Stress transmission and incipient yield flow in dense granular materials

Raphael Blumenfeld

Earth Science and Engineering, Imperial College London, London SW7 2AZ, UK
and
Cavendish Laboratory, J J Thomson Avenue, Cambridge CB3 0HE, UK

Abstract. Jammed granular matter transmits stresses non-uniformly like no conventional solid, especially when it is on the verge of failure. Jamming is caused by self-organization of granular matter under external loads, often giving rise to networks of force chains that support the loads non-uniformly. An ongoing debate in the literature concerns the correct way to model the static stress field in such media: good old elasticity theory or newcomer isostaticity theory. The two differ significantly and, in particular in 2D, isostaticity theory leads naturally to force chain solutions. More recently, it has been proposed that real granular materials are made of mixtures of regions, some behaving elastically and some isostatically. The theory to describe these systems has been named stato-elasticity.

In this paper, I first present the rationale for stato-elasticity theory. An important step towards the construction of this theory is a good understanding of stress transmission in the regions of pure isostatic states. A brief description is given of recently derived general solutions for 2D isostatic regions with nonuniform structures, which go well beyond the over-simplistic picture of force chains.

I then show how the static stress equations are related directly to incipient yield flow and derive the equations that govern yield and creep rheology of dense granular matter at the initial stages of failure. These equations are general and describe strains in granular materials of both rigid and compliant particles.

Keywords: Stress transmission, failure, yield
PACS: 45.70.-n, 83.80.Fg

INTRODUCTION

The ubiquity of granular materials in nature and their overwhelming technological significance have focused much attention on these systems. A coherent fundamental understanding of these materials in general and of their behaviours in specific regimes in particular, is yet to emerge. The lack of such understanding handicaps effective modelling for practical applications. As matters stand currently, any progress towards a first-principles modelling has significant scientific and applied benefits. Particularly problematic to model are dense granular materials (DGM). One reason that constitutive relations in the dense regime have been difficult to obtain is that, although the material may appear uniform to the eye above some scale, the internal stress distribution is not - DGM transmit stresses non-uniformly along force networks[1, 2, 3, 4]. The network re-organises as the material is loaded by external forces, with its structure continually failing and re-consolidating. The configurations of the stress networks determine, in turn, where failure occurs, as well as the dynamics as the material deforms. Thus, ultimately, stress determination and dynamics must be modelled self-consistently; we are currently still far from this goal.

Nevertheless, a significant step in this direction is a fundamental understanding of the way that granular matter transmits stresses when the structure is given. This problem has been debated much in the literature. Conventional models, based on constitutive relations involving strain or strain rate, are hard-pressed to explain the observed nonuniform stress fields. Therefore, it was suggested in the mid-nineties that these could be explained by assuming that DGM are statically determinate, or isostatic[5, 6, 7, 8]. Following this suggestion, phenomenological continuous stress-structure constitutive relations were proposed [9, 10, 11, 12, 13, 14, 15]. It was assumed that there are linear relations between the stress components, with the coefficients somehow related to the local structure. In two-dimensional systems, there is only one relation (there are three in three dimensions):

$$A\sigma_{xx} + B\sigma_{xy} + C\sigma_{yy} = 0 \ . \tag{1}$$

In the treatments that followed this suggestion, the coefficients A, B and C were taken to be constant and uniform throughout the material. It has been proposed that, if these relations lead to hyperbolic set of stress equations - in

contrast to the elliptic nature of the equations of elasticity theory, for example - then the solution would propagate along force chains, thus reproducing the experimental observations in two dimensions. Since the stress coefficients in the stress-structure relations were constant, the characteristic paths, and consequently the force chains, were unavoidably straight.

This treatment remained phenomenological until the end of the twentieth century and quite controversial on many levels - a controversy that will be discussed in slightly more detail below. In 2002, this formalism received a firmer theoretical support when equation (1) was derived from first principles, relating the coefficients A, B and C directly to the local granular structure[16]. This formulation did away with the phenomenological aspect and made the two-dimensional theory parameter-free. The constitutive equation (1) was then rewritten in the form

$$Q : \sigma = 0 \,, \tag{2}$$

where Q is a rank two tensor that depends uniquely on the local granular structure. In a later work [17] it was shown that the geometry constrains the components of the structure tensor Q in such a way that the equations are indeed hyperbolic, as had been conjectured in previous works[8, 9].

That relation (2) can be derived from first principles is intriguing - constitutive closure relations are normally postulated. Nevertheless, the basic derivation in [16] was a not sufficient. To clinch the applicability of this relation it is necessary to be able to coarse-grain it to the macroscale. This proved to be a problem: the components of the fabric tensor Q were shown to fluctuate locally around zero mean, hindering coarse-graining by simple volume averages. However, it was observed that the fluctuations are anti-correlated between neighbouring particles, much in the same way that disordered antiferromagnetic spin systems are between neighbouring spins. This observation was then used to reformulate the problem in [18] in a way that could be coarse-grained. This resolved the problem, finally establishing equation (2) as a general closure relation for ideally isostatic granular materials of rigid particles. It should be noted that, once the fabric tensor Q is given, a unique solution exists for any arbitrary combination of boundary stress loads [19, 20].

Nevertheless, the development of isostaticity theory has been fraught with other doubts and criticism. Granular assemblies in nature are not necessarily isostatic and the usefulness of this concept to general particulate systems has been questioned[21, 22]. In the original formulation the particles were considered perfectly rigid and convex, hence making contacts at points. This led to concerns over the validity of the theory when the particles have finite compliance. This issue was examined in detail in [23], where it was shown that increasing particles compliance does not invalidate isostaticity theory altogether. Rather, the theory needs to be corrected and the corrections increase with particles compliance and contact area between particles. However, significantly, the corrections decrease rapidly with increasing system size.

To be specific, consider a granular assembly of N stiff, but not perfectly rigid, particles. The mean coordination number is such that, had the particles been perfectly rigid, the assembly would be isostatic. The analysis in [23] shows that, under a given external loading, the stress solution in the material would be the isostatic solution, but with corrections. This was shown as follows. First, the centroid point of each contact surface was identified and an alternative approximate system was considered, whose inter-granular contacts were points located at those centroid. To choose the approximate system, it was noted that the exact solution could be found by solving for an equivalent system whose contact forces acted at contact points located judiciously across the contact surfaces. The location of the equivalent contact points, however, needs to be determined by the effective surface contact force distributions, which are unknown. In the approximate system, the contact points are chosen at the centroids of the contact surfaces.

Now, the alternative approximate system is also marginally rigid and, therefore, it has an isostatic stress solution. The error between this solution and the correct one is due to the differences between the approximate and correct equivalent contact points. This error was then shown to decreases as N^{α}, where $\alpha = 3/4$ and $8/9$ in two and three dimensions, respectively[23]. It was therefore concluded that, on macroscopic sizes, the isostatic solutions of the approximate systems converge to the correct stress solutions of compliant-particles systems.

Another concern had been the range of applicability of isostaticity theory for materials that are not precisely at marginally rigid stress states. The problem was that the theory has been developed under the assumption that the mean coordination number is exactly right to make the material statically determinate. The contention was that the status of the theory is unclear when the mean coordination number goes above that value. This is a valid contention - it is clear that the equations can no longer hold as originally formulated. However, it turns out that the way that the solutions deteriorate, whether abruptly or gradually, is an important issue. This issue will be discussed later in the paper, following insight drawn from experimental observations [24]. That discussion will be shown to lead to two

conclusions: (i) that the isostatic, or marginally rigid, state behaves as a critical point and (ii) that the stress behaviour of real materials depends on the proximity to the critical point [17, 25].

