gas-like effects would likely require a granular temperature and heat flux to express how energetic the granular gas is and consequently how much pressure it supports.

Several mechanisms for plastic dilation have been proposed (see section 3.4.4), but a direct 3D discrete element study would be ideal for quantifying the precise dilational dependences. One might hypothesize, based on equation 3.9, that the plastic dilation has a form

$$\frac{d\eta}{d\gamma^p} = \frac{\dot{\eta}}{\tilde{d}^p(\mathbf{M})} = A(\eta)B(\Phi(I) - 0.63e^{-\eta})$$

where $\eta = \log(\det\mathbf{F}^p)$ measures plastic dilation, γ^p is a plastic shear strain, $I = \tilde{d}^p(\mathbf{M})/\frac{1}{d}\sqrt{\frac{P}{\rho_s}}$ is the inertial number, and the functions A and B are empirical, with $B = 1$ when the magnitude of its argument is large and $B \to 0$ as the input goes to zero. The above states that a material originally at random close packing ($\Phi = 0.63$) obeys Bagnold type dependences at steady flow. While the flow is unsteady, it dilates according to A and then ceases dilation as the Bagnold relationship is approached. A relation of this form appears to agree with results of Rycroft *et al.* [122], and collaborative efforts with Rycroft are underway to quantify this evolution law.

Once a form for the plastic dilation has been tested, it could be included in the material model. It would serve to enhance the computation of elasto-statics, where the moduli are known to vary with the packing fraction, as well as give meaningful packing fraction data throughout. However, the technique of artificial density reduction may no longer be valid, since the steady packing fraction, while only a few percent different than at the start, may take more time to develop than is typically allotted in a simulation run.

3.9.3 Flow condition

While the plastic flow rule being utilized here upholds codirectionality, a direct test of the flow condition still needs to be performed to verify whether this is the best candidate. Such a test may be forthcoming with recent DEM data on large conical hopper flow. Some models have had success utilizing double-shearing conditions [3], and Rycroft's silo flow data has shown only marginally codirectional flow. For the purposes of simplicity though, codirectionality has fit the current needs and appears to be sufficient to predict basic flow behavior. However, a fine-tuned flow condition would be essential to model highly asymmetric 3D flows.

Chapter 4

General conclusion

This thesis has presented two very distinct models for granular material flow. The SFR, on the one hand, is able to capture some of the more complicated slow flow phenomena of granular matter. On the other hand, moderate flow rheology and an improved statics law are clearly represented in the nonlinear elasto-plastic model. What remains is to configure a complete model that covers all three categorical regimes of dense granular materials: statics, quasi-static flow, and moderate flow. If the SFR and elasto-plastic model could somehow be combined, this goal would be achieved. While the conclusion sections of the past two chapters deal directly with their respective flow models, we take some time now to comment directly on the prospect of a joint SFR/elasto-plastic formulation.

Seeing as the limit-state hypothesis may not be reliable, it would probably be more realistic to try to integrate SFR principles into the elasto-plastic model rather than vice versa. Observing the elasto-plastic flow solutions of chapter 3, it is indeed quite incidental that where the SFR would most be needed, is precisely where it was derived to work best— at incipient failure. That is, elasto-plastic predictions are the most incorrect in regions immediately surrounding the fluid/solid interface, where material is close to the static yield criterion.

One might hypothesize that the predicted elasto-plastic fluid/solid interface is where spots are *generated*. The rate of spot generation could grow proportional to the strain-rate gradient at the predicted interface, so that spots are born, quite

literally, to fight rapid spatial changes in the flow field. In this sense, spots would modify the elasto-plastic solution in a fashion that ensures the flow does not vary on a scale smaller than a spot width.

After being birthed at the elasto-plastic interface, spots could then wander through the material, so that the final velocity field is the sum of the elasto-plastic solution and the additional spot-related motion. Allowing spots to travel through regions deemed 'static' by the elasto-plastic model is precisely what is desired, since in truth, these region should be undergoing intermittent failure events. The idea of spots moving through solid-like material also resonates well with how the SFR was derived— spot drift is driven by a static-to-kinetic drop in μ at the spot boundaries, which is sensible for a spot moving through solid-like material. The drift formula for spots could remain almost entirely the same, except modified for non-incipient stresses. Since sub-yield material would not have slip-lines per se, the planes of weakest orientation could be substituted, which are the internal planes along which the shear to normal stress ratio is highest. It would make sense if, like in the SFR proper, spots do not have any effect on the stress field that they move through, as they represent rate-independent deformation.

By integrating the SFR into the elasto-plastic model as outlined above, certain other benefits would follow. Originally, spots were considered free-volume carriers. But under this notion, the Spot Model (and the SFR) would predict the packing fraction to be lowest down the center of a draining silo, where the spot concentration is highest. This is in contrast with discrete simulation results of [122] that show the packing fraction is indeed lower off-center, where more rapid shearing occurs. When this was observed, it was cause to remove a direct association of free-volume with spots. However, by instituting a spot formulation within an elasto-plastic model as outlined above, the spot concentration field is likely to align more closely with the free volume profile. Thus, the physical interpretation of spots as free volume carriers might make more sense under this joint approach.

Another benefit would be that spot motion could become universally frame-indifferent. In the SFR formulation of chapter 2, the issue of frame-indifference arises

because there is no way for a spot to know the motion of the "background". For example, in a silo flow, spots always move up. But the definition of "up" is frame dependent— a silo experiment performed in a moving train would have upward spot drift if viewed by an observer in the train, but should have slanted spot drift if viewed by an outside observer. If the spot drift is determined solely by the stress tensor and the body forces, this type of adjustment is not possible. The SFR dodges this issue by constricting to the case of steady flows, where the frame of reference can be adjusted post facto. However, a better fix for this problem may arise if the SFR were integrated into the elasto-plastic model. Each static region in an arbitrary elasto-plastic flow has a unique velocity and rigid-body spin, which can be extracted from the elasto-plastic flow solution. A spot traveling through a static zone could thus be directed to drift against this background motion. In other words, the spot motion occurring within a sub-yield region could be defined uniquely in the frame of reference for which the rigid-body elasto-plastic background motion appears stationary.

These ideas are still preliminary. There are a few non-ignorable problems with splicing these models. How would boundary conditions for the spot concentration be determined? In chapter 2, kinematic boundary conditions for the flow are translated directly into boundary conditions for the spot flux. But in a combined approach, where the velocity is the sum of elasto-plastic and spot-based velocity fields, the answer is not obvious. Perhaps zero spot flux at boundaries would be a plausible guess, so that the elasto-plastic solution carries the burden of fitting all the boundary conditions. In order for the flow to remain smooth across the elasto-plastic interface, spots may have to be allowed to wander into the above yield material. The details of how this may take place are non-trivial, since we would not want spot motion to interfere too much with the rapidly flowing zones, as these are well-represented by the Jop–Pouliquen flow law. It is also not clear if SFR-type spot drift applied within an elasto-plastic stress field will even give remotely correct flow behavior. While the SFR flow predictions of chapter 2 may look valid, changing the stress field and origination of spots as is being suggested here may have an unduly large effect on the outcome. Of course, a 3D version of the SFR would have to be derived if 3D flow predictions

were desired, though it would be understandable to test plane-strain flows at first. Perhaps one of the ways discussed in the conclusion of chapter 2 could be used to extend the SFR into 3D at some later point.

But beyond any of this speculation, the two models presented in this thesis already appear to have a good degree of effectiveness and generality on their own. The kind of versatility each has been shown to have herein is noteworthy especially in a field like dense granular materials, where predictive flow models with general applicability are quite rare. In spite of this, there is no shortage of observations and data in the granular flow literature. It has been my goal from the very beginning to move beyond the role of observer, and attempt to synthesize known results into general models. While this has been no simple task by any stretch of the imagination, the process has been exciting and worth the effort.

Appendix A

Review and comparison of other granular plasticity models

A.1 Critical state soil mechanics

A common precept in plasticity is the notion of *normality* or *associatedness*. Flows that obey normality have a flow rule defined in terms of the yield function Y as follows:

$$\mathbf{D} = \lambda \frac{\partial Y}{\partial \mathbf{T}} \tag{A.1}$$

where λ is a positive multiplier (we presume $\mathbf{D} \approx \mathbf{D}^p$ for now). For a 3D flow, this means that if the yield function were plotted in 6-space as a function of all 6 independent entries in the 3D stress tensor, the strain-rate matrix would be a 'vector' pointing normal to the yield surface oriented toward greater values of Y.

One of the first gripes about the use of friction-based yield criteria in describing granular materials is that the principle of normality gives a flow rule that predicts unstoppable dilatancy. Consider a rough extension of the Coulomb yield criterion into 3D, $Y = \mu(\mathrm{tr}\mathbf{T})/3 + |\mathbf{T_0}|/\sqrt{2}$, which displays the basic property that yield occurs when a certain multiple of the pressure equals the shear stress. Its associated flow

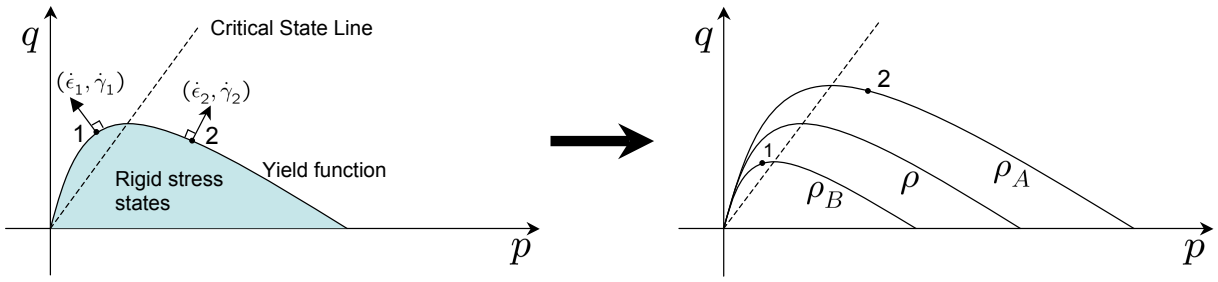


Figure A-1: (left) Critical State Theory's spade-shaped yield function for material at some density ρ_s; (right) Deformations with non-zero volumetric strain will cause the material to settle down on a new yield function.

rule is

$$\mathbf{D} = \lambda \left(\frac{\mathbf{T_0}}{\sqrt{2}|\mathbf{T_0}|} + \frac{\mu}{3}\mathbf{I} \right).$$

The trace of this strain-rate tensor is $\lambda\mu$, implying that material undergoing plastic flow will never stop dilating.

Roscoe and co-workers [117] present a different viewpoint on the issue. In what became known as *Critical State Soil Mechanics*, explained in detail in [127], they argue that normality does hold, but that in fact the Coulomb yield criterion is not technically the correct yield function.

Backed by results from triaxial stress experiments on soil samples, Critical State Theory claims that soils have a yield function that depends on the soil consolidation as measured by the local density ρ. The yield curve for material at a particular density is defined in terms of two stress tensor invariants: the pressure $p = -\frac{1}{3}\text{tr}\mathbf{T}$ and the equivalent shear stress $q = |\mathbf{T_0}|/\sqrt{2}$. Plotted in these variables, the principle of normality is equivalent to the statement that the strain-rate vector $(\dot{\epsilon}, \dot{\gamma})$ is normal to the yield curve and pointing outward, where $\dot{\epsilon} = -\text{tr}\mathbf{D}$ is a volumetric strain-rate which determines changes in density, and $\dot{\gamma} = \frac{\sqrt{2}}{3}|\mathbf{D}_0|$ is a shear strain-rate proportional to the total shear deformation (volume-conserving part of the deformation). Figure A-1 displays the theory's picture of the yield function and how it changes after material deformation. Any stress state underneath the yield curve corrresponds to rigid material. Under normality, material at point 1 in the initial state will undergo a deformation according to the vector $(\dot{\epsilon}_1, \dot{\gamma}_1)$. Since $\dot{\epsilon}_1$ is negative, the material will

dilate and settle down at point 1 on the right on a new yield curve corresponding to $\rho_{_A} < \rho$. The material at stress state 2 will likewise undergo compaction and arrive at point 2 on the yield curve corresponding to $\rho_{_B} > \rho$.

The critical state line is defined as the locus of points for which normality predicts no volumetric changes during deformation— note that wherever a yield curve intersects the critical state line, the curve becomes parallel to the p axis and thus the corresponding strain-rate vector has no volumetric component. The theory reasons that the critical state line is indeed a straight line of the form

$$q = Mp.$$

As flow developes, the stress states throughout the material will continually move toward the critical state line and once local volume changes finally stop, all flowing material stress states should lie on the critical state line. Thus in a steady flow, the critical state line might falsely appear to *be* the yield function when in fact it is only a locus of states from a family of yield functions. So, it is argued then that the reason normality previously failed to describe granular materials was because it was applied mistakenly to the critical state line and not to the true family of yield functions.

A.2 The Anand-Gu model

We now cover a model presented by Anand and Gu [3] which simultaneously accounts for a variety of granular characteristics. To summarize the equations, we begin by defining an additive decomposition of the velocity gradient into elastic and plastic parts

$$\nabla \mathbf{v} = \mathbf{L} = \mathbf{L}^e + \mathbf{L}^p$$

and define corresponding deformation-rate and spin tensors

$$\mathbf{D}^{e,p} = \mathrm{sym}(\mathbf{L}^{e,p}) \quad , \quad \mathbf{W}^{e,p} = \mathrm{skw}(\mathbf{L}^{e,p}).$$

The material stresses are dependent on the elastic strain and thus we write a hypo-elastic stress-rate/strain-rate law utilizing the Jaumann-Zaremba objective stress-rate:

$$\mathbf{T}^{\nabla e} = \dot{\mathbf{T}} - \mathbf{W}^e\mathbf{T} + \mathbf{T}\mathbf{W}^e = \mathscr{C}[\mathbf{D}^e].$$

We will delve more into the meaning and justification of this equation in a later section. The term $\mathscr{C}$ is a fourth order tensor of elasticity moduli which may depend on the stresses and the relative density $\eta = \rho/\rho_s$ where ρ_s is the mass density of solid grains.

The material is assumed to obey a Mohr-Coulomb yield function of the general form

$$Y = \tau - \mu\sigma - c \leq 0$$

where $\sigma = -\hat{n} \cdot \mathbf{T}\hat{n}$ is the compressive stress on a plane with outward normal $\hat{n}$, and for some direction $\hat{n}^\perp$ orthogonal to $\hat{n}$, $\tau = |\hat{n}^\perp \cdot \mathbf{T}\hat{n}|$ is the magnitude of the shear stress in the direction $\hat{n}^\perp$. The parameter c represents cohesion and μ the friction coefficient.

A.2.1 Double-Shearing Flow Rule

The flow rule employed in this model does not have the simple form of a Rudnicki-Rice-type model [119] wherein the deviatoric plastic deformation rate tensor is a scalar multiple of the deviatoric stress tensor

$$\mathbf{D}_0^p = |\mathbf{D_0^p}|\frac{\mathbf{T_0}}{|\mathbf{T_0}|} \tag{A.2}$$

and

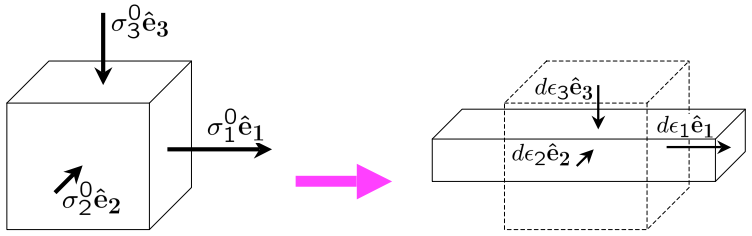
$$\mathrm{tr}\mathbf{D^p} = \sqrt{2}\beta|\mathbf{D_0^p}| \tag{A.3}$$

for β a dilatancy parameter. These equations state that the eignvectors of $\mathbf{D^p}$ are aligned with those of $\mathbf{T}$ (i.e. coaxiality), but furthermore they state that when yielding occurs, the relative sizes of the deviatoric principal plastic strain increments $d\epsilon_i^0$ are a

simple multiple of the deviatoric principal stresses, i.e. in the principal stress frame:

$$\begin{pmatrix} d\epsilon_1^0 & 0 & 0 \\ 0 & d\epsilon_2^0 & 0 \\ 0 & 0 & d\epsilon_3^0 \end{pmatrix} = \lambda \begin{pmatrix} \sigma_1^0 & 0 & 0 \\ 0 & \sigma_2^0 & 0 \\ 0 & 0 & \sigma_3^0 \end{pmatrix}. \tag{A.4}$$

Or, if we temporarily set dilation to 0, this flow rule (now a Mises flow rule) can be understood pictorally with ease:



This is a simple rule for a general deformation process, but we prefer a rule which more closely upholds what we deem the physical basis of granular deformation: internal slip. A Mohr's Circle diagram for a material element at yield (Figure A-2) is instructive in this cause. Each circle indicates the locus of traction stress states (σ, τ) that are possible along some line in the specified plane.

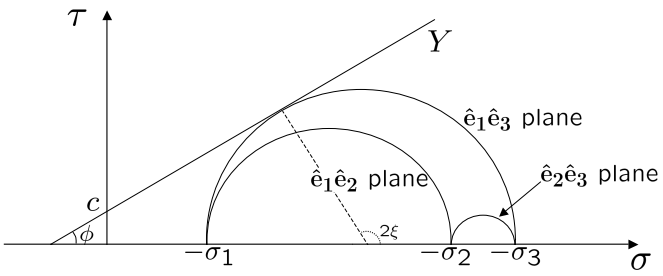


Figure A-2: Mohr's Circle diagram for the three principal planes in the case $\sigma_1 > \sigma_2 > \sigma_3$.

Consider the general case of $\sigma_1 > \sigma_2 > \sigma_3$. Each Mohr's Circle diameter must be equal to the difference of that plane's principal stresses and each circle must intersect the σ axis at the principal stresses. Thus the $\hat{\mathbf{e}}_1\hat{\mathbf{e}}_3$ plane will have the largest Mohr's Circle and the other two circles must be fully enveloped by this one. Since no traction

state can ever lie above the yield function, only the $\hat{\mathbf{e}}_1\hat{\mathbf{e}}_3$ principal plane can ever be tangent to the yield function, and likewise slip-lines will only develope within the $\hat{\mathbf{e}}_1\hat{\mathbf{e}}_3$ plane. To determine how the slip-lines are oriented within this plane, we observe the angle of the Mohr's Circle arc measured anti-clockwise from the σ axis to the radius ending at the yeilding traction state. By trigonometry this angle is $\pi/2 + \phi = 2\xi$ and thus by Mohr's Circle conventions, slip-lines should appear within the material in the $\hat{\mathbf{e}}_1\hat{\mathbf{e}}_3$ plane at angles of $\pm\xi$ off the $\hat{\mathbf{e}}_1$ direction.

