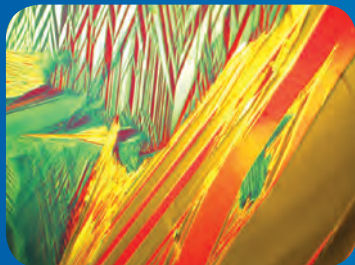# CURRENT EVENTS BULLETIN

Friday, January 12, 2018, 1:00 PM to 4:45 PM

Room 6E, Upper Level, San Diego Convention Center
Joint Mathematics Meeting, San Diego, CA

1:00 PM **Richard D. James**, *University of Minnesota*

Materials from mathematics

Mathematical ideas lead to new shape-shifting materials with unprecedented reversibility.

2:00 PM **Craig L. Huneke**, *University of Virginia*

How complicated are polynomials in many variables?

Polynomials get more complicated as the number of variables increases… or do they?

3:00 PM **Isabelle Gallagher**, *Université Paris Diderot*

From Newton to Navier-Stokes, or how to connect fluid mechanics equations from microscopic to macroscopic scales

Navier-Stokes is supposed to describe fluid flow—how does (or doesn't) the behavior of atoms do that?

4:00 PM **Joshua A. Grochow**, *University of Colorado, Boulder*

The Cap Set Conjecture, the polynomial method, and applications (after Croot-Lev-Pach, Ellenberg-Gijswijt, and others)

The capset conjecture generalizes the question: how frustrating can the popular game of "Set" be—and involves some beautiful and serious mathematics.

AMS
AMERICAN MATHEMATICAL SOCIETY

Organized by David Eisenbud, University of California, Berkeley

Introduction to the Current Events Bulletin

Will the Riemann Hypothesis be proved this week? What is the Geometric Langlands Conjecture about? How could you best exploit a stream of data flowing by too fast to capture? I think we mathematicians are provoked to ask such questions by our sense that underneath the vastness of mathematics is a fundamental unity allowing us to look into many different corners -- though we couldn't possibly work in all of them. I love the idea of having an expert explain such things to me in a brief, accessible way. And I, like most of us, love common-room gossip.

The Current Events Bulletin Session at the Joint Mathematics Meetings, begun in 2003, is an event where the speakers do not report on their own work, but survey some of the most interesting current developments in mathematics, pure and applied. The wonderful tradition of the Bourbaki Seminar is an inspiration, but we aim for more accessible treatments and a wider range of subjects. I've been the organizer of these sessions since they started, but a varying, broadly constituted advisory committee helps select the topics and speakers. Excellence in exposition is a prime consideration.

A written exposition greatly increases the number of people who can enjoy the product of the sessions, so speakers are asked to do the hard work of producing such articles. These are made into a booklet distributed at the meeting. Speakers are then invited to submit papers based on them to the *Bulletin of the AMS*, and this has led to many fine publications.

I hope you'll enjoy the papers produced from these sessions, but there's nothing like being at the talks -- don't miss them!

<div align="right">

David Eisenbud, Organizer
Mathematical Sciences Research Institute
de@msri.org

</div>

For PDF files of talks given in prior years, see
http://www.ams.org/ams/current-events-bulletin.html.
The list of speakers/titles from prior years may be found at the end of this booklet.

Materials from mathematics

Richard D. James

Department of Aerospace Engineering and Mechanics
University of Minnesota
james@umn.edu

Abstract. I survey some examples of materials whose recent discovery was based in an essential way on mathematical ideas. The main idea concerns "compatibility" – the fitting together of the phases of a material. Some of the emerging materials have the ability to change heat directly into electricity, without the need of a separate electrical generator.

Contents

1 Mathematics and materials science

Metallurgy in the early 20th century was mainly concerned with the understanding of the phase diagram of steel and its use in designing

processing treatments. It blossomed into materials science at mid century. From the beginning mathematicians have been interested. A touchstone for the aficionados is von Neumann's one page discussion in 1952 of Cyril Stanley Smith's paper, "Grain shapes and other metallurgical applications of topology"[55], where he discovered the $n - 6$ law for grain growth[1]. C. S. Smith replies, "The discussion of Dr. von Neumann is much appreciated, and his conclusions are as remarkable as they are nonobvious on first consideration of the problem."[2]. But the fascination of using mathematical reasoning to understand materials goes back much further. In his 1745 paper, "Physical investigations on the smallest parts of matter" Euler [25] reasons about how, owing to the presence of elemental molecules, bodies of different material can exhibit different masses and differing responses to gravity.

Today, largely due to the extreme forms of nonlinearity encountered in the behavior of materials, mathematics and materials science enjoy a healthy interaction. The mutual respect for certainty is pleasing. Like some theorems, the discovery of a spectacular new material represents an unmistakable advance, not clouded by shades of meaning.

In this paper we survey some recent developments and open problems in a central subfield of materials science: phase transformations. More precisely, our discussion concerns the mathematical theory that underlies the synthesis of materials that undergo phase transformations. What elements does one use, in what proportion, and with what processing, to achieve unprecedented behavior? Our behavior of interest will concern the hysteresis and reversibility of phase transformations.

The line of research surveyed here draws on, and owes much to, the work of many mathematicians and materials scientists. The author would particularly like to acknowledge the critical contributions of J. Ball, K. Bhattacharya, X. Chen, S. Conti, J. Cui, I. Fonseca, G. Friesecke, D. Kinderlehrer, R. Kohn, A. Ludwig, M. Luskin, S. Müller, F. Otto, E. Quandt, N. Schryvers, H. Seiner, Y. Song, V. Srivastava, I. Takeuchi, M. Wuttig, G. Zanzotto, J. Zhang and B. Zwicknagl.

---

[1]See also Mullins' reinterpretation [44].

[2]To which C. S. Smith adds, "It is greatly to be hoped that he, or some other mathematician, will be able to deduce similar relations in three dimensions...", a hope that would have to wait 55 years for fulfillment [41].

2 Phase transformations, hysteresis and reversibility

The types of phase transformations we consider are called structural or martensitic transformations. These are solid-to-solid phase transformations in which there is a change of crystal structure. The simplest example is a cubic-to-tetragonal phase transformation. In this case the unit cell of the high temperature cubic phase spontaneously elongates (or shrinks) along one of the four-fold axes upon cooling to the transformation temperature θ_c, changing the cubic unit cell to a tetragonal one (Figure 2 below). By symmetry, there are three four-fold axes, and so three *variants* of the tetragonal phase. The high temperature, often high symmetry, phase is called *austenite* and the low temperature, low symmetry, phase *martensite*. The change from the cubic to one of the three tetragonal cells involves a deformation, but no diffusion, i.e., no switching of atom positions, so these transformations can happen quite fast. Also – and this is probably the feature of greatest interest in materials science – the electronic structure or bonding pattern can change drastically during the transformation, because these aspects are sensitive to the geometry of the unit cell. For this reason the two phases can have very different properties. For example, one phase can be a strong magnet while the other phase is nonmagnetic, a feature we will exploit below in Section 8.

For a given material the identification of a particular transformation temperature θ_c is an oversimplification. In fact, one has to cool the material to a temperature $\theta^- < \theta_c$ before transformation occurs, and similarly, upon heating, one has to heat the material to $\theta^+ > \theta_c$. (The meaning of θ_c then becomes unclear, and we shall return to this later.) This phenomenon is called *hysteresis* and is one of the main features we will discuss here.

Figure 1: Hysteresis loop.

Even in transformations that have a big distortion, it can range from 100s of degrees C to 1 degree C. The difference $\theta^+ - \theta^-$ is a measure of the hysteresis. In cases that the transformation is not so abrupt people measure a property such as electrical resistance vs. temperature by steadily heating, then steadily cooling, the material and they fit
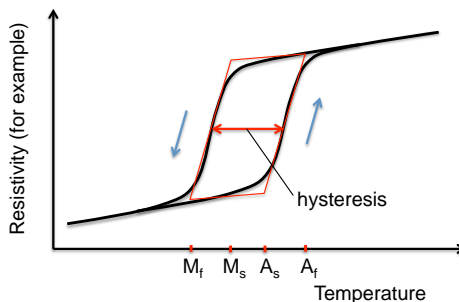
the resulting graph by a parallelogram (Figure 1). The width of the parallelogram is then a measure of the hysteresis.