A third concern is seemingly fundamental. It had been conjectured [8, 12, 13] and then shown more rigorously [17] that (2) leads to hyperbolic stress equations, whose solutions propagate along characteristics paths (CPs) (see below). The CPs are determined by the structure, namely, the tensor Q and they define preferred directions. The question then is how can this happen in isotropic systems. Unfortunately, this question has not been discussed in great detail in the literature. The view of this author is that the structure can only be isotropic under isotropic loading. DGM respond to applied loading by local rearrangement and self-organisation. The material response to anisotropic loading cannot but be anisotropic, leading to a locally anisotropic fabric tensor. Thus, it is rare to find, either in nature or in simulations, DGM that are isotropic. However, we can consider a 'gedunken' isotropic material. To generate such a system one should start from an isotropic structure and apply to it isotropic boundary loads, e.g. purely compressive σ_{rr} along the boundary of a circular disc. Then every point along the boundary is a stress source of two CPs that propagate into the system. This gives rise to an isotropic distribution of CPs in the material and, correspondingly to an isotropic solution. Thus, isostaticity theory should not violate any common sense symmetry.

An interesting simulation to test this issue [26], claiming to the contrary, is nevertheless inconclusive. Such tests should make sure of the isotropy by: (i) loading the system isotropically (i.e. independent of direction rather than equally along the axes, which leads to isotropy in fluids but not in DGM) and (ii) that the isotropy is reflected in the statistics of the fabric tensor Q. Thus, this contention still needs to be substantiated.

It should be noted that the above argument suggests that a complete stress theory of granular materials should ultimately take into consideration the material organisation under the loading, a point that has been alluded to above.

Turning to failure and yield of DGM, it should be commented that the manner in which DGM transmit static stresses is significant to the understanding of these materials well beyond static phenomena. Not only do DGM exhibit simultaneously properties that are normally associated with solids, liquids and gases, but they also display rich behaviour that is uniquely their own and cannot be observed in other conventional phases. Thus, static stress transmission is an obvious first step to study the poorly understood (and therefore inadequately modelled) dynamics of DGM. One particular area, where modelling could benefit much from improvement in fundamental understanding, is of flow and rheology. The importance of flow of DGM to many real-life applications cannot be over-emphasized. For example, in powder metallurgy and the transport of cereals, seeds and coal, it is important that the material fail easily locally so that it can flow without blockage. Yet, in other application, such as construction, it is essential that soils be sufficiently stable so that structures do not collapse. In both these cases we need a good understanding of the basic physics in order to model: (i) the threshold to failure; (ii) the failure mechanisms and (iii) the post-failure dynamics of flow.

The main problem is that, unlike in fluid mechanics, where the continuum equations of motion are well established - the Navier-Stokes equations - there is no agreement on the continuous equations that describe rheology of DGM. A major stumbling block is that the homogenization from the particle-level dynamics to the macro-scale is sensitive both to the physics on the discrete level and to the coarse-graining procedure used to upscale to the continuum. Moreover, unlike in conventional fluids, which have no yield threshold to flow, most real DGM flow only once they have been loaded above some stresses, known generally as the 'yield surface'. Thus, the failure of DGM is affected directly by the way that the material supports static stresses. This further suggests that an essential criterion for a good model is a seamless convergence to the correct stress state when the flow stops.

In this paper, I will first review a simple experiment that leads to intriguing two conclusions: first, that isostatic states can be approached very closely, or even exactly realised, in real systems and, second, that isostatic states act as critical points in the sense that they sport a diverging lengthscale. The physical interpretation of this lengthscale and its relevance to understanding the physics of DGM will be discussed. I will then describe a recent development on modelling failure and yield of DGM. Specifically, I will present a derivation of the equations of flow as the material fails. The relation between these equations and the convergence to isostatic static stress solutions will be shown and discussed, thus establishing a seamless transition from the static to the quasi-static dynamic theories. Many of the results reported here have been obtained in collaboration with R. C. Ball and S. F. Edwards.

169

A TWO-DIMENSIONAL FREE-FALL PILING EXPERIMENT

Mean coordination number and marginal rigidity

In this section I describe the experiment reported in [24] and discuss an interesting interpretation of its results. The main aim of the experiment was to investigate the realizability and nature of marginally rigid states. Before getting down to the experiment, it is constructive to discuss the conditions that define such states.

It is clear that there is a minimal number of contacts per particle (called the mean coordination number), below which the system cannot be mechanically stable, except for a measure zero of metastable configurations. Marginally rigid states are mechanically stable configurations of particles that are minimally connected. In such states the inter-granular forces are statically determinate, a fact that can be used to identify the condition for marginal rigidity. This condition depends on several characteristics of the particles, including whether they are frictional or frictionless, and spherical or not. To some extent, it also depends on system size, but limiting the discussion to macroscopically large assemblies, where this dependence is negligible, I will disregard it in the following.

Consider a d-dimensional assembly of N ($>> 1$) rigid particles of sufficiently high friction coefficient, such that negligibly few particles slipped as it was constructed. Let z_g be the number of force-carrying contacts (defined here as the coordination number) of particle g. The corrections due to size effects are of order $N^{-(d-1)/d}$ in d dimensions and hence negligible. The mean coordination number per particle is

$$\bar{z} = \frac{1}{N} \sum_{g=1}^{N} z_g , \tag{3}$$

and the total number of inter-granular contacts is therefore

$$N_{cont} = \frac{1}{2} N \bar{z} . \tag{4}$$

At each force-carrying contacts there is an inter-granular force to be determined, giving altogether dN_{cont} unknowns. To solve for these we have at our disposal the balance conditions. For each particle there are d force, and $d(d-1)/2$ torque moment, balance equations. The latter is the number of axes of rotations in d dimensions. The total number of equations is then

$$N_{eq} = N[d + d(d-1)/2] = N \frac{d(d+1)}{2} . \tag{5}$$

For the system to be statically determined, the number of equations should equal the number of unknowns. From expressions (4) and (5) we see that this gives a condition for the mean coordination number

$$\bar{z} = z_c^f = d + 1 . \tag{6}$$

In assemblies of frictionless particles, the direction of the force vector at each contact point is determined by the geometry (specifically, by the normal direction at the local contact), which leaves the magnitude of the force as the only unknown. It follows that the number of unknowns is $N_{cont} = \frac{1}{2} N \bar{z}$, while the number of equations is the same as in (4). Again demanding that the number of equations be equal to the number of unknowns, we find that the condition for static determinacy is that the mean coordination number is

$$\bar{z} = z_c^s = d(d+1) . \tag{7}$$

A redundant case to be considered is when the particles are smooth and perfectly spherical. This case is rare in nature but it is used frequently for numerical simulations. In this case, the number of unknowns is again equal to the number of contacts, but the number of equation is not the same as in the previous cases. Due to the shape of the particles, when the forces are balanced on a particle, the torque balance equations for it are also automatically satisfied, since all the force direction lines pass through the particle center. This makes the torque balance equations redundant, leaving only Nd equations to solve. It follows that, for static determinacy in such systems, the mean coordination number should be

$$\bar{z} = z_c^{ss} = 2d . \tag{8}$$

170
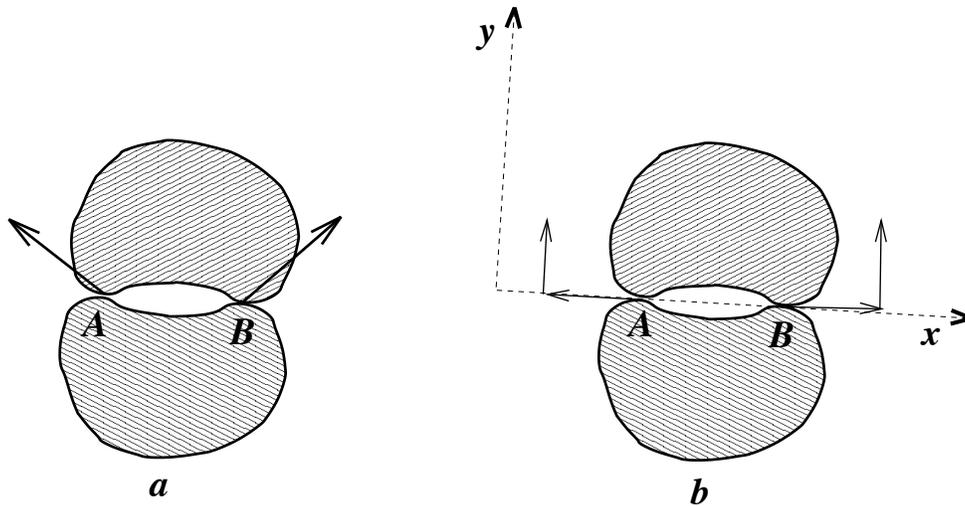
FIGURE 1. Two rigid particles in contact through two points A and B. The two contact forces in (a) can be decomposed into the four components in (b), of which only three can be resolved. Since the two x-components of the forces are equal, opposite, and act along the same line, then the stress around the two particles is unaffected by this indeterminacy. Consequently the mean coordination number is larger than for systems with only convex particles.

