JOURNAL OF DIFFERENTIAL EQUATIONS 66, 19–50 (1987)

Instability in the Evolution Equations Describing Incompressible Granular Flow

DAVID G. SCHAEFFER*

Department of Mathematics, Duke University, Durham, North Carolina 27706

Received September 10, 1985

1. INTRODUCTION

In this paper, equations governing the time dependent flow of granular material under gravity are derived and analyzed. Formally these equations bear a strong resemblance to the Navier–Stokes equations for the flow of an incompressible, viscous fluid. However, the main result of this paper is that, depending on geometric and material parameters, the equations governing granular flow may lead to a violent instability analogous to that for

$$u_t = u_{xx} - u_{yy};$$

i.e., in some directions in Fourier transform space, the linearization of the governing equations resembles the *backwards* heat equation. Moreover instability is to be expected for the parameter values arising in most industrial applications.

In deriving the equations, we assume as constitutive laws a rigid-perfectly plastic, incompressible, cohesionless Coulomb powder with a yield surface of von Mises type, and we assume that the eigenvectors of the strain rate and stress tensors are parallel. (These assumptions are explained in Sect. 2.) The occurrence of instabilities with other constitutive laws is discussed in Section 5. In particular, with the Tresca yield surface favored by industry, the equations exhibit an even more severe instability than with the von Mises yield surface considered here.

Manufacturing industries must handle vast quantities of raw materials which are normally stored in granular form. Usually material is withdrawn from a storage bin by allowing the material to flow under the action of gravity through an outlet in the bottom of the bin. (Cf. Fig. 1.1.) Difficulties in the withdrawal process cause enormous financial losses; even the com-

* Supported by NSF Grant MCS-8401591.

0022-0396/87 $3.00

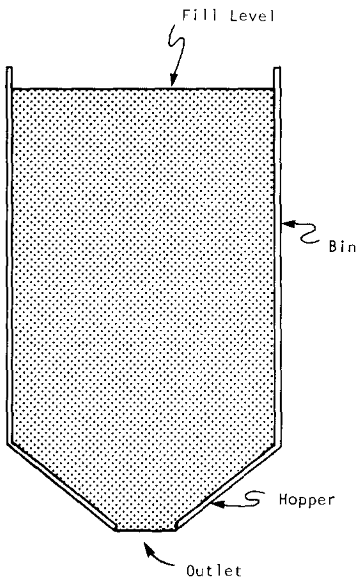

FIG. 1.1. A typical silo.

plete collapse of silos is not rare. Currently, design criteria for silos are derived from the steady state versions of the evolution equations analyzed in this paper. Thus the existence of an instability in these equations raises serious questions about silo design. (*Remark:* Often the flow in a silo is modeled using a technique known as *slice analysis;* i.e., by averaging over the horizontal dimensions of the silo, one obtains simplified equations which depend on only one space variable. Since the instability we study depends on direction, such one-dimensional models cannot shed any light on the phenomena.)

Ordinarily the flow in real silos is highly *unsteady;* indeed, the flow is pulsating with the material returning to rest during part of each cycle. (The period is typically a few seconds or less.) In Section 5 of this paper we conjecture that such pulsating flow results from the instability found in this paper. If the conjecture is true, the next problem is to prove the existence of time periodic solutions in an appropriate wider context (elastic-perfectly plastic rather than rigid-perfectly plastic) and to derive effective equations for the average stress and average velocity. This problem is of great practical importance since it represents a bottleneck for the application of mathematics to silo design.

A secondary result of this paper is that in three dimensions with a von Mises yield surface, the steady state equations governing granular flow are sometimes elliptic. (By contrast, in two dimensions or in three dimensions with a Tresca yield surface, the steady state equations are always hyperbolic.) The steady state equations are elliptic if and only if the evolution equation does *not* exhibit the instability analyzed in this paper; loosely speaking, this happens only for a freely flowing material in a steep hopper with smooth walls.

This paper is organized as follows: the equations to be studied are derived in Section 2; their instability is analyzed in Section 3; the instability is related to Jenike's radial solution in Section 4; and certain other pertinent issues are issues are discussed in Section 5.

2. Derivation of the Equations

(a) *Introduction*

In this section we derive the following equations for the velocity v and the average stress σ (i.e., the trace of the stress tensor, divided by 3) in a flowing granular material:

$$\begin{aligned} &\text{(a)} \quad \rho \frac{\partial v_i}{\partial t} = -k \frac{\partial}{\partial x_j} (\sigma |V|^{-1} V_{ij}) - \frac{\partial \sigma}{\partial x_i} + \rho g_i \\ &\text{(b)} \quad \operatorname{div} v = 0. \end{aligned} \tag{2.1}$$

Here and below we use the summation convention; the density ρ is assumed constant; k is a dimensionless physical constant specific to the material under study and is related to the angle of repose (the angle of the steepest pile that does not collapse; g is the acceleration of gravity; V_{ij} is the strain rate tensor

$$V_{ij} = -\frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) \tag{2.2}$$

(cf. the discussion of sign conventions below); and $|V|$ indicates the Euclidean norm of a matrix

$$|V| = \left\{\sum_{i,j} V_{ij}^2\right\}^{1/2} = \{\operatorname{tr} V^T V\}^{1/2}.$$

Although the derivation of (2.1) is a straightforward combination of established concepts, these equations have not, to our knowledge, appeared in the literature.

Note that (2.1) is meaningful only if $|V| \neq 0$; in other words, (2.1) contains the implicit assumption that the material in actually deforming. Although it is not apparent from the equations, another related assumption is also needed for (2.1) to be applicable—the material must have undergone substantial deformation in the past. (Cf. the discussion in Subsection (d).) Thus regarding the typical silo/hopper configuration illustrated in Fig. 1.1, Eqs. (2.1) describe the flow well inside the converging hopper, but not in the silo or near the top of the hopper.

Equations (2.1) bear an intriguing resemblance to the linearized Navier–Stokes equations for the low speed flow of a viscous, incompressible fluid, with σ playing the role of the pressure. Indeed, the only difference appears in the dissipation term. Specifically in (2.1) the dissipation is

(i) proportional to σ, rather than independent of it, and

(ii) homogeneous of degree zero in the velocity, rather than of degree one.

These differences are a natural consequence of the very different dissipation mechanisms in the two cases. In a viscous fluid, dissipation is due to momentum transfer from collisions; in a granular material, an assembly of many small particles in frictional contact, dissipation is due to friction between sliding particles. At the low speeds considered in this paper, momentum transfer is negligible in a granular material. Thus in this regime, dissipation would not be increased if the flow speed were doubled. (A striking illustration of this point occurred when mechanical plows replaced draught animals on farms: it was found, to everyone's surprise, that plowing at greater speeds does not require greater forces.)

In deriving (2.1) we shall assume that the reader is familiar with the basics of continuum mechanics as described in Prager [13, Chaps. 1–4]. Our notation differs from Prager's only in a sign convention. Specifically, the stress tensor T_{ij} measures *compressive* stresses, so that for a material in compression the eigenvalues σ_i of T_{ij} are positive. Similarly, the eigenvalues of the strain rate tensor V_{ij}, defined by (2.2), give the rates of *compression* of the material. This sign convention is natural for granular materials since such materials disintegrate under tensile stresses.

The unknown quantities in the flow are the velocity v_i and the stress tensor T_{ij}. (As mentioned above, ρ is assumed constant.) These variables are subject to the laws of conservation of mass and momentum and to two constitutive laws:

(i) the yield condition and

(ii) the flow rule.

Conservation of energy does not contribute an equation since the kinetic energy lost to friction is converted to heat and the temperature is not otherwise coupled to the variables under study. Equations for conservation of mass and momentum are

$$\begin{aligned}&\text{(a)}\quad \operatorname{div} v = 0\\&\text{(b)}\quad \rho\frac{\partial v_i}{\partial t}+\frac{\partial T_{ij}}{\partial x_j}=\rho g_i.\end{aligned}\tag{2.3}$$

In (2.3b), we have neglected the convective derivative $(v\cdot\nabla)v_i$ in the acceleration; this will be justified in Subsection (c) by order of magnitude estimates for the various terms. Since the constitutive laws are not standard knowledge among mathematicians, we will discuss equations for these laws at some length.

(b) *Constitutive Laws*

The yield condition may be motivated by thinking of the properties of a granular material as lying somewhere between those of a liquid and those of a true solid. Specifically, even at rest, a granular material can sustain some shearing stress but only an amount proportional to the average stress σ, where

$$\sigma=\frac{\sigma_1+\sigma_2+\sigma_3}{3}=\frac{\operatorname{tr} T}{3}.$$

Ultimately this behavior derives from the law of sliding friction applied to the individual particles in a granular material. A constitutive law formulating this behavior is provided by the von Mises-type yield condition

$$\sum_{i=1}^{3}(\sigma_i-\sigma)^2\leqslant k^2\sigma^2,\tag{2.4}$$

where k is a constant characteristic of the material. Moreover, for the material to deform, *equality must hold in* (2.4)—this point is central since it provides the equation

$$\sum_{i=1}^{3}(\sigma_i-\sigma)^2=k^2\sigma^2\tag{2.5}$$

which is used in the derivation of (2.1). We shall suppose that the material is rigid-perfectly plastic, which means that there is no viscosity; hence (2.5) is unaffected by any motion the material is undergoing, provided of course that there is some deformation. (Regarding motivation, it may be helpful to compare (2.5) with a perfect fluid for which, even in motion, the only allowable stress tensor is a hydrostatic pressure, i.e., $\sigma_i=\sigma$.)