To understand slipping in 3D, suppose we travel down the $\hat{\mathbf{e}}_2$ direction continuously drawing slip-line pairs in the $\hat{\mathbf{e}}_1\hat{\mathbf{e}}_3$ plane. This family of slip-lines will then form two slip-systems which locally appear as a pair of intersecting slip-planes. Excluding dilation, the Double-Shearing flow rule claims that deformations can only be made by superposing two simultaneous shearing motions along these slip-systems. The direction of each shear, as necessitated by the dissipation inequality, is aligned with the shear stress. To add in dilational effects we claim that as material shears an amount $d\gamma$ along a slip-system, the material dilates in the direction normal to the slip surface an amount $\beta d\gamma$. Looking back at the Mohr's Circle diagram, we see that there is only one point of tangency with the yield function implying identical shear and compressive stresses on both slip-systems. Likewise there is no physical reason why the shearing rates along the two slip-systems should differ. Thus we adopt the hypothesis that the shearing rates along both slip-systems are equal and consequently the dilation rates must also be equal. Since all plastic motion is now symmetric about both the major and minor principal axes, we see that there can be no plastic spin: $\mathbf{W}^p = \mathbf{0}$.

Pictorally, we can visualize Double-Shearing (in this general case of all-differing principal stresses) as in Figure A-3. The stress state determines the slip-systems upon which we apply equal shearing and dilation.

One important result is that no deformation occurs in the $\hat{\mathbf{e}}_2$ direction. This is markedly different than a Mises flow rule and places a significant restriction on the space of allowable deformations. Our interpretation of Double-Shearing is actually different than A.J.M. Spencer's original 1964 development of the Double-Shearing

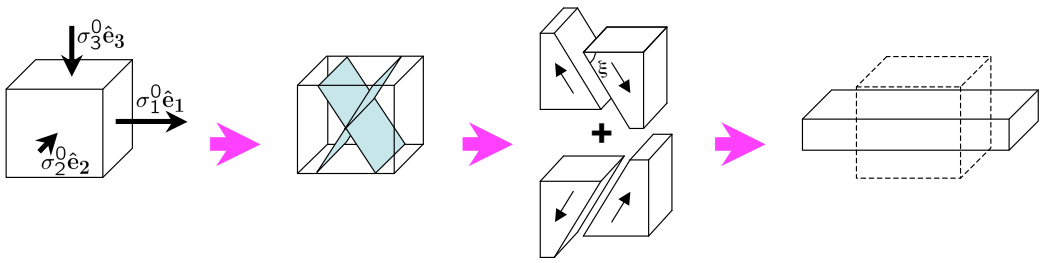


Figure A-3: Cartoon of Double-Shearing in the general case.

flow rule [133]. Spencer applied the rule to 2D geometries in a limit-state stress field and argued that the shearing rates on the two slip-lines need not be the same and in fact should be determined from the condition that the flow remain isochoric. Contrary to Spencer's limit-state assumptions, the model at hand allows an elastic component in the deformation and thus we may simply claim that all necessary spins come from $\mathbf{W}^e$ and that the plastic shear-rate on both slip-systems is equal.

To determine the mathematical form of the flow rule, since the deformation is planar, it is easiest to write the deformation as a superposition of two 2×2 tensors, each tensor representing a shearing/dilation on one of the slip-systems. With the aid of Figure A-3 (for signs) we write a tensor for one slip-system by rotating the coordinates an angle ξ or $-\xi$ off the $\hat{\mathbf{e}}_\mathbf{1}, \hat{\mathbf{e}}_\mathbf{3}$ basis and then writing the deformation as a simple two element tensor for horizontal shear plus vertical dilation. This gives

$$\mathbf{D}^p = \frac{1}{2}\left[\mathbf{R}\begin{pmatrix} 0 & \dot{\gamma} \\ 0 & \beta\dot{\gamma} \end{pmatrix}\mathbf{R}^T + \mathbf{R}^T\begin{pmatrix} 0 & -\dot{\gamma} \\ 0 & \beta\dot{\gamma} \end{pmatrix}\mathbf{R}\right]$$

where

$$\mathbf{R} = \begin{pmatrix} \cos\xi & -\sin\xi \\ \sin\xi & \cos\xi \end{pmatrix}.$$

This simplifies to

$$\mathbf{D}^p = \dot{\gamma}[\mathbf{P} + \beta\mathbf{N}]$$

for

$$\mathbf{P} = \frac{1}{2}\sin(2\xi)[\hat{\mathbf{e}}_\mathbf{1} \otimes \hat{\mathbf{e}}_\mathbf{1} - \hat{\mathbf{e}}_\mathbf{3} \otimes \hat{\mathbf{e}}_\mathbf{3}]$$

and

$$\mathbf{N} = \sin^2 \xi \hat{\mathbf{e}}_\mathbf{1} \otimes \hat{\mathbf{e}}_\mathbf{1} + \cos^2 \xi \hat{\mathbf{e}}_\mathbf{3} \otimes \hat{\mathbf{e}}_\mathbf{3}.$$

We redefine $\dot{\gamma}$ as the magnitude of the plastic deformation enabling us to write

$$\mathbf{D}^p = \dot{\gamma} \frac{\mathbf{P} + \beta \mathbf{N}}{\|\mathbf{P} + \beta \mathbf{N}\|}. \tag{A.5}$$

In the degenerate case where $\sigma_1 = \sigma_2 > \sigma_3$, $-\sigma_2$ approaches $-\sigma_1$ in the Mohr's Circle diagram and thus when yielding occurs, both the $\hat{\mathbf{e}}_\mathbf{1}\hat{\mathbf{e}}_\mathbf{3}$ and the $\hat{\mathbf{e}}_\mathbf{2}\hat{\mathbf{e}}_\mathbf{3}$ Mohr's Circles will be tangent to the yield function causing slip-lines to form in both planes. Furthermore, since the stress tensor has repeated eigenvalue $\sigma_1 = \sigma_2$, the $\hat{\mathbf{e}}_\mathbf{1}$ and $\hat{\mathbf{e}}_\mathbf{2}$ directions are not unique and any vector $\hat{\mathbf{e}}_\mathbf{3}^\perp$ orthogonal to $\hat{\mathbf{e}}_\mathbf{3}$ is an eigenvector corresponding to $\sigma_1 = \sigma_2$. Be that as it may, Mohr's Circle indicates that all planes $\hat{\mathbf{e}}_\mathbf{3}^\perp\hat{\mathbf{e}}_\mathbf{3}$ contain slip-lines and thus the slip-system is a cone symmetric about the $\hat{\mathbf{e}}_\mathbf{3}$ axis.

To shear and dilate equally on this slip-system is less easy to intuitionalize, but if we imagine the cone as being composed of many diagonal bars oriented symmetrically about a vertical axis, we can see the deformation as the superposition of shears along each bar, each shearing motion taking place in a vertical plane. To compare this pictorally to the previous case observe Figure A-4. Ultimately, since the shears and

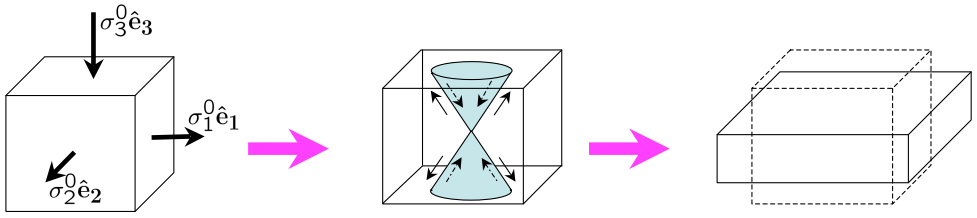


Figure A-4: Cartoon of Double-Shearing in the $\sigma_1 = \sigma_2$ degenerate case.

dilations are all radially symmetric and equal, the superposition will give a triaxial deformation contracting in the $\hat{\mathbf{e}}_\mathbf{3}$ direction and expanding radially in the orthogonal directions.

Mathematically, this case takes the same form as the previous, but with

$$\mathbf{P} = \frac{1}{2}[\hat{\mathbf{e}}_\mathbf{1} \otimes \hat{\mathbf{e}}_\mathbf{1} - (1/2)(\mathbf{I} - \hat{\mathbf{e}}_\mathbf{1} \otimes \hat{\mathbf{e}}_\mathbf{1})]$$

and

$$\mathbf{N} = \frac{1}{2}[\sin^2\xi\hat{\mathbf{e}}_\mathbf{1} \otimes \hat{\mathbf{e}}_\mathbf{1} + (1/2)\cos^2\xi(\mathbf{I} - \hat{\mathbf{e}}_\mathbf{1} \otimes \hat{\mathbf{e}}_\mathbf{1})].$$

The other degenerate case, $\sigma_1 > \sigma_2 = \sigma_3$, is entirely analagous. The motion is a triaxial elongation about the $\hat{\mathbf{e}}_\mathbf{1}$ axis. All that must be done to attain the mathematical form for this case is to replace all "3" subscripts in the previous case with "1" and replace ξ with $\xi + \pi/2$.

An important point about Double-Shearing is that only two classes of deformation are ever allowed: planar and triaxial. This reduction in the space of possible deformations tends to show itself in the form of shear-bands. For example in plane-strain setups, if material ever approaches a degenerate stress state, it cannot undergo plastic deformation— triaxial deformation is prohibited in plane-strain so $\dot{\gamma}$ must go to 0. Thus, non-degenerate zones of planar deforming material are seen to have sharp boundaries where the stress state crosses into the realm of degeneracy.

A.2.2 Hardening/Softening and Dilation

We need a few more constitutive laws to describe the hardening/softening and dilational behaviors of granular matter. According to Critical State Theory, material whoses friction coefficient is above a certain constant volume value μ_{cv} should experience dilation and below that value compaction. We may apply this notion to our problem most simply by writing

$$\beta = h_\beta(\mu - \mu_{cv}). \tag{A.6}$$

Thus as we expect, when $\mu = \mu_{cv}$ the dilation is 0 and the material volume stops changing.

The value of μ should always approach μ_{cv} as the plastic shear becomes large, but

the general shape of the μ vs. γ curve can vary depending on the starting parameters. Specifically, we wish to include the following phenomenological behavior: for some critical relative density η_{cr} and initial friction $\mu(0) = 0$, over-consolidated materials ($\eta_0 > \eta_{cr}$) hit a peak value for μ greater than μ_{cv} before asymptotically approaching μ_{cv} whereas under-consolidated materials ($\eta_0 < \eta_{cr}$) monotonically increase to μ_{cv}.

A general form for the hardening/softening that appears to capture these effects is:

$$\begin{aligned}
\dot{\mu} &= h\dot{\gamma} \\
h &= h_\mu \left| 1 - \frac{\mu}{\mu_s} \right|^p \ \text{sign}\left(1 - \frac{\mu}{\mu_s}\right) \\
\mu_s &= \mu_{cv} + b(\eta - \eta_{cr})^q \ \text{step}(\eta - \eta_{cr})
\end{aligned}$$

where step(x) is 1 if x is positive and otherwise 0.

To check this form more directly, we first need an expression for the evolution of relative density which can then be related to β and likewise μ. We employ conservation of mass to achieve this end:

$$\dot{\eta} = -\eta \ \text{tr}\mathbf{D}^p. \tag{A.7}$$

According to Equation A.5,

$$\text{tr}\mathbf{D}^p = \frac{\dot{\gamma}\beta}{g(\beta, \xi)}$$

where $g(\beta, \xi) = \|\mathbf{P} + \beta\mathbf{N}\|$ is easily determined given the stress case. Applying this in the mass conservation equation and writing β and ξ in terms of μ gives:

$$\dot{\eta} = -\eta \ \frac{\dot{\gamma} \ h_\beta(\mu - \mu_{cv})}{g\left(h_\beta(\mu - \mu_{cv}), \ \frac{\pi}{4} + \frac{(\arctan \mu)}{2}\right)}$$

$$\implies \quad \frac{d\eta}{d\gamma} = \frac{\dot{\eta}}{\dot{\gamma}} = -\eta \ \frac{h_\beta(\mu - \mu_{cv})}{g\left(h_\beta(\mu - \mu_{cv}), \ \frac{\pi}{4} + \frac{(\arctan \mu)}{2}\right)} \equiv f_1(\mu, \eta). \tag{A.8}$$

The previous evolution law for μ gives

$$\frac{d\mu}{d\gamma} = \frac{\dot{\mu}}{\dot{\gamma}} = h \equiv f_2(\mu, \eta). \tag{A.9}$$

Equations A.8, A.9 make up a complete dynamical system for μ and η as they vary over the course of plastic shearing γ. For the sake of concreteness let us focus on the non-degenerate stress case for which

$$g(\beta, \xi) = (1/2)\sqrt{1 + 3\beta^2 + (\beta^2 - 1)\cos(4\xi) - 2\beta\sin(4\xi)}.$$

Using the "ode45" function from MATLAB 7.1 we can plot a phase space diagram of the μ, η system (see Figure A-5). The values of the parameters used match those found in Anand and Gu [3]. Of importance to the diagram, $\eta_{cr} = 0.54$ and $\mu_{cv} = 0.613$. Note in the figure that if the material is over-consolidated and $\mu_0 < \mu_{cv}$ the friction first increases above μ_{cv} and then asymptotes to μ_{cv} as we intended. In the under-consolidated case the friction increases steadily to μ_{cv}. The diagram also indicates that all over-consolidated initial material states will eventually have $\eta \to \eta_{cr}$ whereas under-consolidated material states can be found that approach any desired value of $\eta < \eta_{cr}$. This result can also be attained analytically by solving $f_1 = f_2 = 0$. Such a qualitative difference between the long-term density behaviors of initially over- and under-consolidated materials occurs because the step function used in the evolution of μ places a discontinuity in h that separates the dynamics of the two regimes.

We can even express as an integral the final relative density η_f for an under-consolidated material element initially at η_0. First note that in Equation A.8, we can write $f_1(\mu, \eta) = \eta\ \Phi(\mu)$. Thus:

$$\frac{d\eta}{d\gamma} = \eta\Phi(\mu) \quad \Longrightarrow \frac{d(\ln\eta)}{d\gamma} = \Phi(\mu),$$

$$\Longrightarrow \ln\left(\frac{\eta_f}{\eta_0}\right) = \int \Phi d\gamma = \int \Phi \frac{d\gamma}{d\mu} d\mu = \int_{\mu_0}^{\mu_{cv}} \frac{\Phi}{f_2} d\mu$$

where the final integral is well-posed since the integrand is a function of μ only (f_2

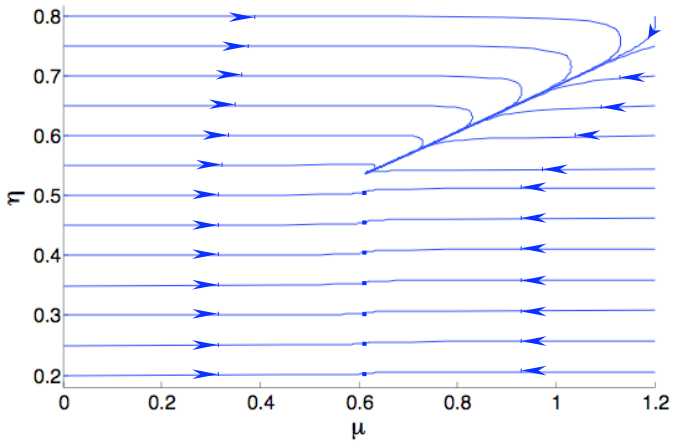


Figure A-5: Phase-space diagram for μ and η. Arrows point in the direction of increasing γ.

has no η dependence in the under-consolidated case).

A.3 Comparison to Critical State Theory

It is interesting to compare this η, μ interdependence with that of a Critical State model. The simplest one, the so-called "Granta-Gravel" model, uses a yield condition of the following form:

$$Y = q + Mp\left(\ln p - \frac{\Gamma - 1/\eta}{\lambda} - 1\right) \leq 0$$

where Γ, λ, and M are material constants, η is the relative density, and $p = -\frac{1}{3}\text{tr}\mathbf{T}$ and $q = |\mathbf{T_0}|/\sqrt{2}$ represent the stress state. The principal of normality is employed for the flow rule, which simplifies in terms of p and q to

$$M\left(\ln p - \frac{\Gamma - 1/\eta}{\lambda}\right) = \frac{-\text{tr}\mathbf{D}^p}{\frac{\sqrt{2}}{3}|\mathbf{D}^p|} = \frac{\dot{\eta}/\eta}{\frac{\sqrt{2}}{3}\dot{\gamma}} = \frac{3}{\eta\sqrt{2}}\frac{d\eta}{d\gamma}.$$

The γ used here is the true plastic shear and so it is slightly different than the γ used previously (Equation A.5) which has a volumetric contribution. Since we will only be using the above equation for large-strain behavior, this difference is to no avail. A major dissimilairty between this equation for $d\eta/d\gamma$ and the equivalent for

the Anand-Gu model (Equation A.8), is that the evolution of η with strain depends explicity on the pressure p which can be varied at will over the course of the deformation via loading boundary conditions. If we enforce a fixed pressure $p = p_0$ throughout the deformation, we may write,

$$\frac{d\eta}{d\gamma} = -A\eta + B$$

for positive constants $A = -\frac{\sqrt{2}M}{3}(\ln p_0 - \Gamma/\lambda)$ and $B = \frac{M\sqrt{2}}{3\lambda}$. This yields,

$$\eta = \frac{B - C_0 \exp(-A\gamma)}{A} \quad \text{for } C_0 \text{ such that} \quad \eta_0 = \frac{B - C_0}{A}.$$

We see that in the Granta-Gravel model with $p = p_0$ fixed, η always approaches $\eta_f = B/A = (\Gamma - \lambda \ln p_0)^{-1}$ regardless of the inital relative density η_0. This is unlike the Anand-Gu model in which η_f depends on η_0 and does not depend on the stress path. So while some ideas are the same among the two models (e.g. the general dependence of dilation on μ, the characteristic shapes of the various μ vs. γ curves) there are also some important qualitative differences.