At first sight it is difficult to notice anything very different about a material with $100°$ C hysteresis vs. one with $1°$ C hysteresis. The first can be a bigger or smaller transformation than the second by any obvious criterion: size of the distortion (measured in various norms), size of the latent heat, stiffness. There is also no obvious correlation with the elements used in making the alloys, or their proportions. One might think that it would correlate with the volume change – bigger volume change means bigger hysteresis – but this is also not true in general.

Hysteresis is interesting partly because in applications, such as the one described in Section 8, it is synonymous with loss, and thus it is desirable to make it as small as possible. It is also fascinating from the viewpoint of mathematical theory. Usually, loss would be associated with the "damping terms", i.e., processes of viscoelasticity or viscosity, as embodied say in the Navier-Stokes equations. The expectation from scaling the energy equation of the Navier-Stokes equations is that, if one shears a fluid back and forth and measures some the overall displacement vs. some overall force, the resulting hysteresis loop will shrink to zero as the rate of shearing (at fixed amplitude) gets lower and lower[3]. But the hysteresis loops in phase transformations, as far as we can tell, do not shrink to zero as the rate of change of temperature or force tends to zero. Rather, there is a limiting loop at zero rate. This is called *rate-independent hysteresis*. So, studying the effect of damping terms is not the right idea.

Reversibility is a general term that is usually quantified experimentally by passing back and forth through the phase transformation many times, by say periodically changing the temperature, and measuring some property each cycle. A good property to measure is latent heat, since it is proportional to the amount of material that actually transforms. Solid-solid phase transformations have a latent heat that is absorbed on heating through the phase transformation (just like water boiling on the stove) and released on cooling. It is measured by calorimetry. A phase transformation is seen to be lacking reversibility if the latent heat decreases each cycle. Often in these cases nonreversibility is also seen more dramatically as a complete failure of the

---

[3]Here, the analogy to phase transformations is closer than it may seem. Phase transformations can often also be induced by applying a cyclic force, leading again to a hysteresis loop as in Figure 1, but with "temperature" replaced by force and "resistivity" by displacement. See Section 7.
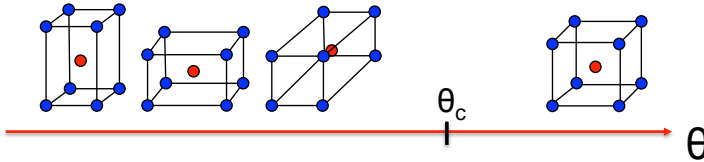
Figure 2: Cubic to tetragonal phase transformation. The red axis represents temperature θ. There is no diffusion, only distortion.

material after a certain number of cycles. A highly nonreversible phase transformation is the β to α transformation in the element tin, that occurs a little below room temperature. Transformation upon cooling just once a shiny bar of β-tin a little below room temperature yields a pile of gray powder of α-tin.

3 Theory of phase transformations

We start with the simple cubic-to-tetragonal phase transformation mentioned above. A crystal (such as $BaTiO_3$, Fe_7Pd_3, In_4Tl, Mn_9Cu, $Ni_{65}Al_{35}$, Ni_2MnGa) transforms upon cooling somewhat below θ_c (because of hysteresis) by elongating along one of the four-fold cubic axes, to yield the three *variants of martensite*. To obtain the variants of martensite from the cubic structure, one applies linear transformations $U_1 = diag(\eta_2, \eta_1, \eta_1)$, $U_2 = diag(\eta_1, \eta_2, \eta_1)$, $U_3 = diag(\eta_1, \eta_1, \eta_2)$ to the cubic lattice, here written in an orthonormal cubic basis.

Two elementary points should be emphasized. First, the red and blue balls in Figure 2 indicate this is an ordered alloy, but, generally speaking, and consistent with the nonstoichiometric composition of some of the alloys given above, there can be a certain probability of finding an atom on a certain site, for example, a body center. For example, the nonstoichiometric alloys Mn_xCu_{1-x} ($0.8 < x < 0.95$) all undergo cubic-to-tetragonal phase transformations, and all the compositions listed at the beginning of this section can be perturbed within limits. Second, let us number, left to right, the variants 1, 2, 3 in Figure 2. All the variants are exactly the same up to rigid rotation. For example, variant 1 can be rotated to look exactly like variant 2 by a rigid rotation $R_{\pi/2}$ of angle $\pi/2$. That does not mean that they are the same: what matters crucially here is the deformation, and these are of course different, $R_{\pi/2}U_1 \neq U_2$. However, for an appropriate rigid rotation $R \in$ SO(3), the two deformations $y_1(x) = U_1x$ and $y_2(x) = RU_2x$ do agree on a lower dimensional set, and this observation will be relevant to our study of hysteresis.

We begin with a lattice model of the phase transformation. In

the simplest case we consider a Bravais lattice such as any one of the lattices shown in Figure 2. This is the set of points $\mathcal{L}(e_1, e_2, e_3) = \{\nu^1 e_1 + \nu^2 e_2 + \nu^3 e_3 : (\nu^1, \nu^2, \nu^3) \in \mathbb{Z}^3\}$ where e_1, e_2, e_3 are given linearly independent vectors in $\mathbb{R}^3$, called *lattice vectors*. In Figure 3 the lattice vectors can be chosen as the vectors from a blue atom to two nearest blue atoms, together with a vector to the red atom (Figure 3). Let $e_1 = \alpha \hat{e}_1, e_2 = \alpha \hat{e}_2$ where $\hat{e}_1, \hat{e}_2, \hat{e}_3 = \hat{e}_1 \times \hat{e}_2$ are orthonormal and $\alpha > 0$. We can consider various tetragonal Bravais lattices defined by lattice vectors $\alpha \hat{e}_1, \alpha \hat{e}_2, \alpha(\hat{e}_1 + \hat{e}_2 + \gamma \hat{e}_3)/2$, with $\gamma > 0$. (The constants α, γ that quantify the distances between atoms are called *lattice parameters*). The value $\gamma = 1$ gives the BCC lattice. A famous observation of Bain is that there is exactly one other choice of $\gamma > 0$ in which the associated Bravais lattice is cubic, that being $\gamma = \sqrt{2}$, which gives the face-centered cubic lattice (FCC). In fact quite a few phase transformations can be viewed as perturbations of the BCC to FCC (or the reverse) transformation, obtained by moving γ from 1 to $\sqrt{2}$ and changing α a bit.
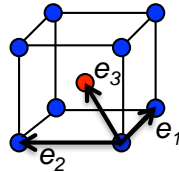

Figure 3: Lattice vectors for the BCC lattice (with so-called B2 ordering).

Many lattices of interest in phase transformations are not simply Bravais lattices. Rather, they are general periodic structures, i.e., the periodic extension of a finite number of atomic positions. These can be viewed as the union of a finite number of Bravais lattices, all made with the same lattice vectors, that is,

$$\mathcal{L}(e_1, e_2, e_3; x_1, \ldots, x_M) = \{x_k + \mathcal{L}(e_1, e_2, e_3) : k = 1, \ldots, M\}, \quad (1)$$

where the *base points* $x_1, \ldots x_M$ are given points in $\mathbb{R}^3$. Conventionally, a description is chosen with smallest M, in which case we speak of e_1, e_2, e_3 as a set of *primitive* lattice vectors. A completely ordered lattice assigns a certain species to all positions with the same value of k. The lattice vectors e_1, e_2, e_3 can now also be interpreted as defining the periodicity. During a phase transformation in a complex lattice, the lattice vectors can change, the base points can change, and typically even the number M of base points changes. In its low temperature phase an important alloy discussed later $Zn_{45}Au_{30}Cu_{25}$ consists of $M = 18$ Bravais lattices.

Of course at positive temperature the atoms are vibrating about average positions, and phenomena such as the release of latent heat are intimately related to these vibrations. Nevertheless, it is convenient

to use a kinematics of periodic lattices based say on these averaged positions.