Incidentally, the deformation of granular material is *plastic* in that if, after deformation, the shearing stress is reduced so that the inequality in (2.4) is strict, the material shows essentially no tendency to return to its original state.

Let us write

$$k=\sqrt{2}\sin\delta. \tag{2.6}$$

We claim that δ is the angle of internal friction [5, 7] measured in plane strain experiments. To prove this, we need the following result that will be derived below using the flow rule: in plane strain, at least after appropriate reindexing, we have

$$\sigma_3\equiv(\sigma_1+\sigma_2)/2. \tag{2.7}$$

Given (2.7), the inequality (2.4) reduces to the standard [5, 7] two dimensional relation

$$|\sigma_1-\sigma_2|\leqslant\sin\delta(\sigma_1+\sigma_2), \tag{2.8}$$

which proves the claim. Typically δ lies between 20 and 60°; values of δ at the low and of this range have received less study since the handling of such materials tends not to cause problems. Let us relate δ to the angle of repose, which may be characterized mathematically as the largest angle θ such that there is a solution of the equilibrium equations $\partial T_{ij}/\partial x_j=\rho g_i$ of the form

$$T_{ij}=a_{ij}[(\sin\theta)\,x_1-(\cos\theta)\,x_2]$$

for which (2.8) holds; here a_{ij} is constant and x_2 is the vertical coordinate. It turns out that this angle equals δ.

The cone (2.5) consists of two nappes, but only one of them is physically relevant. Since a granular material can support only compressive stresses, we want only the nappe on which

$$\sigma_i>0, \qquad i=1,2,3. \tag{2.9}$$

However, as shown in the following lemma, it is possible to satisfy (2.9) only if

$$\delta<60°, \tag{2.10}$$

or in terms of k, only if

$$k<\sqrt{3/2}. \tag{2.11}$$

Thus we shall make (2.10) an explicit assumption.

LEMMA 2.1. *One nappe of the cone* (2.5) *is contained in the positive octant* $\{\sigma_i>0\colon i=1,2,3\}$ *if and only if* $\delta<60°$.

Proof. As $\delta \to 0$, the cone (2.5) degenerates onto the line $\sigma_1 = \sigma_2 = \sigma_3$. Thus for small δ, one nappe is contained in the positive octant. As δ increases, this nappe remains inside the positive octant as long as (2.5) does not intersect any of the coordinate planes $\{\sigma_i = 0\}$. By symmetry it suffices to consider $\{\sigma_3 = 0\}$. Thus if we define a quadratic form

$$Q(\sigma) = \Sigma(\sigma_i - \sigma)^2 - k^2\sigma^2,$$

then one nappe remains inside the positive octant as long as $Q(\sigma_1, \sigma_2, 0)$ is positive definite in σ_1, σ_2. Now

$$Q(\sigma_1, \sigma_2, 0) = \tfrac{1}{9}\{(6 - k^2)(\sigma_1^2 + \sigma_2^2) - (6 + 2k^2)\,\sigma_1\sigma_2\},$$

which is positive definite only if (2.11) holds. The proof is complete.

Now we turn to the flow rule. For a granular material to deform, the stresses in different directions must be different. Intuitively it is clear that the response of the material to such unequal stresses should be to contract in the directions of greater stress and to expand in the directions of smaller stress. As will be elaborated in the following lemma, the flow rule is a quantitative formulation of these ideas. Specifically the flow rule links the strain rate and stress tensors by requiring (in matrix notation)

$$(\exists q > 0) \text{ s.t. } V = q(T - \sigma I). \tag{2.12}$$

We want the scalar q to be positive so that the major stress axis corresponds to an eigendirection of V with positive eigenvalue, i.e., a contracting direction.

LEMMA 2.2. *Let A and B be $n \times n$ symmetric matrices. There exists a scalar q such that $A = qB$ if and only if* (i) *there is an orthonormal basis for $\mathbb{R}^n$ whose members are eigenvectors of both A and B and* (ii) *the eigenvalues a_i and b_i of A and B satisfy*

$$\frac{a_1}{b_1} = \frac{a_2}{b_2} = \cdots = \frac{a_n}{b_n}. \tag{2.13}$$

Proof. If $A = qB$, then conditions (i) and (ii) are obvious. Conversely, if conditions (i) and (ii) hold, then $A = qB$ may be easily deduced in the representation, whose existence follows from (i), of A and B as diagonal matrices. The proof is complete.

By the lemma, the flow rule (2.12) requires that the eigenvectors of T be parallel to the eigenvectors of V. The flow rule also contains the assumption of incompressibility since

$$\operatorname{div} v = -\operatorname{tr} V = -q \operatorname{tr}(T - \sigma I) = 0.$$

According to (2.13), the flow rule contains in addition one scalar relation between the eigenvalues of T and of V. Let us use this latter relation to derive (2.7). For plane strain, one of the eigenvalues of V vanishes identically; say $\lambda_3(V) \equiv 0$. We deduce from (2.12) that $\sigma_3 \equiv \sigma$ from which (2.7) follows.

Although the flow rule determines the *sign* of the scalar q linking V and $T - \sigma I$, it gives no information about the *magnitude* of q. In other words, in a granular material the strain rate cannot be determined pointwise from a knowledge of the stresses; rather, V can only be determined by solving the appropriate partial differential equations. This fact, the source of much nonintuitive behavior of granular materials, leads to the property that dissipation in (2.1) is unaffected by scaling the velocity.

(c) *Derivation and Discussion of* (2.1)

As the main step is deriving (2.1), we will combine the yield condition and flow rule to obtain

$$T = \sigma(k\,|V|^{-1}\,V + I); \tag{2.14}$$

(2.1) follows immediately on substitution of (2.14) into (2.3). The proof of (2.14) is more transparent if we define the *deviator* of an $n \times n$ matrix A:

$$\operatorname{dev} A = A - \frac{1}{n}\,(\operatorname{tr} A)\,I.$$

In this notation, the yield condition (2.5) may be rewritten

$$|\operatorname{dev} T| = k\sigma$$

where the norm of dev T is the Euclidean norm, and the flow rule (2.12) may be rewritten

$$(\exists q > 0) \text{ s.t.} \qquad V = q \operatorname{dev} T.$$

We start the proof of (2.14) from the trivial decomposition

$$T = \operatorname{dev} T + \sigma I.$$

We use the flow rule to write dev $T = q^{-1}V$, and we use the yield condition to show that $q^{-1} = k\sigma\,|V|^{-1}$, thereby obtaining (2.14).

Let us justify, for the slow flows considered in this paper, the neglect of the convective derivative $\rho(v \cdot \nabla)\,v_i$ in (2.1) by making order of magnitude estimates for the various terms. The convective derivative is of the order $\rho v_0^2/l_0$, where v_0 and l_0 are a characteristic velocity and length, respectively. We estimate the average stress σ by the hydrostatic pressure $\rho g l_0$. Since k is dimensionless and of order 1, both terms in (2.1) involving σ are of the order ρg, i.e., of the same order as gravity. If we define an analogue of Reynolds number here,

$$R=\frac{\text{convective derivative}}{\text{dissipative term}},$$

then R is of the order v_0^2/gl_0. In a typical industrial hopper we may use $l_0=10$ ft, $v_0=0.1$ ft/s, which leads to the estimate $R\approx 3\times 10^{-5}$. In other words, the convective derivative may be neglected because all accelerations are much less than gravity.

If a silo is designed so that material, on leaving the silo, goes into free fall, then near the outlet the above estimates are inaccurate; indeed, in this situation the convective acceleration near the outlet is of the same order as gravity. However, in most industrial silos the output rate is set by a feeder device to a value much less than free fall, and consequently the convective derivative is negligible throughout the silo.

Next we deduce from (2.1) that, provided $\delta<60°$, the average stress σ satisfies a second-order elliptic equation. (For the linearized Navier–Stokes equation, the pressure satisfies Laplace's equation.) Taking the the divergence of (2.1a), we find that

$$k\frac{\partial^2}{\partial x_i\,\partial x_j}(\sigma\,|V|^{-1}\,V_{ij})+\frac{\partial^2\sigma}{\partial x_i\,\partial x_i}=0. \tag{2.15}$$

We regard V as given; thus (2.15) is a linear equation for σ whose principal symbol is the quadratic form associated with the matrix $kA+I$, where $A=|V|^{-1}\,V$. By the following lemma, (2.15) is elliptic.

LEMMA 2.3. *If $\delta<60°$, then $I+kA$ is positive definite.*

Proof. Since tr $A=0$ and tr $A^2=1$, the eigenvalues a_i of A satisfy

$$\begin{aligned}&\text{(a)}\quad a_1+a_2+a_3=0\\&\text{(b)}\quad a_1^2+a_2^2+a_3^2=1.\end{aligned} \tag{2.16}$$

Thus $a_3=-(a_1+a_2)$, and by (2.16b),

$$a_1^2+a_1a_2+a_2^2=\tfrac{1}{2}. \tag{2.17}$$

It follows by a calculus argument that the extreme values of a_1 on this ellipse are $\pm\sqrt{2/3}$, and by symmetry the same estimate holds for a_2 and a_3. Therefore by (2.11) the eigenvalues $1+ka_i$ of $I+kA$ satisfy

$$1+ka_i>1-\sqrt{\tfrac{3}{2}}\sqrt{\tfrac{2}{3}}=0.$$

The proof is complete.