The above considerations have been around for a long time, in fact some of these arguments date back to Maxwell[27, 28]. Nevertheless, they do not hold for all packings. In particular, they include a couple of underlying assumptions. One is that in particles of arbitrary shape there is no redundancy of any of the torque moment equations. For example, if the contact force vectors around a particle all happen to pass through one or more of its axes of rotation then the torque around those axes ares automatically zero and there is no need to invoke the relevant couple equations. In the following I assume that such occurrences are very rare, in fact of measure zero. Another assumption is that the particles are rigid and convex and, consequently, that a particle can only make contact with any neighbour at one point. This assumption fails when particles have non-convex shapes, whereupon two particles can make contact at two points. To understand the consequence of lifting the one-contact-point assumption, consider a piling experiment, where particles with high friction coefficients 'fall' very gently, so that they come to rest to form a stationary pile without slipping and without disturbing the rest of the pile. In d dimensions, any solid, and in particular both the falling particle and the consolidated pile, has $d(d+1)/2$ degrees of freedom: motion along one of the d directions and rotation around the $d(d-1)/2$ axes of rotation. Thus, prior to the collision the particle-pile system has $d(d+1)$ degrees of freedom. Once the particle has settled, the new consolidated pile has only $d(d+1)/2$ degrees of freedom. This means that $d(d+1)/2$ degrees of freedom have been 'lost' and these provides $d(d+1)/2$ conditions. From these conditions we can solve for the forces between the recently-arrived particle and the pile. At each contact Newton's third law (action and reaction) is satisfied, which gives that there are $d+1$ unknowns per particle to solve for and, when there is only one point of contact between particles, then we could solve for the inter-granular forces if there are on average $d+1$ contacts per particle. This is an alternative derivation the above value of z_c^f.

However, when the shape of the landing particle has a non-convex part, the particle may lean stably against only one particle, in which case the above argument needs to be reconsidered. For simplicity, the situation is illustrated in two dimensions in figure 1. Particles contacting at two points transmit two inter-granular forces and hence give rise to four unknowns. This is one unknown more than the $d(d+1)/2 = 3$ conditions that can be derived from the 'loss' of the degrees of freedom. Therefore, one unknown must remain indeterminate. Fortunately, this does not affect the global stress field, as we shall see next. Let us decompose the forces along axes x and y, where x coincides with the line that joins the two contact points, as shown in figure 1. We can solve for the y-components and use Newton third law to determine that the x-component of one force is equal and opposite to the x-component of the other. The remaining one unresolved unknown is the magnitude of the x-component. The local stress on either of the two particles can be derived from the area normalized symmetric part of the force moment

171

$$\sigma_{ij}^g = \frac{1}{A} \sum_{g'=1}^{z_g} f_i^{g'g} \rho_j^{g'g} ,$$ (9)

where σ_{ij}^g is the ij-th component of the stress on particle g, g' are the touching neighbours of this particle, $f_i^{gg'}$ is the i-th component of the force that partice g' applies on g, $\rho_j^{g'g}$ is the j-th component of the position vector extending from the center of mass (say) of particle g to its contact with g', and A is the area associated with particle g. Since the unresolved x-components are equal, opposite and act along a line tangent to the contact points, then they do not contribute at all to the forces, nor to the torque moment around g. It follows that the unresolved degree of freedom does not affect the stress around this particle. This argument, which was first presented in [24], leads to two conclusions. First, the stress around two particles touching at two points may be well defined in spite of the fact that not all the inter-granular forces can be determined. Second, in counting the number of contacts we have to allow for multiple contacts. The latter means that, in a system that contains multiple inter-particle contacts, the mean coordination number for marginal rigidity is higher than the one calculated previously. To find how much higher should z_c^f be, for example, in two dimensions, we note that there is one such contact for every pair of double-contacting particles, increasing the mean coordination by $1/N$ for each one of them. It follows that the correct value for marginal rigidity in two dimensions for a system with a fraction w of double contacts out of all inter-granular contacts is

$$\bar{z} = z_c^f = 3 + w .$$ (10)

This observation will be directly relevant to the experiment to be described next.

The experiment

The experiment models free fall and pile-up of particles in two dimensions (see figure 2). Model noncircular particles of approximately $1.5 \pm 0.2 \text{cm}^2$ were punched out from a cardboard sheet, using a stainless steel punch. The cardboard was fibrous, ensuring that the cut edges gave high coefficient of inter-granular friction. The friction coefficient was estimated from slipping tests to be higher than 50. The particles had a slight non-convex region. The experiment was carried out on a horizontal glass plate, with the pile building up within a U-shaped collector, whose surface was made of a similar material as the particles. The internal dimensions of the collector were $18 \times 24 \text{cm}^2$. The particles were initially placed on a thin transparent film, lying on top of the glass plate, at a notionally random distribution subject to the requirement of no contacts and that the initial spatial distribution is reasonably uniform. These particles were then conveyed towards the collector at an approximate speed of 0.5m/min. This was done either by moving the collector towards the particles (at high initial densities) or by moving the particles towards the collector via sliding the transparent plastic sheet under it (at low initial densities). Care was taken to maintain a constant advance rate and particularly to avoid relative transverse motion between the collector surface and the 'free falling' particles. The slow fall rate ensured both negligible particle deformation and minimisation of inertial effects, which could transmit vibrations through the forming pile and rearrange it subsequently to piling. Data was collected in the form of photographs of the growing pile, taken normal to the plane at regular intervals during the process, and a photograph of the 'consolidated' final pile. For each final pile we have determined its bulk and boundary particles, the contacts and the double contacts. The determination of the number of contacts was done both by eye and by taking the zero limit of the cumulative distribution of the spaces between neighbouring particles. Both methods agreed excellently. Following the discussion of the previous subsection, to determine the proximity to the marginal state, we measured coordination number as twice the total number of contacts divided by the total number of particles. This quantity, if calculating coordination from individual particles, means double counting coordinations between boundary particles and the boundary. Relatively few double contacts were observed and these were counted as two separate contacts. In the experiment the double contacts comprised 10% of the total number of inter-granular contacts. The photographs taken during each piling process were also used to identify a yield front (YF) - a 'layer' of particles that have collided with the pile, but which have not yet come to rest. The particles that belong to this layer were identified by comparing the intermediate positions of the particles to their position in the final pile. The comparison was done by superposing the photographs on top of a lightbox and determining overlaps. The force of sliding friction between the particles and the moving base played a role analogous to gravitational force on a stationary mass. Since the approach was at constant (low) velocity, the falling particles did not accelerate and the situation is indeed of free fall in two dimensions.
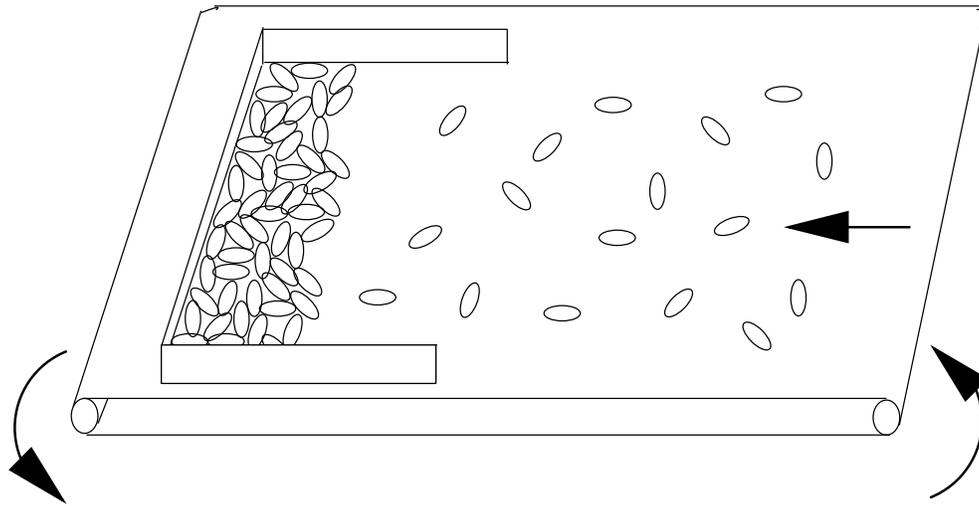
FIGURE 2. A two-dimensional free-fall piling experiment. Particles are conveyed at constant slow speed and pile up against a collector. The particle-particle and particle-collector friction coefficient is very high. The friction between particles and the underlying conveyor mimics gravity to press the pile against the collector. Photographs are taken from below to monitor the piling process.