It would seem to be an easy matter to decide which atom goes where during a phase transformation. Already, that decision has been made tacitly by drawing Figure 2 and assigning U_1, U_2, U_3. But in fact, especially with complex lattices, this is a nontrivial problem. Currently, there is no experimental method that can follow individual atoms during a phase transformation, i.e., the pathway, though there are some possibilities on the horizon. To determine the pathway, Bain [1] favored minimizing a measure of "strain" defined from mappings between the parent and transformed lattices, but he did not specify a norm. Mathematically, the difficulty can be appreciated by noticing that $\mathcal{L}(e_i) = \mathcal{L}(\mu_i^j e_j)$, where $(\mu_i^j) \in GL(3, \mathbb{Z})$. (Here, equality of two Bravais lattices means that they consist of the same infinite set of points, and $GL(3, \mathbb{Z})$ is the set of real 3×3 matrices of integers with determinant ± 1. Also, we use the summation convention: $\mu_i^j e_j = \sum_{j=1}^{3} \mu_i^j e_j$.) In fact, a classical theorem of crystallography (easily proved) says that $\mathcal{L}(e_1, e_2, e_3) = \mathcal{L}(f_1, f_2, f_3)$ for linearly independent vectors f_1, f_2, f_3 if and only if $f_i = \sum_{j=1}^{3} \mu_i^j e_j$ for some $\mu \in GL(3, \mathbb{Z})$. So, many different choices of lattice vectors implies many possible pathways[4].

Many transformations between phases involve complex lattices. Empirically, in the notation of (1), the often accepted mechanism of transformation is that a Bravais sublattice of austenite with lattice vectors $\nu_i^j e_j$, with $\nu_i^j \in \mathbb{Z}^{3 \times 3}$ but with $\det \nu > 1$, is transformed to a primitive lattice of martensite[5]. As in [37], this again gives rise to integer minimization problems for which rigorous algorithms can be devised that converge to a minimizer in a finite number of steps. An example using a particular measure of strain (different than [37]) is given in [16], and software can be found at `http://www.structrans.org`.

Really, determination of the pathway should be the province of first principles calculations, and many examples are being explored in this context [51, 67]. To describe a typical approach, consider the cubic to tetragonal transformation of Figure 2, with cubic lattice vectors e_1^c, e_2^c, e_3^c, such as those shown in Figure 3 (relabeled). Schematically, a typical procedure is the postulation of a one-parameter family of

---

[4]Recently, using a reasonable measure of strain, Muehlemann and Koumatos [37] prove that the Bain mechanism for BCC to FCC gives the smallest strain.

[5]Always, in this case $\det \nu$ has a value such that the volume of the unit cell associated to $\nu_i^j e_j$ is about the same as a primitive unit cell of the martensite phase.

unit cells, say defined by linear transformations of the cubic unit cell $F(\xi), 0 \leq \xi \leq 1$, with $F(0) = I$, $F(1) = U_1$. Then, density functional theory calculations of periodic lattices defined by lattice vectors Fe_1^c, Fe_2^c, Fe_3^c are carried out, and optimized using the nudged elastic band method (see [62] for a mathematical description of this method in a different context). This method in principle gives the lowest saddle on a pathway $F(\xi)$ between cubic and tetragonal phases. But, on further contemplation, it misses a critical aspect of phase transformations which is central to this article: microstructure and compatibility! In fact, even in a near perfect single crystal, transformation never proceeds by a homogeneous deformation. Rather, the new phase nucleates and then grows. Inhomogeneous pathways[6] must have lower – likely much lower – saddles. The experimental evidence is that in many cases the highly inhomogeneous austenite/martensite interface (whose energy cannot be represented by a one-parameter ansatz of the type described here) represents the lowest saddle. It is a big challenge to have a first principles method that could cope with even the simplest microstructures, but well worth investigating. What is a few-parameter first principles anzatz that captures the austenite/martensite interface?

With this atomistic background we give a brief summary of a continuum theory of phase transformations [4, 9] we will use, while pointing out some deficiencies along the way. We do this first in the simplest case of transformations between Bravais lattices.

We assume that lattice vectors for the austenite are e_1^a, e_2^a, e_3^a and, for martensite, e_1^m, e_2^m, e_3^m. We wish to encompass also elastic deformations of both lattices, so we use the notation e_1, e_2, e_3 for generic lattice vectors. Eventually we will have to specify a domain, but for now we just assume preservation of orientation, $(e_1 \times e_2) \cdot e_3 > 0$. A general atomistic model will generate a free energy per unit volume once the lattice and temperature are prescribed, so we assume such free energy $\hat{\varphi}(e_1, e_2, e_3, \theta)$ defined for $(e_1 \times e_2) \cdot e_3 > 0$ and temperature $\theta > 0$.

This free energy $\hat{\varphi}(e_1, e_2, e_3, \theta)$ is subject to basic symmetries. We have frame-indifference, $\hat{\varphi}(Re_1, Re_2, Re_3, \theta) = \hat{\varphi}(e_1, e_2, e_3, \theta)$ for all $R \in$ SO(3) and $\theta > 0$, and the condition that the free energy should only depend on the lattice $\mathcal{L}(e_1, e_2, e_3)$ and not otherwise on the lattice vectors: $\hat{\varphi}(\mu_1^j e_j, \mu_2^j e_j, \mu_3^j e_j, \theta) = \hat{\varphi}(e_1, e_2, e_3, \theta)$ for $(\mu_i^j) \in GL(3, \mathbb{Z})$ and $\theta > 0$.

The function $\hat{\varphi}(e_1, e_2, e_3, \theta)$ assigns a value of free energy to a perfect

---

[6]Given a smooth $F(\xi)$, $F(0) = I, F(1) = U_1$, an interesting relevant mathematical problem that relies on the Cauchy-Born rule described below is to solve, under weak conditions of regularity, $\nabla y(x) = R(x)F(\xi(x))$ for $y : \Omega \to \mathbb{R}^3$, $R : \Omega \to$ SO(3), $\xi : \Omega \to (0, 1)$, where Ω is a domain in $\mathbb{R}^3$.

Bravais lattice $\mathcal{L}(e_1, e_2, e_3)$ at temperature θ. But we would like to treat more complex structures than perfect lattices, such as the microstructures shown in Figures 4, 9, 10. Locally, near most points, they look almost like perfect lattices. This suggests that we think of the austenite lattice as a reference lattice and define, for linear transformations F from $\mathbb{R}^3$ to $\mathbb{R}^3$ with $\det F > 0$,

$$\varphi(F, \theta) = (\det F)\,\hat{\varphi}(Fe_1^a, Fe_2^a, Fe_3^a, \theta). \tag{2}$$

(The presence of $\det F$ converts the free energy per volume of $\mathcal{L}(e_1, e_2, e_3)$ to a free energy per volume of the reference lattice $\mathcal{L}(e_1^a, e_2^a, e_3^a)$.) For a smooth mapping $y : \Omega \to \mathbb{R}^3$ the gradient ∇y is the local linear transformation, and therefore suggests a passage to continuum theory

$$\inf_{y \in \mathcal{A}} \int_\Omega \varphi(\nabla y(x), \theta)\, dx. \tag{3}$$

This cornerstone of the theory (2), (3) is the Cauchy-Born rule [24]. It can be approached in a simple but rigorous way [13] via the "large body limit", e.g., the asymptotics as $\varepsilon \to 0$ of $y_\varepsilon : \Omega/\varepsilon \to \mathbb{R}^3$ given by $y_\varepsilon(x) = (1/\varepsilon)y(\varepsilon x)$, which has the feature of making $\nabla y_\varepsilon(x/\varepsilon)$ more and more constant, and therefore representing a more perfect lattice, on a bigger and bigger collection of atoms near x/ε as $\varepsilon \to 0$. A complete understanding of the Cauchy-Born rule likely involves difficulties beyond the already insanely difficult "crystallization problem" [60, 26]. Nevertheless, by making clever but realistic assumptions, many interesting papers shed light on its successes and failures [28, 14, 66].

