Research

Author for correspondence:
T. Barker
e-mail: thomas.barker@manchester.ac.uk

Well-posed continuum equations for granular flow with compressibility and $\mu(I)$-rheology

T. Barker[1], D. G. Schaeffer[2], M. Shearer[3] and J. M. N. T. Gray[1]

[1]School of Mathematics and Manchester Centre for Nonlinear Dynamics, University of Manchester, Oxford Road, Manchester M13 9PL, UK
[2]Mathematics Department, Duke University, Box 90320, Durham, NC 27708-0320, USA
[3]Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205, USA

TB, 0000-0001-9403-0435

Continuum modelling of granular flow has been plagued with the issue of ill-posed dynamic equations for a long time. Equations for incompressible, two-dimensional flow based on the Coulomb friction law are ill-posed regardless of the deformation, whereas the rate-dependent $\mu(I)$-rheology is ill-posed when the non-dimensional inertial number I is too high or too low. Here, incorporating ideas from critical-state soil mechanics, we derive conditions for well-posedness of partial differential equations that combine compressibility with I-dependent rheology. When the I-dependence comes from a specific friction coefficient $\mu(I)$, our results show that, with compressibility, the equations are well-posed for all deformation rates provided that $\mu(I)$ satisfies certain minimal, physically natural, inequalities.

1. Introduction

Much effort has been devoted to formulating constitutive laws for continuum models of granular materials [1–5]. However, the lack of acceptable dynamic theories, i.e. well-posed equations in the sense of Joseph & Saut [6],

for granular flow has severely hampered progress in modelling many geophysical and industrial problems. In the simplest class of models, flow is described by partial differential equations (PDEs) for the density, the velocity vector and the stress tensor; conceptually, such models are hardly more complicated than the Navier–Stokes equations. The equations represent conservation laws for mass and momentum coupled to constitutive equations to close the system. However, despite the appeal of their simplicity, they have been plagued with ill-posedness, i.e. small perturbations grow at an unbounded rate in the limit that their wavelength tends to zero [6]. Such behaviour is clearly unphysical. However, the immediate practical implication of ill-posedness is that numerical computations either blow-up, even at finite resolution, or do not converge to a well-defined solution as the grid is refined, i.e. the numerical results are grid dependent [7–10].

The first model of this type [2,11,12] specifies constitutive laws that represent a tensorial generalization of the work of de Coulomb [13] on earthwork fortifications. In the language of plasticity theory, it is a rate-independent, rigid/perfectly plastic model with a yield condition based on friction between the grains. However, it was shown to be ill-posed in all two-dimensional contexts and all realistic three-dimensional contexts [2]. Critical-state soil mechanics (CSSM) [1] is a sophisticated elaboration of Coulomb behaviour that allows for compressibility. It also suffers from ill-posedness, depending on the degree of consolidation. This ill-posedness is much less severe than for a Coulomb material [3,12], but is still physically unsatisfactory and introduces potential issues for the numerical simulation of transient granular flows. More recently, the $\mu(I)$-rheology [4,5,14] introduces a modest amount of rate dependence into (incompressible) Coulomb behaviour through the non-dimensional *inertial number*, which is proportional to the shear rate and inversely proportional to the square root of the pressure. As shown in Barker *et al.* [9], this theory leads to well-posed (two dimensional) equations in a significant region of state space, but it is ill-posed at both low and high inertial numbers.

This paper presents an analysis of constitutive equations that extend the incompressible $\mu(I)$-rheology of Jop *et al.* [5] to compressible deformations, through combination with CSSM. The main result is that, in two dimensions, compressible I-dependent equations can be made well-posed for all densities, for all stress states and for all deformation rates. In other words, to obtain well-posedness, Coulomb behaviour is modified by including only two natural, fairly small, perturbations of the theory, namely compressibility and rate dependence. Following this very general treatment, which has implications for many existing formulations [15–17], we elucidate our findings with an illustrative model that includes physically motivated features and reduces to the $\mu(I)$-rheology in the incompressible limit. This has the advantage that it retains the conceptual simplicity of the original theory. Although we consider only two-dimensional flow, it should be noted that in numerous cases it has been found that flow in two dimensions is more prone to ill-posedness than in three [2,3,18]. Thus, we anticipate that the corresponding three-dimensional equations including these effects will also be well-posed.

Currently, a wide range of new constitutive laws for granular materials are being developed including the $\mu(I)$-rheology [4,5], elasto-plastic formulations [19,20], non-local rheologies [21–24], kinetic theory [25], as well as Cosserat [26], micro-structural [27] and hypoplastic theories [28]. Enormous progress has been made over the past decade and there is the realistic and exciting prospect that practical granular flows, which span the solid-like, liquid-like and gaseous regimes, may shortly be described by continuum models. In this paper, we seek to understand one of the conceptually simplest formulations that leads to well-posed equations.

In §2, we introduce the equations to be studied and formulate our well-posedness result for them. This theorem is proved in §§3 and 4. In §5, we solve the new equations for steady, uniform chute flow. In two appendices, we summarize key ideas from CSSM and survey topics regarding ill-posed partial differential equations.

2. Governing equations

Dense granular flow is described by the solids volume fraction ϕ, the velocity vector $\boldsymbol{u}$ and the stress tensor $\boldsymbol{\sigma}$. In two dimensions, this constitutes six scalar unknowns that are spatially

and temporally dependent. These are governed by conservation laws plus constitutive relations. Conservation of mass gives the scalar equation

$$(\partial_t + u_j\partial_j)\phi + \phi\,\mathrm{div}\,\boldsymbol{u} = 0, \tag{2.1}$$

and conservation of momentum gives the vector equation

$$\rho_*\phi(\partial_t + u_j\partial_j)u_i = \partial_j\sigma_{ij} + \rho_*\phi g_i, \tag{2.2}$$

where ρ_* is the constant intrinsic grain density and g is the acceleration due to gravity. Closure of these equations is achieved through three constitutive relations.

(a) The Coulomb constitutive model

For a Coulomb material, which is assumed to be incompressible, the first constitutive relation states that ϕ is a constant. This then reduces (2.1) to the

$$\text{Flow rule: } \mathrm{div}\,\boldsymbol{u} = 0. \tag{2.3}$$

For the next constitutive relation, the stress tensor

$$\sigma_{ij} = -p\delta_{ij} + \tau_{ij} \tag{2.4}$$

is decomposed into a pressure term (where $p = -\sigma_{ii}/2$) plus a trace-free tensor $\boldsymbol{\tau}$ called the deviatoric stress. The second relation is then the

$$\text{Yield condition: } \|\boldsymbol{\tau}\| = \mu p, \tag{2.5}$$

where μ is a constant and for any tensor $\boldsymbol{T}$ the norm is defined by

$$\|\boldsymbol{T}\| = \sqrt{\frac{T_{ij}T_{ij}}{2}}. \tag{2.6}$$

This yield condition expresses the idea that a granular material cannot deform unless the shear stress is sufficient to overcome friction.[1] The third constitutive relation requires that the eigenvectors of the deviatoric stress tensor and the deviatoric strain-rate tensor[2]

$$D_{ij} = \tfrac{1}{2}(\partial_j u_i + \partial_i u_j) - \tfrac{1}{2}(\mathrm{div}\,\boldsymbol{u})\delta_{ij} \tag{2.7}$$

are aligned (see figure 1 for motivation), which may be written

$$\text{Alignment: } \frac{D_{ij}}{\|\boldsymbol{D}\|} = \frac{\tau_{ij}}{\|\boldsymbol{\tau}\|}. \tag{2.8}$$

In words, the above equation may be interpreted as asserting that in the space of trace-free symmetric 2×2 matrices, *which is two-dimensional*, $\boldsymbol{D}$ and $\boldsymbol{\tau}$ are parallel. Thus, this matrix equation entails only one scalar relation. For reference below we record that

$$\boldsymbol{D} = \frac{1}{2}\begin{bmatrix} \partial_1 u_1 - \partial_2 u_2 & \partial_1 u_2 + \partial_2 u_1 \\ \partial_1 u_2 + \partial_2 u_1 & \partial_2 u_2 - \partial_1 u_1 \end{bmatrix}. \tag{2.9}$$

---

[1]Thus, (2.5) contains the implicit assumption that material is actually deforming. Otherwise, (2.5) must be replaced by inequality, $\|\boldsymbol{\tau}\| \leq \mu p$, and the governing equations are underdetermined unless further relations, such as those of elasticity, are included.

[2]Note that, for incompressible flow, the full strain-rate tensor $(\partial_j u_i + \partial_i u_j)/2$ and the deviatoric strain-rate tensor are equal as the second term on the right in (2.7) vanishes.
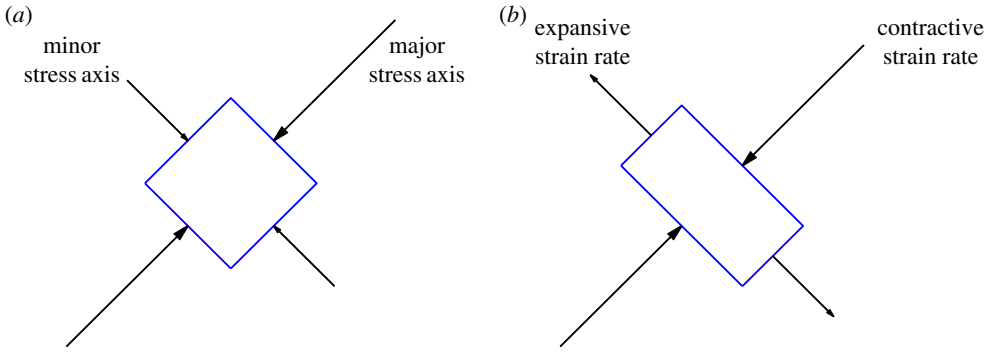
Figure 1. (*a*) Illustrative stress eigenvectors; along the major axis the stress eigenvalue is $-(p + \|\boldsymbol{\tau}\|)$, with the minus sign indicating compression. (*b*) A possible material deformation that is consistent with the stress field in (*a*). (Online version in colour.)

It is customary [2,5] to process these equations by expressing the deviatoric stress $\boldsymbol{\tau}$ in terms of p and the strain rate as follows:

$$\tau_{ij} = \|\boldsymbol{\tau}\| \frac{\tau_{ij}}{\|\boldsymbol{\tau}\|} = \mu p \frac{D_{ij}}{\|\boldsymbol{D}\|}, \tag{2.10}$$

where we have invoked (2.5) and (2.8). We may substitute (2.10) into (2.2) to obtain

$$\rho_* \phi (\partial_t + u_j \partial_j) u_i = \partial_j \left[\frac{\mu p}{\|\boldsymbol{D}\|} D_{ij} \right] - \partial_i p + \rho_* \phi g_i \tag{2.11}$$

and the resulting equation, together with (2.3), gives three equations for pressure p and velocity $\boldsymbol{u}$. In form, at least, these equations resemble the incompressible Navier–Stokes equation. However, in two dimensions (as considered here) they are always ill-posed [2].

(b) Incompressible $\mu(I)$-rheology

Work described by the Groupement De Recherche Milieux Divisés [4] has significantly improved the Coulomb model by including some rate dependence (in the sense of plasticity [29]) in the yield condition while making no changes in the incompressible flow rule (2.3) and the alignment condition (2.8). Specifically, a wide range of experiments are captured by replacing the constant μ in (2.5) by an increasing function $\mu(I)$ of the *inertial number*,

$$I = \frac{2d\|\mathbf{D}\|}{\sqrt{p/\rho_*}}, \tag{2.12}$$

where d is the particle diameter. The expression

$$\mu(I) = \mu_1 + \frac{\mu_2 - \mu_1}{I_0/I + 1}, \tag{2.13}$$

where μ_1, μ_2 and I_0 are constants with $\mu_2 > \mu_1$, is a frequently used form [30]. Below we shall assume that

$$\mu'(I) > 0 \quad \text{and} \quad \mu''(I) < 0. \tag{2.14}$$

The modified yield condition changes (2.11) to read

$$\rho_* \phi (\partial_t + u_j \partial_j) u_i = \partial_j \left[\frac{\mu(I)p}{\|\boldsymbol{D}\|} D_{ij} \right] - \partial_i p + \rho_* \phi g_i. \tag{2.15}$$

Table 1. Table of criteria for linear well-posedness.

model	Coulomb	incomp. $\mu(I)$	CSSM	CIDR
conditions	always ill-posed	(2.16)	always ill-posed	(2.19), (2.20)

The effect of this seemingly small perturbation is profound. Unlike for Coulomb material, equations (2.15) and (2.3) are linearly well-posed for I, μ satisfying

$$4\nu^2 - 4\nu + \mu^2\left(1 - \frac{\nu}{2}\right)^2 \le 0, \quad \text{where } \nu = \frac{I}{\mu}\frac{\mathrm{d}\mu}{\mathrm{d}I}. \tag{2.16}$$

For the specific $\mu(I)$-curve (2.13) this inequality covers a significant range of inertial numbers, specifically when the deformation rate is neither too small nor too large relative to the pressure. Outside of this range, the maximal-order linear stability analysis and numerical simulations show that perturbations grow exponentially with growth rates tending to infinity as their wavelength is reduced [9]. This behaviour is the hallmark of ill-posedness and leads to unphysical numerical solutions that strongly depend on the grid resolution used.