Appendix B

SFR in a hopper

This appendix shall present flow results from solving the SFR in the wedge hopper geometry. The stress field in the hopper is non-trivial. First we solve for flow using Jenike's Radial Stress Field, which provides smooth stresses, but does not utilize a pressure-free top boundary. Pierre Gremaud (University of North Carolina) used the Discontinuous Galerkin method to solve for the incipient-failure stress field under different top traction boundary conditions. The SFR is then solved again using his discontinuous stress data.

The Radial Stress Field was solved in a hopper with apex half-angle 30°, internal friction angle 30°, wall friction 26°, and passive wall stresses. Applying coaxiality gives the corresponding "Radial Velocity Field", which necessarily has no θ component. The SFR was solved using $L_s = 0.1$ units and zero θ velocity along all boundaries. The results are pictured in figure B-1. The SFR flow field does have a small θ component. Also the SFR solution predicts a narrower zone of flow down the center of the hopper.

Pierre Gremaud's p and ψ fields are plotted in figure B-2. The angle of wall friction is assumed to be 15° for $r > 1$ and 10° for $r < 1$. The stress information is assumed to travel down and thus the Radial Stress Field corresponding to the higher angle of wall friction is taken as the "initial condition" at $r = 1$. The region from $r = 0.3686$ to $r = 1$ is pictured. As before $\phi = 30°$ and passive wall stresses are used. The SFR is then solved using zero (unconvolved) θ velocity at all boundaries and

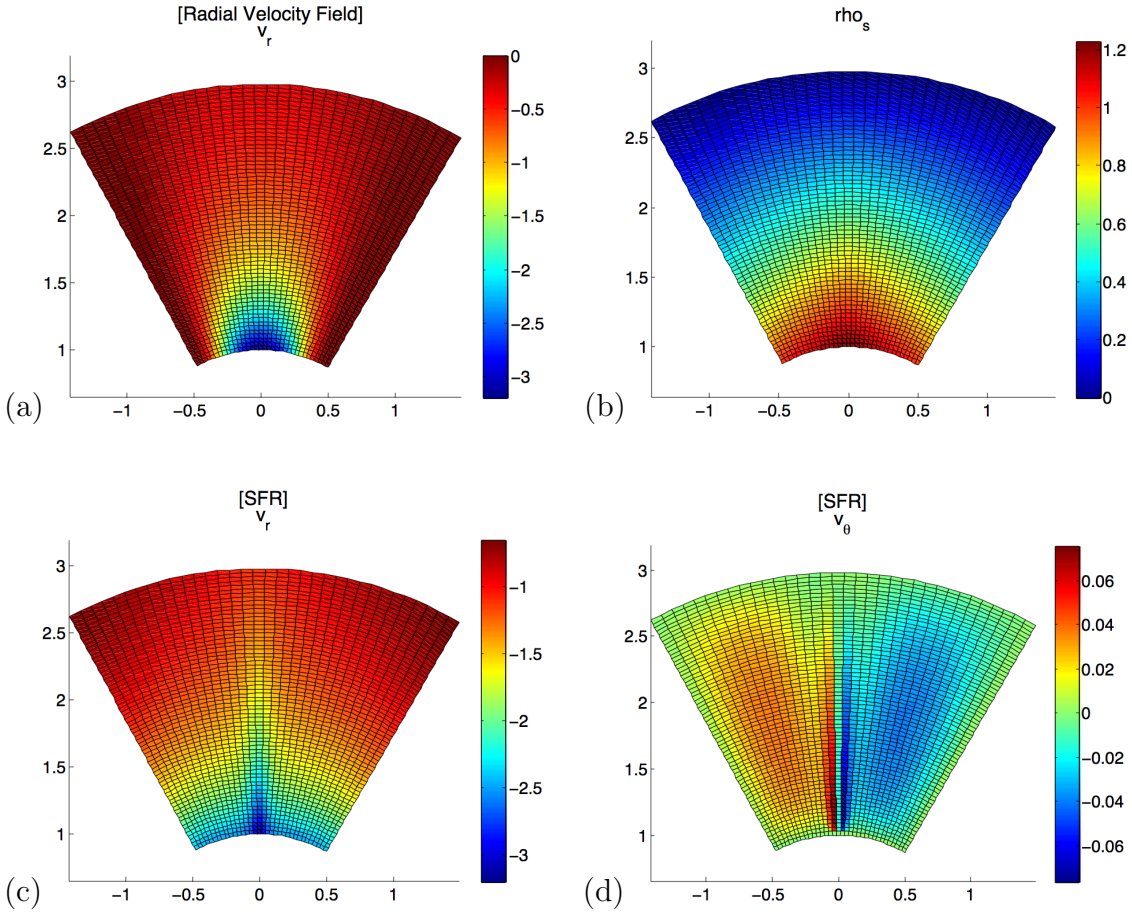


Figure B-1: Flow prediction using coaxiality upon the Radial Stress Field (a). Using the same stresses, the SFR is applied yielding results (b), (c), (d).

$L = 0.08$. The resulting velocity field, thanks to convolution with the spot influence, is necessarly smooth and fluctuates on a scale no smaller than L_s. From the velocity components, it is clear that the trajectories are not straight, nor simple curves, but rather wiggle slightly going from top to bottom of the pictured hopper.

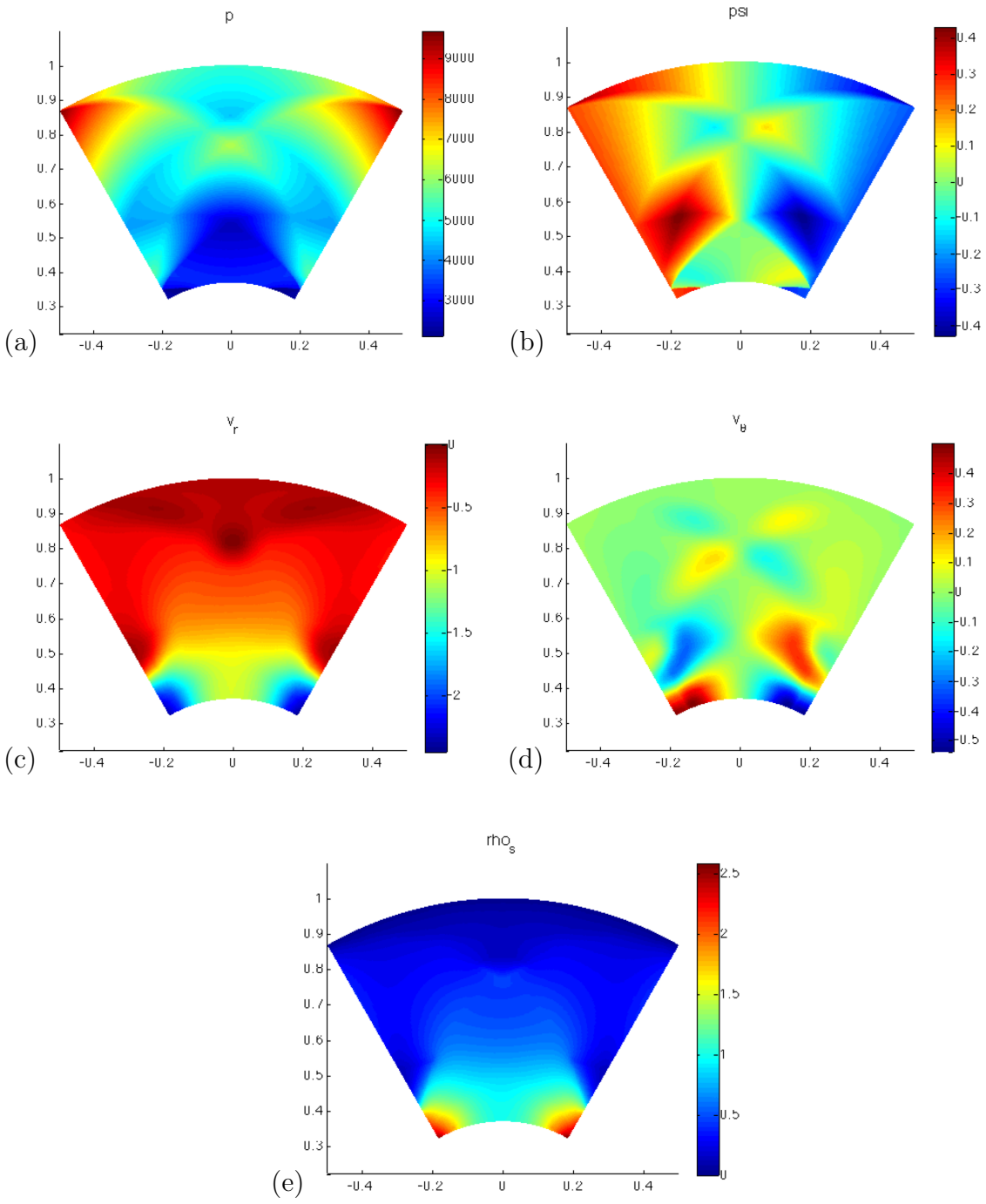


Figure B-2: (a) (b) Represents the discontinuous stress field generated by Pierre Gremaud for a hopper with wall friction jump. The corresponding SFR fields are displayed in (c) (d) and (e). Note the velocity field is smooth.

Appendix C

Deciding the stochastic dilemma

In the original Spot Model derivation [15], the issue of the "phantom 2" was pointed out as a matter of nontrivial mathematical value. In the process of deriving the continuum limit of particle motion from a superposition of spot influences, an extra 2 shows up, which renders the resulting flow field *compressible*. This matter was left unresolved for some time, likely because the institution of the Spot Model as a monte-carlo method for producing real flows did not appear to be crippled by this abstract consequence of stochastic differentials.

In fact the solution to this issue can be understood on both abstract and picturesque levels. In the parlance of stochastic calculus, the lack of incompressibility came from choosing a less optimal interpretation of the *stochastic dilemma*. Let us clarify this point. Suppose the flow region is broken into n distinct volumes. In Bazant's original work, the displacement of a particle during a time-step is determined by all the nearby spots that displace during that increment:

$$\Delta \mathbf{R}_p(\mathbf{r}_p) = -\sum_n \sum_{j=1}^{\Delta N_s^{(n)}} w\left(\mathbf{r}_p, \mathbf{r}_s^{(n)} + \Delta \mathbf{R}_s^{(j)}(\mathbf{r}_s^{(n)})\right) \Delta \mathbf{R}_s^{(j)}(\mathbf{r}_s^n) \tag{C.1}$$

The subscipt "p" indicates particle position and "s" indicates spot position. ΔN_s^n represents the number of spots in the n^{th} volume and w is the spot influence which is non-zero whenever the particle and spot separations are within the spot width.

To interpret the above, the displacement of a particle can be obtained by finding all spots that displace during the time interval into a position that overlaps the particle, and then superposing their reverse displacements weighted by the influence of the displaced spots on the particle.

Note in this interpretation, that the spot influence at the *end* of the spots' displacements is used to produce particle displacement. The resulting stochastic differential equation gives the following mean particle velocity:

$$\mathbf{u}_p(\mathbf{r}_p, t) = -\int dV_s w(\mathbf{r}_p, \mathbf{r}_s)[\rho_s(\mathbf{r}_s, t)\mathbf{u}_s(\mathbf{r}_s, t) - 2\mathbf{D}_s(\mathbf{r}_s, t) \cdot \boldsymbol{\nabla}\rho_s(\mathbf{r}_s, t)] \tag{C.2}$$

where $\mathbf{u}_s$ is the spot drift and $\mathbf{D}_s$ is the spot diffusion tensor. For the sake of simplicity, consider the case where $w = \delta(|\mathbf{r}_s - \mathbf{r}_p|)$. Then the integral above is represented by the term in brackets. Expressing $\mathbf{u}_p$ accordingly and taking the divergence gives the surprising result that volume is not conserved. If it were not for that factor of 2 in front of the spot diffusion tensor, the mean particle velocity field would be incompressible as per the Fokker-Planck equation for the spot distribution.

To "fix" this problem, we observe the following change of stochatsic interpretation. Consider a *Stratonovich* interpretation, in which the beginning and ending of the spot displacement are equally weighted in determining a spot's influence on a particle:

$$\Delta\mathbf{R}_p(\mathbf{r}_p) = -\sum_n \sum_{j=1}^{\Delta N_s^{(n)}} \frac{w(\mathbf{r}_p, \mathbf{r}_s^{(n)}) + w(\mathbf{r}_p, \mathbf{r}_s^{(n)} + \Delta\mathbf{R}_s^{(j)}(\mathbf{r}_s^{(n)}))}{2} \Delta\mathbf{R}_s^{(j)}(\mathbf{r}_s^n) \tag{C.3}$$

The above can be rewritten (supressing arguments for brevity)

$$\Delta\mathbf{R}_p = \frac{1}{2}\left(\underbrace{-\sum_n \sum_{j=1}^{\Delta N_s^{(n)}} w(\mathbf{r}_p, \mathbf{r}_s^{(n)} + \Delta\mathbf{R}_s^{(j,n)})\Delta\mathbf{R}_s^{(j,n)}}_{\text{Bazant's differential}} + \underbrace{-\sum_n \sum_{j=1}^{\Delta N_s^{(n)}} w(\mathbf{r}_p, \mathbf{r}_s^{(n)})\Delta\mathbf{R}_s^{(j,n)}}_{\text{New term}} \right) \tag{C.4}$$

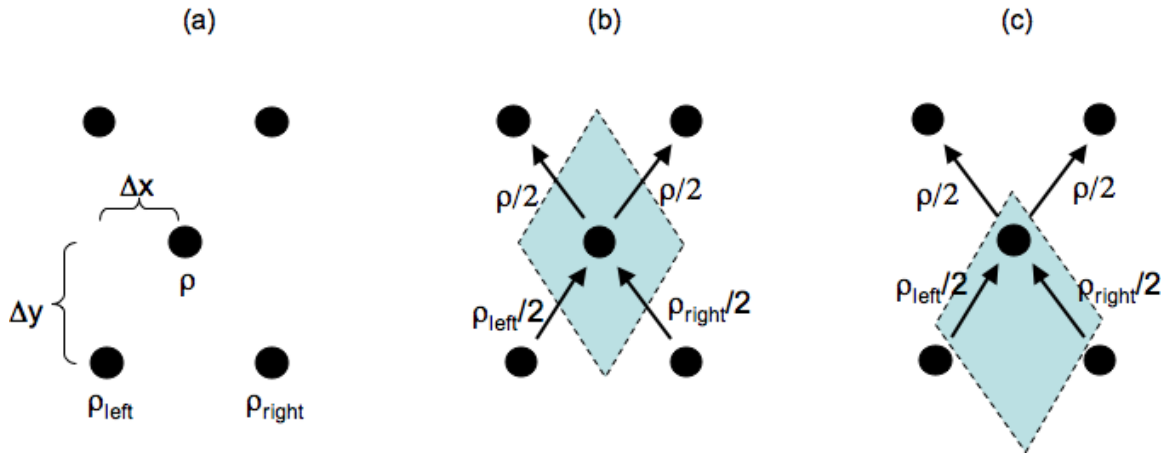


Figure C-1: A simplified example showing the order one effect of changing stochastic interpretation. (a) Spots are constrained to randomly move through a lattice. Steady-state spot concentrations are shown. (b) The time average spot motion through the centered control volume gives an analagous Stratonovich interpretation for motion at the center lattice point. (c) A lowered control volume gives the analagous Bazant interpretation for motion at the center lattice point.

The particle drift under this interpretation is thus one-half the contribution from the Bazant interpretation (that of equation C.2) plus one-half the contribution from the new term. The new term represents a drift of

$$\mathbf{u}_p^{\text{new}}(\mathbf{r}_p, t) = -\int dV_s w(\mathbf{r}_p, \mathbf{r}_s)[\rho_s(\mathbf{r}_s, t)\mathbf{u}_s(\mathbf{r}_s, t)] \tag{C.5}$$

No gradients of ρ appear in the new term because it accounts solely for the beginning of the spots' displacements, whereas Bazant's differential must account for the effect of spots from somewhere else (with possibly different spot concetrations) coming in and ending atop the particle. Putting the two contributions together, gives the new mean particle velocity

$$\mathbf{u}_p(\mathbf{r}_p, t) = -\int dV_s w(\mathbf{r}_p, \mathbf{r}_s)[\rho_s(\mathbf{r}_s, t)\mathbf{u}_s(\mathbf{r}_s, t) - \mathbf{D}_s(\mathbf{r}_s, t) \cdot \boldsymbol{\nabla}\rho_s(\mathbf{r}_s, t)] \tag{C.6}$$

The 2 has vanished and the above form does indeed deliver incompressibility.

Is there a simpler explanation for how this result occurs? Consider an example problem that mirrors the situation in silo flow: spots (of delta-function influence) that

always drift upward an amount Δy each time-step while moving either Δx or $-\Delta x$ horizontally. For further simplicity, suppose the spots are constrained to a lattice with these horizontal and vertical spacings (see figure C-1a). To obtain the particle velocity field from the steady spot distribution, we must closely anayze the happenings near one lattice point during one time interval Δt. Our lattice point of interest is the center point in figure C-1a. For clarity, we assume all spots in the system to begin and end their movements at the beginning and ending of the time step Δt and that during the step, all spots move with the same constant speed $\frac{\sqrt{\Delta x^2+\Delta y^2}}{\Delta t}$.