The unspecified $\mathcal{A}$ in (3) raises another issue. Normally in the calculus of variations, $\mathcal{A}$ would be chosen to match the growth conditions on φ. But the symmetry $\hat{\varphi}(\mu_1^j e_j, \mu_2^j e_j, \mu_3^j e_j, \theta) = \hat{\varphi}(e_1, e_2, e_3, \theta)$ rules out any of the usual growth conditions that would put finite energy deformations y in a reasonable Sobolev space. Note that a relevant choice of (μ_i^j) in $GL(3, \mathbb{Z})$ is, for arbitrary large $m \in \mathbb{Z}$,

$$(\mu_i^j) = \begin{pmatrix} 1 & m & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

Of course, we could brutally assign strong growth conditions for φ, but then we would violate the symmetry, which is supposed to *preserve* the energy. Another disturbing fact is that a deformation gradient F satisfying $Fe_i^a = \mu_i^j e_i^a$ with μ_i^j as in (4), and say $m = 1$, if imposed on a crystal, is likely to cause either failure of the crystal or massive plastic

deformation, phenomena that are not so relevant to the study of phase transformations.

These observations suggest a resolution due to Ericksen [24] and Pitteri [49]: cut down the symmetry $\hat{\varphi}(\mu_1^j e_j, \mu_2^j e_j, \mu_3^j e_j, \theta) = \hat{\varphi}(e_1, e_2, e_3, \theta)$ to a subgroup, excluding exactly those μ_i^j in $GL(3, \mathbb{Z})$ that would correspond to massive plastic deformation, and at the same time cut down the domain of φ to be invariant under exactly these symmetries (and frame-indifferent), and, with luck, include in the domain the tetragonal or other lower symmetry phases that are of interest.

Such a domain $\mathcal{D}$ can be found [4, 9, 50]. In the simplest case the answer is the following. The energy density $\varphi : \mathcal{D} \times (0, \infty)$ satisfies

$$\begin{aligned} R\mathcal{D}H \;&=\; \mathcal{D} \quad \text{and} \quad \varphi(RFH, \theta) = \varphi(F, \theta), \\ & \text{for all} \;\; F \in \mathcal{D}, \; R \in \mathrm{SO}(3), \; H \in G^a, \end{aligned} \tag{5}$$

where

$$G^a = \{Q \in \mathrm{SO}(3) : Q e_i^a = \mu_i^j e_j^a, \; i = 1, 2, 3, \text{for some } \mu_i^j \in GL(3, \mathbb{Z})\}. \tag{6}$$

Hence, as indicated by the notation, the austenite phase dominates the symmetry. From its definition G^a is a finite, crystallographic group of rotations, i.e., one of the 11 Laue groups.

These discrete and continuous symmetries imply an energy-well structure, that also is assumed to change with temperature in a way that is consistent with the phase transformation. A symmetric, positive-definite *transformation stretch matrix* U_1 is given, and its orbit under the symmetries (5), (6) is $\mathrm{SO}(3)U_1 \cup \cdots \cup \mathrm{SO}(3)U_n$. The free energy φ is assumed to be smooth on $\mathcal{D} \times (0, \infty)$ and a typical energy-well structure is

$$\begin{aligned} \theta \;&>\; \theta_c, \quad \varphi(\cdot, \theta) \;\; \text{minimized on } \mathrm{SO}(3) \\ \theta \;&<\; \theta_c, \quad \varphi(\cdot, \theta) \;\; \text{minimized on } \mathrm{SO}(3)U_1 \cup \cdots \cup \mathrm{SO}(3)U_n, \end{aligned} \tag{7}$$

where $\{U_1, \ldots, U_n\} = \{QU_1Q^T : Q \in G^a\}$. In fact, it can be seen that the Laue group of the martensite is $G^m = \{Q \in G^a : QU_1Q^T = U_1\}$, which also shows, by Lagrange's theorem, that $n = \mathrm{order}\, G^a / \mathrm{order}\, G^m$. From a practical viewpoint the restriction $QU_1Q^T = U_1$ for all $Q \in G^m$ is very useful, since these groups are often easily known from an X-ray measurement, but direct determination of U_1 is hampered by the issue raised above about knowing where atoms go.

We shall use the theory in the form above, but in some cases applying it to lattices that are more complex than Bravais lattices. In

the case of a general lattice of the form (1) one expects an atomistic free energy[7] $\hat{\varphi}(e_1, e_2, e_3, x_2 - x_1, x_3 - x_1, \ldots, x_m - x_1, \theta)$, together with a Cauchy-Born rule of the form

$$\varphi(F, s_1, \ldots, s_{m-1}, \theta) = \hat{\varphi}(Fe_1^a, Fe_2^a, Fe_3^a, s_1, s_2, \ldots, s_{m-1}, \theta), \quad (8)$$

with again F replaced by $\nabla y(x)$, and $s_1(x), \ldots, s_{m-1}(x)$, $x \in \Omega$, as unconstrained functions. Under suitable hypotheses for the resulting problem in the calculus of variations, we could minimize out the functions $s_1(x), \ldots, s_{m-1}(x)$, leading back to a theory somewhat like that given above, with potentially a significant lack of smoothness due to intersections of branches of minimizers. These possibilities are interesting, but it is disgraceful that, as of late 2017, we do not have a complete theory of symmetry for the more general multilattice case[8].

For definiteness we will put $\varphi(I, \theta_c) = 0$ in this paper.

4   Microstructure and nonattainment

As can been seen by the many examples presented in the article by C. S. Smith [55] mentioned at the beginning of this paper, small bubbles of a soap froth disappear and big ones grow, and the grains of a polycrystalline metal coarsen over time. As in the simplest linear elliptic and parabolic equations, there is a strong tendency to simplify and smooth. Exactly the opposite often happens in a martensitic phase transformation. One begins with a uniform crystal of austenite and, upon cooling through the transformation, one gets a plethora of fine microstructures of martensite. The mathematical origins of the spontaneous formation of fine structure comprise a fascinating and ongoing chapter of nonlinear analysis that began with the work of L. C. Young [65, 64, 63]. For a broad overview see the article of S. Müller [43].

We should clarify the distinction between grains (which collectively are often also called "microstructure") and the microstructure resulting from a phase transformation. The theory presented above is for a single crystal, when in its austenite phase, modeled by the uniform lattice $\mathcal{L}(e_1^a, e_2^a, e_3^a)$. When it transforms, say by cooling, it forms microstructure, due to the tendency arising from energy minimization of the deformation $y : \Omega \to \mathbb{R}^3$ to have a gradient ∇y near the energy wells

---

[7]The presence of the differences $x_i - x_1$ arises from the translation invariance of $\hat{\varphi}$.

[8]The state-of-the-art is the last chapter of Pitteri and Zanzotto [50], especially Section 11.7.

SO(3)$U_1, \ldots,$ SO(3)U_n. But often the materials (both metals and oxides) that undergo phase transformations are studied as polycrystals, with differently oriented grains. Then one has to rewrite the theory presented in Section 3 for a polycrystal, which is easy to do [11]. The "fighting between the grains" during transformation has consequences: single crystals and polycrystals of the same material do exhibit somewhat different macroscopic response. Usually, phase transformations in polycrystals are studied at lower temperatures, in which case the grains do not coarsen during normal time scales.