(c) Compressibility and *I*-dependent rheology

We refer to CSSM (see appendix A) for guidance in introducing compressibility into the rheology. Thus, we make no change in the alignment condition (2.8); we assume ϕ-dependence in the yield condition,

$$\|\boldsymbol{\tau}\| = Y(p, \phi, I), \tag{2.17}$$

and we allow for volumetric changes by introducing a new function $f(p, \phi, I)$ and modifying the flow rule to

$$\operatorname{div} \boldsymbol{u} = 2f(p, \phi, I)\|\mathbf{D}\|. \tag{2.18}$$

To get well-posed equations, our analysis (see §3) shows that the yield condition and the flow-rule functions must be related by the equation[3]

$$\frac{\partial Y}{\partial p} - \frac{I}{2p}\frac{\partial Y}{\partial I} = f + I\frac{\partial f}{\partial I} \tag{2.19}$$

and satisfy the inequalities

$$\partial_I Y > 0 \tag{2.20a}$$

and

$$\partial_p f - \frac{I}{2p}\partial_I f < 0. \tag{2.20b}$$

We may now state our main result, the well-posedness theorem for the system (2.1), (2.2), (2.8), (2.17), (2.18), which we call the CIDR equations. (Mnemonic: compressible *I*-dependent rheology.)

Theorem 2.1. *Under hypotheses (2.19) and (2.20), the CIDR system is linearly well-posed.*

The term *linearly well-posed* is defined in appendix B, and the result is proved in §§3 and 4. Table 1 contains a comparison between the conditions for linear well-posedness for the different constitutive models that have been discussed here.

Remark 2.2. The *I*-dependence in these equations need not relate to a friction coefficient $\mu(I)$. In §2e, we connect the equations to $\mu(I)$-rheology.

---

[3]If Y and f are independent of I, then (2.19) leads to the CSSM flow rule (A 2) derived from normality.

(d) Derivation of evolution equations

To place the equations in a larger continuum-mechanics context, we show that the CIDR equations of motion can be rewritten as a system of three evolution equations for the velocity $\boldsymbol{u}$ and the solids fraction ϕ. In form, these equations are analogous to the Navier–Stokes equations for a viscous, compressible fluid. We make no use of this form of the equations in our proof of well-posedness.

We want to eliminate stresses from the equations of motion. To this end, we propose to solve for the mean stress p using the flow rule (2.18), which we rewrite as

$$f(p, \phi, I) = \frac{\text{div } \boldsymbol{u}}{2\|\boldsymbol{D}\|}. \tag{2.21}$$

Note that $f(p, \phi, I)$ depends on p both directly in its first argument and indirectly through $I = 2d\|\boldsymbol{D}\|/\sqrt{p/\rho_*}$ in its third argument. However,

$$\frac{\partial}{\partial p}\left[f\left(p, \phi, \frac{2d\|\boldsymbol{D}\|}{\sqrt{p/\rho_*}}\right)\right] = \partial_p f - \frac{I}{2p}\partial_I f, \tag{2.22}$$

which by assumption (2.20b) is non-zero. Thus, we may apply the implicit function theorem to (2.21) to solve $p = P(\nabla \boldsymbol{u}, \phi)$.[4] Given this, we may define

$$T(\nabla \boldsymbol{u}, \phi) = Y(P(\nabla \boldsymbol{u}, \phi), \phi, I(\nabla \boldsymbol{u}, \phi)), \quad \text{where } I(\nabla \boldsymbol{u}, \phi) = \frac{2d\|\boldsymbol{D}\|}{\sqrt{P(\nabla \boldsymbol{u}, \phi)/\rho_*}}$$

and substitute into conservation of momentum to obtain an equation

$$\rho_* \phi(\partial_t + u_j \partial_j)u_i = \partial_j\left[\frac{T(\nabla \boldsymbol{u}, \phi)}{\|\boldsymbol{D}\|}D_{ij}\right] - \partial_i[P(\nabla \boldsymbol{u}, \phi)] + \rho_* \phi g_i. \tag{2.23}$$

This equation, along with (2.1), gives a system of three evolution equations for the velocity $\boldsymbol{u}$ and the solids fraction ϕ. It is possible that previous formulations of compressible $\mu(I)$ equations [15–17] may be seen as CIDR equations with specific constitutive laws specified. In this paper, we choose to elucidate the well-posedness result with more generic choices of f and Y in order to remain impartial.

(e) Connection to $\mu(I)$-rheology

Without making any attempt to be general, we illustrate one example of how $\mu(I)$-rheology may be included in constitutive relations of the form (2.17), (2.18). Motivated by equation (A 3) in appendix A, we make the ansatz

$$Y(p, \phi, I) = \alpha(I)p - \frac{p^2}{C(\phi)} \tag{2.24a}$$

and

$$f(p, \phi, I) = \beta(I) - \frac{2p}{C(\phi)}. \tag{2.24b}$$

In these equations, it is worth emphasizing that p, ϕ, I are treated as independent variables, not to be confused with the dependence of I on p in the previous subsection. The function $C(\phi)$ is an increasing function of ϕ. As ϕ varies (with I fixed) the yield loci $\|\boldsymbol{\tau}\| = Y(p, \phi, I)$ derived from (2.24a) form a nested family of convex curves in stress space (figure 2b). Observe from (2.18) that

---

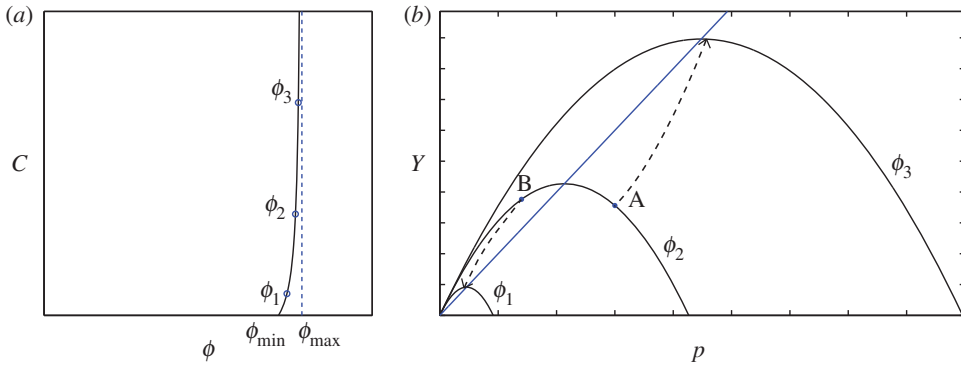[4]Note that P in fact depends only on div $\boldsymbol{u}$, $\|\boldsymbol{D}\|$ and ϕ.

Figure 2. (*a*) An example curve for the function $C(\phi)$ with a minimum solids volume fraction ϕ_{min} and a vertical asymptote at ϕ_{max}. (*b*) Nested yield surfaces of the form (2.24) for a fixed value of I with differing solids volume fractions. (The solid blue line, the dashed arrows and the labels A and B refer to a discussion of CSSM in appendix A.)

deformation without volumetric strain is possible if $f(p,\phi,I) = 0$; i.e. for (2.24b), if $p/C(\phi) = \beta(I)/2$. Substituting this formula into (2.17) and using (2.24a), we derive

$$\|\boldsymbol{\tau}\| = \left[\alpha(I) - \frac{\beta(I)}{2}\right]p,$$

for such isochoric deformation to be possible. Thus, to recover the yield condition $\|\boldsymbol{\tau}\| = \mu(I)p$ of the $\mu(I)$-rheology, let us require that

$$\alpha(I) - \frac{\beta(I)}{2} = \mu(I). \tag{2.25}$$

Lemma 2.3. *Equations (2.19) and (2.25) imply that*

$$\alpha(I) = \frac{4}{5}\mu(I) + \frac{12}{25}I^{-2/5}\int_0^I J^{-3/5}\mu(J)\,\mathrm{d}J \tag{2.26}$$

and

$$\beta(I) = -\frac{2}{5}\mu(I) + \frac{24}{25}I^{-2/5}\int_0^I J^{-3/5}\mu(J)\,\mathrm{d}J. \tag{2.27}$$

Proof. Substituting the relations (2.24) into (2.19), and using (2.25) to eliminate β, we derive the linear ordinary differential equation for $\alpha = \alpha(I)$:

$$\tfrac{5}{2}I\alpha'(I) + \alpha(I) = 2\mu(I) + 2I\mu'(I). \tag{2.28}$$

Solving this linear equation for $\alpha(I)$, with an integrating factor, we obtain

$$I^{2/5}\alpha(I) = \frac{4}{5}\int_0^I J^{-3/5}\mu(J)\,\mathrm{d}J + \frac{4}{5}\int_0^I J^{2/5}\mu'(J)\,\mathrm{d}J,$$

from which the formula (2.26) follows after integrating the second integral by parts. Finally, substituting this formula for $\alpha(I)$ into (2.25), we obtain the formula (2.27) for $\beta(I)$. ∎

Lemma 2.4. *The yield condition and flow-rule function (2.24a,b) that follow from (2.26), (2.27) verify hypotheses (2.19) and (2.20), provided $\mu(I)$ satisfies (2.14).*

Proof. Of course (2.19) is satisfied because this equation was imposed in deriving (2.26), (2.27).

Differentiating (2.24b), we see that $\partial_p f(p, \phi, I) = -2/C(\phi) < 0$. To calculate $\partial_I f(p, \phi, I)$, we first reparametrize the integral in (2.27) to obtain $\beta(I) = -\frac{2}{5}\mu(I) + \frac{24}{25}\int_0^1 s^{-3/5}\mu(sI)\,ds$. Then

$$\partial_I f(p, \phi, I) = \beta'(I) = -\frac{2}{5}\mu'(I) + \frac{24}{25}\int_0^1 s^{2/5}\mu'(sI)\,ds.$$

By (2.14), $\mu''(I) < 0$, so $\mu'(sI) > \mu'(I)$ for $0 < s < 1$. Thus,

$$\beta'(I) > \mu'(I)\left\{-\frac{2}{5} + \frac{24}{25}\int_0^1 s^{2/5}\,ds\right\} = \mu'(I)\left\{\frac{24}{35} - \frac{2}{5}\right\} > 0,$$

the last inequality using (2.14). Consequently,

$$\partial_p f(p, \phi, I) - \frac{I}{2p}\partial_I f(p, \phi, I) < 0, \tag{2.29}$$

proving inequality (2.20b).

For inequality (2.20a), we reparametrize the integral (2.26) and differentiate to obtain

$$\partial_I Y(p, \phi, I) = \alpha'(I)p = p\left(\frac{4}{5}\mu'(I) + \frac{12}{25}\int_0^1 s^{2/5}\mu'(sI)\,ds\right) > 0,$$

as desired. ∎

(f) The incompressible limit

Based on an analogy with CSSM, let us suppose that $C(\phi)$ is a sensitive function of ϕ, say of the form

$$C(\phi) = \frac{\rho_\star g d}{b}\check{C}\left(\frac{\phi - \phi_{\min}}{\phi_{\max} - \phi_{\min}}\right), \tag{2.30}$$

where g is the constant of gravitational acceleration, b is a non-dimensional parameter and the factor $\rho_\star g d$ gives C the dimensions of pressure. The non-dimensional function $\check{C}$ has an argument that is dependent on the minimum solids fraction $\phi_{\min}$ for sustained stress transmission between grains (random loose packing[5]) and $\phi_{\max}$, which is the maximum packing fraction that can be attained. Typically, $\Delta\phi = \phi_{\max} - \phi_{\min}$ is small. For definiteness we may take

$$\check{C}(y) = \frac{y}{1-y}, \tag{2.31}$$

as in figure 2. Note that $C(\phi)$ diverges as $\phi \to \phi_{\max}$; thus, (2.30) requires that ϕ is confined to a narrow range,

$$\phi_{\min} \leq \phi < \phi_{\max}. \tag{2.32}$$

In physical terms, the maximum solids fraction $\phi_{\max}$ represents the jamming threshold. We call the limit $\Delta\phi \to 0$ *incompressible* because, as may be seen from (2.32), the density of the material becomes essentially constant.