Measurements were taken of both the density of the consolidated pile (figure 3) and its mean coordination number (figure 4) as functions of the initial density of the falling particles. It can be observed that, as the initial density increases the final pile density decreases. This is characteristic of a jamming system. As particles fall, contacts are made and broken in the YF. With very dilute initial density, particles join the pile and stabilize by making two contacts with it. As the initial density increases, falling particles interfere with one another and not all of them can make two contacts with the stationary pile. Thus, the decrease in the number of contacts, and hence in the final density, is a collective phenomenon - the evolution of the contact network is highly correlated and is governed by the cooperative stress-driven dynamics.

The experiment has a natural limit density, identified as the point where the density of the initial particles is equal to the density of the final pile (see figure 3). Clearly, there can be no experimental measurement to the right of this point since the density of the initial particles cannot exceed that of the consolidated pile that it generates. This density, ϕ_c, is found by extrapolating a fit to the the pile density measurements as a function of the initial density, and locating the intersection with the equal-densities line. It is plausible that this is the density of loose random packing, since it is the absolute limiting lowest density that one can obtain in this experiment.

One significant quantitative result is that, on extrapolating the fit of the mean coordination number to ϕ_c, $\bar{z}$ was found to approach 3.1 ± 0.1, in excellent agreement with the prediction of the marginal rigidity mean coordination number given in (10), combined with the observed value of $w = 0.1$. This leads to the conclusion that isostatic states can be approached arbitrarily closely. This observation has been supported by an independent repeat experiment, carried out with different, and more, particles[29].

Another significant observation concerns the YF. Figure 5 shows three typical experiments at three different initial densities before all the particles have come to rest. Of particular interest is the amount of particles that belong to the YF (coloured red). In the YF, particles are continuously shifting and moving relative to their neighbours en route to a stable mechanical equilibrium. Specifically, observe that the number of the particles that comprise the YF increases rapidly as the initial density increases toward the limit density ϕ_c (see figure 6). At low initial densities, the YF is about one particle thick while close to ϕ_c it is more than 50% of the pile. This suggests that, for macroscopically large packings, the number of particles participating in the YF is a finite fraction of the total number of particles. This conclusion, which needs to be checked for larger assemblies of particles, suggests that the size of the YF *diverges* as the system tends to the 'thermodynamic limit', i.e. when the total number of particles tends to infinity. These observations and conclusions have also been supported by the experiment of Sibille and Mullin[29], also shown in figure 6.

A diverging lengthscale is a fingerprint of a critical point and its occurrence in marginally rigid packings gives an important clue to the nature of the isostatic state. Specifically, diverging lengthscales are common in second order phase transitions and in critical phenomena [30]. To understand the physical phenomenon that this lengthscale

173
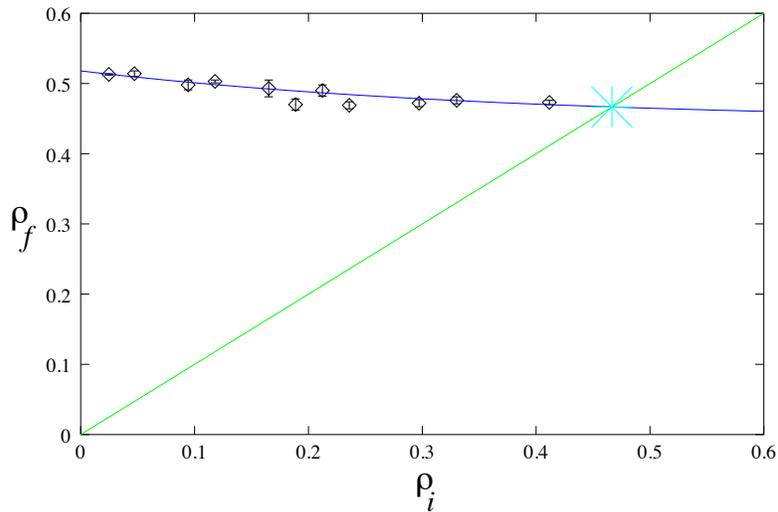
FIGURE 3. The density of the consolidated pile, as a function of the initial density of the free-falling particles. Note that the higher the initial density the lower the density of the final pile. The line fitting the result is a simple exponential and comes both as a guide to the eye and to provide an estimate for the critical density, ρ_c. The value of ρ_c is determined by the intercept of the fitting line and the equal-density line, which passes through the origin.
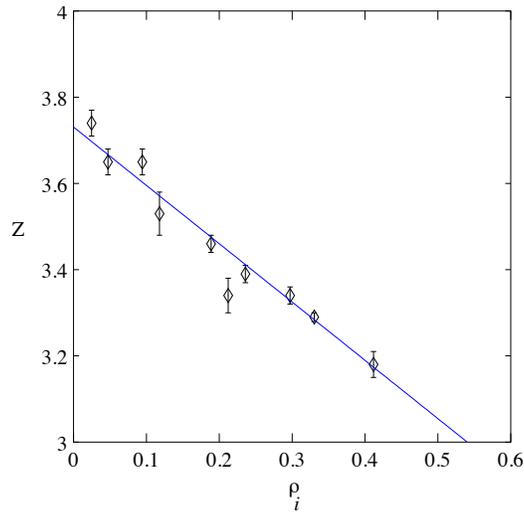
FIGURE 4. The mean coordination number of the consolidated pile, z, as a function of the initial density of the free-falling particles. the number of contacts was found both visually and by taking the zero limit of the cumulative distribution of the space probability density between neighbouring particles. The two method agreed perfectly. A straight line fit was used to extrapolate the value of the mean coordination number to ρ_c. The value of z at ρ_c was found to be 3.1, in excellent agreement with the marginally rigid value for our pile with 10% double contacts.

is associated with, recall that, at the critical point, the mean coordination number is on the verge of becoming mechanically unstable. Perturbing a particle from its position affects the particles in contact with it and they then have to shift position too. This, in turn, affect their neighbours and so on. Thus, the first interpretation is that the diverging lengthscale corresponds to the response length of the marginally rigid pile to small particle displacements. The minimal connectivity of the pile makes every contact breakage major event that typically leads to a long-range rearrangement.