The simplest relevant example of nonattainment is the following. Let $A \neq B \in \mathbb{R}^{3 \times 3}$, with rank$(B - A) = 1$, i.e., $B - A = a \otimes n$. Make the drastic and unphysical simplification that the free energy density is smooth on $\mathbb{R}^{3 \times 3}$, independent of temperature, and satisfies

$$\varphi(A) = \varphi(B) = 0 < \varphi(F), \quad F \notin \{A, B\}. \qquad (9)$$

We have brutally omitted all the symmetries and retained only the structure of having energy wells, and then only two. However, we have made them "rank-1 connected", a feature which is shared by the tetragonal phase in the model described above: there exists $R \in$ SO(3) such that rank$(RU_2 - U_1) = 1$ (see Lemma 5.1 below). Assume φ satisfies the mild growth conditions of being bigger than a positive constant outside a sufficiently large ball $|F| > \rho > 0$, and also assume that $\Omega \subset \mathbb{R}^3$ is open, bounded and has a Lipschitz boundary. Let $0 < \lambda < 1$ and consider

$$\inf_{\substack{y \in W^{1,1}(\Omega, \mathbb{R}^3), \\ y(x) = (\lambda B + (1 - \lambda)A)x, \ x \in \partial\Omega}} \int_\Omega \varphi(\nabla y(x)) \, dx. \qquad (10)$$

This example exhibits nonattainment of the infimum by a mechanism that is common in martensitic crystals, and the proof is simple but typical. We first show that the infimum is 0. We define the 1-periodic function $\chi_\lambda : \mathbb{R} \to \mathbb{R}$ by

$$\chi_\lambda(s) = \begin{cases} 1 - \lambda, & i \leq s < i + \lambda, \\ -\lambda & i + \lambda \leq s < i + 1, \end{cases} \quad i \in \mathbb{Z}. \qquad (11)$$

Note that the integral of χ_λ over one period is zero. Then, for $k = 1, 2, \ldots$ we define

$$y^{(k)}(x) = (\lambda B + (1 - \lambda)A)x + \left(\frac{1}{k} \int_0^{k(x \cdot n)} \chi_\lambda(s) ds \right) a. \qquad (12)$$

By differentiation, $\nabla y^{(k)}(x)$ takes the values B and A on alternating layers of width λ and $1 - \lambda$, and $y^{(k)}$ converges uniformly to the linear map $(\lambda B + (1 - \lambda)A)x$ on $\mathbb{R}^3$ as $k \to \infty$. Thus, $y^{(k)}$ is uniformly close to satisfying the boundary conditions. It can be made to satisfy them exactly by introducing a smooth function $\psi_\varepsilon : \mathbb{R}^3 \to [0,1]$, depending on the small parameter $\varepsilon > 0$, and satisfying

$$\psi_\varepsilon(x) = \left\{ \begin{array}{ll} 1, & \Omega_\varepsilon, \\ 0, & \mathbb{R}^3 \setminus \Omega, \end{array} \right. \qquad |\nabla \psi| < 2/\varepsilon, \tag{13}$$

where $\Omega_\varepsilon = \{x \in \Omega : \mathrm{dist}(x, \partial\Omega) > \varepsilon\}$. (One can replace Ω by a slightly smaller interior domain, choose ψ_ε to be proportional to the distance to the boundary of that domain (or 1), and then mollify, choosing the various scales appropriately.) Then, the function

$$y_{k,\varepsilon}(x) = \psi_\varepsilon(x)y^{(k)}(x) + (1 - \psi_\varepsilon(x))(\lambda B + (1 - \lambda)A)x), \quad x \in \Omega, \tag{14}$$

satisfies the boundary conditions $y_{k,\varepsilon}(x) = (\lambda B + (1 - \lambda)A)x$, $x \in \partial\Omega$. Also, $\nabla y_{k,\varepsilon}(x)$ takes the values A and B on Ω_ε and

$$\begin{aligned} |\nabla y_{k,\varepsilon}(x)| &= |\psi_\varepsilon(\nabla y^{(k)} - (\lambda B + (1 - \lambda)A)) \\ &\quad + (y^{(k)}(x) - (\lambda B + (1 - \lambda)A)x) \otimes \nabla \psi_\varepsilon| \\ &\leq |B - A| + \frac{2|a|}{k\varepsilon} \end{aligned} \tag{15}$$

Choosing, say, $\varepsilon = 1/k$ we have sequence with bounded gradient that satisfies the boundary conditions, whose gradient takes the values A and B on larger and larger fractions of Ω as $k \to \infty$. Thus, the infimum in (10) is zero.

To show nonattainment, we note that by the growth conditions, any minimizer $\hat{y}$ would have to lie in $W^{1,\infty}(\Omega, \mathbb{R}^3)$ and give zero energy density,

$$\nabla \hat{y} \in \{A, B\} \quad a.e. \ \Omega. \tag{16}$$

Extend $\hat{y}$ to all of $\mathbb{R}^3$ by making $\hat{y}(x) = (\lambda B + (1 - \lambda)A)x$ on $\mathbb{R}^3 \setminus \Omega$, and note that $\hat{y}$ is in $W^{1,\infty}(\mathbb{R}^3, \mathbb{R}^3)$. On Ω we can write

$$\nabla \hat{y}(x) = \eta(x)B + (1 - \eta(x))A = A + \eta(x)a \otimes n, \quad \text{where } \eta(x) \in \{\pm 1\}. \tag{17}$$

Thus $z(x) = \hat{y}(x) - Ax$ satisfies $\nabla z(x)n^\perp = 0$ on Ω and, in fact, on all of $\mathbb{R}^3$, and for all $n^\perp \cdot n = 0$. This we can integrate on $\mathbb{R}^3$ to get $z(x) = f(x \cdot n)$. The function f is completely determined by its value *outside* Ω, and therefore it must be $f(x) = \lambda(n \cdot x)a$. Hence,

$$\hat{y}(x) = Ax + z(x) = (A + \lambda a \otimes n)x = (\lambda B + (1 - \lambda)A)x \tag{18}$$

13

Since $0 < \lambda < 1$, $\lambda B + (1 - \lambda)A$ is in no cases equal to A or B, and therefore we have reached a contradiction with (16).

This type of example can be generalized to cases that satisfy the symmetries [4, 12, 8, 40] but still, known results of this type are quite special. In fact the main important examples of this article, and the concept of supercompatibility, illustrate the subtlety in trying to prove some kind of general result on non-attainment. Much more can be said, and in particular the Young measure is a beautifully simple tool to learn more about the structure of minimizing sequences without explicitly calculating them.

One is led by these results to study minimizing sequences, as well as minimizers, and to understand their relation to observed microstructures. The most important microstructure in martensite becomes an immediate target: the classic austenite/martensite interface. This microstructure
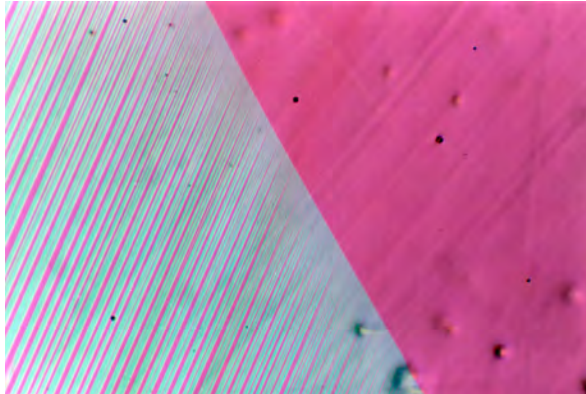

Figure 4: Austenite-martensite interface in $Cu_{69.5}Al_{27}Ni_{3.5}$. The banded structure on the left is martensite, and the uniform phase on the right is austenite (Courtesy of C. Chu).

is shown in Figure 4. It is widely seen as the interface between austenite and martensite at the finest level. When first observed in detail in the 1950s, it was puzzling because, when people measured the normal m to the interface (by viewing one such interface on two inclined free surfaces, such as at a corner), they found that it was, in the language of materials science, "irrational". That is, when m was expressed in the orthonormal cubic[9] basis of austenite, the components were not small integers. How could such an beautifully planar interface be noncrystallographic, i.e., not apparently have any relation to the adjacent crystal lattice of austenite?