Lemma 2.5. *As $\Delta\phi \to 0$, the CIDR equations reduce to the equations of incompressible $\mu(I)$-rheology,* (2.3), (2.15).

Proof. We process the CIDR equations, which have the six unknowns ϕ, u_i and σ_{ij}, as follows. First, we reduce to five unknowns—ϕ, u_i, p and $\tau = \|\tau\|$—by recalling the definition (2.4) and the

---

[5]Note that we are assuming isotropy, which may be a questionable assumption at such low densities.

alignment condition (2.8) to write

$$\sigma_{ij} = -p\delta_{ij} + \tau \frac{D_{ij}}{\|\boldsymbol{D}\|}.$$

Next we use the yield condition to eliminate ϕ, reducing this number to four. Specifically, substituting (2.24a) into (2.17), we write the yield condition

$$\tau = \alpha(I)p - \frac{p^2}{C(\phi)}. \tag{2.33}$$

Solving (2.33) for ϕ we obtain

$$\phi = \Phi(\nabla\boldsymbol{u}, p, \tau) = C^{-1}\left(\frac{p^2}{\alpha(I)p - \tau}\right), \tag{2.34}$$

where the dependence on $\nabla\boldsymbol{u}$ comes from the fact that $I = 2d\|\boldsymbol{D}\|/\sqrt{p/\rho_*}$. Substitution of this formula into the conservation laws (2.1), (2.2) yields the equations

$$(\partial_t + u_j\partial_j)\Phi(\nabla\boldsymbol{u}, p, \tau) + \Phi(\nabla\boldsymbol{u}, p, \tau)\,\mathrm{div}\,\boldsymbol{u} = 0 \tag{2.35a}$$

and

$$\rho_*\Phi(\nabla\boldsymbol{u}, p, \tau)(\partial_t + u_j\partial_j)u_i = \partial_j\left[\frac{\tau}{\|\boldsymbol{D}\|}D_{ij}\right] - \partial_i p + \rho_*\Phi g_i. \tag{2.35b}$$

Finally, we show that the flow rule (2.18) may be rewritten

$$\mathrm{div}\,\boldsymbol{u} = 4\left[\frac{\tau}{p} - \mu(I)\right]\|\mathbf{D}\|. \tag{2.36}$$

To see this, we combine (2.24b) with (2.25) to conclude

$$f(p, \phi, I) = \beta(I) - \frac{2p}{C(\phi)} = 2[\alpha(I) - \mu(I)] - \frac{2p}{C(\phi)}$$

and then substitute the relation $\alpha(I) = \tau/p + p/C(\phi)$ derived by manipulating (2.33). Thus, the system (2.35), (2.36) governs the evolution of the four unknowns u_i, p and τ.

Now we claim that if $C(\phi)$ has the form (2.30), then (2.35), (2.36) is a singular perturbation of (2.3), (2.15). It follows from (2.30) that (2.34) has the expansion

$$\phi = \phi_0 + \Delta\phi\check{\Phi}(\nabla\boldsymbol{u}, p, \tau), \quad \text{where } \check{\Phi}(\nabla\boldsymbol{u}, p, \tau) = \check{C}^{-1}\left(\frac{bp^2}{\rho_*gd[\alpha(I)p - \tau]}\right). \tag{2.37}$$

Substituting (2.37) into the continuity equation (2.35a), we find

$$\Delta\phi(\partial_t + u_j\partial_j)\check{\Phi}(\nabla\boldsymbol{u}, p, \tau) + [\phi_0 + \Delta\phi\check{\Phi}(\nabla\boldsymbol{u}, p, \tau)]\,\mathrm{div}\,\boldsymbol{u} = 0.$$

If $\Delta\phi = 0$, then this equation reduces to $\mathrm{div}\,\boldsymbol{u} = 0$, although this is of course a highly singular limit. Thus, if $\Delta\phi = 0$, the left-hand side of (2.36) vanishes, so this equation simplifies to the yield condition $\tau = \mu(I)p$, and substitution into (2.35b) yields (2.15). This proves the lemma. ∎

3. Proofs, part I: linearization

(a) An alternative formulation of the alignment condition

It is convenient to study the linearized equations with a reformulated alignment condition that describes stress in terms of eigenvectors of, rather than entries of, the stress tensor. Since $\boldsymbol{\tau}$ defined by (2.4) has trace zero, it has eigenvalues[6] $\pm\|\boldsymbol{\tau}\|$. Taking ψ as the angle that the eigenvector with

---

[6] Hence $\boldsymbol{\sigma}$ has eigenvalues $-p \pm \|\boldsymbol{\tau}\|$. Note that $-p - \|\boldsymbol{\tau}\|$ is the major stress eigenvalue—although this eigenvalue is the smaller algebraically, it is the larger in absolute value.

eigenvalue $-\|\boldsymbol{\tau}\|$ makes with the x_1-axis gives

$$\boldsymbol{\tau} = -\|\boldsymbol{\tau}\| \begin{bmatrix} \cos 2\psi & \sin 2\psi \\ \sin 2\psi & -\cos 2\psi \end{bmatrix}, \tag{3.1}$$

which may be verified by checking that $(\cos\psi, \sin\psi)$ is an eigenvector of this matrix with eigenvalue $-\|\boldsymbol{\tau}\|$. Thus, the stress tensor σ_{ij} is completely specified by the three scalars p, $\|\boldsymbol{\tau}\|$ and ψ.

Focusing on the first rows of the strain-rate tensor (2.9) and of (3.1), we extract from the matrix equation (2.8) the vector equation

$$(\partial_1 u_1 - \partial_2 u_2, \partial_1 u_2 + \partial_2 u_1) = k(\cos 2\psi, \sin 2\psi), \tag{3.2}$$

where $k = -2\|\boldsymbol{D}\| < 0$. Since $\boldsymbol{D}$ and $\boldsymbol{\tau}$ lie in the two-dimensional space of trace-free, symmetric matrices, (3.2) is equivalent to (2.8). It follows from (3.2) that

$$\text{Alt. alignment:} \quad (\partial_1 u_2 + \partial_2 u_1)\cos 2\psi - (\partial_1 u_1 - \partial_2 u_2)\sin 2\psi = 0. \tag{3.3}$$

In point of fact, this equation is slightly weaker than the alignment condition since (3.3) is consistent with the possibility that $k > 0$ in (3.2); to rule out the latter possibility we impose the supplemental inequality[7] that

$$(\partial_1 u_1 - \partial_2 u_2)\cos 2\psi \le 0. \tag{3.4}$$

(b) The calculation

Substitution of the stress tensor (3.1) into the momentum balance equations (2.2) allows for the full set of equations to be written as

$$\rho_* \phi(\partial_t + u_1 \partial_1 + u_2 \partial_2)u_1 + \partial_1[p + \tau \cos(2\psi)] + \partial_2[\tau \sin(2\psi)] = \rho_* \phi g_1, \tag{3.5a}$$

$$\rho_* \phi(\partial_t + u_1 \partial_1 + u_2 \partial_2)u_2 + \partial_1[\tau \sin(2\psi)] + \partial_2[p - \tau \cos(2\psi)] = \rho_* \phi g_2, \tag{3.5b}$$

$$(\partial_t + u_1 \partial_1 + u_2 \partial_2)\phi + \phi(\partial_1 u_1 + \partial_2 u_2) = 0, \tag{3.5c}$$

$$\partial_1 u_1 + \partial_2 u_2 = 2f\|\boldsymbol{D}\| \tag{3.5d}$$

and $\qquad (\partial_2 u_1 + \partial_1 u_2)\cos(2\psi) + (\partial_2 u_2 - \partial_1 u_1)\sin(2\psi) = 0. \tag{3.5e}$

This system has five scalar unknowns, $\boldsymbol{U} = (u_1, u_2, \phi, p, \psi)$. In (3.5a), (3.5b), τ is a mnemonically suggestive abbreviation for the yield function $Y(p, \phi, I)$ in (2.17), and in (3.5d), a repetition of (2.18), the function f depends on arguments (p, ϕ, I) that are not written explicitly.

As in appendix B, to linearize the equations we substitute a perturbation of a base solution $\boldsymbol{U}^{(0)}(\boldsymbol{x}, t)$, say

$$\boldsymbol{U} = \boldsymbol{U}^{(0)} + \hat{\boldsymbol{U}}, \tag{3.6}$$

into the equations, retain only terms that are linear in the perturbation $\hat{\boldsymbol{U}}$ and freeze the coefficients at an arbitrary point $(\boldsymbol{x}^*, t^*)$. It is convenient to temporarily drop most terms not of maximal order and estimate their effect in a calculation at the end of the argument. For example, this construction applied to (3.5c) yields the constant-coefficient, linear equation

$$(\partial_t + u_1^* \partial_1 + u_2^* \partial_2)\hat{\phi} + \phi^*(\partial_1 \hat{u}_1 + \partial_2 \hat{u}_2) = 0, \tag{3.7}$$

where $u_j^* = u_j^{(0)}(\boldsymbol{x}^*, t^*)$ and $\phi^* = \phi^{(0)}(\boldsymbol{x}^*, t^*)$. Lower-order terms $\partial_j \phi^* \, \hat{u}_j$ and $\partial_j u_j^* \, \hat{\phi}$ in the full linearization of (3.5c) have been dropped in (3.7).

In expanding the fully nonlinear factor $\|\boldsymbol{D}\|$ in (3.5d), we may take advantage of the rotational invariance of the equations to arrange that $\psi^* = 0$; i.e. we may calculate in a rotated coordinate

---

[7]It is also true that $(\partial_1 u_2 + \partial_2 u_1)\sin 2\psi \le 0$, and if $\cos 2\psi$ were to vanish, we would need to use this inequality to guarantee that $k < 0$. However, this issue will not arise in the analysis below.
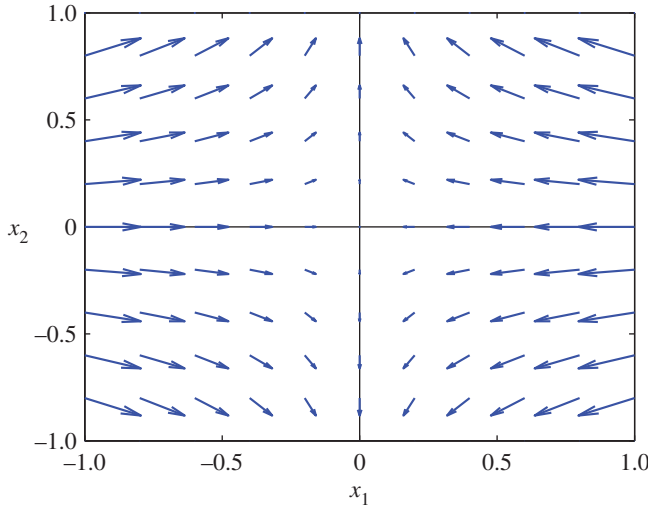
Figure 3. An example of a base-state velocity field for the strain-rate tensor (3.8) with $\partial_1 u_1^{(0)} \equiv -1$ and $\partial_2 u_2^{(0)} \equiv 1/2$. (Online version in colour.)

system for which, at (x^*, t^*), the x_1-axis is the maximal stress axis. Then by the alignment condition (3.3), the base-state deviatoric strain-rate tensor is diagonal at (x^*, t^*)

$$D^* = \begin{bmatrix} \dfrac{(\partial_1 u_1^* - \partial_2 u_2^*)}{2} & 0 \\ 0 & \dfrac{(\partial_2 u_2^* - \partial_1 u_1^*)}{2} \end{bmatrix}, \tag{3.8}$$

and by (3.4), in the 1,1-position of this matrix, $\partial_1 u_1^* - \partial_2 u_2^* < 0$. This corresponds to non-zero compression along the major stress axis, as illustrated in figure 3. Now

$$\|(D^* + \hat{D})\| = \tfrac{1}{2}[(\partial_1 u_1^* - \partial_2 u_2^* + \partial_1 \hat{u}_1 - \partial_2 \hat{u}_2)^2 + (\partial_2 \hat{u}_1 + \partial_1 \hat{u}_2)^2]^{1/2} \tag{3.9a}$$

and

$$\approx \|D^*\| - \frac{(\partial_1 \hat{u}_1 - \partial_2 \hat{u}_2)}{2}, \tag{3.9b}$$

where the approximation follows from the expansion

$$\sqrt{(-A + X)^2 + Y^2} = A - X + \mathcal{O}(X^2 + Y^2),$$

if $A > 0$ and $|X|, |Y| \ll A$. Thus, as given in table 2, the (local) linearization of $\|D\|$ equals $-(\partial_1 \hat{u}_1 - \partial_2 \hat{u}_2)/2$.