(d) *Alternative Constitutive Laws*

In this subsection we attempt to place the constitutive laws in a larger context, particularly in regard to alternative constitutive laws. This information will not be used until Section 5.

Although we have assumed incompressibility in deriving (2.1), in fact all granular materials are at least slightly compressible; moreover the yield surface and the density are linked in a subtle way. The equations which describe granular flow when compressibility is included [5, 9] are significantly more complicated than (2.1). However, if these equations are expanded in an asymptotic series in powers of the compressibility, then (2.1) emerges as the zeroth order term in the expansion. Unfortunately the subsequent terms in the expansion are *singular* perturbations of (2.1). For example, the first-order perturbation of (2.1b) may be written

$$\operatorname{div} v + \varepsilon P\left(\frac{\partial \sigma}{\partial t}, \frac{\partial^2 \sigma}{\partial x_i \, \partial x_j}, \frac{\partial^3 v_i}{\partial x_j \, \partial x_k \, \partial x_l}\right) = 0 \tag{2.18}$$

where the arguments of P indicate the highest order derivatives in the perturbation. (The perturbation of (2.1a) does not contain any singular terms.) A more careful analysis of this expansion, such as is contained in [11] for the analogous problem in fluid flow, is greatly needed.

The derivation of (2.1) by the neglect of higher order terms in the expansion is called the *critical state approximation*. This approximation is probably valid if the material has undergone substantial deformation and is continuing to deform. In effect, by having undergone large deformations the material reaches an asymptotic state which simplifies the equations. However, before the material has undergone such substantial deformation—for example, in the silo in Fig. 1.1—the full equations including compressibility must be used.

Also related to compressibility is another force not included in this paper, viz., the effect of air pressure in the voids between particles. As material descends in the hopper, it expands in response to a decrease in the average stress σ, thereby creating a partial vacuum which opposes the flow. It seems that in a typical industrial installation, this force is substantially larger than the neglected inertial terms in (2.1). However, neither effect is considered in this paper.

Regarding our yield condition (2.4), there is another generalization of the two-dimensional condition (2.8) to three dimensions which we call the *Tresca* condition. In geometric terms, Tresca's yield surface is a hexagonal pyramid rather than a cone. In symbols, if we index the principal stresses so that $\sigma_3 \leqslant \sigma_2 \leqslant \sigma_1$, then Tresca's yield condition requires that

$$\sigma_1 - \sigma_3 \leqslant \sin \delta(\sigma_1 + \sigma_3);$$

moreover, the material may deform only if equality holds. The associated flow rule depends discontinuously on the middle principal stress σ_2; more precisely, there are three different flow rules depending on whether $\sigma_2 = \sigma_3$, $\sigma_3 < \sigma_2 < \sigma_1$, or $\sigma_2 = \sigma_1$. (See [6] for further explanation.) In designing hoppers, industry has used Tresca's yield condition almost exclusively. Experimentally it is difficult to tell which relation is more accurate since most experiments are essentially two dimensional and in two dimensions both yield conditions reduce to (2.8). However, recent examination of the evidence [8] indicates that the von Mises condition gives better agreement with certain kinds of experiments.

Incidentally, both names, von Mises and Tresca, derive by analogy from the plastic flow of metals. In metals, the threshold for plastic yield is essentially independent of the average stress σ. Thus for metals the yield surface is *cylindrical* rather than conical. This apparently minor difference is in fact a profound difference. In particular, for plastic flow of metals, the equations analogous to (2.1) are *monotone* [4], so the instability analyzed in this paper does not arise.

The flow rule (2.12) is challenged by some authors [2, 12]. In requiring that the eigenvectors of T and V be aligned, this condition neglects the rotation of a material element during deformation. In two dimensions [14] or in three dimensions with a Tresca yield surface [15], Spencer has proposed an alternative constitutive law which incorporates this rotation. (Cf. also [12].) Unfortunately accurate experiments are difficult and so far do not favor one theory over the other (although there is definite evidence [3] that under certain conditions the eigenvectors of T and V may be somewhat out of alignment). The steady state equations derived from the two constitutive laws are not terribly different, but the evolution equations derived from them are dramatically different. Let us elaborate. With (2.12) as the flow rule, we eliminated the deviator of T from the evolution equations to obtain (2.1), apparently a parabolic system. Spencer replaces (2.12) by a contitutive law that involves the time derivative of the stress tensor. Because of this time derivative, it is no longer possible to eliminate the deviator of T from the equations; rather the equation must be left as a first-order system. As we will see in Section 5, this system of evolution equations is best described as *elliptic*!

In deriving (2.1), we eliminated the scalar factor q in the flow rule (2.12). However, it is important to remember the requirement that q be positive. If, in the evolution of a flow, it becomes impossible to satisfy (2.12) with $q > 0$, then it is necessary to introduce *elastic–pastic theory*. In this theory the flow rule is generalized to read

$$V = \frac{\varepsilon}{\sigma} \frac{\partial}{\partial t} (\operatorname{dev} T) + q \operatorname{dev} T. \tag{2.19}$$

The first term on the right, in which we have neglected the convective derivative of T and similar rotational terms (cf. [13, Sects. 8.1, 2]), represents the elastic deformation. The factor ε is a small parameter characterizing the eleastic properties of the material—the material is incompressible and has a Young's modulus proportional to the average stress σ. The elastic deformation may be nonzero even if the stresses do not lie on the yield surface (2.5). Thus the yield condition (2.5) must be generalized to read

$$q(k\sigma - |\text{dev } T|) = 0, \tag{2.20}$$

where both factors are nonnegative. In a region where q vanishes identically, Eqs. (2.3), (2.19) constitute a first-order linear hyperbolic system with the side constraint $\text{div } v = 0$. In a region where q is everywhere positive, Eqs. (2.3), (2.19), (2.20) constitute a small, but singular, perturbation of (2.1).

3. Analysis of the Instability

(a) *Determination of the Principal Part of* (2.1)

On performing the indicated differentiations in (2.1a) and (2.2), we derive the equations

$$\begin{aligned} &\text{(a)} \quad \frac{\partial v_i}{\partial t} = \frac{k\sigma}{\rho\,|V|}\left\{\frac{1}{2}\frac{\partial^2 v_i}{\partial x_j\,\partial x_j} - A_{ij}A_{kl}\frac{\partial^2 v_k}{\partial x_j\,\partial x_l}\right\} \\ &\qquad\qquad - \frac{1}{\rho}(\delta_{ij} + kA_{ij})\frac{\partial \sigma}{\partial x_j} + g_i \\ &\text{(b)} \quad \text{div } v = 0 \end{aligned} \tag{3.1}$$

where the matrix A is defined by

$$A = |V|^{-1}\,V. \tag{3.2}$$

To define the principal part of the right-hand side of system (3.1), we introduce a system of weights as in Leray [10] or Agmon, Douglis and Nirenberg [1]. Specifically, we assign the weights $s_1 = s_2 = s_3 = 2$, $s_4 = 1$ to the four equations (3.1), and we assign the weights $t_1 = t_2 = t_3 = 0$, $t_1 = 1$ to the four unknowns v_1, v_2, v_3, σ, respectively. The contribution of the jth unknown to the principal part of the ith equation consists of the derivatives of order $s_i - t_j$. With this definition, *every* term in (3.1), except the zeroth order term g_i, belongs to the principal part.

This definition of the principal part is necessary in order to avoid trivialities. For example, it is known [7] that in two dimensions the steady state equations derived from (3.1) may be rewritten, by introduction of a new unknown, as a first-order, 4×4 hyperbolic system. We will show in Subsection (d) that with our definition of principal part, the two dimensional steady state equations derived from (3.1) are hyperbolic (in the sense of Leray) with the same characteristic directions as the 4×4 system. By contrast, according to the naive definition of principal part (i.e., just second-order derivatives), these steady state equations would be completely degenerate since the principal part of the fourth equations would vanish identically and σ would not appear in the principal part of any equation.