The question is, what is the average particle velocity at the center lattice point, calculated over the time interval t to $t+\Delta t$, that results when the spots at all lattice points perform a displacement? To answer this one must first *choose* how to orient a control volume around the center lattice point. Figure C-1b displays the Stratonovich interpretation, where the time-average spot motion is computed within a centered, diamond-shaped control volume. During the first half of the time interval, all spots originating at the lattice point of interest are in the process of travelling away but have not yet left the control volume. Half-way through the time interval, those spots have left the volume, but spots from the bottom two lattice points enter. During the latter half of the interval, $\rho_{\text{left}}/2$ spots from the lower-left lattice point and $\rho_{\text{right}}/2$ from the lower-right lattice point are moving through the control volume. We Taylor approximate that

$$\rho_{\text{left,right}} \cong \rho + (\mp\Delta x, -\Delta y)\cdot\nabla\rho$$

and thus the time average velocity of spots moving through the control volume is

$$\begin{aligned}
\mathbf{v}_{\text{spot}} \propto\ & \rho/2(-\Delta x,\Delta y) + \rho/2(\Delta x,\Delta y) \\
& + \rho_{\text{left}}/2(\Delta x,\Delta y) + \rho_{\text{right}}/2(-\Delta x,\Delta y) \\
& = (0,\rho\Delta y) + (\frac{\rho_{\text{left}}-\rho_{\text{right}}}{2}\Delta x, \frac{\rho_{\text{left}}+\rho_{\text{right}}}{2}\Delta y) \\
& \cong (0,\rho\Delta y) + (-\Delta x^2\rho_x, \rho\Delta y - \Delta y^2\rho_y) \\
& = 2(0,\rho\Delta y) - (\Delta x^2\rho_x, \Delta y^2\rho_y)
\end{aligned}$$

$$\propto \mathbf{u}_s\rho - \mathbf{D}_s \cdot \nabla\rho.$$

Analagous to the stochastic calculus result above, no added 2 appears in front of the diffusion term when using this, the Stratonovich interpretation.

On the other hand, suppose as in figure C-1c, that the control volume used to compute averages is not centerd about the lattice point of interest, but rather is dropped lower so that the point is barely within the volume. This picture corresponds to the Bazant interpretation. Time averaging the spot motion during one interval using this control volume will indeed yield that extra 2, since the effect of the spots originating at the center lattice point will be neglected. Under this interpretation, all of the motion attributed to a particle at the center lattice point is due to those spots coming in from below. While neither interpretation is mathematically the more "correct", it does indeed appear that the Stratonovich interpretation is the more valid physical hypothesis. This kind of dilemma, where order one changes to the result can occur by microscopically moving a control volume, is a special byproduct of stochatsic processes. The change in predicted mean behavior induced by switching interpretations (i.e. control volume locations) is deemed "noise-induced".

Appendix D

Eulerian numerical method for fluid/solid interaction

D.1 Past work on fluid/fluid interaction

The work of Kang, Fedkiw, and Liu [75] reveals a method for solving the incompressible Navier-Stokes equations in the presence of a discontinuous viscosity and/or density field. These discontinuities segregate the flow into two distinct, interacting fluids which make contact along some interface. The method is fully Eulerian and tracks the interface not by moving points that reside on it, but using a first order scheme that determines where the interface lies between the grid points at each time step.

As a quick review, the stress tensor $\mathbf{T}$ is symmetric and has a unique additive decomposition into two parts: the deviatoric stress $\boldsymbol{\tau}$ which is trace-free and describes shear stresses, and the pressure tensor $-p\mathbf{1}$. Once a functional form has been selected for $\boldsymbol{\tau}$, a solution is found by solving for linear momentum balance and incompressibility, that is:

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{v} + \frac{1}{\rho}\frac{\partial p}{\partial \mathbf{x}} = \frac{1}{\rho}\left(\frac{\partial}{\partial \mathbf{x}} \cdot \boldsymbol{\tau}\right) + \mathbf{g} \tag{D.1}$$

$$\frac{\partial}{\partial \mathbf{x}} \cdot \mathbf{v} = 0 \tag{D.2}$$

where, $\mathbf{v} = (u, v, w)^T$ is the velocity vector and $\mathbf{x} = (x, y, z)^T$ is the spatial location. To avoid confusion later, we avoid use of ∇ instead specifying derivative operators in general format— for example,

$$\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)_{32} = \frac{\partial w}{\partial y}.$$

We claim as in the Navier-Stokes that the fluid obeys a linearly viscous law, which imposes the following form on $\boldsymbol{\tau}$:

$$\boldsymbol{\tau} = \mu\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)^T\right). \tag{D.3}$$

In the simplest case, the projection method can be invoked to solve these equations. However, the presence of the interface entails certain modifications to the numerics that must take place before taking projections.

The parameters μ and ρ can change values across the interface. We treat the interface as the 0 level set of some function ϕ. The standard advection law

$$\dot{\phi} = \frac{\partial \phi}{\partial t} + \mathbf{v} \cdot \frac{\partial \phi}{\partial \mathbf{x}} = 0 \tag{D.4}$$

is used to track the motion of the interface— that is, the values of the function ϕ are forced to move with the flow so that the level set $\phi(\mathbf{x}) = 0$ always describes the current configuration of the interface. It is important that $\left|\frac{\partial \phi}{\partial \mathbf{x}}\right|$ never becomes too large or else our ability to accurately interpolate ϕ between grid points is jeopardized. As such, we also invoke the reinitialization condition

$$\frac{\partial \phi}{\partial t} + sign(\phi)\left(\left|\frac{\partial \phi}{\partial \mathbf{x}}\right| - 1\right) = 0 \tag{D.5}$$

each step to reshape ϕ without affecting its zero level set, given that ϕ begins as the signed distance function with respect to the interface.

The velocity is continuous across the interface and discretization creates, effectively, a continuous velocity gradient under central differencing. However, the second derivative will be discontinuous and inaccurately represented by the usual discretization thereby necessitating jump conditions. If there were two parallel, adjacent fluid filaments on either side of the interface, both tangent to the inteface, then the rate of deformation of one can be made arbitraily close to the rate of deformation of the other by bringing the filaments closer and closer to the interface. If we define $\mathbf{t_1}$ and $\mathbf{t_2}$ as orthognal unit vectors that are tangent to the interface and $\mathbf{n}$ the unit normal, this fact implies:

$$\left[\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)(\mathbf{t_1}, \mathbf{t_2})\right] = \mathbf{0} \tag{D.6}$$

where $[\cdot]$ is the jump in the quantity $\cdot$ across the interface. The normal and tangent vectors can be defined everywhere using the gradient of ϕ, though the meaning of these vectors only applies at the interface. Applying the incompressibility condition to this last result gives

$$\left[\mathbf{n}^T \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \mathbf{n}\right] = 0 \tag{D.7}$$

Certain components of the stress tensor must remain continuous across the interface (in the absence of any surface tension). The shear stress on the interface and the normal compression on the interface cannot jump or else there would be an infinite acceleration. We can state these conditions as:

$$\left[(\mathbf{n}, \mathbf{t_1}, \mathbf{t_2})^T \mathbf{T} \mathbf{n}\right] = \mathbf{0}. \tag{D.8}$$

Equations D.6, D.7, and D.8 can be expanded and reorganized to yield the important jump relationship, which we shall revisit in more depth later:

$$\left[\mu \frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right] = [\mu] \frac{\partial \mathbf{v}}{\partial \mathbf{x}} (\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})(\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})^T + [\mu] \mathbf{n}\mathbf{n}^T \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \mathbf{n}\mathbf{n}^T - [\mu] (\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})(\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})^T \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)^T \mathbf{n}\mathbf{n}^T \tag{D.9}$$

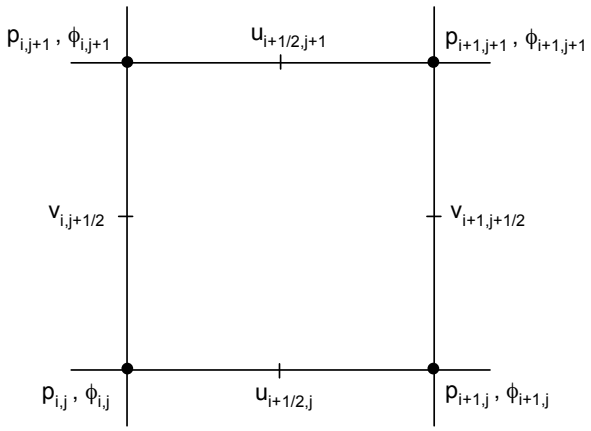


Figure D-1: A MAC grid in 2D.

The right side of the above can be evaluated at any grid point as long as we always use the value that $[\mu]$ takes at the interface. Through Eqs D.6, D.7, and D.8 it can be shown that the right side is actually continuous everywhere even though it only carries meaning at the interface. This fact will be useful in our Eulerian treatment because it enables us to interpolate to the appropriate jump conditions when the interface is between grid points. We refer to this long tensor expression as $\mathbf{J}$.

The first step when using the projection method is to omit the pressure term in Equation D.1 and name the solution to this after one time step $\mathbf{v}^*$:

$$\frac{\mathbf{v}^* - \mathbf{v}}{\Delta t} + \frac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{v} = \frac{1}{\rho}\left(\frac{\partial}{\partial \mathbf{x}} \cdot \boldsymbol{\tau}\right) + \mathbf{g} \tag{D.10}$$

This work uses a MAC grid where p and ϕ are stored on the integer grid points and components of the velocity are stored on the "cell edges". Figure D-1 illustrates analagously in 2D which fields are stored where. Equation D.10 is equivalently three scalar equations, each of which is solved numerically at different locations with respect to the grid. For instance, the first scalar component is solved only at locations of the form $(i + 1/2, j, k)$ for i, j, k integers.

As long as the interface is not crossed, we can apply the simplification

$$\frac{\partial}{\partial \mathbf{x}} \cdot \boldsymbol{\tau} = \mu\left(\frac{\partial}{\partial \mathbf{x}} \cdot \frac{\partial}{\partial \mathbf{x}}\right)\mathbf{v}$$

thanks to incompressibility. So, for example, to compute $u^*_{i+1/2,j,k}$, the Laplacian of u at $(i+1/2, j, k)$ is the only second derivative operation needing to be taken. To compute this, the point must be flanked by two "good" values of u in each of the three directions. On the occasion that one or more of the needed u values are across the interface, the values at the interface must be constructed and used to attain an accurate Laplacian. In the next section we shall delve into the precise technique behind constructing these better values. For now, we summarize the method:

1. Locate precisely where the interface lies by interpolating to the location where $\phi = 0$ between the grid points.

2. Since $\mathbf{J}$ is continuous, approximate the value of $\mathbf{J}$ at the interface via interpolation.

3. Equate the resulting $\mathbf{J}$ to $[\mu \frac{\partial \mathbf{v}}{\partial \mathbf{x}}]$ using one-sided derivatives that connect to an unknown value at the interface.

4. Solve this for the value of the needed velocity component at the interface.

D.2 Introductory large deformation elasticity

While the shear stress in a fluid depends on the rate of deformation, shear stresses in an elastic solid are a function of *total* deformation; that is, how much the solid object has been deformed from its original configuration. For small deformations, the total strain is measured using a displacement gradient. This regime is common knowledge and simple, but unfortunately we are concerned with large deformations for which the infinitessimal definition becomes an invalid strain measure.

We begin with the motion function $\boldsymbol{\chi}$ which maps every starting position $\mathbf{X}$ in the solid body $\mathbf{B}$, to the location $\mathbf{x}$ that it has moved to t time units later (see Figure D.2). In other words,

$$\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t).$$

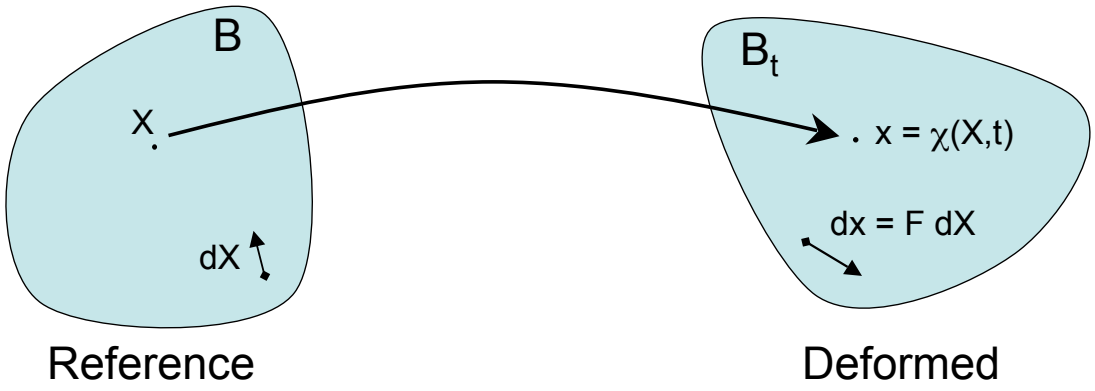


Figure D-2: An illustration of the motion function $\boldsymbol{\chi}$ and the deformation gradient $\mathbf{F}$ as they relate to a solid which deforms from the reference body $\mathbf{B}$ to a deformed body $\mathbf{B_t}$ over the course of t time units.

For a given $\boldsymbol{\chi}$ we define the deformation gradient:

$$\mathbf{F} = \frac{\partial \boldsymbol{\chi}}{\partial \mathbf{X}}. \tag{D.11}$$

The deformation gradient describes local stretching and rotation. For a small material element $\mathbf{dX}$ in the reference body, its image $\mathbf{dx}$ in the deformed body can be seen through the chain rule to be equivalent to $\mathbf{F}\ \mathbf{dX}$. Incompressibility means $\det \mathbf{F} = 1$.

To isolate the stretching from the rotation, we perform the polar decomposition enabling us to write $\mathbf{F} = \mathbf{VR}$ for rotation tensor $\mathbf{R}$ and symmetric positive definite tensor $\mathbf{V}$ known as the "left stretch". This decomposition is unique and we may quickly note that $\mathbf{V} = \sqrt{\mathbf{F}\mathbf{F}^T}$.

Strain locally measures relative changes in length and inter-filament angles. Thus, a valid finite strain measure cannot involve $\mathbf{R}$. It also must asymptote to the infinitesimal definition of strain as $\mathbf{F} \to \mathbf{1}$. There are many functional forms that uphold these restrictions, and consequently there are many valid ways of defining finite strain. We *choose* the Hencky strain measure

$$\mathbf{E^H} = \log(\mathbf{V})$$

because it is additive under simple extension/compression and also because the in-

compressibility condition forces $\mathrm{tr}\mathbf{E}^{\mathbf{H}} = 0$ simplifying the upcoming steps. Let us create a typical quadratic free energy based on the Hencky strain:

$$\psi = G\left|\mathbf{E}^{\mathbf{H}}\right|^2.$$

Incompressible elasticity theory states that in order for such a material to upold isotropy, frame indifference, and the second law of thermodynamics, the stress tensor must relate to the free energy by

$$\mathbf{T} = -p\mathbf{1} + 2\ \mathrm{dev}\left(\frac{\partial\psi}{\partial\mathbf{B}}\mathbf{B}\right) \quad \text{for} \quad \mathbf{B} = \mathbf{F}\mathbf{F}^T = \mathbf{V}^2$$

where the pressure p is independent of the strain and will be calculated using projections. Applying the above law to our stated free energy gives the deviatoric stress

$$\boldsymbol{\tau} = 2G\mathbf{E}^{\mathbf{H}} \tag{D.12}$$

which maintains such a simple form because of our choice of the Hencky strain measure.

D.3 Eulerian solid mechanics

We now wish to extend the algorithm from Section D.1 to apply when one material is an elastic incompressible solid. This notion is rare among solid algorithms because the grid points are fixed. Solids, due to their dependence on strain, are normally discretized in a Lagrangian vantagepoint where computation is performed on a lattice that advects with the material therefore enabling a direct update for $\mathbf{F}$ each time step.

Throughout, we will treat the entire material as having a general constitutive law

$$\boldsymbol{\tau} = \boldsymbol{\tau}^f + \boldsymbol{\tau}^s \tag{D.13}$$

where $\boldsymbol{\tau}^f$ takes the form in Eq D.3 and $\boldsymbol{\tau}^s$ follows Eq D.12. The shear modulus G is zero in the fluid and the viscosity μ is zero in the solid, so the above form correctly

segregates the flow.

We must develop an Eulerian method for constructing $\boldsymbol{\tau}^s$. Define a vector field $\boldsymbol{\xi}(\mathbf{x}, t)$ by $\boldsymbol{\xi}(\mathbf{x}, 0) = \mathbf{x} = \mathbf{X}$ and the evolution law

$$\dot{\boldsymbol{\xi}} = \frac{\partial \boldsymbol{\xi}}{\partial t} + \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{x}} \mathbf{v} = \mathbf{0}. \tag{D.14}$$

Then at any time later, since the initial values of $\boldsymbol{\xi}$ are being advected through the flow, we may claim that

$$\mathbf{X} = \boldsymbol{\chi}^{-1}(\mathbf{x}, t) = \boldsymbol{\xi}(\mathbf{x}, t)$$

thus giving us a way of knowing where the material began that is now occupying some location $\mathbf{x}$. This provides us the needed link to the reference configuration and enables us to determine $\mathbf{F}$ via

$$\mathbf{F}^{-1} = \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{x}}. \tag{D.15}$$

The full $\boldsymbol{\xi}$ vector is computed and stored on *each* of the cell edges. Unlike the velocity, every component of $\boldsymbol{\xi}$ must be stored at every cell edge since the divergence of τ^s does not simplify to Laplacians and requires a full $\mathbf{F}$ in all directions every time.

Away from an interface, $(\mathbf{F}^{-1})_{i,j,k}$ can be found using centered differences, for example the $\{23\}$ component is:

$$\left(F_{23}^{-1}\right)_{i,j,k} = \frac{(\xi_2)_{i.j,k+1/2} - (\xi_2)_{i,j,k-1/2}}{\Delta z}. \tag{D.16}$$

Inverting gives $\mathbf{F}_{i,j,k}$ and consequently $\boldsymbol{\tau}^s_{i,j,k}$.

D.4 The solid/fluid jump relation

In analogy with Section D.1, we note that $\boldsymbol{\xi}$ and, numerically, its gradient $\mathbf{F}$ are both continuous. However, $\frac{\partial}{\partial \mathbf{x}} \cdot \boldsymbol{\tau}^s$ is discontinuous. Like before when a velocity was needed at the interface to discretize the second derivative in Eq D.10, we must know $\boldsymbol{\xi}$ to first-order at the interface to properly discretize this divergence.