In (3.5d), the function f contains p, ϕ and I as implicit arguments. As reflected in the table, the dependence on p and ϕ contributes zeroth-order terms in these variables to the linearization.

In (3.5a), (3.5b), τ also depends on p, ϕ and I, and the terms involving τ are differentiated; hence new issues arise in linearizing them. For example, by the chain rule,

$$\partial_j[\tau \cos(2\psi)] = \cos(2\psi) \left\{ \partial_p \tau \, \partial_j p + \partial_\phi \tau \, \partial_j \phi + \partial_I \tau \left[\frac{2d}{\sqrt{p/\rho_*}} \partial_j \|D\| - \frac{d\|D\|}{\sqrt{p^3/\rho_*}} \partial_j p \right] \right\}$$

$$- 2\tau \sin(2\psi) \partial_j \psi.$$

Table 2. List of maximal-order linearizations of terms in (3.5a)–(3.5e), to assist in deriving (3.10a)–(3.10e). $\hat{D}_{11} = (\partial_1 \hat{u}_1 - \partial_2 \hat{u}_2)/2$ (only in this table).

term in (3.5a)–(3.5e)	contribution to (3.10a)–(3.10e)
$\|\boldsymbol{D}\|$	$-\hat{D}_{11}$
I	$-\dfrac{I^*}{\|\boldsymbol{D}^*\|}\hat{D}_{11} - \dfrac{I^*}{2p^*}\hat{p}$
$\partial_j[\tau \cos(2\psi)]$	$(\partial_p \tau)^* \partial_j \hat{p} + (\partial_\phi \tau)^* \partial_j \hat{\phi}$ $+ (\partial_I \tau)^* \left\{ -\dfrac{I^*}{\|\boldsymbol{D}^*\|} \partial_j \hat{D}_{11} - \dfrac{I^*}{2p^*} \partial_j \hat{p} \right\}$
$\partial_j[\tau \sin(2\psi)]$	$2\tau^* \partial_j \hat{\psi}$
$f\|\boldsymbol{D}\|$	$-f^* \hat{D}_{11} + \|\boldsymbol{D}^*\|(\partial_p f)^* \hat{p} + \|\boldsymbol{D}^*\|(\partial_\phi f)^* \hat{\phi}$ $+ \|\boldsymbol{D}^*\|(\partial_I f)^* \left\{ -\dfrac{I^*}{\|\boldsymbol{D}^*\|} \hat{D}_{11} - \dfrac{I^*}{2p^*} \hat{p} \right\}$

Since $\psi^* = 0$, the full linearization of, say, the first term here equals $(\partial_p \tau)^* \partial_j \hat{p}$, a term given in the table, plus lower-order terms

$$(\partial_j p)^* \left\{ (\partial_{pp} \tau)^* \hat{p} + (\partial_{\phi p} \tau)^* \hat{\phi} + (\partial_{Ip} \tau)^* \left[-\frac{I^*}{\|\boldsymbol{D}^*\|} \hat{D}_{11} - \frac{I^*}{2p^*} \hat{p} \right] \right\}.$$

All of these terms, as well as numerous other analogous terms in the full linearization of (3.5a) that are not of maximal order, have been dropped in (3.10a)–(3.10e).

Putting all the pieces together, we obtain the linearization[8] of the system (3.5a)–(3.5e)

$$\rho_* \phi^* d_t^* \hat{u}_1 + A(-\partial_{11} \hat{u}_1 + \partial_{12} \hat{u}_2) + (\partial_\phi \tau)^* \partial_1 \hat{\phi} + (1 + B)\partial_1 \hat{p} + 2\tau^* \partial_2 \hat{\psi} = 0, \quad (3.10a)$$

$$\rho_* \phi^* d_t^* \hat{u}_2 + A(\partial_{12} \hat{u}_1 - \partial_{22} \hat{u}_2) - (\partial_\phi \tau)^* \partial_2 \hat{\phi} + (1 - B)\partial_2 \hat{p} + 2\tau^* \partial_1 \hat{\psi} = 0, \quad (3.10b)$$

$$d_t^* \hat{\phi} + \phi^*(\partial_1 \hat{u}_1 + \partial_2 \hat{u}_2) = 0, \quad (3.10c)$$

$$(1 + C)\partial_1 \hat{u}_1 + (1 - C)\partial_2 \hat{u}_2 - 2\|\boldsymbol{D}^*\|(\partial_\phi f)^* \hat{\phi} + \Gamma \hat{p} = 0 \quad (3.10d)$$

and $\quad \partial_2 \hat{u}_1 + \partial_1 \hat{u}_2 + 4\|\boldsymbol{D}^*\|\hat{\psi} = 0, \quad (3.10e)$

where

$$d_t^* = \partial_t + u_1^* \partial_1 + u_2^* \partial_2, \quad A = \frac{I^*}{2\|\boldsymbol{D}^*\|}(\partial_I \tau)^*, \quad B = (\partial_p \tau)^* - \frac{I^*}{2p^*}(\partial_I \tau)^* \quad (3.11)$$

and

$$C = f^* + I^*(\partial_I f)^* \quad \text{and} \quad \Gamma = -2\|\boldsymbol{D}^*\| \left((\partial_p f)^* - \frac{I^*}{2p^*}(\partial_I f)^* \right). \quad (3.12)$$

Observe that, by hypothesis (2.19), $B = C$, a fact that we use in (4.3) and below.

---

[8]These equations are maximal order except that in (3.10a) and (3.10b) the term $d_t^* \hat{u}_j$ retains first-order spatial derivatives even though these equations also contain second-order derivatives of $\hat{u}_j$.

4. Proofs, part II: calculation of growth rates

(a) The eigenvalue problem

We now look for exponential solutions of (3.10a)–(3.10e),

$$\hat{U}(\mathbf{x}, t) = e^{i\langle \boldsymbol{\xi}, \mathbf{x}\rangle + \lambda t}\tilde{U}, \tag{4.1}$$

where $\tilde{U} = (\tilde{u}_1, \tilde{u}_2, \tilde{\phi}, \tilde{p}, \tilde{\psi})$ is a 5-vector of scalars, $\boldsymbol{\xi} = (\xi_1, \xi_2)$ is a vector wavenumber, $\langle , \rangle$ indicates the inner product and λ is the growth rate. The function (4.1) is a solution of (3.10a)–(3.10e) iff λ, $\tilde{U}$ satisfies the generalized eigenvalue problem

$$S\tilde{U} = -(\lambda + i\langle \boldsymbol{u}^*, \boldsymbol{\xi}\rangle)E\tilde{U}, \tag{4.2}$$

where $\boldsymbol{u}^* = (u_1^*, u_2^*)$,

$$S = \begin{bmatrix} A\xi_1^2 & -A\xi_1\xi_2 & i(\partial_\phi\tau)^*\xi_1 & (1+B)i\xi_1 & 2i\tau^*\xi_2 \\ -A\xi_1\xi_2 & A\xi_2^2 & -i(\partial_\phi\tau)^*\xi_2 & (1-B)i\xi_2 & 2i\tau^*\xi_1 \\ i\phi^*\xi_1 & i\phi^*\xi_2 & 0 & 0 & 0 \\ (1+B)i\xi_1 & (1-B)i\xi_2 & -2\|D^*\|(\partial_\phi f)^* & \Gamma & 0 \\ i\xi_2 & i\xi_1 & 0 & 0 & 4\|D^*\| \end{bmatrix} \tag{4.3}$$

and

$$E = \begin{bmatrix} \rho_*\phi^* & & & & \\ & \rho_*\phi^* & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix}. \tag{4.4}$$

On the right side of (4.2), the modified eigenvalue parameter is $\lambda + i\langle \boldsymbol{u}^*, \boldsymbol{\xi}\rangle$ because

$$d_t^* e^{i\langle \boldsymbol{\xi}, \mathbf{x}\rangle + \lambda t} = (\lambda + i\langle \boldsymbol{u}^*, \boldsymbol{\xi}\rangle) e^{i\langle \boldsymbol{\xi}, \mathbf{x}\rangle + \lambda t}.$$

Equation (4.2) is a *generalized* eigenvalue problem because E, the matrix of coefficients of time-derivative terms, is not invertible. To extract an ordinary eigenvalue problem, we decompose S into blocks

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}, \tag{4.5}$$

where

$$S_{11} = \begin{bmatrix} A\xi_1^2 & -A\xi_1\xi_2 & i(\partial_\phi\tau)^*\xi_1 \\ -A\xi_1\xi_2 & A\xi_2^2 & -i(\partial_\phi\tau)^*\xi_2 \\ i\phi^*\xi_1 & i\phi^*\xi_2 & 0 \end{bmatrix} \tag{4.6}$$

and S_{12}, S_{21} and S_{22} fill out the rest of the matrix. Defining $\tilde{U}_1 = (\tilde{u}_1, \tilde{u}_2, \tilde{\phi})$ and $\tilde{U}_2 = (\tilde{p}, \tilde{\psi})$, we rewrite (4.2) as

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}\begin{bmatrix} \tilde{U}_1 \\ \tilde{U}_2 \end{bmatrix} = -(\lambda + i\langle \boldsymbol{u}^*, \boldsymbol{\xi}\rangle)E\begin{bmatrix} \tilde{U}_1 \\ \tilde{U}_2 \end{bmatrix}. \tag{4.7}$$

The zero entries in the last two rows of E mean that $S_{21}\tilde{U}_1 + S_{22}\tilde{U}_2 = 0$ so we can solve for

$$\tilde{U}_2 = -S_{22}^{-1}S_{21}\tilde{U}_1. \tag{4.8}$$

Substitution of $\tilde{U}_2$ into (4.7) then reduces this problem[9] to the ordinary 3×3 eigenvalue problem,

$$E_{11}^{-1}[S_{11} - S_{12}S_{22}^{-1}S_{21}]\tilde{U}_1 = -(\lambda + i\langle \boldsymbol{u}^*, \boldsymbol{\xi}\rangle)\tilde{U}_1, \tag{4.9}$$

where E_{11} is the 3×3 block in the upper left of E.

---

[9]In other words, we are performing on the symbol level the reduction that we performed on the operator level in §2d.