Furthermore, applying a small force to a particle in a marginally rigid pile, one expects the force to be sufficient to break a contact. However, to prevent the contact from breaking we need to adjust the force on the neighbour particle. This in turn would give rise to further adjustments due to the tenuous connectivity. This suggests that the force response

174
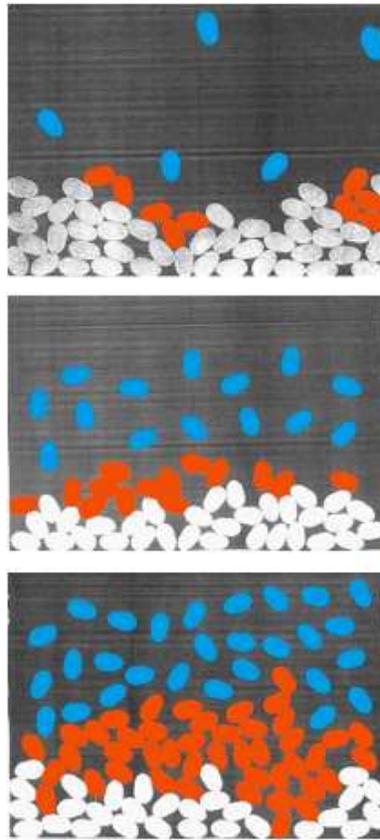
FIGURE 5. The behaviour of the yield front. The particles belonging to the front (coloured red) have encountered the pile but have not yet reached the final consolidated positions. The three different pictures show measurements from three different experiments at initial densities that increase from top to bottom. At the low density the front is less than one particle deep, while at the high density it comprises more than half the pile. As the density increases the consolidation process becomes more cooperative.

length should also become comparable to the system size, i.e. diverge. Note that such a thought experiment probes nothing else but the Green function of the continuous stress theory that applies to these material.

IMPLICATIONS OF THE EXPERIMENT FOR CONTINUUM STRESS THEORY

The conclusions and interpretations of the previous section are significant since they point directly to a resolution of the long-standing debate over the correct theory for real granular materials. It is the critical nature the marginally rigid state and the divergence of the response of the system to stress perturbations, which gives us a big clue.

As mentioned, critical points are well known from the study of traditional second-order phase transitions. A system undergoing such a transition is in one phase initially and, as it approaches the critical point, regions of another phase form and develop. A well-known example is the transition in metals from a normal conducting state to a superconducting one upon reducing the temperature. The temperature where the transition occurs is the critical point. As the temperature decreases toward this point, superconducting regions form, and then grow, within the initially normal phase. At the critical temperature, the superconducting regions span the entire material.

The typical size of the superconducting clusters can be quantified by measuring the the density-density correlation function of the superconducting phase and identifying the correlation length over which this function decays. As the critical point is approached from above, the correlation length increases until it become comparable to the system size at the critical temperature. In systems that go to the thermodynamic limit, i.e. become infinitely (or macroscopically) big, the correlation length diverges. It should be emphasized that, even when the system is only close to the critical point, there are sufficiently large superconducting regions to affect experimental measurements of the conductivity.

175
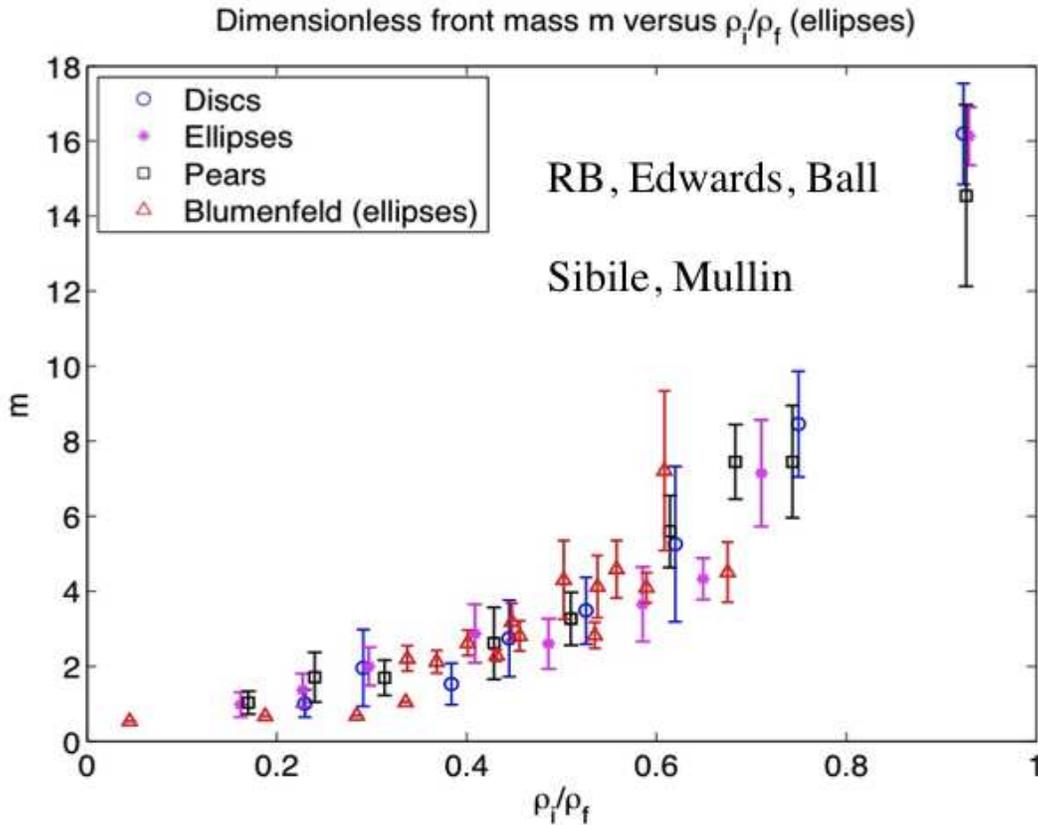
FIGURE 6. The measured front (arbitrarily normalized) mass as a function of the initial density. The plot shows a diverging trend as the ratio of the initial and final densities approach unity at the critical density. (plot courtesy of Sibille and Mullin [29])

Indeed, macroscopic measurements depend sensitively on the proximity of the material to a critical point.

This physical interpretation of the diverging lengthscale is very similar to the situation in DGM. At the critical density ρ_c, the material is minimally connected and hence fully isostatic. The proximity of the granular system to the critical point can be parameterized by $\delta z = \bar{z} - z_c$, the difference of the mean coordination number from the critical value z_c . Finite positive value of δz correspond to existence of over-connected regions in the material. A steady increase in δz from zero corresponds to growth of the over-connected regions at the expense of the isostatic medium.

Thus, we can regard real DGM as two-phase composites: part isostatic and part over-connected. The stress field in the isostatic regions is governed by the isostaticity equations, namely, the balance conditions and (2), whilst the over-connected regions follow a different closure relation.

In principle, it is straightforward to construct a stato-elasticity theory. Consider a material containing a particular distribution of isostatic and over-connected regions. To determine which phase a region belongs to and the boundaries of a region can be done by the local mean coordination number. The exact algorithm for such a determination is downstream from the main thrust of this paper. In a simple version, we can assume that elasticity theory holds in all the over-connected regions, while isostaticity theory holds only in the marginally rigid regions. Then, in each region the stress solutions are written in general form, using the different sets of equations, and the solutions in neighbouring regions are tailored to match at common boundaries. While there are practical problems in this procedure, exact solutions are currently possible for simple geometries [31]. .

This resolves the long-standing controversy in the field on the correct way to model stresses in granular materials - whether with the theory of elasticity or isostaticity. In a sense, both approaches are valid, but in different regimes. Real DGM generically contain both phases: marginally rigid regions where the hyperbolic [17] equations of isostaticity apply and over-connected regions where the elliptic equations of elasticity do. Thus, granular media should be modelled by a combined theory.