(b) *Growth Rates for Exponential Solutions*

The system (3.1) is quasilinear since the highest order derivatives enter linearly, i.e., multiplied by a coefficient which depends only on lower order derivatives. Let us "freeze" these coefficients, discard the zeroth order term in (3.1a), and look for an exponential solution

$$\begin{pmatrix} v \\ \sigma \end{pmatrix} = e^{i(\xi, x) + \lambda t} \begin{pmatrix} a \\ i\alpha \end{pmatrix},$$

where $a \in \mathbb{R}^3$, $\alpha \in \mathbb{R}$, and $(\xi, x) = \Sigma_1^3 \xi_i x_i$. This leads to the eigenvalue problem

$$\begin{aligned} &\text{(a)} \quad \frac{k\sigma}{\rho\,|V|}\left[-\frac{1}{2}\,|\xi|^2 I + (A\xi)(A\xi)^T\right] a + \frac{\alpha}{\rho}(I + kA)\,\xi = \lambda a \\ &\text{(b)} \qquad\qquad\qquad\qquad\qquad\qquad (\xi, a) = 0. \end{aligned} \tag{3.3}$$

To solve for α, we take the inner product of (3.3a) with ξ and use (3.3b), finding

$$\alpha = -\frac{k\sigma}{|V|}\left\{\frac{(A\xi, \xi)(A\xi, a)}{|\xi|^2 + k(A\xi, \xi)}\right\}; \tag{3.4}$$

by Lemma 2.3, the denominator in (3.4) is positive. Substituting (3.4) into (3.3a), we obtain

$$\frac{k\sigma}{\rho\,|V|}\, L(\xi)\, a = \lambda a$$

where

$$L(\xi) = -\frac{1}{2}\,|\xi|^2 I + (A\xi)(A\xi)^T - \frac{(A\xi, \xi)}{|\xi|^2 + k(A\xi, \xi)}\,(\xi + kA\xi)(A\xi)^T. \tag{3.5}$$

Now $L(\xi)$ defines a linear transformation from the two-dimensional subspace $\{\xi\}^\perp \subset \mathbb{R}^3$ into itself—the choice of α in (3.4) guarantees that $L(\xi)\,a \in \{\xi\}^\perp$. If E is the orthogonal projection onto $\{\xi\}^\perp$, then using $E\xi = 0$ we compute

$$L(\xi) = EL(\xi)\,E = -\frac{1}{2}\,|\xi|^2\,I + \frac{|\xi|^2}{|\xi|^2 + k(A\xi, \xi)}\,(EA\xi)(EA\xi)^T.$$

In other words, $L(\xi)$ is a rank 1 perturbation of $-\frac{1}{2}\,|\xi|^2\,I$; hence the eigenvalues of $L(\xi)$ are

$$\begin{aligned} &\text{(a)} \quad \lambda_1(\xi) = -\frac{1}{2}\,|\xi|^2 \\ &\text{(b)} \quad \lambda_2(\xi) = -\frac{1}{2}\,|\xi|^2 + \frac{|\xi|^2}{|\xi|^2 + k(A\xi, \xi)}\,|EA\xi|^2. \end{aligned} \tag{3.6}$$

Regarding $\lambda_2(\xi)$, we have

$$|EA\xi|^2 = |A\xi|^2 - \frac{(A\xi, \xi)^2}{|\xi|^2},$$

so that (3.6b) may be rewritten

$$\lambda_2(\xi) = \frac{|\xi|^2\,|A\xi|^2 - (A\xi, \xi)^2 - \frac{1}{2}\,|\xi|^4 - \frac{1}{2}k\,|\xi|^2\,(A\xi, \xi)}{|\xi|^2 + k(A\xi, \xi)}. \tag{3.7}$$

To conclude, if an exponential solution of (3.1) (with frozen coefficients) has spatial dependence $e^{i(\xi, x)}$, then it grows at the rate

$$\frac{k\sigma}{\rho\,|V|}\,\lambda_i(\xi), \qquad i = 1 \text{ or } 2.$$

(c) *Identification of the Unstable Directions*

In this subsection we show that, depending on parameters, $\lambda_2(\xi)$ may be positive for some directions ξ, thereby proving our main result that solutions of (3.1) are subject to an instability. The controlling parameters are k and A, where A is defined by (3.2). Without loss of generality we may investigate $\lambda_2(\xi)$ in a coordinate system such that $A = \operatorname{diag}(a_1, a_2, a_3)$; since the eigenvalues of A satisfy (2.16), there is only one essential parameter in A, say a_2.

Let U be the set of unstable directions; in symbols,

$$U = \{\xi \in \mathbb{R}^3 : P(\xi) > 0\}$$

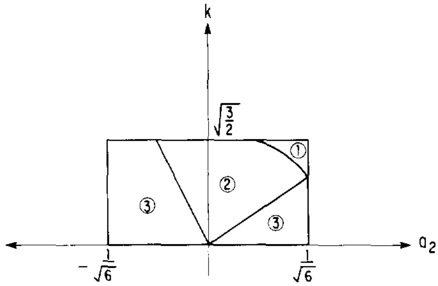

FIG. 3.1. Regions in parameter space.

where $P(\xi)$ is the homogeneous quartic form appearing in the numerator of (3.7). The main task of this subsection is to determine how U depends on k and a_2. Specifically we will show that if k, a_2 lie in region 1, 2, or 3 of Fig. 3.1, then U consists of the region between two cones (cf. Fig. 3.2a) or the interiors of two disjoint cones (cf. Fig. 3.2b), or is empty respectively. Thus the initial value problem for (3.1) is well posed only for k, a_2 in region 3.

Regarding the boundaries in Fig. 3.1, the limits on k come from (2.11) and the limits on a_2 come from the following considerations. To eliminate redundancy we index the eigenvalues of A so that

$$a_3 \leqslant a_2 \leqslant a_1. \tag{3.8}$$

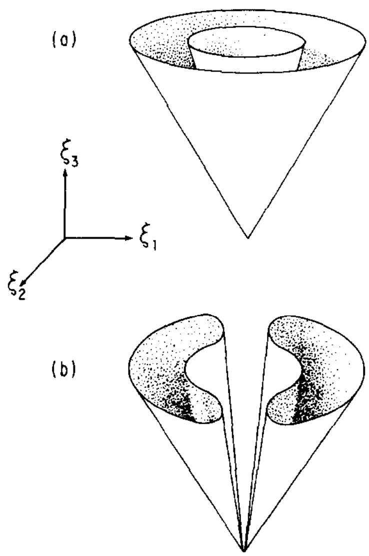

FIG. 3.2. Cones of unstable directions. Notes: (i) Only the top half of the cone is shown. (ii) A third case where U is empty is not shown.

We claim that with this convention

$$-\frac{1}{\sqrt{6}} \leqslant a_2 \leqslant \frac{1}{\sqrt{6}}. \tag{3.9a}$$

First consider $a_2 \leqslant 1/\sqrt{6}$. There is no issue if $a_2 \leqslant 0$, so suppose $a_2 > 0$. By (3.8) we have $a_2 \leqslant a_1$, by (2.16a) we have $|a_3| \geqslant 2a_2$, and by (2.16b) we have

$$1 = a_1^2 + a_2^2 + a_3^2 \geqslant 6a_2^2.$$

Consideration of $a_2 \geqslant -1/\sqrt{6}$ is similar, and the claim is proved. Since we showed in the proof of Lemma 2.3 that $|a_i| \leqslant 2/\sqrt{6}$, the corresponding limits for a_1 and a_3 are

$$-\frac{2}{\sqrt{6}} \leqslant a_3 \leqslant -\frac{1}{\sqrt{6}}, \qquad \frac{1}{\sqrt{6}} \leqslant a_1 \leqslant \frac{2}{\sqrt{6}}. \tag{3.9b}$$

In Section 4 we will discuss how this stability information applies to Jenike's radial solution in two and three dimensions. Here we simply interpret the following distinguished values for a_2:

$$\begin{array}{llll} \text{(a)} & a_2 = 0, & a_1 = \frac{1}{\sqrt{2}}, & a_3 = -\frac{1}{\sqrt{2}} \\ \text{(b)} & a_2 = \frac{1}{\sqrt{6}}, & a_1 = \frac{1}{\sqrt{6}}, & a_3 = -\frac{2}{\sqrt{6}} \\ \text{(c)} & a_2 = -\frac{1}{\sqrt{6}}, & a_1 = \frac{2}{\sqrt{6}}, & a_3 = -\frac{1}{\sqrt{6}}. \end{array} \tag{3.10}$$

Formula (3.10a) applies in plane strain; formulas (3.10b) and (3.10c) apply to the flow along the axis of three-dimensional axisymmetric converging and diverging hoppers, respectively.

We now set the framework for the proof. By homogeneity it suffices to analyze

$$U_* = \{(\xi_1, \xi_2): P(\xi_1, \xi_2, 1) > 0\}.$$

Since we have chosen coordinate so that A is diagonal, $P(\xi)$ may be written

$$P(\xi_1, \xi_2, \xi_3) = Q(\xi_1^2, \xi_2^2, \xi_3^2) \tag{3.11}$$

for a certain quadratic form $Q(\eta)$. Let

$$\begin{array}{ll} \text{(a)} & \Gamma = \{\eta \in \mathbb{R}^3: Q(\eta) > 0\}, \\ \text{(b)} & \Gamma_* = \{(\eta_1, \eta_2): Q(\eta_1, \eta_2, 1) > 0\}. \end{array} \tag{3.12}$$

Allowing for the eight equivalent signs $\pm\xi_i$ in (3.11), we see that U_* is in eight-to-one correspondence with

$$\Gamma_* \cap \{(\eta_1, \eta_2): \eta_1 \geq 0, \eta_2 \geq 0\}.$$

Thus the analysis of U_* hinges on how Γ_* meets the first quadrant in the η_1, η_2-plane. We shall prove that $Q(\eta_1, \eta_2, 0)$ is a negative definite form in η_1, η_2, which implies that Γ_*, if nonempty, is the interior of an ellipse. There are three qualitative different ways Γ_* meets the first quadrant, and these are illustrated in Fig. 3.3. For the intersections shown in Figs. 3.3a and b, the resulting cone U is illustrated in Figs. 3.1a and b, respectively; for the intersection shown in Fig. 3.1c, U is empty. We will distinguish between the three cases in Fig. 3.3 by investigating the restriction of $Q(\eta)$ to the coordinate planes $\{\eta_i = 0\}$.