The study of minimizing sequences for the austenite/martensite interface explains this irrationality and, quantitatively, the observed normal m as well as other features of Figure 4. Given the minimizing sequence constructed above, is is very easy to find a related one modeling the austenite/martensite interface. Consider an energy density of

---

[9]In many cases, including that of Figure 4, the austenite is cubic

the type given in (5-(7), and choose A, B from the martensite wells, e.g. $A = \hat{R}U_2$ and $B = U_1$, with $\hat{R} \in SO(3)$. Here we are using the freedom of an overall rotation of say Figure 4 to omit a possible rotation matrix in front of U_1. In fact, as already mentioned, it will be seen from results presented in Section 6 that there is often a choice of $\hat{R}$ such that

$$\hat{R}U_2 - U_1 = a \otimes n, \tag{19}$$

$a, n \in \mathbb{R}^3$. Equation 19 is called the *twinning equation*. We assume (19), choose $A = U_1$ and $B = \hat{R}U_2$ and repeat the construction (11)-(12) verbatim. The steps (13) to (14) can also be repeated, except now, say,

$$\psi_\varepsilon(x) = \begin{cases} 1, & x \cdot m < -\varepsilon, \\ 0, & x \cdot m > 0, \end{cases} \qquad |\nabla\psi| < 2/\varepsilon, \tag{20}$$

where, without loss of generality, we have put the origin on the austenite-martensite interface. On the austenite $x \cdot m > 0$ we choose an arbitrary deformation gradient from the austenite well $R \in SO(3)$, so (14), (15) are replaced by

$$\begin{aligned} y_{k,\varepsilon}(x) &= \psi_\varepsilon(x)y^{(k)}(x) + (1 - \psi_\varepsilon(x))Rx, \\ |\nabla y_{k,\varepsilon}(x)| &= |\psi_\varepsilon(\nabla y^{(k)} - R) \\ &+ ((\lambda B + (1 - \lambda)A)x - R^T x) \otimes \nabla\psi_\varepsilon|, \end{aligned} \tag{21}$$

but now, in the crucial last term, Rx has only to agree approximately with $(\lambda B + (1 - \lambda)A)x$ when $-\varepsilon < x \cdot m < 0$. This gives the sufficient (and necessary, for any reasonable choice of ψ_ε) condition

$$R^T(\lambda B + (1 - \lambda)A) = I + b \otimes m. \tag{22}$$

Rewriting (22) and (19) in a common notation and, without loss of generality, replacing $R^T \in SO(3)$, by R, we obtain the equations of the *crystallographic theory of martensite*:

$$\hat{R}U_2 - U_1 = a \otimes n, \quad R(\lambda\hat{R}U_2 + (1 - \lambda)U_1) = I + b \otimes m \tag{23}$$

Since (19) will be solved below, it is easiest to think of $\hat{R}, U_1, U_2$ as given, consistent with the first of (23), and so the unknowns are the volume fraction λ, $R \in SO(3)$ and $b, m \in \mathbb{R}^3$. As noted, m usually comes out irrational in the sense of materials science. The reason is clear: (23) embodies a geometric as well as a crystallographic restriction. Using (19), (23) can also be written

$$\hat{R}U_2 - U_1 = a \otimes n, \quad R(U_1 + \lambda a \otimes n) = I + b \otimes m \tag{24}$$

The solutions $(R, \lambda, b \otimes m)$ of (24) have been checked against observed values of these quantities many times, with amazing success [61].

This is one kind of minimizing sequence, but there are many more. If one looks at almost any martensitic microstructure, one can guess a minimizing sequence or, in some cases, a minimizer, and one can learn about this microstructure. One can get very good at this kind of guessing, then filling in many details by rather simple calculations, and it is certainly very informative. However, it is not really predictive of the microstructure that will result under such-and-such conditions.

The austenite-martensite interface, as for example shown in Figure 4, is produced by cooling a free crystal (no loads). Thus, nominally, the associated minimization problem is (3) with no imposed boundary conditions and $\theta = \theta_c$. But this minimization problem has a much simpler minimizer: $y(x) = x$, i.e., all austenite. Thus, one can say that there exists a minimizer, but *Nature prefers a minimizing sequence*, Figure 4. Of course, somehow, during cooling, the material has to go from austenite to martensite, and the crystallographic theory of martensite provides an obvious low energy pathway ("a zero limiting energy pathway"). But, at least in the normal pursuit of solutions of the calculus of variations, if one finds an absolute minimizer, one usually quits and does not then seek a minimizing sequence!

There are several features of Figure 4 that are not predicted by the crystallographic theory of martensite. Obviously, the bands on the left are not infinitely fine. There are also more bands near the interface than far from the interface – a branching phenomenon can be noticed, and this is quite common in martensites. By careful examination, 4-5 generations of branching can be seen in Figure 4. However, the measured the volume fraction λ predicted by the crystallographic theory[10] agrees closely (e.g., within 2-3%) with that measured on Figure 4 either close or far from the austenite interface. These observations are believed to be a consequence of a small regularization which has been studied mathematically in simpler models (see [34] for a survey, and [52, 15]). The consequences of this regularization will be critical for our understanding of hysteresis. We return to this in Section 5.

---

[10]suitably modified to represent the volume fraction of bands on the image $y(\Omega)$ and measured on a window that contains at least about 10 bands.

5 Hysteresis

What causes hysteresis? There are diverse thoughts about this in a large literature [11], some of which are in fact inconsistent with the observation of rate-independent hysteresis. Ideas also include the pinning of interfaces by defects [59] and spin flips in an Ising model mediated by disorder [54]. There is also a deep mathematical literature on the *modeling* of hysteresis, which is not so much concerned with its origins, but rather with the accurate simulation of hysteresis loops [42]. These kinds of simulations have been used in industry, for example to quantify whether the hysteresis of a new batch of material is the same as that of the last batch, but they are not so useful for the discovery of materials.

The concept we will explore is related to the ubiquitous austenite/martensite interface, the crystallographic theory of martensite, and regularized models. As explained in Section (4), the austenite/martensite interface is a low energy structure modeled by a minimizing sequence, leading to the algebraic problem (24). We also noted some discrepancies between theory and experiment, notably the nonzero scale of the twin bands on the left of Figure 4. This is believed to be related to a small interfacial energy on the boundaries of these bands. Finer and finer bands (at fixed volume fraction λ) reduce the elastic energy in the transition layer near the austenite/martensite interface, but have more and more interfacial energy. Coarse bands, on the other hand, have little interfacial energy but unacceptable elastic energy in the transition layer. People believe that what you see in Figure 4 is a compromise between these two tendencies, that in fact is captured by regularized models [34, 36, 35]. The story is subtle: many regularized models assert that the transition layer is delocalized, and is accompanied by branching of the bands as seen in Figure 4. "Delocalized"' means that the elastic energy arises not from a layer parallel to the austenite/martensite interface, as in (20), but rather arises from the martensite bands having normals slightly perturbed away from n, due to the branching. In martensites in hard materials, one pays a big energetic penalty for perturbing n. More on this below.

So, there are two sources of energy, both positive, that are missed by the minimizing sequence: the interfacial energy on the boundaries of the martensite bands, and the elastic energy in the transition layer. We can imagine that this sets up an *energy barrier*. On cooling from the austenite

---

[11]see Section 1.4 of [10] for a broad summary

phase, as soon as a nucleus of martensite (such as that seen in Figure 5) forms, it is accompanied by extra interfacial and transition-layer energy. For it to grow, a lowering of the temperature below θ_c is required in order that the lowering of the martensite well with temperature compensates for the barrier. A similar process could happen upon heating. As soon as austenite forms, it must be accompanied by transition-layer and interfacial energy, and another barrier is set up, requiring a increase of temperature to above θ_c. Hence, hysteresis.

Figure 5: Schematic nucleus of martensite, bounded by two austenite/martensite interfaces.

Mathematically, the identification of energy barriers is not part of the usual calculus of variations. Linear stability analysis, e.g., the study of the loss of positivity of the second variation as a parameter θ is varied, identifies energy barriers, but clearly the study of the second variation would miss the barrier identified here. The "nucleus" shown in Figure 5 is missed by linearization about the undistorted austenite phase: it is a large, localized disturbance. More precisely, at the measured temperature $\theta < \theta_c$, on the shoulder of the hysteresis loop (Figure 1) where the martensite begins to grow, the second variation at the ambient austenite phase is strongly positive-definite. Promising emerging methods [32], [33] not based on linearization exploit the fact that the nucleus has small support, together with the energy well structure and compatibility. Otherwise, the time honored method is: make a clever ansatz. Numerically, one can collect a zoo of ansatzes, but, without insight from experiment, this does not seem hopeful. Fortunately, in the present case, one can make a reasonable ansatz, Figure 5.