176

However, to implement a stato-elasticity theory, one must have a good grasp of each of the separate theories. While a good understanding of elasticity exists, that of isostaticity is still rudimentary. In particular, for a long time, isostaticity theory was associated only with force chains and long response lengths. This changed recently when a more thorough study of nonuniform isostatic media has been carried out [19, 20], showing that the idea of well isolated force chain needs to be revised when the components of the fabric tensor Q are not constant across the medium.

Force chains have been observed experimentally and they occur naturally when equations (1) and (2) are hyperbolic [8, 12, 13, 14]. That these equations are indeed hyperbolic on the smallest scale was shown from first-principles in [17]). The solutions of hyperbolic equations propagate along characteristic paths (CPs) and in the original works the forces were finite along the CPs and vanished elsewhere. Thus, the CPs were equated with the observed force chains. However, this picture is based on analyses of systems where the components of Q were uniform. The recent work of [19] and [20] showed that the situation is more complex because, when Q varies with position, the CPs of the solutions 'couple' and 'interact', with the strength of the coupling proportional to the gradients of the components of Q. The coupling led to several effects. First, and not surprisingly, the CPs meander rather than be straight. Second, the forces along the CPs decay, suggesting that force chains eventually 'evaporate'. Third, forces can 'leak' to the region between the CPs, with the leakage increasing with the magnitudes of the gradients of Q. Fourth, for locally sharp gradients of Q, the CPs can branch. This result was significant because it explained observed branching without having to invoke additional stochastic mechanism [32]. These results form a basis for better understanding of stresses in isostatic regions within the two-phase picture described above.

FAILURE AND CREEP FLOW

In this section, I focus on the physics of failure of DGM and on the initial dynamics immediately after failure. In the engineering community, the term failure is understood on the macroscopic continuum level, where the reference is to observables measured on such scales. However, failure is caused by local rearrangement of individual particles. Some such rearrangement may only lead to local deformation that do not translate to global failure and result only in irreversible behaviour, i.e. reversing the loading conditions does not reproduce the original microstructure. Some local rearrangement may spread out to span large parts of the material, in which case they are regarded as global failure. The conditions for each of these scenarios are complex and beyond the scope of this discussion. Here, I focus on the local failure of the contact network between particles and the initial motion that it results in. The main advantage of this discussion is that it allows to define an exact yield surface and criterion, as will be shown below.

Generically, a granular system fails when the external loading on it exceeds a certain threshold, commonly known as the 'yield surface'. This is a surface in the space spanned by the individual components of the stress tensor and, in natural granular systems, it normally encloses a finite region. However, one can imagine granular systems, whose yield surface degenerates to a point at the origin, namely, it would yield under the smallest external loading.

For failure to occur, regions inside the material must be at, or close to, marginally rigid states. Moreover, macroscopic failure would be preceded by an incipient marginally rigid region extending from boundary to boundary. An example of such a region is an incipient shear band. Therefore, it makes sense to focus on failures in regions that are initially in isostatic stress states. For the purpose of this discussion, I will assume that the granular system consists of convex and rigid particles of high friction coefficient to inter-granular slipping.

Generically, particle displacements occur by two mechanisms: slippage of particles in contact and rotation. In generally random granular assemblies the two mechanisms normally take place simultaneously. In the following, the displacement of a particle centroid is identified with that of a point of the continuous representative material. On the continuous level we can define a continuous displacement field, from which a strain and strain rate fields can be derived. Thus,

$$\dot{e}_{ij} \equiv \frac{1}{2}\left(\partial_j \dot{u}_i + \partial_i \dot{u}_j\right) = \dot{e}_{ij}^{rot} + \dot{e}_{ij}^{sl}, \tag{11}$$

where u_i is the displacement in the ith direction, $\partial_i \equiv \partial/\partial x_i$ and x_i is the ith Cartesian coordinate. The quantities in this equation and in the rest of this section refer to centroids of particles, but particle indices have been omitted to avoid cluttering the notation. For analysis purposes it is useful to treat the rotational and slippage contributions independently. This is not only an exercise in idealization. It is interesting to note that there are special circumstances, where the two mechanisms are indeed separable in the sense that macroscopic strain rate can be achieved, at least momentarily, either by pure slippage or by pure particle rotations. This happens in systems where all loops have even numbers of particles around them. An example of a part of a two-dimensional such a system is shown in figure 7.
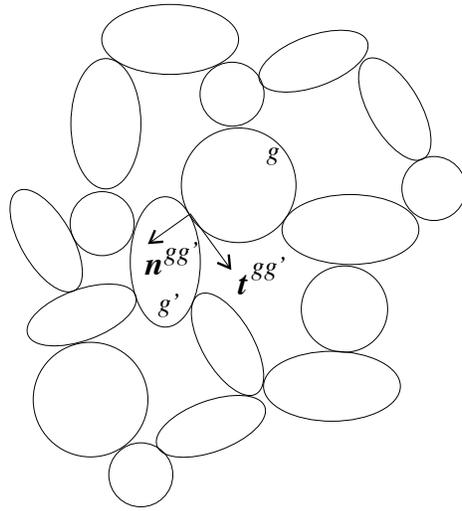
FIGURE 7. An 'unfrustrated' part of a granular assembly in two dimensions. The even number of particles around every void loop makes it possible to divide the structure into two sub-structures, black and white in the figure, such that every particle is in contact only with particles of the opposite colour. Associating spin up with one colour and down with the other, maps this exactly to an unfrustrated antiferromagnetic ground state.

In these systems, the particles can co-rotate happily, with every particle rotating oppositely to its neighbours and no slippage taking place. Such structures I term 'unfrustrated' [16], in analogy to antiferromagnetic spin systems, as explained in the caption of figure 7. It is in unfrustrated systems that the region enclosed by the yield surface shrinks to zero, for such systems can strain without dissipating energy. Exciting mechanically equilibrated unfrustrated systems by perturbative oscillatory loading, they experience what is known as 'floppy modes' [33].

Energy is dissipated when two particles in contact rotate in the same direction, whereupon they must experience mutual solid friction. The situation is analogous to two antiferromagnetically interacting spins, forced to align in parallel - hence the term frustration. As particles of unfrustrated systems rotate, they displace and contacts may be broken and new ones made. Whether the structure remains unfrustrated or not is an interesting and relevant issue that is not addressed in this presentation. To avoid this issue, I will focus on the dynamics before the connectivity changes.

A particular example of unfrustrated systems is the deformed Honeycomb structure sketched in figures 8 and 9. We can generate displacement of the entire right hand side of the structure relative to the left hand side about the shear line B in two ways. One is by making all the particles on the right hand side of B slip over those on the left hand side, as shown in figure 8. The slippage then occurs only along the shear line. Another way is to rotate all the particles on the right hand side, as shown in figure 9, and leave the particle on the left hand side stationary.

Focusing first on the slipping strain rate, let the shear line B pass through the contact between particles g and g' (figure 9) and let the tangent unit vector to their surfaces at the contact point be $\hat{t}^{gg'} = \vec{S}_B$, as in figure 7. The normal to the surfaces of the particles at the contact point is $\hat{n}^{gg'} = \varepsilon \hat{t}^{gg'}$, where ε is a $\pi/2$-rotation operator. In terms of these, the symmetric strain is proportional to

$$\gamma_{ij} = \frac{1}{2} \left(\partial_i \hat{t}_j^{gg'} + \partial_j \hat{t}_i^{gg'} \right) , \tag{12}$$

where $\hat{t}_j$ is the jth component of the unit vector $\hat{t}$. It follows that the slippage strain rate at this contact is

$$\dot{e}_{ij}^{sl} = A\gamma_{ij} , \tag{13}$$

where A is the displacement rate in the tangential direction. As we shall see below, this first-principles formulation of the slippage contribution on the granular level has advantages over traditional plasticity-based models. While γ_{ij} can be determined, in principle, from the local structure, the rate A is an unknown scalar field that needs to be solved from the flow equations.