Substituting into (3.7), we find

$$\begin{aligned} P(\xi) = & -\tfrac{1}{2}\{(1+ka_1)\,\xi_1^4 + (1+ka_2)\,\xi_2^4 + (1+ka_3)\,\xi_3^4\} \\ & + \left[a_1^2 - 2a_1a_2 + a_2^2 - 1 - \frac{k}{2}(a_1+a_2)\right]\xi_1^2\xi_2^2 \\ & + \left[a_1^2 - 2a_1a_3 + a_3^2 - 1 - \frac{k}{2}(a_1+a_3)\right]\xi_1^2\xi_3^2 \\ & + \left[a_2^2 - 2a_2a_3 + a_3^2 - 1 - \frac{k}{2}(a_2+a_3)\right]\xi_2^2\xi_3^2. \end{aligned}$$

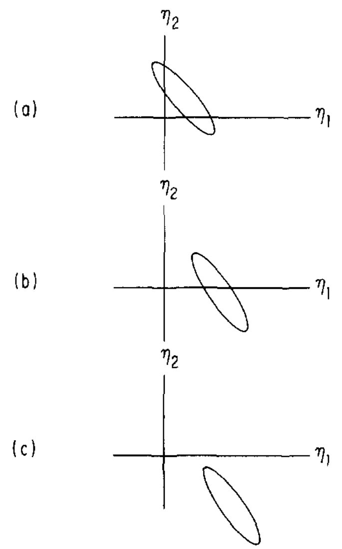

FIG. 3.3. Intersections of Γ_* with the first quadrant.

We use (2.16) and its corollary (Cf. (2.17).)

$$a_i^2 + a_i a_j + a_j^2 = \tfrac{1}{2} \qquad (i \neq j)$$

to simplify these coefficients; this leads to (3.11) with $Q(\eta) = -\frac{1}{2}(\eta, M\eta)$, where $M = \{m_{ij}\}$ is given by

$$M = \begin{pmatrix} 1 + ka_1 & 3a_3^2 - 1 - \dfrac{k}{2}\, a_3 & 3a_2^2 - 1 - \dfrac{k}{2}\, a_2 \\ 3a_3^2 - 1 - \dfrac{k}{2}\, a_3 & 1 + ka_2 & 3a_1^2 - 1 - \dfrac{k}{2}\, a_1 \\ 3a_2^2 - 1 - \dfrac{k}{2}\, a_2 & 3a_1^2 - 1 - \dfrac{k}{2}\, a_1 & 1 + ka_3 \end{pmatrix}.$$

Let M_1 be the 2×2 submatrix of M,

$$M_1 = \begin{pmatrix} m_{22} & m_{23} \\ m_{32} & m_{33} \end{pmatrix},$$

associated with the restriction of $Q(\eta)$ to the coordinate plane $\{\eta_1 = 0\}$; similarly let M_2 and M_3 be the 2×2 submatrices of M associated with the restriction of $Q(\eta)$ to $\{\eta_2 = 0\}$ and $\{\eta_3 = 0\}$, respectively.

LEMMA 3.1. *M_i is positive semidefinite if either*

(i) *$a_i > 0$ and $k < 2a_i$ or*

(ii) *$a_i < 0$ and $k < 6\,|a_i|$;*

indeed, under these circumstances M_i is positive definite unless $|a_i| = 2/\sqrt{6}$. If neither (i) *nor* (ii) *is satisfied, then M_i has a negative eigenvector which moreover lies in the first quadrant.*

Remark. The regions in the k, a_i-plane where conditions (i) and (ii) hold are indicated in Fig. 3.4. Note that by (3.9a, b) we have $|a_i| \leqslant 2/\sqrt{6}$.

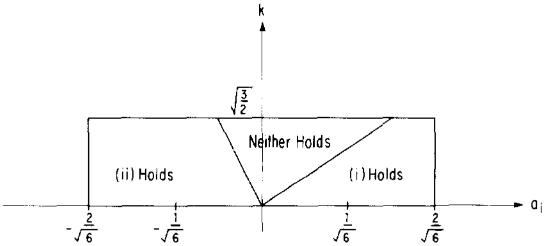

FIG. 3.4. Regions in parameter space.

Proof. The matrix M_i is positive definite if and only if tr $M_i > 0$ and det $M_i > 0$. Using (2.16a) we see that

$$\operatorname{tr} M_i = 1 - ka_i$$

which is positive by (2.11) and (3.9). A simple calculation shows that

$$\det M_i = (2 - 3a_i^2)\left(3a_i^2 - ka_i - \frac{k^2}{4}\right). \tag{3.13}$$

The first factor in (3.13) is positive unless $|a_i| = 2/\sqrt{6}$, the exceptional, extreme case. The second factor vanishes if $k = 2a_i$ or $k = -6a_i$; conditions (i) and (ii) identify the circumstances under which this factor is positive.

When neither (i) nor (ii) is satisfied, M_i has a negative eigenvector. This eigenvector will lie in the first quadrant if and only if $3a_i^2 - 1 - (k/2)\, a_i$, the off diagonal element of M_i, is negative. This off diagonal element is negative throughout the triangle, where (i) and (ii) fail; viz.,

$$k \leqslant \sqrt{\frac{3}{2}}, \qquad -\frac{k}{6} \leqslant a_i \leqslant \frac{k}{2}. \tag{3.14}$$

The proof is complete.

As a first application of the lemma, we claim that M_3 is positive definite. By (3.9b), we have $a_3 \leqslant -1/\sqrt{6}$, and condition (ii) of the lemma is satisfied for all such values of a_3, which proves the claim. Thus Γ_*, if nonempty, is the interior of an ellipse. (Warning: Because $Q(\eta) = -\frac{1}{2}(\eta, M\eta)$, the cone Γ is the set for which $(\eta, M\eta)$ is *negative.*)

As a second application of the lemma, we study the intersection of Γ_* with $\{\eta_2 = 0,\ \eta_1 > 0\}$. This intersection is empty if M_2 is positive definite, and it is nonempty if M_2 has a negative eigenvector in the first quadrant $\{\eta_1 > 0,\ \eta_3 > 0\}$. Thus by the lemma this intersection is nonempty if and only if k, a_2 lies in the triangle (3.14). Note that one corner of (3.14) is forbidden by (3.9a) and that the remaining, allowed portion of the triangle is precisely the union of regions 1 and 2 in Fig. 3.1. Thus Γ_* intersects $\{\eta_2 = 0,\ \eta_1 > 0\}$ if and only if k, a_2 lies in either region 1 or region 2 of Fig. 3.1.

As a final application of the lemma, we conclude similarly that Γ_* intersects $\{\eta_1 = 0,\ \eta_2 > 0\}$ if and only if k, a_1 lies in the triangle (3.14). Only one corner of (3.14) is allowed by (3.9b). The boundary of the allowed part of the triangle (viz., $k = 2a_1$, $a_1 \geqslant 1/\sqrt{6}$) appears in Fig. 3.1 as the boundary between regions 1 and 2. Although in Fig. 3.1, a_2 is the independent variable, a_1 and a_2 are linked by (2.17), so that the formula $k = 2a_1$ defines a curve in the k, a_2-plane. If k, a_2 lie above this curve, then Γ_* intersects $\{\eta_1=0, \eta_2>0\}$.

To summarize, we have shown that Γ_* intersects both $\{\eta_2=0, \eta_1>0\}$ and $\{\eta_1=0, \eta_2>0\}$ when k, a_2 lie in region 1, that Γ_* intersects only $\{\eta_2=0, \eta_1>0\}$ when k, a_2 lie in region 2, and that Γ_* intersects neither set when k, a_2 lie in region 3. Moreover, the origin in the η_1, η_2-plane does not belong to Γ_* since for $\eta=(0,0,1)=e_3$ we have

$$(e_3, Me_3)=m_{33}=1+ka_3>0.$$

It follows that Fig. 3.3a or b describe Γ_* when k, a_2 belongs to regions 1 or 2, respectively; as discussed above, analysis of (3.11) then shows that U has the form indicated in Figs. 3.2a or b. To show that Fig. 3.3c describes Γ_* when k, a_2 belongs to region 3, we must rule out the possibility that Γ_* might be contained in the interior of the first quadrant, $\{\eta_1>0, \eta_2>0\}$; this will complete the proof.

We will show in Lemma 3.2 that $\det M<0$ for all k, a_2 such that $|a_2|<1/\sqrt{6}$. Since $\operatorname{tr} M=3>0$, we conclude that M has two positive and one negative eigenvalues. It follows that Γ_* is never empty, and since M_3 is positive definite, Γ_* varies continuously with the parameters k, a_2.

Next we determine Γ_* for an explicit choice of parameters: $k=0$, $a_2=0$. In this case

$$M=\begin{pmatrix} 1 & \frac{1}{2} & -1 \\ \frac{1}{2} & 1 & \frac{1}{2} \\ -1 & \frac{1}{2} & 1 \end{pmatrix},$$

and the corresponding region $\Gamma_*^{(0)}$ is an ellipse contained in the fourth quadrant; the boundary of $\Gamma_*^{(0)}$ is tangent to the positive η_1-axis. By continuous dependence on parameters, if k, a_2 is close to (0, 0), then Γ_* must be close to $\Gamma_*^{(0)}$. However if k, a_2 belong to region 3, then Γ_* cannot intersect the positive η_1-axis. Therefore, for such k, a_2,

$$\Gamma_* \cap \{\eta_1 \geqslant 0, \eta_2 \geqslant 0\}=\varnothing. \tag{3.15}$$

Moreover we may extend the conclusion (3.15) to all of region 3—by continuity, for (3.15) to fail, Γ_* would have to meet the boundary of the first quadrant, and we have shown that this cannot happen in region 3.