We decompose the 3×3 matrix in (4.9) into smaller blocks,

$$\begin{bmatrix} \dfrac{(M+N)}{\rho_*\phi^*} & \dfrac{iV}{\rho_*\phi^*} \\ i\phi^*\xi^T & 0 \end{bmatrix} \tilde{U}_1 = -(\lambda + i\langle u^*, \xi \rangle)\tilde{U}_1, \tag{4.10}$$

where we calculate

$$M = A \begin{bmatrix} \xi_1^2 & -\xi_1\xi_2 \\ -\xi_1\xi_2 & \xi_2^2 \end{bmatrix} \tag{4.11}$$

as the contribution of S_{11},

$$N = \begin{bmatrix} \dfrac{(1+B)^2}{\Gamma}\xi_1^2 + \dfrac{\tau^*}{2\|D^*\|}\xi_2^2 & \dfrac{(1-B^2)}{\Gamma}\xi_1\xi_2 + \dfrac{\tau^*}{2\|D^*\|}\xi_1\xi_2 \\ \dfrac{(1-B^2)}{\Gamma}\xi_1\xi_2 + \dfrac{\tau^*}{2\|D^*\|}\xi_1\xi_2 & \dfrac{(1-B)^2}{\Gamma}\xi_2^2 + \dfrac{\tau^*}{2\|D^*\|}\xi_1^2 \end{bmatrix} \tag{4.12}$$

as the contribution of $-S_{12}S_{22}^{-1}S_{21}$, which is symmetric, and

$$V = \begin{bmatrix} \left((\partial_\phi\tau)^* + \dfrac{2(1+B)\|D^*\|(\partial_\phi f)^*}{\Gamma}\right)\xi_1 \\ \left(-(\partial_\phi\tau)^* + \dfrac{2(1-B)\|D^*\|(\partial_\phi f)^*}{\Gamma}\right)\xi_2 \end{bmatrix}. \tag{4.13}$$

(b) Estimation of the eigenvalues

We claim that the growth-rate eigenvalues (4.10) satisfy

$$\max_{j=1,2,3} \sup_{\xi\in\mathbb{R}^2} \Re\lambda_j(\xi) < \infty.$$

By compactness, it suffices to prove that

$$\max_{j=1,2,3} \limsup_{|\xi|\to\infty} \Re\lambda_j(\xi) < \infty. \tag{4.14}$$

Since only the real parts of eigenvalues matter, we may drop the term $i\langle u^*, \xi \rangle$ in (4.10) and verify (4.14) for the eigenvalue problem[10]

$$P\tilde{U} = -\lambda\tilde{U}, \tag{4.15}$$

where we write

$$P = \begin{bmatrix} \dfrac{(M+N)}{\rho_*\phi^*} & \dfrac{iV}{\rho_*\phi^*} \\ i\phi^*\xi^T & 0 \end{bmatrix} \tag{4.16}$$

for the matrix in (4.10) and we shorten the notation by dropping the subscript 1 on $\tilde{U}$. For large ξ, it is instructive to use perturbation theory to compare the eigenvalues (4.15) with the eigenvalues $P_0\tilde{U} = -\Lambda\tilde{U}$, where

$$P_0 = (\rho_*\phi^*)^{-1}\begin{bmatrix} M+N & 0 \\ 0 & 0 \end{bmatrix}. \tag{4.17}$$

Lemma 4.1. *Provided $\xi \neq 0$, the 2×2 matrix $M + N$ is positive definite.*

Proof. Since M and N are symmetric, it suffices to show that the trace and determinant of $M + N$ are positive. According to (2.20), $A > 0$ and $\Gamma > 0$, from which it follows immediately that $\mathrm{tr}(M + N) > 0$.

---

[10]Do not forget the minus sign in this equation—the growth rates are *negative* eigenvalues of P.

Regarding the determinant, for any 2×2 matrices

$$\det(\boldsymbol{M} + \boldsymbol{N}) = \det \boldsymbol{M} + \det \boldsymbol{N} + \chi(\boldsymbol{M}, \boldsymbol{N}), \tag{4.18}$$

where

$$\chi(\boldsymbol{M}, \boldsymbol{N}) = M_{22}N_{11} + M_{11}N_{22} - M_{12}N_{21} - M_{21}N_{12} \tag{4.19}$$

accounts for the cross terms. For the specific matrices (4.11) and (4.12), $\det \boldsymbol{M} = 0$,

$$\det \boldsymbol{N} = \frac{2\tau^*}{4\Gamma \|\boldsymbol{D}^*\|}[(1 + B)^2 \xi_1^4 - 2(1 - B^2)\xi_1^2 \xi_2^2 + (1 - B)^2 \xi_2^4] \tag{4.20a}$$

$$= \frac{2\tau^*}{4\Gamma \|\boldsymbol{D}^*\|}[(1 + B)\xi_1^2 - (1 - B)\xi_2^2]^2 \geq 0 \tag{4.20b}$$

and

$$\chi(\boldsymbol{M}, \boldsymbol{N}) = \frac{\tau^*}{2\|\boldsymbol{D}^*\|}\xi_1^4 + \left(\frac{4}{\Gamma} + \frac{\tau^*}{\|\boldsymbol{D}^*\|}\right)\xi_1^2 \xi_2^2 + \frac{\tau^*}{2\|\boldsymbol{D}^*\|}\xi_2^4 > 0. \tag{4.21}$$

This proves the lemma. ∎

Remark 4.2. It is noteworthy that $\det \boldsymbol{N} > 0$ except for the two directions

$$\frac{\xi_1}{\xi_2} = \pm\sqrt{\frac{1 - B}{1 + B}}. \tag{4.22}$$

Effectively, this calculation rederives the result of Pitman & Schaeffer [12] that the equations of CSSM, even without I-dependence, are well-posed for all directions except possibly those defined by (4.22).

It follows from lemma 4.1 that $\boldsymbol{P}_0 \tilde{\boldsymbol{U}} = -\Lambda \tilde{\boldsymbol{U}}$ has two eigenvalues, say Λ_1, Λ_2, where $\Lambda_1, \Lambda_2 < 0$ and is homogeneous of degree 2 in $\boldsymbol{\xi}$. Since $\boldsymbol{P}$ is an $\mathcal{O}(|\boldsymbol{\xi}|)$-perturbation of $\boldsymbol{P}_0$, two of the growth-rate eigenvalues of (4.15) satisfy

$$\lambda_j = \Lambda_j + \mathcal{O}(|\boldsymbol{\xi}|), \quad j = 1, 2,$$

both of which are negative in the limit $|\boldsymbol{\xi}| \to \infty$; i.e. they are bounded above by zero in this limit. The third growth rate is given by

$$\lambda_3 = -\frac{\det \boldsymbol{P}}{\lambda_1 \lambda_2} = -\frac{\det \boldsymbol{P}}{\Lambda_1 \Lambda_2} + \mathcal{O}(|\boldsymbol{\xi}|^{-1}).$$

The first term on the extreme right is the ratio of two quartics, the denominator being non-zero, so it is bounded, and the perturbation decays at infinity. This verifies (4.14) for all three eigenvalues derived from (3.10a) to (3.10e).

It remains to consider the effect of the lower-order terms that were neglected in (3.10a)–(3.10e). Inclusion of these terms would lead, after a calculation as above, to an eigenvalue problem (4.15) for a perturbed matrix

$$\begin{bmatrix} \dfrac{\boldsymbol{M} + \boldsymbol{N}}{\rho_* \phi^*} + \mathcal{O}(\boldsymbol{\xi}) & \dfrac{i\boldsymbol{V}}{\rho_* \phi^*} + \mathcal{O}(1) \\ i\phi^* \boldsymbol{\xi}^{\mathrm{T}} + \mathcal{O}(1) & \mathcal{O}(1) \end{bmatrix}.$$

As above, two of the eigenvalues of this matrix are negative and $\mathcal{O}(|\boldsymbol{\xi}|^2)$, and invoking the determinant shows that the third is bounded. This verifies (4.14) for eigenvalues of the full linearization of (3.5a)–(3.5e) and hence shows that the system is linearly well-posed. It is, therefore, expected that numerical solutions of the full two-dimensional nonlinear transient equations will not exhibit the exponential blow-up of perturbations or the dependence on grid resolution seen in the incompressible $\mu(I)$ equations [9].

5. Chute flow in compressible *I*-dependent rheology

Let us recall the steady-uniform 'Bagnold' solution [4,31] to the (incompressible) $\mu(I)$-rheology [4,5] for chute flow. Assuming that $\mu(I)$ is given by (2.13), these solutions exist for inclination angles ζ between the maximum and minimum angles $\zeta_2 = \tan^{-1}(\mu_2)$ and $\zeta_1 = \tan^{-1}(\mu_1)$, respectively. Letting Oxz be a Cartesian coordinate system, with the x-axis pointing downslope and the z-axis being the upward pointing normal, the Bagnold solution is

$$p_{\text{bag}} = \rho_\star \phi_\star g(h-z) \cos \zeta, \quad \phi_{\text{bag}} = \phi_\star \quad \text{and} \quad u_{\text{bag}} = \frac{2I_\zeta}{3d} \sqrt{\phi_\star g \cos \zeta} (h^{3/2} - (h-z)^{3/2}), \quad (5.1)$$

where h is the flow depth, $\phi_\star$ is a constant solids volume fraction and u is the downslope velocity. At a fixed inclination, the inertial number is equal to the positive constant

$$I_\zeta = I_0 \left(\frac{\tan \zeta - \mu_1}{\mu_2 - \tan \zeta} \right). \quad (5.2)$$

There is strong experimental and discrete element method (DEM) evidence [4,5,31,32] for both the lithostatic pressure distribution and the three-halves power in the dependence of the velocity on the thickness in (5.1). Fortunately, the CIDR solution is close to this, as we now show.

Assume that all variables depend only on z and only the x-component of the velocity is non-zero, say $u = u(z)$. Motivated by Bagnold flow, the scalings

$$p = (\rho_\star gh \cos \zeta)\check{p}, \quad z = h\check{z} \quad \text{and} \quad u = \left(\frac{I_\zeta}{d} \sqrt{g \cos \zeta} h^{3/2} \right) \check{u} \quad (5.3)$$

are used to non-dimensionalize the variables and the following ordinary differential equations (ODEs) are derived for $\check{p}$ and $\check{u}$ below.

Lemma 5.1. *Under the above assumptions, the non-dimensional pressure $\check{p}(z)$ and the non-dimensional downslope velocity $\check{u}(z)$ predicted by the CIDR rheology satisfy*

$$\frac{d\check{p}}{d\check{z}} = -\phi \quad (5.4)$$

and

$$\frac{d\check{u}}{d\check{z}} = \sqrt{\check{p}}, \quad (5.5)$$

where the pressure is zero at the free surface $\check{p}(1) = 0$, the velocity is zero at the base $\check{u}(0) = 0$ and for the function $C(\phi)$ proposed in (2.30) and (2.31) the solids fraction $\phi = \phi(\check{p})$ is

$$\phi = \phi_{\min} + \Delta\phi \frac{2\chi\check{p}}{1 + 2\chi\check{p}}, \quad \text{with} \quad \chi = \frac{bh}{\beta(I_\zeta)d}. \quad (5.6)$$

Steady uniform solutions to the CIDR model are shown in figure 4 for $\chi = 0.5, 2, 5, 20, 100$ and with $\phi_{\max} = 0.6$ and $\phi_{\min} = 0.5$. The nature of the solution is controlled by the non-dimensional parameter χ, which is larger for thicker flows. This parameter is also inversely proportional to $\beta(I_\zeta)$, which by (5.2) and (2.27) is a weakly increasing function of the inclination angle as shown in figure 5. For large χ, the concentration graphed in figure 4a is close to $\phi_{\max}$ over a significant proportion of the flow depth, with a narrow boundary layer near the free surface, where the concentration decreases to $\phi_{\min}$. In the limit as $\chi \to \infty$, the solution tends to the red line in figure 4a, which corresponds to Bagnold flow with $\phi_\star = \phi_{\max}$. As χ is decreased, the surface boundary layer becomes thicker and the flow becomes progressively more dilute, tending to the blue line in figure 4a as $\chi \to 0$, which corresponds to Bagnold flow with $\phi_\star = \phi_{\min}$. In figure 4b,c, the non-dimensional pressure and the velocity are shown, which are also bounded between the red and blue lines corresponding to Bagnold flow with $\phi_\star$ equal to $\phi_{\max}$ and $\phi_{\min}$, respectively. As well as the profiles being bounded by two non-dimensional Bagnold solutions, the steady-uniform CIDR solutions also have exactly the same scaling properties (5.3), on the flow density $\rho_\star$, gravity g, the flow depth h, the chute inclination ζ, the particle diameter d and the inertial
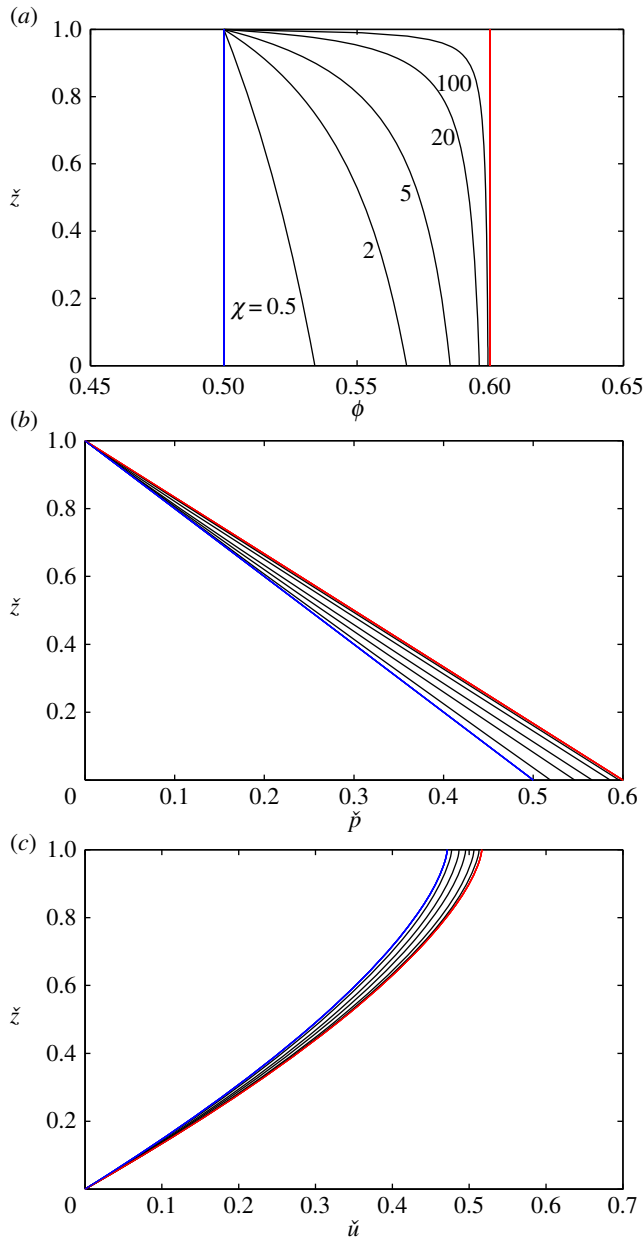
Figure 4. Solutions for (*a*) the solids volume fraction ϕ, (*b*) the non-dimensional pressure $\check{p}$ and (*c*) the non-dimensional velocity $\check{u}$ as a function of the non-dimensional thickness $\check{z}$ for five different values of the non-dimensional parameter $\chi = 0.5$, 2, 5, 20, 100 (black lines). The red line corresponds to $\chi \to \infty$ and Bagnold flow with $\phi_\star = \phi_{max}$, while the blue line corresponds to $\chi \to 0$ and Bagnold flow with $\phi_\star = \phi_{min}$. In all cases, $\phi_{min} = 0.5$ and $\phi_{max} = 0.6$.