This ansatz has been studied in some detail [69, 70]. First, one notes that there are no singularities at the ends and the curvature of the interface plays an insignificant role: one may as well study the energy of two parallel austenite/martensite interfaces, as the width w of the nucleus is varied. Of course, one needs to include interfacial energy per unit area on the boundaries of the bands as well as the transition-layer energy, both of which can be included in the context of an ansatz. Sure enough, there is a barrier when $\theta < \theta_c$. When w is small, the interfacial/transition-layer energies dominate leading to an increase of energy with w. When w is large, the bulk energy dominates, due to $\varphi(U_1, \theta) = \varphi(\hat{R}U_2, \theta) < \varphi(R, \theta)$ in the notation of (19)-(24), leading to

a linear decrease of energy with w. For the state-of-the-art, see [70].

If this barrier is indeed responsible for hysteresis, there is obvious experimental test. It is related to an elementary lemma.

Lemma 5.1 *[3] Let the symmetric 3×3 matrix have ordered eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3$ and corresponding orthonormal eigenvectors e_1, e_2, e_3. A necessary and sufficient condition that there exists $Q \in SO(3)$ such that $QU_1 - I = b \otimes m$ is $\lambda_2 = 1$. The solutions are expressible in the form*

$$b = \rho \left(\sqrt{\frac{\lambda_3(1 - \lambda_1)}{\lambda_3 - \lambda_1}} \, e_1 + \sigma \sqrt{\frac{\lambda_1(\lambda_3 - 1)}{\lambda_3 - \lambda_1}} \, e_3 \right),$$

$$m = \frac{1}{\rho} \left(\frac{\sqrt{\lambda_3} - \sqrt{\lambda_1}}{\sqrt{\lambda_3 - \lambda_1}} \right) \left(-\sqrt{1 - \lambda_1} \, e_1 + \sigma \sqrt{\lambda_3 - 1} \, e_3 \right). \quad (25)$$

where $\sigma \in \pm 1$ and $\rho \neq 0$.

Proof. Operate $(QU_1)^T(QU_1) = U_1^2 = (I + m \otimes b)(I + b \otimes m)$ on $b \times m$ to see that one eigenvalue of U_1 is 1, then write $U_1^2 = (I + m \otimes b)(I + b \otimes m)$ in the eigenvector basis of U_1 to show it is the middle eigenvalue, and also that $\lambda_2 = 1$ is sufficient.

The experimental test is at hand, when one realizes that $QU_1 - I = b \otimes m$ is necessary and sufficient that there is a continuous function $y : \Omega \to \mathbb{R}^3$ taking the values QU_1 (martensite) and I (austenite). (Note also that the matrices $U_1, \ldots, U_n$ all have the same eigenvalues.) For examples, see the pictures at $f = 0, 1$ in Figure 7. So, if $\lambda_2 = 1$, there is no need of the fine bands, no need of the transition layer. We can transform by passing a single plane separating austenite and martensite through the material, at least for a single crystal. In principle, there could be a small interfacial energy on this plane but, essentially, the barrier is gone. This is also seen by specialization of the results in [69, 70] to the case $\lambda_2 = 1$.

How can this be an experimental test? The matrix U_1 (and therefore $U_2, \ldots U_n$) are properties of the material. While it is difficult to change the symmetries (5) and (6) of a material, the transformation stretch matrix U_1 does change with composition. So, start with a material having U_1 with a middle eigenvalue reasonably close to 1, and tune the composition to make λ_2 exactly 1. When this was done[12], the results

---

[12]first, primitively, in the lab of the author

were astonishing. Later, people who do combinatorial synthesis[13] tried this and, in fact, these studies very much highlighted the importance of "combi-methods". A collection of measurements done with both combi and bulk methods is shown in Figure 6. Each marker is a different alloy. One can see that the hysteresis can be reduced to near zero by tuning the composition to make $\lambda_2 = 1$, some of these alloys being exceptional. The perfect interfaces of "$\lambda_2 = 1$ alloys" have been seen in transmission electron microscopy [23], and in fact the angle in Q and b, m can be measured from the micrographs, showing nice agreement with Lemma 5.1. This is not at all restricted to NiTiX alloys. Today, there is a lively ongoing effort to make new low hysteresis alloys this way.
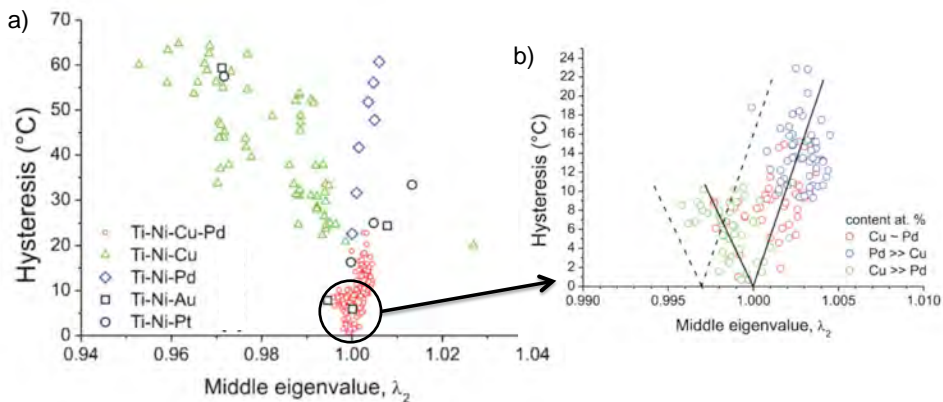

Figure 6: Measured width of the hysteresis of alloys in the NiTiX system ([21, 69, 68]). Each marker is a different alloy. Panel b) is a close-up of a) near $\lambda_2 = 1$. Note that the width of the hysteresis loop can be reduced to near zero.

There is another interesting consequence that is revealed by combi methods [21]. The X-ray measurement that gives values of λ_2 for the many alloys, also gives measurements of λ_1, λ_3. The product $\lambda_1\lambda_2\lambda_3$ is the volume ratio of the two phases. If $\lambda_1\lambda_2\lambda_3 \neq 1$ and the new phase nucleates on the interior, it grow up in a hole of the wrong volume. Changing the volume of the hard material requires a lot of energy, and one would think this would set up an energy barrier. However, when hysteresis is plotted vs. $\lambda_1\lambda_2\lambda_3$ for combinatorial library, unlike in Figure 6, there is no clear correlation. In fact, some of the alloys with the biggest hysteresis and some with the smallest hysteresis have $\lambda_1\lambda_2\lambda_3 = 1$. Perhaps nucleation does occur from the boundary?

---

[13]notably I. Takeuchi and A. Ludwig and their groups

The NiTiX system is the most intensely studied transforming material system. In recent years it has been the most intensely studied system among all metal alloys [47]. People have made many thousands of NiTiX alloys and characterized them. How did they miss the obvious sharp drop of Figure 6? The reason is that λ_2 is extremely sensitive to composition (see Fig. 1 of [17]). When people previously made their alloy series, they always jumped over the composition at which $\lambda_2 = 1$. Without a theory, it can be hard to find a singularity.

6   Supercompatibility

We can appreciate from these results on hysteresis, that a) for some aspects of material behavior, there really is a difference between minimizing sequences and minimizers in the problem (3), b) regularized models reveal this difference quite clearly, and c) degeneracies like $\lambda_2 = 1$ have an important effect on hysteretic behavior. By satisfying $\lambda_2 = 1$ we also disrupt the balance between bulk and interfacial energy. From a practical viewpoint, in a "$\lambda_2 = 1$ material" we might still expect to see a compatible interface in a 100nm (or even 10nm) crystal (or grain) , but if λ_2 is not extremely close to 1 we are unlikely to see an austenite/martensite interface at these scales.