The rotation-based strain rate is interesting. Consider the local microstructure around a particle g, in contact with particles $g_1, g_2, \ldots, g_n$ (e.g. $n = 3$ in figure 10). We wish to derive the shift in the centroid of the contact points (i.e.

178
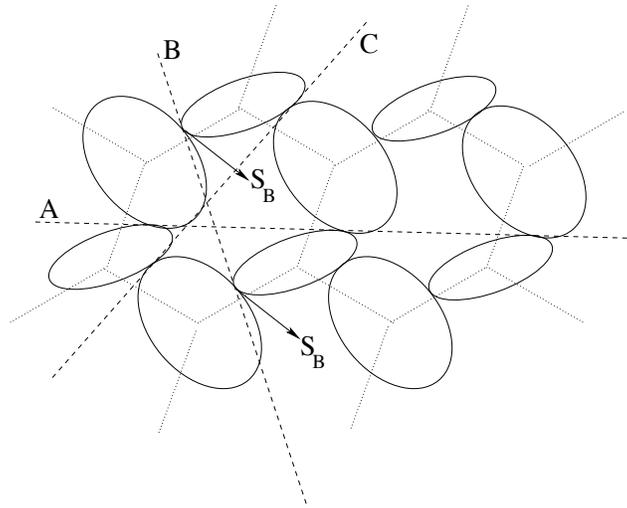
FIGURE 8. An unfrustrated deformed Honeycomb structure of particles can be sheared along three shear lines: A, B, and C. For example, displacing the half-system to the right of shear line B relative to the left hand side, can be achieved by slipping the contacts along the shear lines in the S_B direction, which is determined by the geometry at these contacts.
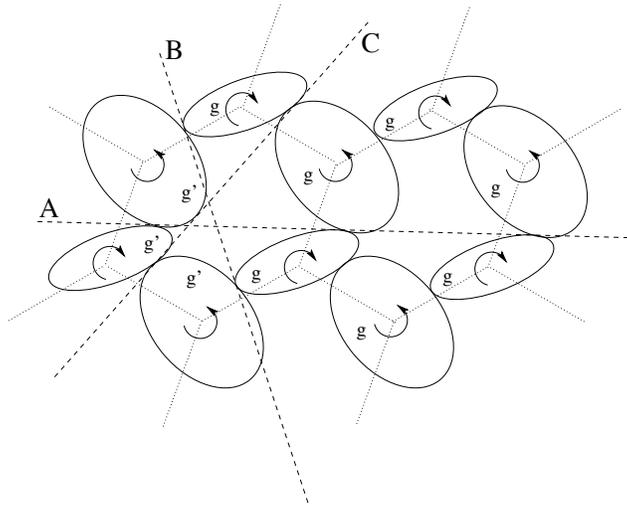
FIGURE 9. Pure rotation strain in an unfrustrated deformed Honeycomb structure of particles. The system can be strained by rotating the particles to the right of the shear line B (the g particles) by a small angle θ in the shown directions. If we keep the g' particles stationary, the entire right hand side of the system would displace relative to the left hand side.

the mean position vector of these points) due to an infinitesimally small rotation θ. Since for an infinitesimal rotation this is a linear-order calculation, then we can consider a small rotation around each of the contact points and superpose the contributions. This calculation leads to the result
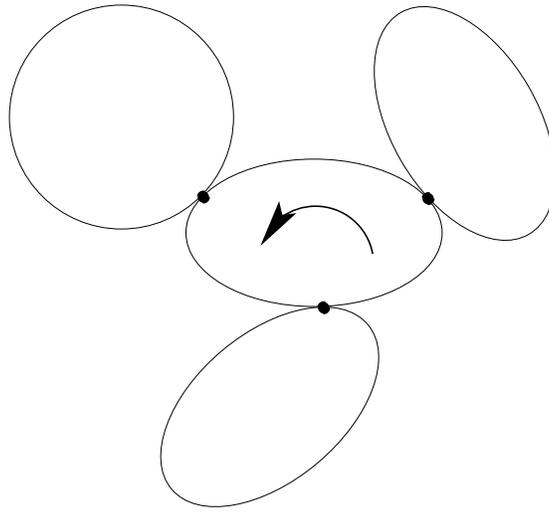
$$e_{ij}^{rot} = Q_{ij}^T \theta \ . \tag{14}$$

The tensor Q^T turns out to be the transpose of the tensor Q in eq. (2) (which is symmetric anyway, $Q^T = Q$).

Thus, we can write now the strain rate for general two-dimensional systems, whether frustrated or not,

$$\dot{e}_{ij} = A\gamma_{ij} + Q_{ij}^T \omega \ , \tag{15}$$

where $\omega = \dot{\theta}$ is the local angular velocity field, measured through particle rotations.

179

FIGURE 10. To calculate the strain that results from pure rotation, rotate the central particle by a small angle θ. To linear order this is the same as rotating the centroid of the central particle by this angle around ever contact point and superposing the results to find the displacement of that centroid. The result is that the symmetric strain is as given in equation (14).

Next, let us write explicitly the local yield criterion. Basing the condition for slippage between particles g and g' on da Vinci - Amontons - Coulomb friction threshold criterion, we have

$$| \sigma \cdot \hat{t} | > \mu \sigma \cdot \varepsilon \cdot \hat{t} , \tag{16}$$

where we have omitted the superscripts gg' from $\hat{t}$. The macroscopic yield threshold can then be written in the exact form

$$Y(\sigma) = 1 - \prod_{gg'} \{ 1 - H(| \sigma \cdot \hat{t} | - \mu \sigma \cdot \varepsilon \cdot \hat{t}) \} . \tag{17}$$

When the tangential force at each and every contact is below the yield threshold $Y(\sigma) = 0$. As soon as any of these forces increases above the threshold the particular contact slips and $Y = 1$. This is then a legitimate yield function that describes failure when its value increases above zero.

The above strain rate equation is quasi-static, thus satisfying the static stress equations at all times. These are the continuous form of force balance equations

$$\mathbf{div} \cdot \sigma = F_{ext} \tag{18}$$

and of the torque balance equations

$$\sigma = \sigma^T . \tag{19}$$

In addition, the stress-structure relation (2) is also part of the set of equations.

It should be noted that we need no independent energy equation - the energy dissipation can be determined directly from the existing equations. It is constructive to examine the energy dissipation. To this end, multiply equation (15) from the right by the stress tensor σ and take the trace of the resulting tensor. Since the process is quasi-static, the stress tensor is static at every moment and the quantity on the left hand side, $\dot{e} : \sigma$, is exactly the energy dissipation rate. Considering the two contributions on the right hand side, we anticipate dissipation only from friction due to slippage between particles in contact; rolling of rigid particles over one another should not give rise to dissipation. Thus, we expect all the dissipation to come from the first term

$$\dot{e} : \sigma = A\gamma : \sigma . \tag{20}$$

Note that this equation can be used to measure the scalar field A experimentally. From the left-hand-side term describing strain rate due to rotation we obtain

180

$$Q^T : \sigma\omega = 0 . \tag{21}$$

We now note that equation (21) should hold for any arbitrary angular velocity ω and therefore the relation should be independent of it. We then obtain that, in two dimensions $Q^T : \sigma = Q : \sigma = 0$, which, fittingly, is exactly the stress-structure relation (2) for the static stress field. Not only is this consistent with the static equations, but it also provides the bridge between the quasi-static dynamics and the static solution - it shows that, when the motion described by the above yield equation stops, the system 'lands' on an isostatic stress state - the correct state to land on because the yield surface corresponds to marginally rigid states, as discussed above.