number I_ζ, as the Bagnold solution (5.1). The steady-uniform CIDR solutions are, therefore, very closely related to the classical Bagnold solution and are almost indistinguishable for large χ.

Proof of the lemma. The form of the assumed solution implies that $D_{xx} = D_{zz} = 0$ and $D_{xz} = (1/2)\,\mathrm{d}u/\mathrm{d}z > 0$; in particular, div $\boldsymbol{u} = 0$. It follows from the alignment condition (2.8) that $\sigma_{xx} = \sigma_{zz} = -p$ and $\sigma_{xz} = \tau$, where $\tau = ||\boldsymbol{\tau}||$. Thus, the downslope and normal momentum balances (2.2)
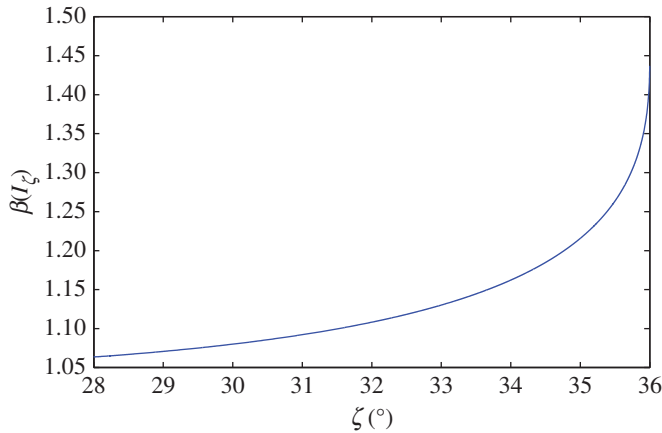
Figure 5. The function $\beta(I_\zeta)$ as a function of the inclination angle ζ for the case $\zeta_1 = 28°$, $\zeta_2 = 36°$ and $I_0 = 0.279$. (Online version in colour.)

reduce to

$$\frac{d\tau}{dz} + \rho_* \phi g \sin \zeta = 0 \tag{5.7}$$

and

$$\frac{dp}{dz} + \rho_* \phi g \cos \zeta = 0. \tag{5.8}$$

We may apply the chain rule to (5.7) and (5.8) to derive $d\tau/dp = \tan \zeta$, and, since tractions at the surface vanish,

$$\tau = p \tan \zeta. \tag{5.9}$$

On the other hand, since div $\boldsymbol{u} = 0$, the flow rule (2.18) implies $f(p, \phi, I) = 0$, and, given the ansatz (2.24b), we conclude that

$$p = \tfrac{1}{2}\beta(I)C(\phi). \tag{5.10}$$

From (5.10), we have $p/C(\phi) = \beta(I)/2$, so by (2.24b) and (2.25)

$$\tau = \left[\alpha(I) - \frac{\beta(I)}{2}\right] p = \mu(I)p. \tag{5.11}$$

Comparing (5.9) with (5.11), it follows that $\mu(I) = \tan \zeta$, and hence the inertial number I is equal to the constant I_ζ everywhere. The value of I_ζ is given by the formula (5.2), which is derived by inverting (2.13) with $\mu(I_\zeta) = \tan \zeta$. Since the inertial number is constant and $C(\phi)$ satisfies (2.30), equation (5.10) can be inverted and then non-dimensionalized, using the scalings (5.3), to give the solids volume fraction $\phi = \phi(\check{p})$ in (5.6). Using the same scalings (5.3), the normal momentum component (5.8) yields the ODE (5.4), which can be solved for the pressure $\check{p}$ by integrating down from the free surface using (5.6). The concentration $\phi(\check{p})$ then follows from (5.6). From the definition of the inertial number (2.12) and that $2\|\boldsymbol{D}\| = |du/dz|$, the downslope velocity satisfies the ODE

$$\frac{du}{dz} = \frac{I_\zeta}{d} \sqrt{\frac{p}{\rho_*}} \tag{5.12}$$

and the scalings (5.3) then give the ODE (5.5), which, since $\check{p}$ is known, can be integrated upwards from the no-slip condition $\check{u}(0) = 0$ at the base. ∎

6. Conclusion and discussion

In this paper, we have analysed a generalization of the $\mu(I)$-rheology that allows for changes in the granular solids volume fraction. The equations of motion, the CIDR model, are found to be

linearly well-posed when the constitutive laws satisfy certain criteria. We indicate (in §2e) how the specific *I*-dependence of the $\mu(I)$-rheology can be fitted into the theory, and we have shown that the inclusion of compressibility removes the ill-posedness of the incompressible rheology [9].

CIDR equations for steady, uniform flow down an inclined chute are solved in §5, where we observe that the solution is comparable to the classical Bagnold solution under conditions similar to the many experimental and DEM results [4,5,31,32] for the velocity profile in steady chute flows. Crucially, the equations are *well-posed at all inertial numbers*, which corresponds to the full range of inclination angles for steady chute flow, in contrast with the limited range of angles in which the incompressible theory is well-posed [9]. At the same time, the new theory captures the dilatant behaviour of over-consolidated granular material as the structure of the equations is motivated by CSSM.

Incidentally, the chute flow calculation relates to Forterre & Pouliquen [33], who modify incompressible $\mu(I)$-rheology by postulating a constitutive law in which ϕ is explicitly specified as a function of the inertial number, say $\Phi(I)$. In the CIDR model, there is no constitutive law relating these two variables; nevertheless, dependence of ϕ on I is implicit in the solution. It is also quite possible that a more subtle dependence of volume fraction on inertial number will be needed to reproduce some phenomena reported recently in experiments and numerical simulations, while retaining the property of well-posedness.

Our primary viewpoint in this paper has been to regard the CIDR model as modifying $\mu(I)$-rheology with compressibility. However, it is equally valid to regard CIDR as modifying CSSM with rate dependence. To recapitulate, our result shows well-posed equations result from such modification provided that the yield locus and flow rule satisfy (2.19) and (2.20). Unlike in the incompressible $\mu(I)$-rheology and the rate-independent CSSM equations, the CIDR equations are linearly well-posed for all deformations and for perturbations in all directions in Fourier space.

An important next step will be to specify constitutive laws, satisfying the general conditions for linear well-posedness of this paper, and formulated to accurately match the available experimental results and discrete numerical simulations for granular flows such as two-dimensional steady flow in a Couette geometry and time-dependent chute flow. Then the CIDR model and well-posedness result can be tested with fully two-dimensional nonlinear transient numerical computations of these flows.

Appendix A. Ideas from critical-state soil mechanics

(a) Constitutive equations

CSSM is an ingeniously constructed version of plasticity that includes compressibility but reduces to a singular perturbation of Coulomb material, which is incompressible, in an appropriate limit. In two-dimensional CSSM, flow is described by the usual six variables, ϕ, $\boldsymbol{u}$ and $\boldsymbol{\sigma}$. Since flow is compressible, the solids fraction ϕ remains as a genuine variable. The governing equations consist of the conservation laws (2.1), (2.2) plus three constitutive laws. One of the constitutive equations is the alignment condition (2.8), with no changes required. The second constitutive equation, like (2.5), specifies the norm of the deviatoric stress,

$$\text{Yield condition: } \|\boldsymbol{\tau}\| = Y(p, \phi), \qquad (A\,1)$$

but as indicated the function Y depends on the solids fraction ϕ as well as on the mean stress p. The final constitutive relation, the flow rule, relates expansion and contraction of material to the slope of the yield surface,

$$\text{Flow rule: div } \boldsymbol{u} = 2\frac{\partial Y}{\partial p}(p, \phi) \, \|\boldsymbol{D}\|. \tag{A 2}$$

We refer to Jackson [1] for a derivation of (A 2) from the normality condition of plasticity.

By way of example, a simple, physically acceptable, yield locus is given by

$$Y(p, \phi) = 2\mu p - \frac{p^2}{C(\phi)}, \tag{A 3}$$

where μ is a coefficient of friction, as in (2.5), and $C(\phi)$ is an increasing function of the solids fraction of the form (2.30). For such a yield condition, it follows from the proof of lemma 2.5, restricted to the case where $\mu(I)$ is independent of I, that equations (2.1), (2.2), (A 1), (A 2), (2.8) reduce to the Coulomb model in the limit $\Delta\phi \to 0$.

(b) Consequences of the flow rule

The behaviour discussed in this subsection occurs under fairly general hypotheses—see Jackson [1]. However, to explain the theory with a minimum of technicalities, we confine the discussion to the specific yield condition (A 3).

The phrase *critical state*, from which CSSM derives its name, refers to a state p, ϕ such that

$$\frac{\partial Y}{\partial p}(p, \phi) = 0. \tag{A 4}$$

For the example yield condition (A 3), condition (A 4) means that

$$2\left(\mu - \frac{p}{C(\phi)}\right) = 0. \tag{A 5}$$

Rewriting the yield condition as $Y(p, \tau) = [2\mu - p/C(\phi)]p$ and invoking (A 5), we deduce that at a critical state

$$\|\boldsymbol{\tau}\| = \mu p. \tag{A 6}$$

The set where (A 6) is satisfied is called the critical state line. Thus, *along the critical state line, the stress satisfies the Coulomb yield condition.*

According to the flow rule (A 2), at a critical state, deformation is not accompanied by any change in ϕ. Let us examine behaviour away from the critical state line. Suppose that, for example, initially the (uniform) state of material is at yield at point A in figure 2b. At this point, $\partial Y/\partial p < 0$, so according to the flow rule div $\boldsymbol{u} < 0$; i.e. material compactifies and becomes stronger, so τ must increase for deformation to continue. Indeed, the stress will continue to increase until a critical state on a larger yield surface is reached, as suggested in figure 2b by the ϕ_3-yield surface. Moreover, if $\Delta\phi$ in (2.30) is small, a very slight increase in ϕ is sufficient to accommodate this evolution. That is, we expect stress to be quickly driven from point A to a critical state on a larger yield surface where the Coulomb yield condition (A 6) is satisfied.

Conversely, at point B in figure 2b, $\partial Y/\partial p > 0$, so under deformation div $\boldsymbol{u} > 0$; i.e. material expands and becomes weaker. It is natural to imagine that the stress is driven to a critical state on a smaller yield surface, as suggested by the arrow in the figure. This would indeed be the case *if material deformed uniformly*, but this assumption is unrealistic for stresses above the critical state line, $\tau > \mu p$. For such stresses, because material expands under deformation and therefore weakens, instability often causes localized deformation—if deformation near one point happens to be slightly larger than elsewhere, the associated expansion lowers the yield condition more near this point, and subsequent deformation tends to concentrate near this point.