It remains only to prove the following lemma.

LEMMA 3.2. *If* $|a_2|<1/\sqrt{6}$, *then* $\det M<0$.

Proof. For the proof of this lemma we temporarily revoke the convention (3.8) and consider all the eigenvalues a_i symmetrically. Since det M is a symmetric polynomial in a_1, a_2, a_3 of degree 6 or less

$$\det M = \sum_{j=0}^{2} p_j(a_1 + a_2 + a_3, a_1^2 + a_2^2 + a_3^2)(a_1 a_2 a_3)^j \tag{3.16}$$

for some polynomials p_j. However, in view of (2.16), the coefficients in (3.16) are actually constant. To determine these constants, we evaluate det M for the three special cases in (3.10), finding

$$\begin{aligned} &\text{(a)} \quad a_1 a_2 a_3 = 0, && \det M = -1 \\ &\text{(b)} \quad a_1 a_2 a_3 = -\frac{1}{\sqrt{54}}, && \det M = 0 \\ &\text{(c)} \quad a_1 a_2 a_3 = \frac{1}{\sqrt{54}}, && \det M = 0, \end{aligned} \tag{3.17}$$

respectively. (Remark: Note that these values for det M do not depend on k. Concerning (3.17b), det M vanishes when (3.10b) holds because of axial symmetry—$P(\xi)$ is a function of $\xi_1^2 + \xi_2^2$ and ξ_3^2, so $Q(\eta)$ depends on $\eta_1 + \eta_2$ and η_3 but not $\eta_1 - \eta_2$. Case (3.17c) may be similarly understood.) It follows from (3.16) and (3.17) that

$$\det M = 54(a_1 a_2 a_3)^2 - 1. \tag{3.18}$$

We deduce from (2.16) that

$$a_1 a_2 a_3 = a_2(a_2^2 - \tfrac{1}{2}).$$

Hence, for $a_2 \in [-1/\sqrt{6}, 1/\sqrt{6}]$,

$$|a_1 a_2 a_3| \leqslant 1/\sqrt{54},$$

with equality holding only at the endpoints. This estimate, conbined with (3.18), completes the proof.

(d) *Discussion of the Steady State Equations*

It is clear that the boundary of U, the cone of unstable directions, is related to characteristic directions of the steady state equations derived from (3.1). To formalize this relation, we will prove in Lemma 3.3 below that the characteristic polynomial of the steady state equations (i.e., the determinant of the principal symbol) is proportional to $|\xi|^2 P(\xi)$. Thus the characteristic normal cone has the form shown in Fig. 3.2. Before proving the lemma, let us use this information to determine how the type of the steady state equations depends on the region in Fig. 3.1 to which k, a_2 belong. For fully three-dimensional flow, we have the following classification:

Region in Fig. 3.1	Type of steady state equations
1	Product of second-order elliptic with fourth-order hyperbolic
2	No type—the number of real characteristics varies with direction
3	Elliptic

In the restricted contexts of plane strain or three-dimensional flow with axial symmetry, the context of most work in granular flow, the classification is different. For example, in plane strain $a_2 \equiv 0$, so k, a_2 always belong to region 2 of Fig. 3.1; however, the restricted symbol $P(\xi_1, 0, \xi_3)$ is a hyperbolic polynomial in two variables. Thus in two dimensions the steady state equations are hyperbolic. (Because of our weights in the definition of the principal part, this system is hyperbolic in the sense of Leray [10]; cf. [7] where the steady state equations are written as a first-order hyperbolic system.) It is easily checked that the characteristic directions make angles $\pm 45°$, $\pm(45-\delta/2)°$ with the major stress axis, corresponding to velocity and stress characteristics, respectively. (Warning: Zeros of $P(\xi)$ are the *normals* to characteristic directions, not the characteristic directions themselves; the normals make angles $\pm 45°$, $\pm(45-\delta/2)°$ with the *minor* stress axis.)

In three dimensions with axial symmetry, we will see from our analysis of the radial solution in Section 4 that all three regions in Fig. 3.1 are possible; which regions actually occur depend on k and the hopper parameters. In regions 1 and 2, the restricted steady state equations are hyperbolic; in region 3, they are elliptic. (By way of contrast, if Tresca's yield condition is assumed in place of von Mises' condition, then these restricted steady state equations are always hyperbolic [7].) In the two hyperbolic regions of Fig. 3.1, the angles between the characteristics and the major stress axis vary with k, a_2.

It remains to compute the characteristic polynomial of the steady state equations. The steady state equations are a quasi-linear system whose principal symbol, by (3.3), equals $C_1 S(\xi) C_2$ where $S(\xi)$ is the 4×4 matrix (in block form)

$$S(\xi) = \begin{bmatrix} -\frac{1}{2}|\xi|^2 I + (A\xi)(A\xi)^T & (I+kA)\xi \\ \xi^T & 0 \end{bmatrix}$$

and C_1, C_2 are constant 4×4 matrices

$$C_1=\begin{pmatrix} \dfrac{k\sigma}{\rho\,|V|}\,I & 0 \\ 0 & 1 \end{pmatrix}, \qquad C_2=\begin{pmatrix} I & 0 \\ 0 & \left(\dfrac{k\sigma}{|V|}\right)^{-1} \end{pmatrix}.$$

LEMMA 3.3. *The determinant* $\det S(\xi)$ *equals* $\frac{1}{2}\,|\xi|^2\,P(\xi)$.

Proof. We may compute the determinant in any convenient set of coordinates, even one which varies with ξ. Given ξ, choose an orthonormal basis e_1, e_2, e_3 for $\mathbb{R}^3$ such that $\xi=|\xi|\,e_1$ and $A\xi$ is a linear combination of e_1 and e_2. Then several entries of $S(\xi)$ vanish or simplify, as follows:

$$S(\xi)=\begin{pmatrix} S_1(\xi) & S_{12}(\xi) & 0 & S_{14}(\xi) \\ S_{21}(\xi) & S_{22}(\xi) & 0 & S_{24}(\xi) \\ 0 & 0 & -\frac{1}{2}\,|\xi|^2 & 0 \\ |\xi| & 0 & 0 & 0 \end{pmatrix}$$

Hence, expanding in the minors of the last two rows, we find

$$\det S(\xi)=\tfrac{1}{2}\,|\xi|^3\,S_{14}(\xi)\left\{S_{22}(\xi)-\frac{S_{12}(\xi)\,S_{24}(\xi)}{S_{14}(\xi)}\right\}. \tag{3.19}$$

In our coordinate system

$$S_{14}(\xi)=\frac{(\xi,\xi)+k(\xi,A\xi)}{|\xi|},$$

which is nonzero if $|\xi|\neq 0$. Moreover the factor in brackets in (3.19) may be interpreted as the 2, 2-entry of $S(\xi)$ resulting from an elementary column operation; viz., subtracting a multiple of the fourth column of $S(\xi)$ from the second in order to annihilate the 1, 2-entry of $S(\xi)$. But this is exactly how the linear transformation $L(\xi)$ in (3.5) was constructed, i.e., an appropriate multiple of $S(\xi)\,e_4$ was subtracted in order to make the range of $L(\xi)$ be orthogonal to ξ. Therefore

$$S_{22}(\xi)-\frac{S_{12}(\xi)\,S_{24}(\xi)}{S_{14}(\xi)}=\lambda_2(\xi)$$

where $\lambda_2(\xi)$ is given by (3.7). The lemma follows on multiplying the factors in (3.19). (Remark: The other eigenvalue of $L(\xi)$ appears as the 3, 3-entry of $S(\xi)$.)

4. Implications for the Radial Solution

In this section we relate the instability calculations of Section 3 to Jenike's radial solution. We find that in most practical situation, the initial value problem for (2.1) is likely to be linearly ill posed. More precisely, for two-dimensional flow the radial solution is always in region 2 of Fig. 3.1, which is unstable (cf. Sect. 3(d)); and for three-dimensional flow with axial symmetry, unless the material under study is free flowing and unless the hopper walls are both steep and smooth, then in at least part of the hopper the radial solution will enter one (or both) of the unstable regions of Fig. 3.1.

(a) *The Radial Solution*

Jenike's radial solution is a similarity solution in an infinite hopper of the steady state equations associated to (2.1). This particular solution is a fundamental component in the design of industrial hoppers. It is studied in detail in [6] for a Tresca yield surface. Here we shall review only the essential features of this solution, emphasizing the differences which result from the use of a von Mises yield surface.