First we must rederive Eq D.9 for the fluid/solid interface. Eqs D.6, D.7 and D.8

still apply, but the deviatoric stress $\boldsymbol{\tau}$ must be treated in the form of Eq D.13. The jump condition described in Eq D.8 now gives

$$\left[(\mathbf{n}, \mathbf{t_1}, \mathbf{t_2})^T(-p\mathbf{1} + \boldsymbol{\tau}^f + \boldsymbol{\tau}^s)\mathbf{n}\right] = \mathbf{0}. \tag{D.17}$$

Expanding $\boldsymbol{\tau}^f$ and utilizing Eqs D.6 and D.7 to pull terms outside of $[\cdot]$'s that do not jump leaves the following three important conditions:

$$\begin{aligned}
-[p] + 2[\mu]\mathbf{n}^T\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{n} + \mathbf{n}^T[\boldsymbol{\tau}^s]\mathbf{n} &= 0 && \text{(D.18)}\\
\mathbf{t_1}^T\left[\mu\,\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right]\mathbf{n} + [\mu]\mathbf{n}^T\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{t_1} + \mathbf{t_1}^T[\boldsymbol{\tau}^s]\mathbf{n} &= 0 && \text{(D.19)}\\
\mathbf{t_2}^T\left[\mu\,\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right]\mathbf{n} + [\mu]\mathbf{n}^T\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{t_2} + \mathbf{t_2}^T[\boldsymbol{\tau}^s]\mathbf{n} &= 0 && \text{(D.20)}
\end{aligned}$$

Eqs D.19 and D.20 enable us to write

$$(\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})^T\left[\mu\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right]\mathbf{n} = -[\mu](\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})^T\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)^T\mathbf{n} - (\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})^T[\boldsymbol{\tau}^s]\mathbf{n}.$$

All other components of $[\mu\frac{\partial \mathbf{v}}{\partial \mathbf{x}}]$ reference components of the velocity gradient that are continuous across the interface. Therefore we have the full jump relationship

$$\begin{aligned}
&\left[\mu\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right] + (\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})(\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})^T[\boldsymbol{\tau}^s]\mathbf{n}\mathbf{n}^T\\
&\quad = [\mu]\frac{\partial \mathbf{v}}{\partial \mathbf{x}}(\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})(\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})^T + [\mu]\mathbf{n}\mathbf{n}^T\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{n}\mathbf{n}^T - [\mu](\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})(\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})^T\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)^T\mathbf{n}\mathbf{n}^T
\end{aligned} \tag{D.21}$$

Since μ and G are never simultaneously nonzero, the jumps indicated on the left side of the above simplify— for example if the solid is on the left of the interface, $[\boldsymbol{\tau}^s] = \boldsymbol{\tau}^s_R - \boldsymbol{\tau}^s_L = -\boldsymbol{\tau}^s_L$. Eq D.21 is identical to Eq D.9 except for the additional term on the left side. The right side expression is what we called $\mathbf{J}$ in Section D.1.

D.5 Process of the numerical solution

In the original fluid/fluid approach, Eq D.10 is solved at each cell edge in a standard way except for when the computation uses velocities that come from locations with differing signs of the level set function. When this occurs, an on-the-spot remedy is avaliable; the jump relation $\mathbf{J}$ is interpolated to the interface and an improved velocity component is generated which is in turn used in computing the Laplacian of the velocity.

Our problem is much more complicated. In order to apply the jump relation, we need $\boldsymbol{\tau}^s$ at the interface which in turn requires knowing $\boldsymbol{\xi}$ at the interface. To maintain the added accuracy that invoking the jump condition gives us, we cannot simply extrapolate the $\boldsymbol{\xi}$ field from inside the solid region to obtain $\boldsymbol{\xi}$ at an interface as this would be as accurate as linearly interpolating the velocity field for the interfacial velocity. To keep high accuracy, we must advect $\boldsymbol{\xi}$ near the interface using velocities found via the jump relation. We could literally follow the interface as a Lagrangian set, but this would mar the simplicity of our fully on-grid approach.

Instead, we will carry along two sets of data to be stored and updated each time step on the cell edges. The first is the extended solid velocity field $\bar{\mathbf{v}}$ which has the initial condition $\bar{\mathbf{v}} = \mathbf{v}$ and the second is the extended solid reference field $\bar{\boldsymbol{\xi}}$ with initial condition $\bar{\boldsymbol{\xi}} = \boldsymbol{\xi}$. As with their unbarred counterparts, the components of $\bar{\mathbf{v}}$ are stored at specific edges and the full $\bar{\boldsymbol{\xi}}$ vector is kept at every cell edge. These fields are smooth continuations of the $\mathbf{v}$ and $\boldsymbol{\xi}$ fields extended out from the solid region with the key feature that they give first order accurate results when interpolated to the interface. Since $\boldsymbol{\xi}$ is not necessary to update variables in the fluid zone, we need only concern ourselves with the extrapolated field $\bar{\boldsymbol{\xi}}$ which is equivalent to $\boldsymbol{\xi}$ in the solid.

Suppose we are at the n^{th} time step and are given the fields $\bar{\boldsymbol{\xi}}$, $\mathbf{v}$, $\bar{\mathbf{v}}$, and ϕ. We shall now describe how to update all of these fields to the $n + 1^{th}$ time step.

Step 1: Advect $\bar{\boldsymbol{\xi}}$, and ϕ to the next time step

Advect $\bar{\boldsymbol{\xi}}$ under the $\bar{\mathbf{v}}$ field numerically along the cell edges invoking local averages

to obtain the needed additional velocity components. For example, at the location $(i+1/2, j, k)$, one would solve

$$\frac{\bar{\boldsymbol{\xi}}^{n+1}_{i+1/2,j,k} - \bar{\boldsymbol{\xi}}^{n}_{i+1/2,j,k}}{\Delta t} + \left(\frac{\partial \bar{\boldsymbol{\xi}}}{\partial \mathbf{x}}\right)^{n}_{i+1/2,j,k} \bar{\mathbf{v}}^{n}_{i+1/2,j,k} = \mathbf{0} \tag{D.22}$$

where

$$\frac{\bar{v}^{n}_{i,j+1/2,k} + \bar{v}^{n}_{i+1,j+1/2,k} + \bar{v}^{n}_{i.j-1/2,k} + \bar{v}^{n}_{i+1,j-1/2,k}}{4} \equiv \bar{v}^{n}_{i+1/2,j,k} \tag{D.23}$$

$$\frac{\bar{w}^{n}_{i,j,k+1/2} + \bar{w}^{n}_{i+1,j,k+1/2} + \bar{w}^{n}_{i.j,k-1/2} + \bar{w}^{n}_{i+1,j,k-1/2}}{4} \equiv \bar{w}^{n}_{i+1/2,j,k} \tag{D.24}$$

and the gradient of $\bar{\boldsymbol{\xi}}$ is computed with centered differences, for example

$$\left(\left\{\frac{\partial \bar{\boldsymbol{\xi}}}{\partial \mathbf{x}}\right\}_{23}\right)_{i+1/2,j,k} = \frac{(\bar{\xi}_2)_{i+1/2,j,k+1} - (\bar{\xi}_2)_{i+1/2,j,k-1}}{2\Delta z}.$$

To update the level set function, we advect it under the actual velocity field; that is,

$$\frac{\phi^{n+1}_{i,j,k} - \phi^{n}_{i,j,k}}{\Delta t} + (u^{n}_{i,j,k}, v^{n}_{i,j,k}, w^{n}_{i,j,k}) \cdot \left(\frac{\partial \phi}{\partial \mathbf{x}}\right)^{n}_{i,j,k} = 0$$

where the velocity components above are obtained at (i, j, k) by simple averaging of the two neighboring off-grid components. The gradient of ϕ is obtained using centered differences. Since ϕ remains differentiable as long as it begins as such, there are no qualms about performing centered differences across the interface. To reinitialize the resulting ϕ field, we apply Eq D.5 until steady-state is reached using centered differences for the gradient of ϕ.

Step 2: Update the $\mathbf{v}$ *field*

We now solve Eq D.10. This relies on our ability to compute the necessary component of $\frac{\partial}{\partial \mathbf{x}} \cdot \boldsymbol{\tau}$ at any cell edge.

Suppose we wish to compute u^* at a fluid point $(i+3/2, j, k)$. The first component of $\frac{\partial}{\partial \mathbf{x}} \cdot \boldsymbol{\tau}^f$ is equivalent to $\left(\frac{\partial}{\partial \mathbf{x}} \cdot \frac{\partial}{\partial \mathbf{x}}\right) u$ so in most cases it is not numerically difficult to find this Laplacian. But suppose we have the problem that one of the u values needed

to compute the second x derivative, say at $(i+1/2, j, k)$, lies across the interface. We obtain this second derivative in a manner analagous to Kang *et al.*. We first locate the point $\mathbf{I} = (i + 1/2 + \alpha, j, k)$, $0 < \alpha < 1$, where the interface crosses the grid by linearly interpolating ϕ^n to determine where $\phi^n = 0$. Recall that Eq D.21 gives us $\mathbf{J}$. This quantity is continuous, so an approximate value of J_{11} at the interface can be found by linearly interpolation; letting J_I represent the interfacial value of J_{11}, compute $J_L = (J_{11})_{i+1/2,j,k}$ and $J_R = (J_{11})_{i+3/2,j,k}$ from the right side of Eq D.21 and interpolate linearly to obtain a value for J_I. Then equate the result to the discretized left side of Eq D.21. That is,

$$\mu \frac{u_{i+3/2,j,k} - u_I}{\Delta x(1-\alpha)} - \left((\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})(\mathbf{0}, \mathbf{t_1}, \mathbf{t_2})^T \boldsymbol{\tau}_I^s \mathbf{n}\mathbf{n}^T\right)_{11} = J_I. \tag{D.25}$$

Before this can be solved, we must determine $\boldsymbol{\tau}_I^s$. To accomplish this, we need a good value for $\mathbf{F}_I^{-1}$ on the solid side of the interface. *For this sole reason, we have invented the solid extended reference variable* $\bar{\boldsymbol{\xi}}$.

As per Eq D.16, replacing $\boldsymbol{\xi}$ with $\bar{\boldsymbol{\xi}}$, solve for $\mathbf{F}_L^{-1} = (\mathbf{F}^{-1})^n_{\lfloor i+1/2+\alpha \rfloor, j, k}$ and $\mathbf{F}_R^{-1} = (\mathbf{F}^{-1})^n_{\lceil i+1/2+\alpha \rceil, j, k}$ and then apply the simple interpolation

$$\frac{\mathbf{F}_I^{-1} - \mathbf{F}_L^{-1}}{i + 1/2 + \alpha - \lfloor i + 1/2 + \alpha \rfloor} = \mathbf{F}_R^{-1} - \mathbf{F}_L^{-1}$$

to obtain $\mathbf{F}_I^{-1}$. From this result, we construct $\mathbf{E}_I^H$ which is used to give $\boldsymbol{\tau}_I^s$.

Now, Eq D.25 may be solved for u_I. With u_I in hand, we may accomplish our original objective here by writing

$$\left(\frac{\partial^2 u}{\partial x^2}\right)^n_{i+3/2,j,k} = \frac{1}{\Delta x}\left(\frac{u^n_{i+5/2,j,k} - u^n_{i+3/2,j,k}}{\Delta x} - \frac{u^n_{i+3/2,j,k} - u_I}{\Delta x(1-\alpha)}\right).$$

Before continuing this procedure anywhere else, we should store a quantity that will be of interest in a future step. Define a field $\overline{u}_x$ stored wherever u is stored and let $(\bar{u}_x)^n_{i+1/2,j,k} = (u_I - u_L)/(\alpha \Delta x)$.

We have described up to now how to obtain u^* at a fluid point. Suppose that we wish to compute u^* at a solid point. Replacing $\boldsymbol{\xi}$ with $\bar{\boldsymbol{\xi}}$ in Eq D.16, we easily

calculate $\mathbf{F}^{-1}$ and consequently $\boldsymbol{\tau}^s$ at all neighboring grid points and compute the divergence using centered differences. Since $\bar{\boldsymbol{\xi}}$ is a solid extention that is differentiable over the interface, no special routine must be employed near the interface.

Now, the algorithm up to this point can be applied at all grid points and all directions yielding a $\mathbf{v}^*$ field fit for the projection method. That is to say, following Kang *et al.*, we obtain the pressure p, by numerically solving

$$\left(\frac{\partial}{\partial \mathbf{x}} \cdot \frac{\partial}{\partial \mathbf{x}}\right) p = \frac{\rho}{\Delta t} \frac{\partial}{\partial \mathbf{x}} \cdot \mathbf{v}^*$$

Then $\mathbf{v}^*$ and p are both used to produce the velocity field at time $n+1$ via

$$\frac{\mathbf{v}^{n+1} - \mathbf{v}^*}{\Delta t} + \frac{1}{\rho} \frac{\partial p}{\partial \mathbf{x}} = \mathbf{0}$$

which is solved numerically in the same fashion as Kang *et al.*.

Step 3: Update the field $\bar{\mathbf{v}}$.

To obtain $\bar{\mathbf{v}}^{n+1}$, we must perform extrapolation. Since we have a level set function on-hand and wish to extrapolate the solid region into the fluid, we utilize the PDE-based extrapolation method described in Aslam [10].

From the last step, we have $\mathbf{v}^{n+1}$ throughout and the components of $\frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{x}}$ at solid points adjacent to the interface. We now fill in the remaining components of $\frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{x}}$ in a fashion fit for solid extrapolation; consider the representative example of the component $\overline{u}_x$ at some arbitrary location $(i+1/2, j, k)$:

$$(\overline{u}_x)_{i+1/2,j,k} = \begin{cases} 0 & \text{if } \phi_{i+1/2,j,k} < 0 \\ (u^n_{i+3/2,j,k} - u^n_{i-1/2,j,k})/(2\Delta x) & \text{if } \phi_{i+1/2,j,k} > 0 \text{ and } \phi_{i+1/2\pm1,j,k} > 0. \end{cases} \tag{D.26}$$

We assume in the above that $\phi > 0$ represents solid and $\phi < 0$ fluid. At locations where neither of the two above conditions apply, $\overline{u}_x$ is already known from Step 2. We perform an analagous procedure to construct every component of $\frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{x}}$.

To linearly extrapolate the velocity field from the solid region, we find a steady-

state solution for

$$\frac{\partial \bar{\mathbf{v}}_n}{\partial t} + H(-\phi)\frac{\partial \bar{\mathbf{v}}_n}{\partial \mathbf{x}}\mathbf{n} = 0$$

where H is the Heavyside function. The variable $\bar{\mathbf{v}}_n$ is the directional derivative

$$\bar{\mathbf{v}}_n = \frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{x}}\mathbf{n} \quad \text{for} \quad \mathbf{n} = \frac{\partial \phi}{\partial \mathbf{x}} \Big/ \left|\frac{\partial \phi}{\partial \mathbf{x}}\right|.$$

The extrapolated velocity is then the steady-state solution to

$$\frac{\partial \bar{\mathbf{v}}}{\partial t} + H(-\phi)\left(\frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{x}}\mathbf{n} - \bar{\mathbf{v}}_n\right) = 0.$$

To illustrate the numerical method for the above procedure, without loss of generality, we perform the calculation for $\bar{u}$. First we compute $\mathbf{n}^{n+1}_{i+1/2,j,k}$ using

$$\left(\frac{\partial \phi}{\partial \mathbf{x}}\right)^{n+1}_{i+1/2,j,k} = (1/2)\left(\left(\frac{\partial \phi}{\partial \mathbf{x}}\right)^{n+1}_{i,j,k} + \left(\frac{\partial \phi}{\partial \mathbf{x}}\right)^{n+1}_{i+1,j,k}\right)$$

and iterate

$$(\bar{u}_n)_{i+1/2,j,k} := (\bar{u}_n)_{i+1/2,j,k} - \Delta t H(-\phi^{n+1}_{i+1/2,j,k})\mathbf{n}^{n+1}_{i+1/2,j,k} \cdot \left(\frac{\partial \bar{u}_n}{\partial \mathbf{x}}\right)_{i+1/2,j,k}$$

using centered differences for the gradient of $\bar{u}_n$. Once $\bar{u}_n$ is sufficiently steady, we solve for $\bar{u}$ by iterating

$$\bar{u}_{i+1/2,j,k} := \bar{u}_{i+1/2,j,k} - \Delta t H(-\phi^{n+1}_{i+1/2,j,k})\left(\mathbf{n}^{n+1}_{i+1/2,j,k} \cdot \left(\frac{\partial \bar{u}}{\partial \mathbf{x}}\right)_{i+1/2,j,k} - (\bar{u}_n)_{i+1/2,j,k}\right)$$

using centered derivatives for the gradient of $\bar{u}$ and the initial condition

$$\bar{u}_{i+1/2,j,k} = \begin{cases} 0 & \text{if } \phi_{i+1/2,j,k} < 0 \\ u^{n+1}_{i+1/2,j,k} & \text{if } \phi_{i+1/2,j,k} > 0. \end{cases} \tag{D.27}$$

The steady solution of this is presicely $\bar{u}^{n+1}$. This process is executed analgously to obtain $\bar{v}^{n+1}$ and $\bar{w}^{n+1}$.