Appendix B. A primer on ill-posed partial differential equations

The following appendix gives a self-contained, elementary summary of key issues regarding ill-posed PDEs. A much more detailed treatment can be found in [6].

(a) Testing for ill-posedness

The initial value problem for a PDE is called *well-posed* in the sense of Hadamard if for general initial data a solution (1) exists, (2) is unique and (3) varies continuously under perturbations of the initial conditions[11] (see also Pinchover & Rubinstein [34]). If one or more of these criteria is not satisfied then the problem is called *ill-posed*. A classic example of an ill-posed problem is the backward heat equation

$$\partial_t u = -\partial_{xx} u. \tag{B 1}$$

In appendix Bb, we show that condition (1) fails; here we show that condition (3) also fails. Taking the Fourier transform reveals that the equation admits solutions

$$u_\xi(x,t) = \sin(\xi x)\, e^{\xi^2 t}, \tag{B 2}$$

for any $\xi \in \mathbb{R}$. Consider the scaled solutions $|\xi|^{-p} u_\xi(x,t)$, where $p > 0$, as perturbations of the trivial solution $u(x,t) \equiv 0$. The initial conditions of the scaled solution—i.e. $|\xi|^{-p} \sin(\xi x)$—tend to zero in the sup norm as $\xi \to \infty$; indeed, if $p > k$ these initial conditions tend to zero in the $\mathcal{C}^k$ norm. On the other hand, for any $t > 0$, the norm

$$\sup_{x \in \mathbb{R}} |\xi|^{-p} u_\xi(x,t) \tag{B 3}$$

tends to infinity in this limit. Thus, an arbitrarily small perturbation of initial conditions for (B 1) can lead to an arbitrarily large solution in an arbitrarily short time.

For more general PDEs, there is a test for ill-posedness based on Fourier analysis of the *linearization* of the equations. The process is summarized as:

  (i) linearize the equations about a base-state solution;
 (ii) freeze the coefficients at some point (x^*, t^*);
(iii) look for solutions with exponential dependence $e^{i\langle \xi, x\rangle + \lambda(\xi)t}$.

We shall say the original PDE is linearly ill-posed (with respect to the base-state solution at the given point) if

$$\limsup_{|\xi| \to \infty} \lambda(\xi) = +\infty.$$

For most examples, if a PDE is linearly ill-posed, it is ill-posed in the sense of Hadamard. (But see Kreiss [35] for exceptional examples.)

An equation is called *linearly well-posed* with respect to a given base solution if the growth rate is bounded from above for all points (x^*, t^*). Linear well-posedness does not imply well-posedness in the sense of Hadamard. For example, it is trivially verified that the Navier–Stokes equations are linearly well-posed, but a major effort is required to show that, even just for a finite time, they are well-posed in the sense of Hadamard, and it is not known whether they are well-posed for all time.

---

[11]More precisely regarding condition (1): we choose a positive integer k and require that the initial value problem has a solution for any initial conditions in $\mathcal{BC}^k(\mathbb{R}^n)$, i.e. for k-times continuously differentiable functions such that all derivatives of order k or less are bounded. Likewise regarding condition (3), we require that, for the same integer k and for any positive time T, the solution operator is continuous as a map from $\mathcal{BC}^k(\mathbb{R}^n)$ into continuous functions on $[0, T] \times \mathbb{R}^n$. We refer to Joseph & Saut [6] for elaboration of these issues.

We illustrate the above test on the following made-up nonlinear system that has some similarity to the PDEs analysed in this paper:

$$\left.\begin{aligned} \partial_t u &= \partial_x v \\ \partial_t v &= \varepsilon u \partial_{xx} v + \partial_x u - v - \sin(x). \end{aligned}\right\} \tag{B 4}$$

and

The linearized equations with frozen coefficients are

$$\left.\begin{aligned} \partial_t \hat{u} &= \partial_x \hat{v} \\ \partial_t \hat{v} &= \varepsilon [u^* \partial_{xx} \hat{v} + (\partial_{xx} v)^* \hat{u}] + \partial_x \hat{u} - \hat{v}, \end{aligned}\right\} \tag{B 5}$$

and

where u^*, v^* is the base-state solution evaluated at the point (x^*, t^*) and $\hat{u}, \hat{v}$ are the perturbations solutions

$$\hat{U} = \begin{bmatrix} \hat{u} \\ \hat{v} \end{bmatrix} = e^{i\xi x + \lambda t} \tilde{U}, \tag{B 6}$$

where $\tilde{U} \in \mathbb{R}^2$ satisfies the eigenvalue problem

$$\begin{bmatrix} 0 & i\xi \\ i\xi + \varepsilon \partial_{xx} v^* & -\varepsilon u^* \xi^2 - 1 \end{bmatrix} \tilde{U} = \lambda \tilde{U}. \tag{B 7}$$

The eigenvalues of (B 7) could be easily calculated exactly but, provided that $u^* \neq 0$, they can be estimated more easily from their asymptotic behaviour as $|\xi| \to \infty$,

$$\lambda_1 = -\varepsilon u^* \xi^2 + \mathcal{O}(\xi) \quad \text{and} \quad \lambda_2 = \frac{\det S(\xi)}{\lambda_1} = -\frac{1}{\varepsilon u^*} + \mathcal{O}(\xi^{-1}), \tag{B 8}$$

where $S(\xi)$ is the 2×2 matrix in (B 7). If $u^* > 0$, then the eigenvalues satisfy

$$\max_{j=1,2} \limsup_{|\xi| \to \infty} \Re \lambda_j(\xi) < \infty, \tag{B 9}$$

so (B 4) is linearly well-posed. On the other hand, if $u^* < 0$, then λ_1 is unbounded, so (B 4) is ill-posed.

Note that, in analysing linear ill-posedness of (B 4), we consider the *full* linearization of the equations, i.e. (B 5). One might be tempted to discard terms with lower-order derivatives in the expectation that the growth of exponential solutions as $|\xi| \to \infty$ ought to be dominated by the highest-order derivatives in the equation. However, the counter-example

$$\partial_t u = \partial_{xxx} u - \partial_{xx} u$$

shows that this expectation is not valid, in general.

Nevertheless, for this example we may in fact analyse exponential solutions of (B 5) by first considering the *maximal-order* linearized equations

$$\left.\begin{aligned} \partial_t \hat{u} &= \partial_x \hat{v} \\ \partial_t \hat{v} &= \varepsilon u^* \partial_{xx} \hat{v} + \partial_x \hat{u}. \end{aligned}\right\} \tag{B 10}$$

and

In each of the above equations, only terms of maximal order are retained, i.e. the terms $(\partial_{xx} v)^* \hat{u}$ and $\hat{v}$ have been dropped from the second equation because it contains the higher-order terms $\partial_x \hat{u}$ and $u^* \partial_{xx} \hat{v}$, respectively. The growth rate of exponential solutions of (B 10) satisfy the same estimates (B 8), and the neglected lower-order terms do not change the leading-order behaviour. Often calculations may be simplified by studying the maximal-order linearization as an intermediate step.

(b) Consequences of ill-posedness

(i) Restrictions on the existence of solutions

In order for an ill-posed initial value problem to have a solution, usually the initial conditions must satisfy an extreme smoothness requirement, stronger than is physically acceptable in most applications. It may be difficult to demonstrate this behaviour, in general, but, for the backwards heat equation, we illustrate the behaviour with the following.

Proposition B.1. *If $\varepsilon \neq 0$, the initial value problem for (B1) with initial data $u(x, 0) = \varepsilon |\sin x|^p$ has no solution for any positive time interval unless the power p is an even non-negative integer.*[12]

Proof. Suppose the 2π-periodic function $f(x)$ has Fourier series $f(x) \sim \sum c_n e^{inx}$. If equation (B 1) with initial condition $u(x, 0) = f(x)$ has a continuous solution for $0 \leq t < \eta$, it has the Fourier series representation

$$u(x, t) = \sum_{n=-\infty}^{\infty} c_n e^{inx + n^2 t}.$$

Moreover, for $0 \leq t < \eta$

$$\sum_{n=-\infty}^{\infty} e^{2n^2 t} |c_n|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |u(x, t)|^2 \, dx < \infty. \tag{B 11}$$

Now if the Fourier coefficients of a function $f(x)$ satisfy $\sum n^{2k} |c_n|^2 < \infty$, then the derivatives $(d/dx)^j f(x)$ are square integrable for $j = 0, 1, \ldots, k$. But, provided p is not an even non-negative integer, the proposed initial data $|\sin x|^p$ has singular behaviour near $x = 0$. Specifically, $(d/dx)^k |\sin x|^p$ is square integrable only if $k < p + 1/2$. It follows for the Fourier coefficients of $|\sin x|^p$ that, if $k > p + 1/2$,

$$\sum_{n=-\infty}^{\infty} n^{2k} |c_n|^2 = \infty.$$

This inequality is incompatible with (B 11), so the initial value problem cannot be solved on any positive time interval. ∎

(ii) Grid-dependent computations

The attempt to solve an ill-posed PDE numerically produces unreliable, grid-dependent, results. Such behaviour has been observed in various physical problems [7–10] where the formulation was based on an ill-posed system of equations. However, in complicated problems like these, usually computational resources are stretched to the limit, meaning behaviour under grid refinement cannot be readily probed. Let us illustrate grid dependence on a much less demanding problem, the toy problem (B 4) above.

If $\varepsilon = 0$ and with initial conditions

$$u(x, 0) = a \quad \text{and} \quad v(x, 0) = \frac{-\sin(x)}{2}, \tag{B 12}$$

the (linear) equations (B 4) have the exact solution

$$u(x, t) = \frac{e^{-t/2}}{2} \left[\cos\left(x + \frac{\sqrt{3}t}{2}\right) + \cos\left(x - \frac{\sqrt{3}t}{2}\right) \right] + a - \cos x$$

$$v(x, t) = \frac{e^{-t/2}}{2} \left[\cos\left(x + \frac{\sqrt{3}t}{2} + \frac{\pi}{6}\right) - \cos\left(x - \frac{\sqrt{3}t}{2} - \frac{\pi}{6}\right) \right]. \tag{B 13}$$

and

The large-time limit of these solutions,

$$u(x, \infty) = a - \cos x \quad \text{and} \quad v(x, \infty) = 0, \tag{B 14}$$

---

[12]Of course, the general solution of (B 1) is a linear superposition of the solutions (B 2). We have no need for the general solution since one counter-example is sufficient to invalidate condition (1) above.
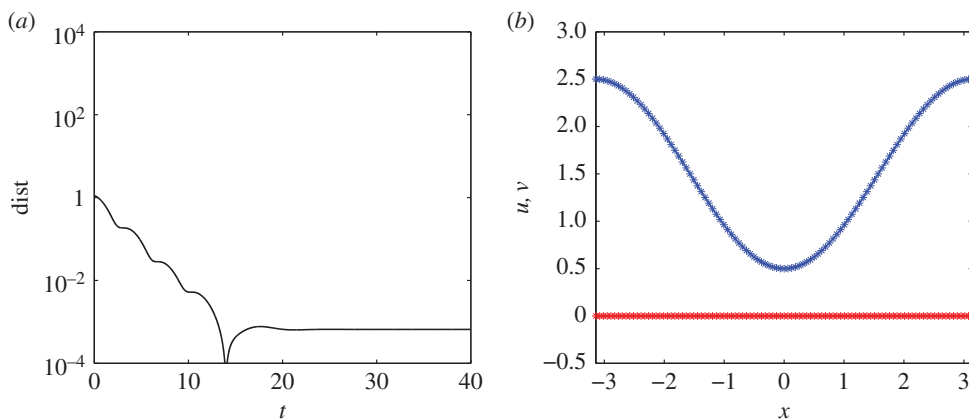
Figure 6. Numerical solutions of (B 4) with the distance from the asymptotic solution (B 15) in (*a*) and the fields at $t = 100$ in (*b*). Here $a = 1.5$, $\varepsilon = 0.01$ and the discretization is $\Delta x = 2\pi/100$ and $\Delta t = 1 \times 10^{-3}$.
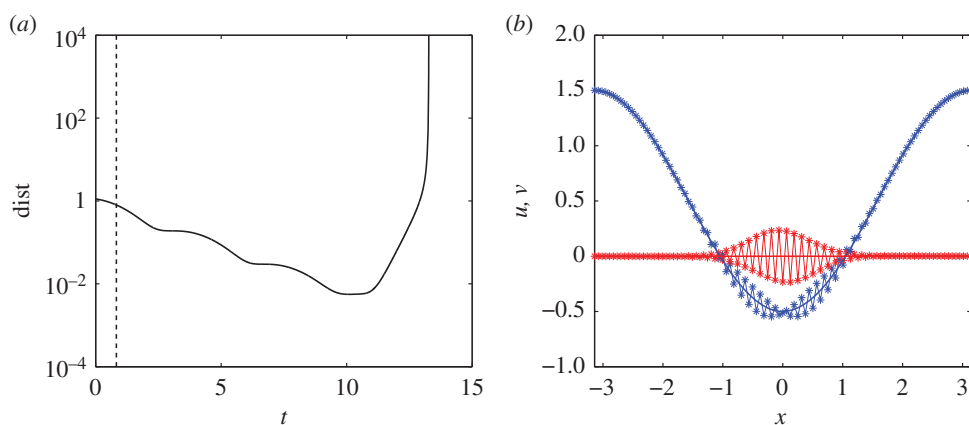

Figure 7. Numerical solutions of (B 4) with the distance from the asymptotic solution (B 15) in (*a*) and the fields at $t = 12.5$ in (*b*). Here $a = 0.5$, $\varepsilon = 0.01$ and the discretization is $\Delta x = 2\pi/100$ and $\Delta t = 1 \times 10^{-3}$. The vertical dashed line in (*a*) is the first time that $u = 0$.

is also a steady-state solution of the nonlinear system (with $\varepsilon > 0$). If $a > 1$, then $u(x, \infty) > 0$, and the calculations above suggest that the equations will be linearly well-posed. However, if $a < 1$, then $u(x, \infty)$ dips below zero over an interval, which suggests that the equations will be ill-posed.