We shall discuss the radial solution only in three dimensions with axial symmetry since, as noted above, the parameters of any two dimensional flow belong to an unstable region of Fig. 3.1. Rather than solve (2.1) directly, we shall return to the original equations (2.3), (2.12) since this is convenient in formulating boundary conditions. Consider a conical domain, say in spherical polar coordinates

$$\{(r, \theta, \varphi): r > 0, \theta < \theta_w\}, \tag{4.1}$$

which corresponds to an infinite, converging hopper. The radial solution exploits the fact that the domain (4.1) is invariant under the scaling transformation

$$(r, \theta, \varphi) \to (cr, \theta, \varphi)$$

where $c > 0$; thus one seeks a solution of (2.3), (2.12) in the form

$$T_{ij}(r, \theta) = rT_{ij}(\theta), \qquad v(r, \theta) = -r^{-2}u(\theta)\, \mathbf{e}_r \tag{4.2}$$

where $\mathbf{e}_r$ is a unit vector in the radial direction. We shall require that

$$u(\theta) > 0, \tag{4.3}$$

i.e., that the flow is *downward*. The stress tensor may be expressed [8] in terms of two scalar functions, the average stress σ and an angle ξ; the latter is related to a_2 by (4.13). Specifically, in spherical polar coordinates with the r, φ, φ-components indexed by $j=1, 2, 3$, respectively, the stress tensor equals

$$T=\sigma I+\frac{k\sigma}{\sqrt{2}}\begin{pmatrix} -\frac{2}{\sqrt{3}}\cos\xi & \sin\xi & 0 \\ \sin\xi & \frac{1}{\sqrt{3}}\cos\xi & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}}\cos\xi \end{pmatrix}. \tag{4.4}$$

This representation derives from the yield condition (2.5), from the fact that in axial symmetry $\mathbf{e}_\varphi$ must be an eigenvector of T, and from the following lemma.

LEMMA 4.1. $T_{\theta\theta}=T_{\varphi\varphi}$.

Proof. It follows from the flow rule (2.12) that

$$\frac{T_{\theta\theta}-\sigma}{T_{\varphi\varphi}-\sigma}=\frac{V_{\theta\theta}}{V_{\varphi\varphi}}.$$

By (4.2) the velocity in the radial solution is purely radial so that

$$V_{\theta\theta}=V_{\varphi\varphi}=\frac{u}{r^3}. \tag{4.5}$$

This proves the lemma.

For use below, let us deduce from (4.3) that

$$|\xi(\theta)|<90^\circ. \tag{4.6}$$

Combining the flow rule (2.12) and (4.4), we see that

$$V=qk\sigma A \tag{4.7}$$

where $q>0$ and

$$A=\frac{1}{\sqrt{2}}\begin{pmatrix} -\frac{2}{\sqrt{3}}\cos\xi & \sin\xi & 0 \\ \sin\xi & \frac{1}{\sqrt{3}}\cos\xi & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}}\cos\xi \end{pmatrix}. \tag{4.8}$$

By equating the two representations for $V_{\varphi\varphi}$ provided by (4.5) and (4.7), we conclude that $u(\theta)$ is a positive multiple of $\cos \xi(\theta)$; thus (4.6) follows from (4.3).

Substitution of (4.2), (4.4) into (2.3), (2.12) leads to a system of first-order ODEs with σ, ξ, and u as the unknown functions of θ. The independent variable θ ranges over the interval $(0, \theta_w)$, and boundary conditions are imposed at both endpoints. The boundary condition at the centerline $\theta = 0$, which derives from symmetry, is $T_{r\theta} = 0$. By (4.4), this means that $\sin \xi$ vanishes for $\theta = 0$; in view of (4.6), we require that

$$\xi(0) = 0. \tag{4.9}$$

The boundary condition at the wall $\theta = \theta_w$, which derives from the law of sliding friction, is

$$\frac{T_{r\theta}}{T_{\theta\theta}} = \tan \delta_w \tag{4.10}$$

where δ_w is a constant (called the angle of wall friction). Substituting (4.4) into (4.10), we obtain the boundary condition

$$\xi(\theta_w) = \xi_w \tag{4.11}$$

where ξ_w satisfies

$$\frac{k \sin \xi_w}{\sqrt{2} + \dfrac{k}{\sqrt{3}} \cos \xi_w} = \tan \delta_w. \tag{4.12}$$

Although there are two distinct solutions ξ_w of (4.12), only one of them is consistent with (4.6).

(b) *Conditions Needed to Avoid the Unstable Regions*

In this subsection we derive the conditions needed to guarantee that for the radial solution, the parameters k, a_2 lie entirely inside region 3 of Fig. 3.1, the stable region. This derivation is based on the following formula expressing a_2 in terms of ξ (cf. (3.9a)):

$$a_2 = \frac{1}{\sqrt{6}} \cos \xi. \tag{4.13}$$

To justify (4.13), recall that a_2 is the middle eigenvalue of $|V|^{-1} V$. Since the matrix A in (4.8) satisfies $|A| = 1$, we conclude from (4.7) that $|V|^{-1} V = A$. It may be seen from (4.8) that the 3, 3-entry of A, the RHS of (4.13), is the middle eigenvalue of A.

We claim that for the radial solution to lie inside region 3 of Fig. 3.2, we must have

$$\delta \leqslant \sin^{-1} \frac{1}{\sqrt{3}} \approx 35.3°. \tag{4.14}$$

Note from (4.6) and (4.13) that $a_2 > 0$. Throughout the component of region 3 of Fig. 3.1 which sits in the right-half plane $\{a_2 > 0\}$, we have $k \leqslant \sqrt{2/3}$. Formula (4.14) follows on recalling that $k = \sqrt{2} \sin \delta$.

Suppose (4.14) holds. The radial solution remains inside region 3 if and only if $a_2 \geqslant k/2$. Using (4.13) to express this condition in terms of ξ, we derive the condition

$$(\forall \theta)\ \xi(\theta) \leqslant \xi_{\max} \tag{4.15}$$

where

$$\xi_{\max} = \cos^{-1}(\sqrt{3} \sin \delta).$$

We will use (4.15) to derive bounds for δ_w and θ_w. However, first note from (4.9) that (4.15) is necessarily satisfied near the axis; i.e., near the axis the radial solution lies in region 3. (By contrast, if (4.14) is *not* satisfied, the radial solution never enters region 3.)

Applying (4.15) at the wall, we deduce that ξ_w must satisfy $\xi_w \leqslant \xi_{\max}$, which by (4.12) leads to an upper bound on δ_w. Representative values for this upper bound are tabulated in Table I. These values show how smooth the walls must be for the radial solution to belong to region 3.

The two point boundary value problem for $\sigma(\theta)$, $\xi(\theta)$, $u(\theta)$ with boundary conditions (4.9), (4.11) may be solved by the shooting method. For $\delta = 30°$, typical trajectories generated by this method are illustrated in Fig. 4.1. Note that $\xi(\theta)$ increases for small θ but decreases for larger θ. Because of this lack of monotonicity, (4.15) leads to an upper bound on θ_w. For example, suppose $\delta = 30°$, in which case $\xi_{\max} = 30°$. In Fig. 4.1, the trajectory tangent to the line $\xi = \xi_{\max}$ assumes its maximum at $\theta \approx 24°$.

TABLE I

Upper Bounds for δ_w

δ	Largest value for δ_w such that radial solution lies in region 3
20°	13.8°
25°	13.7°
30°	11.3°
35°	2.8°

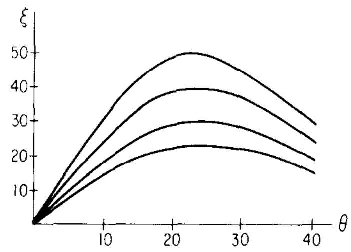

FIG. 4.1. Trajectories in the shooting method (θ and ξ in degrees).

Therefore, if δ_w equals its maximum value of 11.3° (so that $\xi(\theta_w) = \xi_{max}$), condition (4.15) is satisfied if and only if $\theta_w \lesssim 24°$. Of course if $\delta_w < 11.3°$, then somewhat larger values of θ_w are permissible.

5. CONCLUDING REMARKS

(a) *Instabilities with Other Constitutive Laws*

In this subsection we discuss without proofs the instabilities which arise in the evolution equations based on some of the alternative constitutive laws mentioned in Section 2(d). We consider in sequence (i) Tresca's yield surface, (ii) Spencer's flow rule, and (iii) equations which include the effect of compressibility. (Elastic-plastic theory is discussed in Subsection (b).)

Tresca Yield Condition. We analyze the evolution equations only for the flow rule associated to stresses satisfying

$$\sigma_3 < \sigma_2 = \sigma_1 = \frac{1 + \sin\delta}{1 - \sin\delta}\,\sigma_3; \tag{5.1}$$

for example, (5.1) would hold for the flow in an axisymmetric converging hopper [7]. Freeze the coefficients in the equations of motion and retain only the principal part; consider solutions of the resulting equations which have exponential dependence $e^{i(\xi,x) + \lambda(\xi)t}$. A direction $\xi \in \mathbb{R}^3$ belongs to the unstable cone U_{Tr} if and only if there is such an exponential solution with $\operatorname{Re}\lambda(\xi) > 0$. In a coordinate system such that the principal stress axis with eigenvalue σ_3 is (0, 0, 1), we have

$$U_{Tr} = \left\{\xi \in \mathbb{R}^3 : \frac{1 - \sin\delta}{1 + \sin\delta}\,\xi_3^2 \leqslant \xi_1^2 + \xi_2^2 \leqslant \xi_3^2\right\}.$$

(Warning: Although U_{Tr} is invariant under rotations about the ξ_3-axis, the individual eigenvalues $\lambda(\xi)$ are not.) Thus the topology of U_{Tr} is the same as that of the von Mises cone U_{vM} when k, a_2 lie in region 1 of Fig. 3.1; in particular, U_{Tr} is always nonempty, so the Tresca evolution equations are always linearly ill posed. The boundary of U_{Tr} corresponds to the well-known [7] stress and velocity characteristics of the steady state equations.