Appendix E

Explicit VUMAT

The Fortran 90 code below is the material model (VUMAT) under explicit integration. The internally called subroutines "matinv" and "spectral" are not included for the sake of brevity. The "matinv" subroutine computes the inverse and determinant of the inputted matrix, and "spectral" computes the eigenvalues and eigenvectors of the inputted matrix. Both come from *Numerical Recipes*.

```
      SUBROUTINE VUMAT (
!       Read only (unmodifiable) variables :-
     +                    NBLOCK, NDIR, NSHR, NSTATEV, NFIELDV,
     +                    NPROPS, LANNEAL, STEP_TIME, TOTAL_TIME,
     +                    DT, CMNAME, COORD_MP, CHAR_LENGTH, PROPS,
     +                    DENSITY, STRAIN_INC, REL_SPIN_INC,
     +                    TEMP_OLD, STRETCH_OLD, DEFGRAD_OLD,
     +                    FIELD_OLD, STRESS_OLD, STATE_OLD,
     +                    ENER_INTERN_OLD, ENER_INELAS_OLD, TEMP_NEW,
     +                    STRETCH_NEW, DEFGRAD_NEW, FIELD_NEW,
!       Read and ! write (modifiable) variables :
     +                    STRESS_NEW, STATE_NEW, ENER_INTERN_NEW,
     +                    ENER_INELAS_NEW)

      INCLUDE 'VABA_PARAM.INC'

      DIMENSION COORD_MP(NBLOCK,*),CHAR_LENGTH(NBLOCK), PROPS(NPROPS),
```

```
+       DENSITY(NBLOCK), STRAIN_INC(NBLOCK,NDIR+NSHR),
+       REL_SPIN_INC(NBLOCK,NSHR), TEMP_OLD(NBLOCK),
+       STRETCH_OLD(NBLOCK,NDIR+NSHR),
+       DEFGRAD_OLD(NBLOCK,NDIR+NSHR+NSHR),
+       FIELD_OLD(NBLOCK,NFIELDV), STRESS_OLD(NBLOCK,NDIR+NSHR),
+       STATE_OLD(NBLOCK,NSTATEV), ENER_INTERN_OLD(NBLOCK),
+       ENER_INELAS_OLD(NBLOCK), TEMP_NEW(NBLOCK),
+       STRETCH_NEW(NBLOCK,NDIR+NSHR),
+       DEFGRAD_NEW(NBLOCK,NDIR+NSHR+NSHR),
+       FIELD_NEW(NBLOCK,NFIELDV), STRESS_NEW(NBLOCK,NDIR+NSHR),
+       STATE_NEW(NBLOCK,NSTATEV), ENER_INTERN_NEW(NBLOCK),
+       ENER_INELAS_NEW(NBLOCK)

 character*8 CMNAME

 integer  km
 real*8 F_t(3,3),F_tau(3,3),U_tau(3,3),U_inv(3,3), R_tau(3,3)
 real*8 T_tau(3,3), Ee(3,3), det_F, E_tau(3,3), Re_tau(3,3)
 real*8 Fp_t(3,3), Fp_tau(3,3), Fe(3,3), eigvecs(3,3), eigvals(3)
 real*8 nu_p, p, tau,  I_1(3,3),Dp(3,3),Te_0(3,3),stretches(3,3)
 real*8 Fp_inv(3,3), Ee0(3,3), tr_Ee, mus, mu2, diam, I0, dtmp
 real*8 det_Fe, det_Fp_tau, zero_m(3,3),  snake, z, Ue_inv(3,3)
 real*8 mu, kappa, dens_p,  excess, B,  Te_sph, det_U
 real*8 D(3,3), F_inv(3,3)

 I_1 = reshape((/1, 0, 0, 0, 1, 0, 0, 0, 1 /), (/3,3/))
 Zero_m = reshape((/0, 0, 0, 0, 0, 0, 0, 0, 0 /), (/3,3/))
 z=0.d0

 ! Get properties defined in input file

 B = props(01)
 mus = props(02)
 I0 = props(03)
 Diam = props(04)
 mu2 = props(05)
 dens_p = props(06)

 mus = dtan(mus*3.14159265358979/180.)
 mu2 = dtan(mu2*3.14159265358979/180.)
```

```
snake = 1.d0 / mu2**2

! Main loop over the block of material pts supplied by ABAQUS

do km = 1,NBLOCK

  !Copy the old and new Deformation gradients into F_T and F_tau
  !respectively

  F_T(1,1) = DEFGRAD_OLD(KM,1)
  F_T(2,2) = DEFGRAD_OLD(KM,2)
  F_T(3,3) = DEFGRAD_OLD(KM,3)
  F_T(1,2) = DEFGRAD_OLD(KM,4)

  F_tau(1,1) = DEFGRAD_NEW(KM,1)
  F_tau(2,2) = DEFGRAD_NEW(KM,2)
  F_tau(3,3) = DEFGRAD_NEW(KM,3)
  F_tau(1,2) = DEFGRAD_NEW(KM,4)

  U_tau(1,1) = STRETCH_NEW(KM,1)
  U_tau(2,2) = STRETCH_NEW(KM,2)
  U_tau(3,3) = STRETCH_NEW(KM,3)
  U_tau(1,2) = STRETCH_NEW(KM,4)

  if (NSHR .eq. 1) then

    F_T(2,3) = ZERO
    F_T(3,1) = ZERO
    F_T(2,1) = DEFGRAD_OLD(KM,5)
    F_T(3,2) = ZERO
    F_T(1,3) = ZERO

    F_tau(2,3) = ZERO
    F_tau(3,1) = ZERO
    F_tau(2,1) = DEFGRAD_NEW(KM,5)
    F_tau(3,2) = ZERO
    F_tau(1,3) = ZERO

    U_tau(2,3) = ZERO
    U_tau(3,1) = ZERO
```

```
          U_tau(2,1) = U_tau(1,2)
          U_tau(3,2) = ZERO
          U_tau(1,3) = ZERO

        else

          F_T(2,3) = DEFGRAD_OLD(KM,5)
          F_T(3,1) = DEFGRAD_OLD(KM,6)
          F_T(2,1) = DEFGRAD_OLD(KM,7)
          F_T(3,2) = DEFGRAD_OLD(KM,8)
          F_T(1,3) = DEFGRAD_OLD(KM,9)

          F_tau(2,3) = DEFGRAD_NEW(KM,5)
          F_tau(3,1) = DEFGRAD_NEW(KM,6)
          F_tau(2,1) = DEFGRAD_NEW(KM,7)
          F_tau(3,2) = DEFGRAD_NEW(KM,8)
          F_tau(1,3) = DEFGRAD_NEW(KM,9)

          U_tau(2,3) = STRETCH_NEW(KM,5)
          U_tau(3,1) = STRETCH_NEW(KM,6)
          U_tau(2,1) = U_tau(1,2)
          U_tau(3,2) = U_tau(2,3)
          U_tau(1,3) = U_tau(3,1)

        endif

  if(total_time.eq.0 .and. step_time.eq.0) then

          ! If first dummy step: Initialize state vars

          State_old(km,1:9)   = (/1, 0, 0, 0, 1, 0, 0, 0, 1 /) ! Fp
          State_old(km,10:18) = (/0, 0, 0, 0, 0, 0, 0, 0, 0 /)  ! M_0
          State_old(km,19) = 0.d0 !  Sph Mandel

call SPECTRAL(matmul(transpose(F_tau), F_tau), eigvals, eigvecs)

        E_tau = matmul(matmul(eigvecs,
     + 0.5d0*reshape((/ dlog(eigvals(1)),z,z,z,dlog(eigvals(2)),z,z,z,
     + dlog(eigvals(3)) /), (/3, 3/))), transpose(eigvecs))
```

```
    ! Elastic repsonse in dummy step:

         tr_Ee = E_tau(1,1)+E_tau(2,2)+E_tau(3,3)
         Ee0 = E_tau - (tr_Ee/3.d0) * I_1

if(tr_Ee .lt. z)then
  kappa = B*dsqrt(-tr_Ee)*(1.d0 +
     +                   0.5d0*sum(Ee0*Ee0)/(snake*(tr_Ee**2)))
  mu = B*dsqrt(-tr_Ee)/snake
else
kappa=z
mu=z
end if

T_tau = 2.d0 * mu * Ee0 + kappa * tr_Ee * I_1
state_old(km,20:28) = reshape(eigvecs,(/9/))
state_old(km,29) = 0.d0
state_old(km,30) = 0.d0

      else

   det_F=F_tau(1,1)*(F_tau(2,2)*F_tau(3,3)-F_tau(3,2)*F_tau(2,3))
     +      -F_tau(2,1)*(F_tau(1,2)*F_tau(3,3)-F_tau(3,2)*F_tau(1,3))
     +      +F_tau(3,1)*(F_tau(1,2)*F_tau(2,3)-F_tau(2,2)*F_tau(1,3))


         ! Get state vars from last step

         Fp_t = reshape(state_old(km,1:9), (/3, 3/))
         Te_0 = reshape(state_old(km,10:18), (/3, 3/))   ! Dev Mandel
         Te_sph = state_old(km,19) ! Sphere Mandel
         eigvecs = reshape(state_old(km, 20:28), (/3, 3/)) ! Evecs Ce

      p = - Te_sph
tau = dsqrt(0.5d0*sum(Te_0*Te_0))

excess = tau - mus*p - 1.d-3

state_new(km, 31) = 0.d0
state_new(km,33) = -10.d0
```

```
          if(excess .gt. 0.d0)then

state_new(km, 31) = 1.d0
                nu_p = (I0/DIAM)*DSQRT(p / dens_p)*(excess)
     +               /(mu2*p - tau)

        state_new(km,33) = dlog(nu_p)
                Dp = nu_p * Te_0/(2.d0*tau)

       stretches = matmul(matmul(transpose(eigvecs), Dp), eigvecs)

       Fp_tau = matmul(matmul(
     +  matmul(eigvecs, reshape((/
     +  dexp(stretches(1,1)*dt), z, z,
     +  z, dexp(stretches(2,2)*dt), z,
     +  z, z, dexp(stretches(3,3)*dt) /), (/3, 3/))),
     +  transpose(eigvecs)), Fp_t)

det_Fp_tau = Fp_tau(1,1)*(Fp_tau(2,2)*Fp_tau(3,3)-
     +    Fp_tau(3,2)*Fp_tau(2,3))-Fp_tau(2,1)*(Fp_tau(1,2)*Fp_tau(3,3)
     +    -Fp_tau(3,2)*Fp_tau(1,3)) +Fp_tau(3,1)*(Fp_tau(1,2)*
     +    Fp_tau(2,3)-Fp_tau(2,2)*Fp_tau(1,3))

       Fp_tau = Fp_tau * (det_Fp_tau)**(-1.d0/3.d0)

             else

    nu_p = 0.d0
    Dp = zero_m
                    Fp_tau = Fp_t

             det_Fp_tau = Fp_tau(1,1)*(Fp_tau(2,2)*Fp_tau(3,3)-
     +    Fp_tau(3,2)*Fp_tau(2,3))-Fp_tau(2,1)*(Fp_tau(1,2)*Fp_tau(3,3)
     +    -Fp_tau(3,2)*Fp_tau(1,3)) +Fp_tau(3,1)*(Fp_tau(1,2)*
     +    Fp_tau(2,3)-Fp_tau(2,2)*Fp_tau(1,3))

             end if

call matinv(Fp_tau, Fp_inv, dtmp)
```

```
          ! Calculate Fe:

         Fe=matmul(F_tau, Fp_inv)   ! This is off-rotated from Fe

       call SPECTRAL(matmul(transpose(Fe), Fe), eigvals, eigvecs)

       Ee = matmul(matmul(eigvecs,
     + 0.5d0*reshape((/dlog(eigvals(1)),z,z,z,dlog(eigvals(2)),z,z,z,
     + dlog(eigvals(3)) /), (/3, 3/))), transpose(eigvecs))

         tr_Ee = Ee(1,1)+Ee(2,2)+Ee(3,3)
Ee0 = Ee - (tr_Ee/3.d0) * I_1
tr_Ee = dlog(det_F)

if(tr_Ee .lt. z)then

  kappa = B*dsqrt(-tr_Ee)*(1.d0 +
     +         0.5d0*sum(Ee0*Ee0)/(snake*(tr_Ee**2)))

  mu = B*dsqrt(-tr_Ee)/snake

state_new(km,32)=0.
else

kappa=z
mu=z

state_new(km,32)=1.

end if

Te_sph = kappa * tr_Ee
Te_0 = 2.d0 * mu * Ee0

Ue_inv = matmul(matmul(eigvecs,
     + reshape((/ eigvals(1)**(-.5d0),z,z,z,eigvals(2)**(-.5d0),z,z,z,
     + eigvals(3)**(-.5d0) /), (/3, 3/))), transpose(eigvecs))

         Re_tau = matmul(matmul(F_tau, Fp_inv), Ue_inv)
```

```
T_tau = matmul(Re_tau, matmul(Te_0+I_1*Te_sph,
     +            transpose(Re_tau)))/det_F

state_new(km,10:18) =  reshape(Te_0, (/9/))     ! Dev Mandel
     state_new(km,19) = Te_sph    ! Sph Mandel
state_new(km,1:9) = reshape(Fp_tau, (/9/))
state_new(km, 20:28) = reshape(eigvecs, (/9/))  ! Eigvecs of Mandel
state_new(km, 29) = nu_p
state_new(km, 30) = tau/(dabs(p) + 1.d-4)     ! mu

      end if ! if(total_time.eq.0 .and. step_time.eq.0)


!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
         ! Rotate Cauchy stress to ABAQUS stress

        call MATINV(U_tau, U_inv, det_U)
        R_tau = matmul(F_tau, U_inv)
        T_tau = matmul(transpose(R_tau),matmul(T_tau,R_tau))

  ! Update ABAQUS stresses

        do i = 1,NDIR

            STRESS_NEW(KM,i) = T_tau(i,i)

        end do

        if (NSHR .ne. 0) then

          STRESS_NEW(KM,NDIR+1) = T_tau(1,2)

          if (NSHR .ne. 1) then

            STRESS_NEW(KM,NDIR+2) = T_tau(2,3)

            if (NSHR .ne. 2) then

              STRESS_NEW(KM,NDIR+3) = T_tau(1,3)
```

```
      endif
    endif
  endif

  ! Specific internal energy

  stress_power = ONE_HALF * (
+       ( stress_old(km,1)+stress_new(km,1) )*strain_inc(km,1) +
+       ( stress_old(km,2)+stress_new(km,2) )*strain_inc(km,2) +
+       ( stress_old(km,3)+stress_new(km,3) )*strain_inc(km,3))

    if(nshr .eq. 1) then

  stress_power = stress_power + one_half*(
+   TWO*(stress_old(km,4)+stress_new(km,4))*strain_inc(km,4) )

    else

  stress_power = stress_power + one_half*(
+   TWO*(stress_old(km,4)+stress_new(km,4))*strain_inc(km,4) +
+   TWO*(stress_old(km,5)+stress_new(km,5))*strain_inc(km,5) +
+   TWO*(stress_old(km,6)+stress_new(km,6))*strain_inc(km,6) )

    endif

    ener_intern_new(km) = ener_intern_old(km)
+                         + stress_power/density(km)

 end do  ! loop over km

 return

 end
```