Figure 6, where $a = 1.5$, and figure 7, where $a = 0.5$, confirm these expectations. They show numerical solutions using a central-space forward-time explicit scheme on the periodic domain $x \in [-\pi, \pi]$ with a spatial resolution of $\Delta x = 2\pi/100$. Each figure has two panels, one showing the temporal evolution of the distance from the asymptotic solution,

$$\text{Dist} = \max\left\{\sqrt{|u - u(x, \infty)|^2 + |v - v(x, \infty)|^2}\right\}, \tag{B 15}$$

and the other plotting the two variables u, v at a specific (late) time during the computation. In figure 6, the well-posed case, the numerical solution converges to the predicted steady-state solution (B 14) within numerical accuracy. By contrast, in figure 7, after an initial decay, ill-posedness asserts itself and causes the solution to blow up.

Regarding grid dependence, figure 8 shows another computation in the ill-posed case with a coarser grid, $\Delta x = 2\pi/30$. The solution appears to converge to the steady-state solution, just
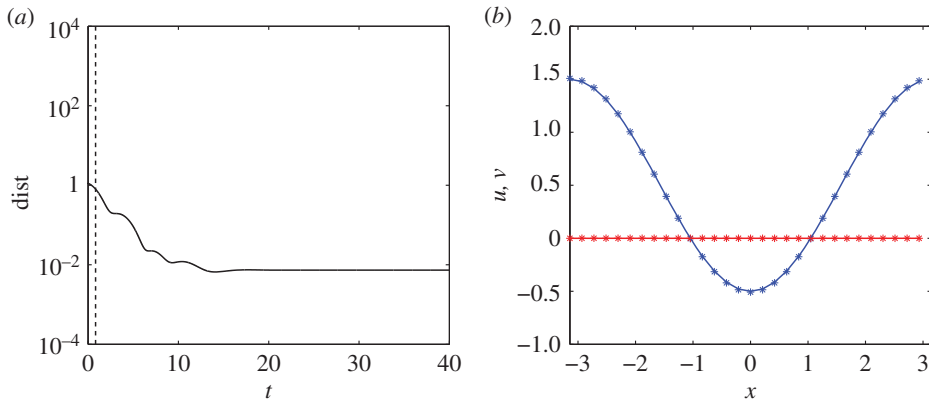
Figure 8. Numerical solutions of (B 4) with the distance from the asymptotic solution (B 15) in (*a*) and the fields at $t = 100$ in (*b*). Here $a = 0.5$, $\varepsilon = 0.01$ and the discretization is $\Delta x = 2\pi/30$ and $\Delta t = 1 \times 10^{-3}$. The vertical dashed line is the first time that $u = 0$.

like in the well-posed case. In other words, the computations on the coarse grid hide the ill-posed character of the underlying PDEs. This highlights that, in order to extract meaningful information from numerical computations, a proper study of grid convergence must be first carried out.

Incidentally, if the grid is made finer than in figures 6 and 7, in the well-posed case $a > 1$ the numerical solution converges to the steady-state solution with a smaller numerical error, while in the ill-posed case it blows up *sooner*, as expected.

References

1. Jackson R. 1983 Some mathematical and physical aspects of continuum models for the motion of the granular materials. In *Theory of dispersed multiphase flow* (ed. R Meyer). New York, NY: Academic Press.
2. Schaeffer D. 1987 Instability in the evolution-equations describing incompressible granular flow. *J. Differ. Equ.* **66**, 19–50. (doi:10.1016/0022-0396(87)90038-6)
3. Schaeffer DG, Pitman EB. 1988 Ill-posedness in three-dimensional plastic flow. *Commun. Pure Appl. Math.* **41**, 879–890. (doi:10.1002/cpa.3160410703)
4. GDR MiDi. 2004 On dense granular flows. *Eur. Phys. J. E* **14**, 341–365. (doi:10.1140/epje/i2003-10153-0)
5. Jop P, Forterre Y, Pouliquen O. 2006 A constitutive relation for dense granular flows. *Nature* **44**, 727–730. (doi:10.1038/nature04801)
6. Joseph DD, Saut JC. 1990 Short-wave instabilities and ill-posed initial-value problems. *Theor. Comput. Fluid Dyn.* **1**, 191–227. (doi:10.1007/BF00418002)
7. Gray JMNT. 1999 Loss of hyperbolicity and ill-posedness of the viscous-plastic sea ice rheology in uniaxial divergent flow. *J. Phys. Oceanogr.* **29**, 2920–2929. (doi:10.1175/1520-0485(1999) 029<2920:LOHAIP>2.0.CO;2)
8. Woodhouse MJ, Thornton AR, Johnson CG, Kokelaar BP, Gray JMNT. 2012 Segregation-induced fingering instabilities in granular free-surface flows. *J. Fluid Mech.* **709**, 543–580. (doi:10.1017/jfm.2012.348)
9. Barker T, Schaeffer DG, Bohorquez P, Gray JMNT. 2015 Well-posed and ill-posed behaviour of the mu(I)-rheology for granular flow. *J. Fluid Mech.* **779**, 794–818. (doi:10.1017/jfm.2015.412)
10. Baker JL, Johnson CG, Gray JMNT. 2016 Segregation-induced finger formation in granular free-surface flows. *J. Fluid Mech.* **809**, 168–212. (doi:10.1017/jfm.2016.673)
11. Mandel J. 1964 Conditions de stabilité et postulate de Drucker. In: *Rheology and soil mechanics* (eds G Kravtchenko and Sirieys). pp. 58–68. London, UK: Taylor & Francis.
12. Pitman EB, Schaeffer DG. 1987 Stability of time dependent compressible granular flow in two dimensions. *Commun. Pure Appl. Math.* **40**, 421–447. (doi:10.1002/cpa.3160400403)

13. de Coulomb CA. 1773 Essai sur une application des règles de maximis & minimis à quelques problèmes de statique, relatifs à l'architecture. *Mem. Math. Acad. R. Sci.* **7**, 343–382.

14. da Cruz F, Emam S, Prochnow M, Roux J, Chevoir F. 2005 Rheophysics of dense granular materials: discrete simulation of plane shear flows. *Phys. Rev. E* **72**, 021309. (doi:10.1103/PhysRevE.72.021309)

15. Boyer F, Guazzelli E, Pouliquen O. 2011 Unifying suspension and granular rheology. *Phys. Rev. Lett.* **107**, 188301. (doi:10.1103/PhysRevLett.107.188301)

16. Trulsson M, Bouzid M, Claudin P, Andreotti B. 2013 Dynamic compressibility of dense granular shear flows. *Eur. Phys. Lett.* **103**, 38002. (doi:10.1209/0295-5075/103/38002)

17. Andrade JE, Chen Q, Le PH, Avila CF, Evans TM. 2012 On the rheology of dilative granular media: bridging solid-and fluid-like behavior. *J. Mech. Phys. Solids* **60**, 1122–1136. (doi:10.1016/j.jmps.2012.02.011)

18. Pitman EB. 1988 The stability of granular flow in converging hoppers. *SIAM J. Appl. Math.* **48**, 1033–1053. (doi:10.1137/0148060)

19. Kamrin K. 2010 Nonlinear elasto-plastic model for dense granular flow. *Int. J. Plast.* **26**, 167–188. (doi:10.1016/j.ijplas.2009.06.007)

20. Jiang Y, Liu M. 2007 From elasticity to hypoplasticity: dynamics of granular solids. *Phys. Rev. Lett.* **99**, 105501. (doi:10.1103/PhysRevLett.99.105501)

21. Pouliquen O, Forterre Y. 2009 A non-local rheology for dense granular flows. *Phil. Trans. R. Soc. A* **367**, 5091–5107. (doi:10.1098/rsta.2009.0171)

22. Kamrin K, Koval G. 2012 Nonlocal constitutive relation for steady granular flow. *Phys. Rev. Lett.* **108**, 178301. (doi:10.1103/PhysRevLett.108.178301)

23. Kamrin K, Henann D. 2015 Nonlocal modeling of granular flows down inclines. *Soft Matter* **11**, 179–185. (doi:10.1039/C4SM01838A)

24. Bouzid M, Trulsson M, Claudin P, Clément E, Andreotti B. 2013 Nonlocal rheology of granular flows across yield conditions. *Phys. Rev. Lett.* **111**, 238301. (doi:10.1103/PhysRevLett.111.238301)

25. Jenkins JT, Savage SB. 1983 A theory for the rapid flow of identical, smooth, nearly elastic, spherical-particles. *J. Fluid Mech.* **130**, 187–202. (doi:10.1017/S0022112083001044)

26. Harris D, Grekova EF. 2005 A hyperbolic well-posed model for the flow of granular materials. *J. Eng. Math.* **52**, 107–135. (doi:10.1007/s10665-004-3717-6)

27. Sun J, Sundaresan S. 2011 A constitutive model with microstructure evolution for flow of rate-independent granular materials. *J. Fluid Mech.* **682**, 590–616. (doi:10.1017/jfm.2011.251)

28. Wu W, Bauer E, Kolymbas D. 1996 Hypoplastic constitutive model with critical state for granular materials. *Mech. Mater.* **23**, 45–69. (doi:10.1016/0167-6636(96)00006-3)

29. Perzyna P. 1966 Fundamental problems in viscoplasticity. *Adv. Appl. Mech.* **9**, 243–377. (doi:10.1016/S0065-2156(08)70009-7)

30. Jop P, Forterre Y, Pouliquen O. 2005 Crucial role of sidewalls in granular surface flows: consequences for the rheology. *J. Fluid Mech.* **541**, 167. (doi:10.1017/S0022112005005987)

31. Silbert LE, Ertas D, Grest GS, Halsey TC, Levine D, Plimpton SJ. 2001 Granular flow down an inclined plane: Bagnold scaling and rheology. *Phys. Rev. E* **64**, 051302. doi:10.1103/PhysRevE.64.051302)(1–14).

32. Gray JMNT, Edwards AN. 2014 A depth-averaged $\mu(I)$-rheology for shallow granular free-surface flows. *J. Fluid Mech.* **755**, 503–534. (doi:10.1017/jfm.2014.450)

33. Forterre Y, Pouliquen O. 2008 Flows of dense granular media. *Annu. Rev. Fluid Mech.* **40**, 1–24. (doi:10.1146/annurev.fluid.40.111406.102142)

34. Pinchover Y, Rubinstein J. 2005 *An introduction to partial differential equations*. Cambridge, UK: Cambridge University Press.

35. Kreiss HO. 1978 *Numerical methods for solving time-dependent problems for partial differential equations*. Montreal, Canada: Presses de l'Université de Montréal.