I conclude this section by pointing out that the set of equations described above, and in particular the strain rate equation (15), can be extended to a more general description of assemblies of non-rigid particles. In such assemblies, equation (15) has to be modified with an additional term on the right hand side, describing strain rate due to the compliance of particles. This is a purely elastic term and, adding it, the equation reads

$$\dot{e}_{ij} = A\gamma_{ij} + Q^T_{ij}\omega + C_{ijkl}\sigma_{kl} , \tag{22}$$

where C_{ijkl} is a conventional compliance tensor that depends locally on the stiffness of the particles in the absence of slippage and rotation. Note that the additional term gives rise to elastic energy, in addition to the dissipation due to slippage friction.

CONCLUSIONS AND DISCUSSION

To conclude, I have reviewed in this paper recent advances on two controversial issues. One concerns the debate in the community on the correct modelling of the static stress transmission in general real dense granular materials (DGM) and the other concerns the equations that govern the failure, subsequent initiation of flow, of DGM. The recent history of isostaticity theory and its basics have been sketched briefly and, in particular, the sensitive dependence of the theory's validity on the marginally rigid nature of the material.

A key experiment has been described, which showed that: (i) marginally rigid states are realizable in DGM and (ii) that the isostatic state behaves as a critical point in the sense that a particular response length diverges in the 'thermodynamic' limit - the response to localized displacements and application of forces. The realizability of marginally rigid states gives initial support to the relevance of isostaticity theory to real systems.

The second result of the experiment has been shown to lead to a far-reaching conclusion: that real DGM should be regarded as two-phase composites - one phase made of marginally rigid regions, where stress transmission is governed by the isostaticity equations, and another phase of over-connected regions, governed by more conventional equations, possibly the elasticity equations. The diverging lengthscale at the marginally rigid state has been argued to correspond to the typical response length to stress perturbations (basically, the response length to the Green function) and, consequently, it corresponds to the typical length of force chains in two dimensions. Indeed solutions to the general isostaticity equations in two dimensions consist of *infinitely* long force chains (albeit with the caveat that force chains may eventually disperse, as shown in [19, 20]) and mentioned above.

The two-phase model of real granular materials, described here, which has been called stato-elasticity, leads to an interesting situation that has no parallel in traditional studies of composite materials. In conventional composites, e.g. of two different electrical conductors or two substances with different elastic properties, the phases satisfy the *same* field equations (e.g. Laplace's equation) but with different constitutive parameters. In the above examples, the electrical conductivity or the elastic constants would differ between the two phases. In granular materials, it is the very form of the field equations that is different. To emphasize the significance of this difference, suppose that the over-connected regions can be modelled by elasticity theory. Then the closure of the stress equations in those regions is via the compatibility conditions and a stress-strain relation, while in the isostatic regions the closure is through sterss-structure relations. This has crucial consequence - while the elasticity equations are elliptic, those of isostaticity are hyperbolic, which means that global solutions over the entire medium are difficult to obtain. This is an issue that is currently studied by this author.

Finally, I have reviewed a recent model for the failure of a DGM and the rheology that follow immediately after the failure. A set of equations has been derived for such flow. The equations are parameter-free - a significant advantage over most many models in the literature. Another advantage is that the set of equations has transparent unknown fields that need to be solved for and there are just enough equations to solve for these fields. The equations hold for DGM that are at marginal rigidity before failure and after flow has ceased and, in particular, they make sure that the system

is in a correct static stress state once the motion stops. The strain rate equations, which is at the heart of the formalism, includes three mechanisms: slippage, rotation and particle compliance.

Interestingly, equation (22), with only the last two terms on the right hand side, also describes iso-auxetic materials [34]. It is expected that this equation would find more applications, another issue that this author is looking into.

REFERENCES

1. J. Smid, and J. Novosad, *Ind. Chem. Eng. Symp.* **63**, V/1–12 (1981).
2. B. Brockbank, J. M. Huntley, and R. C. Ball, *J. Phys. II (France)* **7**, 1521–1532 (1997).
3. L. Vanel, D. W. Howell, D. Clark, R. P. Behringer, and E. Clement, *Phys. Rev. E* **60**, R5040–R5043 (1999).
4. A. P. F. Atman, P. Brunet, J. Geng, G. Reydellet, P. Claudin, R. P. Behringer, and E. Clement, *Eur. Phys. J. E* **17**, 93–100 (2005).
5. S. F. Edwards, and R. B. Oakeshott, *Physica D* **38**, 88–92 (1989).
6. S. F. Edwards, and R. B. Oakeshott, *Physica A* **157**, 1080–1090 (1989).
7. A. Mehta, and S. F. Edwards, *Physica A* **157**, 1091–1100 (1989).
8. J.-P. Bouchaud, M. E. Cates, and P. Claudin, *J. Phys. I (France)* **5**, 639–656 (1995).
9. S. F. Edwards, and C. C. Mounfield, *Physica A* **226**, 1–11 (1996).
10. S. F. Edwards, and C. C. Mounfield, *Physica A* **226**, 12–24 (1996).
11. S. F. Edwards, and C. C. Mounfield, *Physica A* **226**, 25–33 (1996).
12. J. P. Wittmer, P. Claudin, and J.-P. Cates, M. E.and Bouchaud, *Nature* **382**, 336–338 (1996).
13. J. P. Wittmer, M. E. Cates, and P. Claudin, *J. Phys. I (France)* **5**, 39–80 (1995).
14. J. P. Cates, M. E.and Wittmer, J.-P. Bouchaud, and P. Claudin, *Phys. Rev. Lett.* **81**, 1841–1844 (1996).
15. S. F. Edwards, and D. V. Grinev, *Phys. Rev. Lett.* **82**, 5397–5400 (1999).
16. R. C. Ball, and R. Blumenfeld, *Phys. Rev. Lett.* **88**, 115505–115508 (2002).
17. R. Blumenfeld, *Phys. Rev. Lett.* **93**, 108301–108304 (2004).
18. R. Blumenfeld, *Physica A* **336**, 361–368 (2004).
19. M. Gerritsen, G. Kreiss, and R. Blumenfeld, *Phys. Rev. Lett.* **101**, 098001–098004 (2008).
20. M. Gerritsen, G. Kreiss, and R. Blumenfeld, *Physica A* **387**, 6263–6276 (2008).
21. S. B. Savage, "," in *Physics of Dry Granular Media*, edited by H. J. Herrmann, J. P. Hovi, and S. Luding, NATO ASI Series, Kluwer, (Amsterdam, 2002, pp. 25–95.
22. C. Goldenberg, and I. Goldhirsch, *Phys. Rev. Lett.* **89**, 084302–084305 (2002).
23. R. Blumenfeld, "Stress transmission and isostatic states of non-rigid particulate systems," in *Modeling of Soft Matter*, edited by M.-C. T. Calderer, and E. M. Terentjev, Springer-Verlag, Berlin, 2005, vol. 141 of *IMA Volume in Mathematics and its Applications*, pp. 235–246.
24. R. Blumenfeld, S. F. Edwards, and R. C. Ball, *J. Phys.: Cond. Mat.* **17**, S2481–S2487 (2005).
25. R. Blumenfeld, *New J. Phys.* **9**, 160 (2007).
26. G. Combe, and J.-N. Roux, *Phys. Rev. Lett.* **85**, 3628 (2000).
27. J. C. Maxwell, *Phil. Mag.* **27**, 250–261 (1864).
28. J. C. Maxwell, *Edinb. Roy. Soc. Trans.* **26**, 1–40 (1872).
29. L. Sibille, T. Mullin, and P. Poullain, *EuroPhys. Lett.* **86**, 44003–44008 (2009).
30. C. Domb, and M. S. Green, *Phase Transitions and Critical Phenomena*, Academic Press, London, 1972.
31. R. Blumenfeld, *in preparation* (2009).
32. J.-P. Bouchaud, P. Claudin, and M. Levine, D. E.and Otto, *Eur. Phys. J. E* **4**, 451–457 (2001).
33. M. Wyart, S. R. Nagel, and T. Witten, *Euro. Phys. Letters* **72**, 486–492 (2005).
34. R. Blumenfeld, *Molecular Simulation* **31**, 867–871 (2005).