Appendix F

Implicit VUMAT

The Fortran 90 code below is the material model (VUMAT) under implicit integration. The internally called subroutines "matinv" and "spectral" are not included for the sake of brevity. The "matinv" subroutine computes the inverse and determinant of the inputted matrix, and "spectral" computes the eigenvalues and eigenvectors of the inputted matrix. Both come from *Numerical Recipes*.

```
      SUBROUTINE VUMAT (

!     Read only (unmodifiable) variables :-
     +                    NBLOCK, NDIR, NSHR, NSTATEV, NFIELDV,
     +                    NPROPS, LANNEAL, STEP_TIME, TOTAL_TIME,
     +                    DT, CMNAME, COORD_MP, CHAR_LENGTH, PROPS,
     +                    DENSITY, STRAIN_INC, REL_SPIN_INC,
     +                    TEMP_OLD, STRETCH_OLD, DEFGRAD_OLD,
     +                    FIELD_OLD, STRESS_OLD, STATE_OLD,
     +                    ENER_INTERN_OLD, ENER_INELAS_OLD, TEMP_NEW,
     +                    STRETCH_NEW, DEFGRAD_NEW, FIELD_NEW,
!     Read and write (modifiable) variables :-
     +                    STRESS_NEW, STATE_NEW, ENER_INTERN_NEW,
     +                    ENER_INELAS_NEW)

      INCLUDE 'VABA_PARAM.INC'

      DIMENSION COORD_MP(NBLOCK,*),CHAR_LENGTH(NBLOCK), PROPS(NPROPS),
```

```
+       DENSITY(NBLOCK), STRAIN_INC(NBLOCK,NDIR+NSHR),
+       REL_SPIN_INC(NBLOCK,NSHR), TEMP_OLD(NBLOCK),
+       STRETCH_OLD(NBLOCK,NDIR+NSHR),
+       DEFGRAD_OLD(NBLOCK,NDIR+NSHR+NSHR),
+       FIELD_OLD(NBLOCK,NFIELDV), STRESS_OLD(NBLOCK,NDIR+NSHR),
+       STATE_OLD(NBLOCK,NSTATEV), ENER_INTERN_OLD(NBLOCK),
+       ENER_INELAS_OLD(NBLOCK), TEMP_NEW(NBLOCK),
+       STRETCH_NEW(NBLOCK,NDIR+NSHR),
+       DEFGRAD_NEW(NBLOCK,NDIR+NSHR+NSHR),
+       FIELD_NEW(NBLOCK,NFIELDV), STRESS_NEW(NBLOCK,NDIR+NSHR),
+       STATE_NEW(NBLOCK,NSTATEV), ENER_INTERN_NEW(NBLOCK),
+       ENER_INELAS_NEW(NBLOCK)

 character*8 CMNAME

 integer  km, n
 real*8 F_t(3,3),F_tau(3,3),U_tau(3,3),U_inv(3,3),R_tau(3,3)
 real*8 T_tau(3,3), Ee(3,3), det_F, E_tau(3,3), Fe_tr(3,3)
 real*8 Re_tr(3,3), Ee_tr(3,3), trace_Ee_tr, Ee_tr_0(3,3)
 real*8 mag_Ee_tr_0, delta, G, C, dp, J(2,2), Jinv(2,2)
 real*8 Fp_t(3,3), Fp_tau(3,3), Fe(3,3), eigvecs(3,3), eigvals(3)
 real*8 p_old, I_1(3,3), dp_old, Te_0(3,3), stretches(3,3)
 real*8 Ee0(3,3), tr_Ee, mus, mu2, diam, I0, dtmp, det_U
 real*8 zero_m(3,3),  snake, z, Fp_inv(3,3), det_Fp_tau
 real*8 mu, kappa, dens_p,  tau,  Te_sph, x_prev, y_prev
 real*8 f1, f2, coh, x, y, B

 I_1 = reshape((/1, 0, 0, 0, 1, 0, 0, 0, 1 /), (/3,3/))
 Zero_m = reshape((/0, 0, 0, 0, 0, 0, 0, 0, 0 /), (/3,3/))
 z=0.d0

 ! Get properties defined in input file

 B = props(01)
 mus = props(02)
 I0 = props(03)
 Diam = props(04)
 mu2 = props(05)
 dens_p = props(06)
```

```
mus=dtan(mus*3.141459265/180)
mu2=dtan(mu2*3.141459265/180)
snake = 1.d0 /(mu2**2)
coh = 1.d-5

! Main loop over the block of material pts supplied by ABAQUS

do km = 1,NBLOCK

  !Copy the old and new Deformation gradients into F_T and F_tau

  !respectively

  !

  F_T(1,1) = DEFGRAD_OLD(KM,1)
  F_T(2,2) = DEFGRAD_OLD(KM,2)
  F_T(3,3) = DEFGRAD_OLD(KM,3)
  F_T(1,2) = DEFGRAD_OLD(KM,4)

  F_tau(1,1) = DEFGRAD_NEW(KM,1)
  F_tau(2,2) = DEFGRAD_NEW(KM,2)
  F_tau(3,3) = DEFGRAD_NEW(KM,3)
  F_tau(1,2) = DEFGRAD_NEW(KM,4)

  U_tau(1,1) = STRETCH_NEW(KM,1)
  U_tau(2,2) = STRETCH_NEW(KM,2)
  U_tau(3,3) = STRETCH_NEW(KM,3)
  U_tau(1,2) = STRETCH_NEW(KM,4)

  if (NSHR .eq. 1) then
    F_T(2,3) = ZERO
    F_T(3,1) = ZERO
    F_T(2,1) = DEFGRAD_OLD(KM,5)
    F_T(3,2) = ZERO
    F_T(1,3) = ZERO

    F_tau(2,3) = ZERO
    F_tau(3,1) = ZERO
    F_tau(2,1) = DEFGRAD_NEW(KM,5)
```

```
          F_tau(3,2) = ZERO
          F_tau(1,3) = ZERO

          U_tau(2,3) = ZERO
          U_tau(3,1) = ZERO
          U_tau(2,1) = U_tau(1,2)
          U_tau(3,2) = ZERO
          U_tau(1,3) = ZERO

        else
          F_T(2,3) = DEFGRAD_OLD(KM,5)
          F_T(3,1) = DEFGRAD_OLD(KM,6)
          F_T(2,1) = DEFGRAD_OLD(KM,7)
          F_T(3,2) = DEFGRAD_OLD(KM,8)
          F_T(1,3) = DEFGRAD_OLD(KM,9)

          F_tau(2,3) = DEFGRAD_NEW(KM,5)
          F_tau(3,1) = DEFGRAD_NEW(KM,6)
          F_tau(2,1) = DEFGRAD_NEW(KM,7)
          F_tau(3,2) = DEFGRAD_NEW(KM,8)
          F_tau(1,3) = DEFGRAD_NEW(KM,9)

          U_tau(2,3) = STRETCH_NEW(KM,5)
          U_tau(3,1) = STRETCH_NEW(KM,6)
          U_tau(2,1) = U_tau(1,2)
          U_tau(3,2) = U_tau(2,3)
          U_tau(1,3) = U_tau(3,1)

        endif


if(total_time.eq.0 .and. step_time.eq.0) then

          ! If first dummy step: Initialize state vars

        State_old(km,1:9)   = (/1, 0, 0, 0, 1, 0, 0, 0, 1 /) ! Fp
        State_old(km,10:18) = (/0, 0, 0, 0, 0, 0, 0, 0, 0 /)  ! M_0
        State_old(km,19) = 0.d0 !  Sphere part of Mandel

        call SPECTRAL(matmul(transpose(F_tau),F_tau),eigvals,eigvecs)
```

```
 E_tau = matmul(matmul(eigvecs,
     + 0.5d0*reshape((/ dlog(eigvals(1)),z,z,z,dlog(eigvals(2)),z,z,z,
     + dlog(eigvals(3)) /), (/3, 3/))), transpose(eigvecs))

    ! Elastic repsonse in dummy step:

         tr_Ee = E_tau(1,1)+E_tau(2,2)+E_tau(3,3)
         Ee0 = E_tau - (tr_Ee/3.d0) * I_1

          if(tr_Ee . lt . 0.d0) then
               kappa = B*dsqrt(-tr_Ee)*(1 +
     +                 one_half*sum(Ee0*Ee0)/(snake*tr_Ee**2))
               mu = dsqrt(-tr_Ee)*B/snake

          else
       kappa = 0.d0
               mu = 0.d0

          end if

T_tau = 2.d0 * mu * Ee0 + kappa * tr_Ee * I_1
state_old(km,20:28) = reshape(eigvecs,(/9/))
        state_old(km,29) = 0.d0     ! dp
        state_old(km,30) = 0.d0     ! excess


      else


          ! Get state vars from last step

          Fp_t = reshape(state_old(km,1:9), (/3, 3/))
          Te_0 = reshape(state_old(km,10:18), (/3, 3/))   ! Dev Mandel
  Te_sph = state_old(km,19) ! Sph Mandel
          eigvecs = reshape(state_old(km, 20:28), (/3, 3/)) ! Evecs Ce
          dp_old = state_old(km,29)  ! Old plastic shearing rate
          p_old = - Te_sph

call matinv(Fp_t, Fp_inv, dtmp)
```

```
Fe_tr = matmul(F_tau, Fp_inv)

call spectral(matmul(transpose(Fe_tr),Fe_tr),
     +                eigvals, eigvecs)

Re_tr = matmul(Fe_tr,
     + matmul(matmul(eigvecs,
     + reshape((/eigvals(1)**-0.5d0,z,z,z,eigvals(2)**-0.5d0,z,z,z,
     + eigvals(3)**-0.5d0 /), (/3, 3/))), transpose(eigvecs)))

Ee_tr = matmul(matmul(eigvecs,
     + 0.5d0*reshape((/dlog(eigvals(1)),z,z,z,dlog(eigvals(2)),z,z,z,
     + dlog(eigvals(3)) /), (/3, 3/))), transpose(eigvecs))

trace_Ee_tr = Ee_tr(1,1) + Ee_tr(2,2) + Ee_tr(3,3)
Ee_tr_0 = Ee_tr - (trace_Ee_tr/3.d0)*I_1
mag_Ee_tr_0 = dsqrt(sum(Ee_tr_0*Ee_tr_0))

if(trace_Ee_tr . ge. 0.d0)then
dp = 0.d0
Fp_tau = Fp_t
Te_0 = zero_m
Te_sph = 0.d0

else

delta = -trace_Ee_tr

if(dsqrt(2.d0)*mag_Ee_tr_0/snake - mus*delta*(1.d0+mag_Ee_tr_0**2/
     +   (2.d0*snake*delta**2)) - coh/(B*dsqrt(delta)) .lt. 0.d0) then

dp = 0.d0
Fp_tau = Fp_t
Te_sph =-B*delta**(1.5d0)*(1.d0+
     +           mag_Ee_tr_0**2/(2.*snake*delta**2))
Te_0 = (B*2.*dsqrt(delta)/snake)*Ee_tr_0

else
```

```
!  Newton-Raphson iteration

G = B*dsqrt(delta)/snake
C = I0/(diam*dsqrt(dens_p))

if(dp_old .lt. 1.d-9)then

          x = dsqrt(dp_old) + 1.d0

else

         x=dsqrt(dp_old)

end if

y = dsqrt(dsqrt(p_old))

f1 = mu2*(y**4)*(x**2) + C*mus*y**6 -
     + (dsqrt(2.d0)*G*mag_Ee_tr_0 - G*dt*x**2 - coh)*(x**2+C*y**2)

f2 = -y**4 + B*delta**(1.5d0)*(1.d0+((mag_Ee_tr_0-
     + dt*(x**2)/dsqrt(2.d0))**2)/(2.*snake*delta**2))

x_prev = 10.d10
y_prev = 10.d10

n = 1

do while (dabs((x-x_prev)/x)+ dabs((y-y_prev)/y) .gt. 1.d-4)

        x_prev = x
         y_prev = y

J(1,1) = 2*x*(coh + 2*dt*G*x**2 +
     + C*dt*G*y**2 + mu2*y**4 - dsqrt(2.d0)*G*mag_Ee_tr_0)

J(1,2) = 4*mu2*(x**2)*(y**3) + 6*C*mus*y**5
     + + 2*C*y*(coh + dt*G*x**2 - dsqrt(2.d0)*G*mag_Ee_tr_0)

J(2,1) = -dsqrt(2.d0/delta)*(B*dt*x/snake)*
```

```
     + (-(dt*x**2)/dsqrt(2.d0) + mag_Ee_tr_0)

J(2,2) = -4*y**3

   dtmp = J(1,1)*J(2,2)-J(1,2)*J(2,1)

Jinv(1,1) = J(2,2)
Jinv(1,2) = -J(1,2)
Jinv(2,1) = -J(2,1)
Jinv(2,2) = J(1,1)
Jinv = Jinv/dtmp

x = x - (Jinv(1,1)*f1 + Jinv(1,2)*f2)
y = y - (Jinv(2,1)*f1 + Jinv(2,2)*f2)

f1 = mu2*(y**4)*(x**2) + C*mus*y**6 -
     + (dsqrt(2.d0)*G*mag_Ee_tr_0 - G*dt*x**2 - coh)*(x**2 + C*y**2)

f2 = -y**4 + B*delta**(1.5d0)*(1.d0+((mag_Ee_tr_0-
     + dt*(x**2)/dsqrt(2.d0))**2)/(2.*snake*delta**2))

if(n .gt. 1.d10)then

        stop

endif

n = n+1

end do

dp = x**2
Te_sph = - y**4
Te_0= 2.*G*Ee_tr_0-G*dt*dsqrt(2.d0)*dp*Ee_tr_0/mag_Ee_tr_0
tau = dsqrt(sum(Te_0*Te_0)/2.d0)

stretches = matmul(matmul(transpose(eigvecs),
     +       (1.d0/dsqrt(2.d0))*dp*Ee_tr_0/mag_Ee_tr_0), eigvecs)
```

```
      Fp_tau = matmul(matmul(
     +  matmul(eigvecs, reshape((/
     +  dexp(stretches(1,1)*dt), z, z,
     +  z, dexp(stretches(2,2)*dt), z,
     +  z, z, dexp(stretches(3,3)*dt) /), (/3, 3/))),
     +  transpose(eigvecs)), Fp_t)

      det_Fp_tau = Fp_tau(1,1)*(Fp_tau(2,2)*Fp_tau(3,3)-
     +  Fp_tau(3,2)*Fp_tau(2,3))-Fp_tau(2,1)*(Fp_tau(1,2)*Fp_tau(3,3)
     +  -Fp_tau(3,2)*Fp_tau(1,3)) +Fp_tau(3,1)*(Fp_tau(1,2)*
     +  Fp_tau(2,3)-Fp_tau(2,2)*Fp_tau(1,3))

Fp_tau = Fp_tau/det_fp_tau**(1.d0/3.d0)

end if

end if

det_F = F_tau(1,1)*(F_tau(2,2)*F_tau(3,3)-
     +    F_tau(3,2)*F_tau(2,3))-F_tau(2,1)*(F_tau(1,2)*F_tau(3,3)
     +    -F_tau(3,2)*F_tau(1,3)) +F_tau(3,1)*(F_tau(1,2)*
     +    F_tau(2,3)-F_tau(2,2)*F_tau(1,3))

      T_tau = matmul(Re_tr, matmul(Te_0+I_1*Te_sph,
     +           transpose(Re_tr)))/det_F

      state_new(km,10:18) =  reshape(Te_0, (/9/))   ! M_0
      state_new(km,19) = Te_sph    ! Sph Mandel
      state_new(km,1:9) = reshape(Fp_tau, (/9/))
      state_new(km, 20:28) = reshape(eigvecs, (/9/))  ! Evecs M
      state_new(km, 29) = dp ! New shearing rate
      state_new(km, 30) = -tau/Te_sph   ! mu
      state_new(km,33) = dlog(dp+3.4d-4)

      end if ! if(total_time.eq.0 .and. step_time.eq.0)

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
         ! Rotate Cauchy stress to ABAQUS stress

        call MATINV(U_tau, U_inv, det_U)
```

```
      R_tau = matmul(F_tau, U_inv)
      T_tau = matmul(transpose(R_tau),matmul(T_tau,R_tau))

       ! Update ABAQUS stresses

      do i = 1,NDIR

          STRESS_NEW(KM,i) = T_tau(i,i)

      end do

      if (NSHR .ne. 0) then

        STRESS_NEW(KM,NDIR+1) = T_tau(1,2)

        if (NSHR .ne. 1) then

          STRESS_NEW(KM,NDIR+2) = T_tau(2,3)

          if (NSHR .ne. 2) then

            STRESS_NEW(KM,NDIR+3) = T_tau(1,3)

          endif
        endif
      endif

      ! Specific internal energy

      stress_power = ONE_HALF * (
     +       ( stress_old(km,1)+stress_new(km,1) )*strain_inc(km,1) +
     +       ( stress_old(km,2)+stress_new(km,2) )*strain_inc(km,2) +
     +       ( stress_old(km,3)+stress_new(km,3) )*strain_inc(km,3))

        if(nshr .eq. 1) then

       stress_power = stress_power + one_half*(
     +    TWO*(stress_old(km,4)+stress_new(km,4))*strain_inc(km,4) )

        else
```

```
    stress_power = stress_power + one_half*(
+    TWO*(stress_old(km,4)+stress_new(km,4))*strain_inc(km,4) +
+    TWO*(stress_old(km,5)+stress_new(km,5))*strain_inc(km,5) +
+    TWO*(stress_old(km,6)+stress_new(km,6))*strain_inc(km,6) )

     endif

     ener_intern_new(km) = ener_intern_old(km)
+                         + stress_power/density(km)

 end do  ! loop over km

 return

 end
```

Bibliography

[1] cond-mat/0503682.

[2] L. Anand. 2.072 Notes on Continuum Mechanics, 2006.

[3] L. Anand and C. Gu. Granular materials: constitutive equations and strain localization. *J. Mech. Phys. Solids*, 28:1701, 2000.

[4] L. Anand and C. Su. A theory for amorphous viscoplastic materials undergoing finite deformations, with application to metallic glasses. *J. Mech. Phys. Solids*, 53:1362–1396, 2005.

[5] C. Ancey. Dry granular flows down an inclined channel: Experimental investigations on the frictional-collisional regime. *Phys. Rev. E*, 65:011304, 2002.

[6] I. S. Aranson, F. Malloggi, and E. Clement. Transverse instability of avalanches in granular flows down an incline. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 73(5):050302, 2006.

[7] I. S. Aranson and L. S. Tsimring. Continuum description of avalanches in granular media. *Phys. Rev. E*, 64:020301, 2001.

[8] I. S. Aranson and L. S. Tsimring. Continuum theory of partially fluidized granular flows. *Phys. Rev. E*, 65:061303, 2002.

[9] I. S. Aranson and L. S. Tsimring. Patterns and collective behavior in granular media: Theoretical concepts. *Rev. Mod. Phys.*, 78:641–692, 2006.

[10] T. Aslam. A partial differential equation approach to multidimensional extrapolation. *J. Comput. Phys.*, 193:349–355, 2003.

[11] E. Azanza. *Ecoulements granulaires bidimensionnels sur un plan incliné*. PhD thesis, Ecole des Ponts et Chaussées, Marne la Vallée, France, 1997.

[12] R. A. Bagnold. Experiments on a gravity free dispersion of large solid spheres in a newtonian fluid under shear. *Proc. Roy. Soc. London Ser. A*, 225, 1954.

[13] N. Balmforth and A. Provenzale. Patterns of dirt. *Geomorph. Fluid Mech.*, 582:164–187, 2001.

[14] G. W. Baxter, R. P. Behringer, T. Fagert, and G. A. Johnson. Pattern formation in flowing sand. *Phys. Rev. Lett.*, 62:2825, 1989.

[15] M. Z. Bazant. The spot model for random-packing dynamics. *Mechanics of Materials*, 38:717–731, 2006.

[16] R. P. Behringer. Contact force measurements and stress-induced anisotropy in granular materials. *Nature*, 435:1079–1082, 2005.

[17] G. Berton, R. Delannay, P. Richard, N. Taberlet, and A. Valance. Two-dimensional inclined chute flows: Transverse motion and segregation. *Phys. Rev. E*, 68:051303, 2003.

[18] E. C. Bingham. *Fluidity and Plasticity.* McGraw-Hill, 1922.

[19] L. Bocquet, W. Losert, D. Schalk, T. C. Lubensky, and J. P. Gollub. Granular shear flow dynamics and forces: Experiment and continuum theory. *Phys. Rev. E*, 65:011307, 2002.

[20] J.-P. Bouchaud, M. E. Cates, and P. Claudin. Stress distribution in granular media and nonlinear wave equation. *J. Phys.*, 5:639–656, 1995.

[21] J.-P. Bouchaud, M. E. Cates, J. R. Prakash, and S. F. Edwards. A model for the dynamics of sandpile surfaces. *J. Phys. I (France)*, 4:1383, 1994.

[22] J.-P. Bouchaud, M. E. Cates, J. R. Prakash, and S. F. Edwards. Hysteresis and metastability in a continuum sandpile model. *Phys. Rev. Lett.*, 74:1982, 1995.

[23] T. Boutreux, H. A. Makse, and P.-G. de Gennes. Surface flows of granular mixtures. *Euro. Phys. J. B*, 9:105–115, 1999.

[24] T. Boutreux, E. Raphaël, and P.-G. de Gennes. Surface flows of granular materials: A modified picture for thick avalanches. *Phys. Rev. E*, 58:4692–4700, 1998.

[25] C. S. Campbell. Rapid granular flows. *Ann. Rev. Fluid Mech.*, 22:57, 1990.

[26] C. S. Campbell, P. W. Cleary, and M. Hopkins. Large scale landslide simulations: global deformation, velocities and basal friction. *J. Geophys. Rev.*, 100:8267, 1995.

[27] G. Chambon, J. Schmittbuhl, A. Corfdir, J. P. Vilotte, and S. Roux. Shear with comminution of a granular material: Microscopic deformations outside the shear band. *Phys. Rev. E*, 68:011304, 2003.

[28] J. Choi, A. Kudrolli, and M. Z. Bazant. Velocity profile of gravity-driven dense granular flow. *J. Phys.: Condensed Matter*, 17:S2533–S2548, 2005.

[29] J. Choi, A. Kudrolli, R. R. Rosales, and M. Z. Bazant. Diffusion and mixing in gravity driven dense granular flows. *Phys. Rev. Lett.*, 92:174301, 2004.

[30] J. Christoffersen, M. M. Mehrabadi, and S. Nemat-Nasser. A micromechanical description of granular material. *J. Appl. Mech.*, 48:339–344, 1981.

[31] B. D. Coleman and W. Noll. The thermodynamics of elastic materials with heat conduction and viscosity. *Arch. Ration. Mech. Anal.*, 13:167–178, 1963.

[32] F. da Cruz. *Friction and jamming in granular flows.* PhD thesis, Ecole Nationale des Ponts et Chaussees, Marne a la vallée, France, 2004.

[33] F. da Cruz, S. Emam, M. Prochnow, J. Roux, and F. Chevoir. Rheophysics of dense granular materials: Discrete simulation of plane shear flows. *Phys. Rev. E.*, 72:021309, 2005.