Spencer's Flow Rule. We consider only two dimensions. We claim that the evolution equations derived from Spencer's flow rule are degenerate elliptic. Thus for almost every direction ξ, there is a growing exponential solution of the frozen equations. Let us elaborate. For a given spatial dependence $e^{i(\xi,x)}$, there are two associated eigenvalues $\lambda_i(\xi)$, $i = 1, 2$. (Remark: For (2.1) restricted to two dimensions, there is only one such eigenvalue. The extra eigenvalue here comes from the fact that Spencer's replacement for (2.12) contains the time derivative of the stress tensor, i.e., is an evolution equation itself.) The evolution equations, a first-order system, are elliptic in the sense that (i) $\lambda_i(\xi)$ is homogeneous in ξ of degree one, (ii) $\operatorname{Re} \lambda_1(\xi) = -\operatorname{Re} \lambda_2(\xi)$, and (iii) $\operatorname{Re} \lambda_i(\xi)$ is nonzero for all ξ except for the two directions which make an angle $\pm(45° - \delta/2)$ with the minor stress axis. The two degenerate directions correspond to (normals to) the characteristics of the steady state equations. (The steady state equations are hyperbolic and have double characteristics.)

Equations Including Compressibility. When compressibility is included in the equations of motion, (2.1b) is perturbed as in (2.18); in other words, the equations including compressibility are a small, but singular, perturbation of (2.1). (By contrast, the evolution equations derived from both Tresca's yield condition and Spencer's flow rule must be regarded as large perturbations of (2.1).) When freezing coefficients in (2.18), we want to retain *both* the term div v, which dominates at low wave numbers, and the principal part, which dominates at high wave numbers. Thus, unlike the preceding two cases, here the growth rates $\lambda_i(\xi)$ of exponential solutions will *not* be homogeneous in ξ.

As with Spencer's flow rule, we consider only two dimensions. Since we have not yet calculated the growth rates $\lambda_i(\xi)$, we are forced to rely on plausibility arguments. With this qualification, we claim that

(i) at low wave numbers ξ, exponential solutions grow as predicted by (2.1); and

(ii) at high wave numbers ξ, the uncontrolled growth of (2.1) is cut off by the perturbation P in (2.18).

The basis for (i) is the fact, already mentioned, that at low wave numbers, (2.18) is a small perturbation of (2.1b). The basis for (ii) comes from consideration of the steady state equations. For (2.1) in two dimensions, the cone of unstable directions is bounded by (normals to) the stress and velocity characteristics of the steady state equations. However, in the steady state theory including compressibility, the stress and velocity characteristics *coincide*! This fact suggests that for the evolution equations including compressibility, the cone of unstable directions shrinks to a single line which is neutrally stable. Of course this loose plausibility argument needs to be replaced by precise analysis of the growth rates $\lambda_i(\xi)$.

(b) *Relation of the Instability to Oscillatory Motion*

The results of this paper suggest that solutions of (2.1) will deviate dramatically from solutions of the steady state equations. Indeed, on the linear level, one would expect the solution of (2.1) to grow uncontrollably. The following estimate indicates that on the nonlinear level, the behavior is less extreme.

LEMMA 5.1. *Modulo boundary terms in an integration by parts, smooth solutions of* (2.1) *satisfy*

$$\frac{d}{dt}\left\{\frac{\rho}{2}\int v^2\,dx\right\} \leqslant \rho g_i \int v_i\,dx. \tag{5.2}$$

Remark. Physically, (5.2) says that the rate of increase of kinetic energy is bounded by the rate of decrease of potential energy. (The difference is dissipated by friction.)

Proof. We compute from (2.1a) that

$$\frac{d}{dt}\left\{\frac{\rho}{2}\int v^2\,dx\right\} = -k\int v_i \frac{\partial}{\partial x_j}(\sigma\,|V|^{-1}\,V_{ij})\,dx$$
$$-\int v_i \frac{\partial \sigma}{\partial x_i}\,dx + \rho g_i \int v_i\,dx.$$

By (2.1b), the middle term on the right integrates to zero, modulo boundary terms. Integrating by parts in the first term and recalling the minus sign in (2.2), we obtain

$$-k\int \sigma\,|V|^{-1}\,V_{ij}V_{ij}\,dx = -k\int \sigma\,|V|\,dx.$$

Since this quantity is nonpositive, the proof is complete.

Remark. Recall that in a silo such as illustrated in Fig. 1.1, Eq. (2.1) describes the flow only in the converging hopper. Since the flow in the hopper is coupled to the flow in the bin, the physically relevant boundary value problem involves more than (2.1) alone. Therefore we have neglected boundary terms in the above estimate. However, we mention that the appropriate boundary conditions on the hopper walls dissipate energy.

Even if the initial value problem for (2.1) is well posed, solutions of this equation will probably behave erratically. In particular, it seems likely to us that as time evolves, some of the assumptions in the derivation of (2.1) may cease to hold. Let us elaborate. In (2.1), the function σ must be positive; likewise, although the function q in (2.12) was eliminated in deriving (2.1), this function must also be positive. However, since the instability amplifies Fourier modes at high wave numbers, one expects the solution to develop a highly oscillatory profile. The oscillations may grow until the minimum of either σ or q is forced to zero, thereby invalidating the derivation of (2.1). If σ is so forced, voids will develop in the material; indeed, such voids may be seen in certain plane strain hoppers. If q is so forced, elastic properties of the material will become relevant and the flow must be studied with elastic–plastic theory.

Based on the experimental fact that granular flow in silos is typically pulsating, we conjecture that there are time periodic solutions of the elastic–plastic equations. We expect that, starting from rest, such a solution would be plastic (i.e., satisfy (2.20) because $|\text{dev } T| = k\sigma$) during part of a period and would be elastic (i.e., satisfy (2.20) because $q = 0$) during the remainder of a period. When $|\text{dev } T| = k\sigma$, the elastic–plastic equations are a small, but singular, perturbation of (2.1); thus during the plastic part of the period, such a solution would tend to grow as predicted by (2.1). (However, at high wave number the singular nature of the perturbation will probably limit the growth rate. A similar cutoff was found for the compressible equations. There is a need to investigate which theory provides the greater stabilizing influence.) During the elastic part of the period, frictional contact with the walls would bring the material to rest again.

The conjecture suggests a host of other problems—principally to use homogenization to derive effective equations for the time averaged stress and velocity. It would also be desirable to calculate the amplitude and period of the oscillations; this information would be useful in designing silos and the calculation would provide a good test of the theory. It seems that these problems must be solved before mathematics can be effectively applied to silo design.

ACKNOWLEDGMENTS

I am grateful to Roy Jackson for introducing me to granular flow; to Walter Strauss for permission to publish separately the derivation of (2.1), much of which was worked out in collaboration; to Dennis DeTurck for calling my attention to reference [10]; and to Bruce Pitman for supplying the calculations on which Fig. 4.1 is based. My debt to Andy Jenike, who contributed many hours patiently explaining the issues in granular flow, is hard to exaggerate; to the extent that I have succeeded in focusing on questions of genuine practical significance, the credit lies with him.

REFERENCES

1. S. AGMON, A. DOUGLIS, AND L. NIRENBERG, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions, II. *Comm. Pure Appl. Math.* **17** (1964), 35–92.
2. G. DE JOSSELIN DE JONG, The double sliding, free rotating model for granular assemblies, *Geotechnique* **21** (1971), 155–163.
3. A. DRESCHER, An experimental investigation of flow rules for granular materials using optically sensitive glass particles, *Geotechnique* **26** (1976), 591–601.
4. G. DUVANT AND J. L. LIONS, "Inequalities in Mechanics and Physics," Springer-Verlag, New York/Berlin, 1976.
5. R. JACKSON, Some mathematical and physical aspects of continuum models for the motion of granular materials, pp. 291–337 *in* "Theory of Dispersed Multiphase Flow" (R. Meyer, Ed.), Proc. of Adv. Sem. Math. Research Center, Univ. Wisconsin, May, 1982. Academic Press, New York/London, 1982.
6. A. JENIKE, "Gravity Flow of Bulk Solids," Bulletin No. 108, Utah Eng. Expt. Station, Univ. of Utah, Salt Lake City, 1961.
7. A. JENIKE, Steady gravity flow of frictional–cohesive solids in converging channels, *J. Appl. Mech.* **31** (1964), 5–11.
8. A. JENIKE, manuscript in preparation.
9. A. JENIKE AND R. SHIELD, On the plastic flow of Coulomb solids beyond original failure, *J. Appl. Mech.* **26** (1959), 599–602.
10. J. LERAY, "Hyperbolic Differential Equations," Lectures at the Inst. for Adv. Study, Princeton, N. J., 1953.
11. A. MAJDA, "Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables," Springer-Verlag, New York/Berlin, 1984.
12. M. MEHRABADI AND S. COWIN, On the double-sliding, free rotating model for the deformation of granular materials, *J. Mech. Phys. Solids* **29** (1981), 269–282.
13. W. PRAGER, "Introduction to the Mechanics of Continua," Ginn, Waltham, Mass., 1961.
14. A. J. M. SPENCER, A theory of the kinematics of ideal soils under plane strain conditions, *J. Mech. Phys. Solids* **12** (1964), 337–351.
15. A. J. M. SPENCER, Deformation of ideal granular materials, *in* "Mechanics of Solids," Rodney Hill 60th Anniversary Volume (H. Hopkins and M. Sewell, Eds.), pp. 607–652, Pergamon, Elmsford, N.Y., 1981.