[34] P. G. de Gennes. Granular matter: a tentative view. *Rev. Mod. Phys.*, 71:S374–S382, 1999.

[35] M. Depken, J. B. Lechman, M. van Hecke, W. van Saarloos, and G. S. Grest. Stresses in smooth flows of dense granular media. *Europhys. Lett.*, 78:58001, 2007.

[36] M. Depken, W. van Saarloos, and M. van Hecke. Continuum approach to wide shear zones in quasistatic granular matter. *Phys. Rev. E*, 73:031302, 2006.

[37] P. J. Digby. The effective elastic modulus of porous granular rocks. *J. Appl. Mech.*, 48:803, 1981.

[38] A. Drescher. *Analytical Methods in Bin-Load Analysis.* Elsevier, 1991.

[39] A. Drescher. Some aspects of flow of granular materials in hoppers. *Phil. Trans. R. Soc. Lond. A*, 356:2649, 1998.

[40] J. Duffy and R. D. Mindlin. Stress–strain relation and vibrations of granular medium. *J. Appl. Mech.*, 24:585, 1957.

[41] J. Duran. *Sands, Powders, and Grains.* Springer-Verlag, 2000.

[42] S. F. Edwards. In Blackman and Taguena, editors, *Disorder in Condensed Matter Physics.* Oxford, 1991.

[43] S. F. Edwards and D. V. Grinev. Granular media as a physics problem. *Advances in Complex Systems*, 4:451–467, 2001. (also reprinted in Ref. [58]).

[44] D. Ertaş and T. C. Halsey. Granular gravitational collapse and chute flow. *Europhys. Lett.*, 60:931–937, 2002.

[45] P. Evesque and P. G. de Gennes. On the statics of silos. *C. R. Acad. Sci. (PARIS), Ser. II*, 326:761, 1998.

[46] M. L. Falk and J. S. Langer. Dynamics of viscoplastic deformation in amorphous solids. *Phys. Rev. E*, 57:7192–7205, 1998.

[47] D. Fenistein and M. van Hecke. Wide shear zones in granular bulk flow. *Nature*, 425:256, 2003.

[48] J. Geng, D. Howell, E. Longhi, R. P. Behringer, G. Reydellet, L. Vanel, E. Clèment, and S. Luding. Footprints in sand: The response of a granular material to local perturbations. *Phys. Rev. Lett.*, 87:035506, 2001.

[49] B. J. Glasser and I. Goldhirsch. Scale dependence, correlations, and fluctuations of stresses in rapid granular flows. *Phys. Fluids*, 13:407, 2001.

[50] R. Glowinski and P. L. Tallec. Augmented lagrangian and operator splitting method in non-linear mechanics. *SIAM Studies Appl. Math*, 1989.

[51] J. D. Goddard. Nonlinear elasticity and pressure-dependent wave speeds in granular media. *Proc. R. Soc. London, Ser. A*, 430:105, 1990.

[52] C. Goldenberg, A. P. F. Atman, P. Claudin, G. Combe, and I. Goldhirsch. Scale separation in granular packings: Stress plateaus and fluctuations. *Phys. Rev. Lett.*, 96:168001, 2006.

[53] C. Goldenberg and I. Goldhirsch. Force chains, microelasticity, and macroelasticity. *Phys. Rev. Lett.*, 89:084302, 2002.

[54] C. Goldenberg and I. Goldhirsch. Friction enhances elasticity in granular solids. *Nature*, 435:188–191, 2005.

[55] P. A. Gremaud and J. V. Matthews. On the computation of steady hopper flows i. stress determination for coulomb materials. *J. Comput. Phys.*, 166:63–83, 2001.

[56] M. E. Gurtin and L. Anand. A theory of strain-gradient plasticity for isotropic, plastically irrotational materials. part i: Small deformations. *J. Mech. Phys. Solids*, 53:1642–1649, 2005.

[57] M. E. Gurtin and L. Anand. A theory of strain-gradient plasticity for isotropic, plastically irrotational materials. part ii: Finite deformations. *J. Mech. Phys. Solids*, 21:2297–2318, 2005.

[58] T. Halsey and A. Mehta, editors. *Challenges in Granular Physics.* World Scientific, 2002.

[59] Z. Hashin. Analysis of composite-materials - A survey. *J. Appl. Mech.*, 50(3):481–505, 1983.

[60] R. Hill. *The Mathematical Theory of Plasticity.* Oxford at the Clarendon Press, 1950.

[61] R. Hill. The essential structure of constitutive laws for metal composites and polycrystals. *J. Mech. Phys. Solids*, 15:79, 1967.

[62] R. M. Horne and R. M. Nedderman. Analysis of the stress distribution in two-dimensional bins by the method of characteristics. *Powder Technol.*, 14:93, 1976.

[63] I. Iordanoff and M. M. Khonsari. Granular lubrication: toward an understanding between kinetic and fluid regime. *ASME J. Tribol.*, 126:137–145.

[64] H. M. Jaeger, S. R. Nagel, and R. P. Behringer. Granular solids, liquids, and gases. *Rev. Mod. Phys.*, 68:1259–1273, 1996.

[65] P. Jalali, W. Polashenski, T. Tynjälä, and P. Zamankhan. Particle interactions in a dense monosized granular flow. *Physica D*, 162:188, 2002.

[66] A. Jenike. Steady gravity flow of frictional-cohesive solids in converging channels. *Journ. Appl. Mech.*, 31:499, 1964.

[67] A. W. Jenike. *Utah Engineering Experiment Station, Bulletin 108, University of Utah, Salt Lake City*, 1961.

[68] Y. Jiang and M. Liu. Granular elasticity without the coulomb condition. *Phys. Rev. Lett.*, 91:144301, 2003.

[69] M. Jirásek and Z. P. Bažant. *Inelastic Analysis of Structures*. John Wiley and Sons Ltd., 2002.

[70] P. Jop, Y. Forterre, and O. Pouliquen. Crucial role of side walls for granular surface flows: consequences for the rheology. *J. Fluid, Mech.*, 541:21–50, 2005.

[71] P. Jop, Y. Forterre, and O. Pouliquen. A constitutive law for dense granular flows. *Nature*, 441:727, 2006.

[72] L. P. Kadanoff. Built upon sand: Theoretical ideas inspired by the flow of granular materials. *Rev. Mod. Phys.*, 71:435–444, 1999.

[73] K. Kamrin and M. Z. Bazant. Stochastic flow rule for granular materials. *Phys. Rev. E*, 75:041301, 2007.

[74] K. Kamrin, C. H. Rycroft, and M. Z. Bazant. The stochastic flow rule: A multi–scale model for granular plasticity. *Modelling Simul. Mater. Sci. Eng.*, 15:S449, 2007.

[75] M. Kang, R. P. Fedkiw, and X. D. Liu. A boundary condition capturing method for multiphase incompressible flow. *J. Sci. Comput.*, 15:323–360, 2000.

[76] T. S. Komatsu, S. Inagaki, N. Nakagawa, and S. Nasuno. Creep motion in a granular pile exhibiting steady surface flow. *Phys. Rev. Lett.*, 86:1757, 2001.

[77] G. Koval. *Compotement d'interface des matériaux granulaires*. PhD thesis, École Nationale des Ponts et Chaussées, 2008.

[78] E. Kröner. Allgemeine kontinuumstheorie der versetzungen und eigenspannungen. *Arch. Ration. Mech. Anal.*, 4:273–334, 1960.

[79] R. Kuwano and R. J. Jardine. On the applicability of cross-anisotropic elasticity to granular materials at very small strains. *Géotechnique*, 52:10:727–749, 2002.

[80] M. Latzel, S. Luding, H. J. Herrmann, D. W. Howell, and R. Behringer. Comparing simulation and experiment of a 2d granular couette shear device. *Euro. Phys. Journ. E.*, 11:325–333, 2003.

[81] E. H. Lee. Elastic plastic deformation at finite strain. *J. Appl. Mech.*, 36:1–6, 1969.

[82] A. Lemaître. Origin of a repose angle: Kinetics of rearrangements for granular materials. *Phys. Rev. Lett.*, 89:064303, 2002.

[83] A. Lemaître. Rearrangements and dilatency for sheared dense materials. *Phys. Rev. Lett.*, 89:195503, 2002.

[84] P. A. Lemieux and D. J. Durian. From avalanches to fluid flow: A continuous picture of grain dynamics down a heap. *Phys. Rev. Lett.*, 85:4273, 2000.

[85] J. Litwiniszyn. Statistical methods in the mechanics of granular bodies. *Rheol. Acta*, 2/3:146, 1958.

[86] J. Litwiniszyn. The model of a random walk of particles adapted to researches on problems of mechanics of loose media. *Bull. Acad. Pol. Sci.*, 11:593, 1963.

[87] G. Lois, A. Lemaitre, and J. M. Carlson. Numerical tests of constitutive laws for dense granular flows. *Phys. Rev. E.*, 72:051303, 2005.

[88] W. Losert, L. Bocquet, T. C. Lubensky, and J. P. Gollub. Particle dynamics in sheared granular matter. *Phys. Rev. Lett.*, 85:1428–1431, 2000.

[89] P. Lu and M. Zhang. Resistance properties of coal-water paste flowing in pipes. *Fuel*, 81:877–881, 2002.

[90] H. A. Makse, N. Gland, D. L. Johnson, and L. Schwartz. Granular packings: Nonlinear elasticity, sound propagation, and collective relaxation dynamics. *Phys. Rev. E*, 70:061302, 2004.

[91] C. E. Maloney and A. Lemaitre. Amorphous systems in athermal, quasistatic shear. *Phys. Rev. E*, 74:016118, 2006.

[92] A. Medina, J. A. Cordova, E. Luna, and C. Trevino. Velocity field measurements in granular gravity flow in a near 2d silo. *Physics Letters A*, 220:111–116, 1998.

[93] C. C. Mei and K. F. Liu. A bingham-plastic model for a muddy seabed under long waves. *J. Geophys. Res.*, 92:14581–15594, 1987.

[94] G. D. R. Midi. On dense granular flows. *Euro. Phys. Journ. E.*, 14:341–365, 2004.

[95] P. Mills, D. Loggia, and M. Tixier. Model for a stationary dense granular flow along an inclined wall. *Europhys. Lett.*, 45:733–738, 1999.

[96] R. D. Mindlin and H. Deresiewicz. Elastic spheres in contact under varying oblique forces. *J. Appl. Mech.*, 20:327–244, 1953.

[97] J. K. Morgan. Numerical simulations of granular shear zones using the distinct element method 2. effects of particle size distribution and interparticle friction on mechanical behavior. *J. Geophys. Res.*, 104:2721, 1999.

[98] D. E. Mueth, G. F. Debregeas, G. S. Karczmar, P. J. Eng, S. R. Nagel, and H. M. Jaeger. Signatures of granular microstructure in dense shear flows. *Nature*, 406:385–388, 2000.

[99] D. M. Mueth. Measurements of particle dynamics in slow, dense granular couette flow. *Phys. Rev. E.*, 67:011304, 2003.

[100] J. Mullins. Stochastic theory of particle flow under gravity. *J. Appl. Phys.*, 43:665, 1972.

[101] R. M. Nedderman. *Statics and Kinematics of Granular Materials*. Cambridge University Press, 1992.

[102] R. M. Nedderman and U. Tüzün. Kinematic model for the flow of granular materials. *Powder Technology*, 22:243, 1979.

[103] A. N. Norris and D. L. Johnson. Nonlinear elasticity of granular media. *J. Appl. Mech.*, 64:39, 1997.

[104] M. Ostoja-Starzewski. Scale effects in plasticity of random media: status and challenges. *Int. J. Plasticity*, 21:1119–1160, 2005.

[105] E. B. Pitman. Stress and velocity fields in two- and three-dimensional hoppers. *Powder Technol.*, 47:219, 1986.

[106] E. B. Pitman and D. G. Schaeffer. Stability of time dependent compressible granular flow in two dimensions. *Commun. Pure. Appl. Math.*, 40:421–447, 1987.

[107] T. Pöschel. Recurrent clogging and density waves in granular material flowing through a narrow pipe. *J. Phys. (France) I*, 4:499, 1994.

[108] O. Pouliquen. Scaling laws in granular flows down rough inclined planes. *Phys. Fluids*, 11:542, 1999.

[109] O. Pouliquen. Velocity correlations in dense granular flows. *Phys. Rev. Lett.*, 93:248001, 2004.

[110] O. Pouliquen, Y. Forterre, and S. L. Dizes. Slow dense granular flows as a self-induced process. *Advances in Complex Systems*, 4:441–450, 2001. (also reprinted in Ref. [58]).

[111] O. Pouliquen and R. Gutfraind. Stress fluctuations and shear zones in quasistatic granular chute flows. *Phys. Rev. E*, 53:552–561, 1996.

[112] W. Prager and D. C. Drucker. Soil mechanics and plastic analysis or limit design. *Q. Appl. Mathematics*, 10:2:157, 1952.

[113] M. Prochnow. *Ecoulemnets denses de grains secs.* PhD thesis, Ecole Nationale de Ponts et Chaussées, Marne la Vallée, France, 2002.

[114] H. Risken. *The Fokker-Planck Equation.* Springer, 1996.

[115] P. G. Rognon, M. Naaim, J. N. Roux, and F. Chevoir. *in Powders and Grains 2005.* A. A. Balkiema Publishers.

[116] N. Roquet and P. Saramito. An adaptive finite element method for bingham fluid flows around a cylinder. *Comput. Method Appl. M.*, 192:3317–3341, 2003.

[117] K. H. Roscoe, A. N. Schoefield, and A. Thurairajah. Yielding of clays in states wetter than critical. *Geotechnique*, 15:127, 1965.

[118] L. Rothenberg and N. P. Kruyt. Critical state and evolution of coordination number in simulated granular materials. *I. J. Solids Struct.*, 41:5763–5774, 2004.

[119] J. W. Rudnicki and J. R. Rice. Conditions for the localization of deformation in pressure-sensitive and dilatant materials. *J. Mech. Phys. Solids*, 23:371, 1975.

[120] C. H. Rycroft, M. Z. Bazant, G. S. Grest, and J. W. Landry. Dynamics of random packings in granular flow. *Physical Review E*, 73:051306, 2006.

[121] C. H. Rycroft, G. S. Grest, M. Z. Bazant, and J. W. Landry. Analysis of granular flow in a pebble-bed nuclear reactor. *Physical Review E*, 74:021306, 2006.

[122] C. H. Rycroft, K. Kamrin, and M. Z. Bazant. Defining and testing a granular continuum element. unpublished, 2008.

[123] A. Samadani, A. Pradhan, and A. Kudrolli. Size segregation of granular matter in silo discharges. *Phys. Rev. E*, 60:7203–7209, 1999.

[124] S. B. Savage. Analyses of slow high-concentration flows of granular materials. *J. Fluid Mech.*, 377:1, 1998.

[125] S. B. Savage and M. Sayed. Stresses developed by dry cohesionless granular-materials sheared in an annular shear cell. *J. Fluid Mech.*, 142:391, 1984.

[126] D. G. Schaeffer. Instability in the evolution equations decsribing incompressible granular flow. *J. Diff. Eq.*, 66:19–50, 1987.

[127] A. Schoefield and P. Wroth. *Critical State Soil Mechanics*. McGraw-Hill, 1968.

[128] S. Schöllmnann. Simulation of a two-dimensional shear cell. *Phys. Rev. E*, 59:889, 1998.

[129] S. Siavoshi, A. V. Orpe, and A. Kudrolli. Friction of a slider on a granular layer: nonmonotonic thickness dependence and effect of boundary conditions. *Phys. Rev. E*, 73:010301(R), 2006.

[130] L. E. Silbert, D. Ertas, G. S. Grest, T. C. Halsey, D. Levine, and S. J. Plimpton. Granular flow down an inclined plane: Bagnold scaling and rheology. *Phys. Rev. E.*, 64:051302, 2001.

[131] L. E. Silbert, J. W. Landry, and G. S. Grest. Granular flow down a rough inclined plane: transition between thin and thick piles. *Phys. Fluids*, 15:1, 2003.

[132] V. V. Sokolovskii. *Statics of Granular Materials*. Pergamon/Oxford, 1965.

[133] A. J. M. Spencer. A theory of the kinematics of ideal soils under plane strain conditions. *J. Mech. Physics*, 12:337351, 1964.

[134] G. I. Tardos, M. I. Khan, and D. G. Schaeffer. Forces on a slowly rotating, rough cylinder in a couette device containing a dry, frictional powder. *Phys. Fluids*, 10:335, 1998.

[135] P. A. Thompson and G. S. Grest. Granular flow - friction and the dilatancy transition. *Phys Rev Lett*, 67:1751–1754, 1991.

[136] C. Thornton and L. Zhang. A numerical examination of shear banding and simple shear non-coaxial flow rules. *Phil. Mag.*, 86:3425–3452, 2006.

[137] J.-C. Tsai and J. P. Gollub. Slowly sheared dense granular flows: Crystallization and nonunique final states. *Phys. Rev. E*, 70:031303, 2004.

[138] J.-C. Tsai and J. P. Gollub. Granular packings sheared in an annular channel: Flow localization and grain size dependence. *Phys. Rev. E*, 72:051304, 2005.

[139] J.-C. Tsai, G. A. Voth, and J. P. Gollub. Internal granular dynamics, shear-induced crystallization, and compaction steps. *Phys. Rev. Lett.*, 91:064301, 2003.

[140] U. Tüzün and R. M. Nedderman. Experimental evidence supporting the kinematic modelling of the flow of granular media in the absence of air drag. *Powder Technolgy*, 23:257, 1979.

[141] D. Volfson, L. S. Tsimring, and I. S. Aranson. Partially fluidized shear granular flows: Continuum theory and molecular dynamics simulations. *Physical Review E*, 68:021301, 2003.

[142] K. Walton. The effective elastic moduli of a random packing of spheres. *J. Mech. Phys. Solids*, 35:213, 1987.

[143] C.-C. Wang. A new representation theorem for isotropic functions: An answer to professor g. f. smith's criticism of my papers on representations for isotropic functions. *Arch. Ration. Mech. Anal.*, 36:166–197, 1970.

[144] M. Yuhi and C. C. Mei. Slow spreading of fluid mud over a conical surface. *J. Fluid Mech.*, 519:337–358, 2004.