It seems unlikely that there is any further lowering of the energy barrier between austenite and martensite that is possible than by having a perfect, untwinned austenite/martensite interface implied by $\lambda_2 = 1$. But there is the potential to find degeneracies that allow many possible low energy ways to mix austenite and martensite. Such conditions could remove barriers that form when, say, several $\lambda_2 = 1$ interfaces are forced to meet, such as at a grain boundary. One such degeneracy is embodied in the *cofactor conditions* [30].

The cofactor conditions are degeneracy conditions of the crystallographic theory of martensite, which we have reduced to (24). We suppose as above that the twinning equation $\hat{R}U_2 - U_1 = a \otimes n$ has been satisfied. To solve this for $\hat{R} \in SO(3)$, $a, n \in \mathbb{R}^3$ one can recast it in the form of Lemma 5.1 or use the Proposition 12 from [17]. This proposition states, under the conditions that the 3×3 matrix U_1 is symmetric and positive-definite and $U_2 = QU_1Q^T$ for some $Q \in SO(3)$ (all of which are assumed above), there is a solution $\hat{R} \in SO(3)$, $a, n \in \mathbb{R}^3$ of $\hat{R}U_2 - U_1 = a \otimes n$ if and only if there is $\hat{e} \in \mathbb{R}^3, |\hat{e}| = 1$, such that

$$U_2 = (-I + 2\,\hat{e} \otimes \hat{e})U_1(-I + 2\,\hat{e} \otimes \hat{e}), \qquad (26)$$

Formulas for $\hat{e}, a, n$ are given in [17] (see (10) and (A.1)-(A.6) there). We note that there are usually lots of 180 degree rotations $(-I+2\,\hat{e}\otimes\hat{e})$ in the Laue group G^a, so lots of pairs of matrices U_i and U_j, not just $i = 2, j = 1$, satisfy the twinning equation, and can be the basis of constructing austenite/martensite interfaces.

Given $\hat{R}, a, n$ we now turn to the second condition $R(U_1+\lambda a\otimes n) = I + b\otimes m$ of the crystallographic theory of martensite, to be solved for $R \in \mathrm{SO}(3)$, $0 \le \lambda \le 1$, and $b, m \in \mathbb{R}^3$. Following [3] we eliminate R by calculating $(I + b\otimes m)^T(I + b\otimes m)$ to get the necessary condition

$$G(\lambda) := (U_1 + \lambda n \otimes a)(U_1 + \lambda a \otimes n) = (I + m \otimes b)(I + b \otimes m). \quad (27)$$

This condition is sufficient for the existence of $R \in \mathrm{SO}(3)$ if $\det(I + b\otimes m) = 1+b\cdot m > 0$. This follows from the polar decomposition theorem and the fact that $\det(\cdot)$ is rank-1 affine, and so $\det(U_1 + \lambda a \otimes n) = \det(\lambda\hat{R}U_2 + (1 - \lambda)U_1) = \det U_1 > 0$. One can further notice that (27) is related to Lemma 5.1, and therefore is solvable with $1 + b \cdot m > 0$ if and only if the $0 \le \lambda \le 1$ can be chosen so that the middle eigenvalue of $G(\lambda)$ is 1.

Hence, we seek $0 \le \lambda \le 1$ such that the middle eigenvalue of $G(\lambda)$ is 1. If so, necessarily, $\det(G(\lambda) - I) = 0$. This looks like a 6th order polynomial in λ, but, due again to the fact that $\det(\cdot)$ is rank-one affine, it is actually quadratic and symmetric about $1/2$. Thus, aside from the issue of whether the resulting eigenvalue $= 1$ of $G(\lambda)$ is the middle one (an inequality given below), the crystallographic theory of martensite reduces to the question of whether a particular symmetric, quadratic function has roots $0 \le \lambda^* \le 1$ and $1 - \lambda^*$. In fact, if one uses the matrices U_1 and U_2 (appropriately selected!) for the alloy $Cu_{69.5}Al_{27}Ni_{3.5}$ shown in Figure 4, one does in fact have such roots and the resulting solution agrees nicely with Figure 4, with the qualifications mentioned above.

The quantity $\det(G(\lambda) - I)$ can of course be evaluated for any material that undergoes a phase transformation and has an energy well structure with the symmetries of Section 3. Materials with an appreciable value of $|U_1 - I|$ for which $\det(G(\lambda) - I)$ has no roots are usually not reversible.

For a "$\lambda_2 = 1$ material" as discussed in Section 5, necessarily we have solutions of the crystallographic theory because the second of (24) is satisfied at $\lambda = 0, 1$. That is, the symmetric, quadratic function $\det(G(\lambda) - I)$ satisfies $\det(G(0) - I) = \det(G(1) - I) = 0$, and, of course, the roots $0, 1$ give middle eigenvalues of $G(\lambda)$.

In this framework an obvious degeneracy presents itself: the function $\det(G(\lambda) - I)$ could be *identically zero*. Then, assuming the roots obtained give *middle* eigenvalues of $G(\lambda)$, we would have solutions of
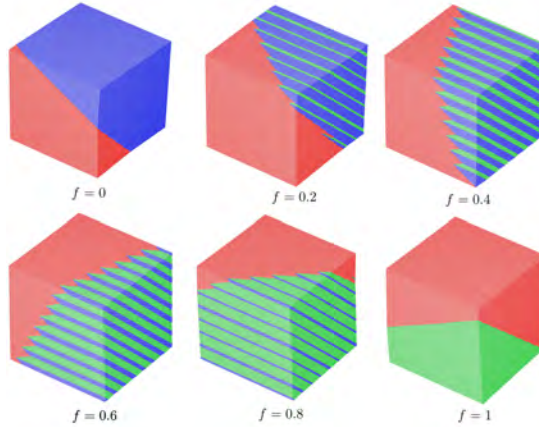
(a) austenite/martensite with Type I twins
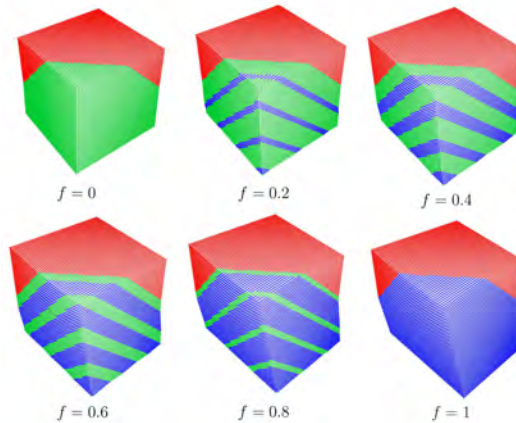
(b) austenite/martensite with Type II twins

Figure 7: Zero elastic energy austenite/martensite interfaces possible under the cofactor conditions, from [17]. Red is austenite and blue/green are two variants of martensite. These pictures exhibit large deformations, zero elastic energy and perfect fitting of the phases, under continuous variation of the volume fraction f. For Type II twins the cofactor conditions imply that the twin boundaries are parallel to the austenite/martensite interface, which clearly makes for easy construction of zero elastic energy microstructures.

the crystallographic theory for every $0 \leq \lambda \leq 1$. A quadratic function, symmetric about $1/2$, is identically zero if and only if its value at 0 and its derivative at 0 are zero. These two conditions, together with an inequality that is necessary and sufficient that the eigenvalues $= 1$

obtained are the middle ones, are the *cofactor conditions* [30, 17]:

$$\lambda_2 = 1, \quad a \cdot U_1 cof(U_1^2 - I)n = 0, \quad trU_1^2 - \det U_1^2 - \frac{|a|^2|n|^2}{4} \geq 2. \quad (28)$$

We review the known alloys that have been tuned to satisfy the cofactor conditions in Section 7.

The cofactor conditions depend on the "twin system" a, n. It is easily seen by operating $Q \ldots Q^T$ on (27), $Q \in G^a$, that its satisfaction for one twin system implies its satisfaction for other twin systems, and there can be further multiplicities of this type [17], depending on the symmetries.

The cofactor conditions imply a plethora of additional austenite/ martensite interfaces modeled by minimizing sequences, but it is not obvious that they provide additional zero elastic energy structures beyond those guaranteed by $\lambda_2 = 1$. It seems from the results of Section 6 that elimination of the elastic energy altogether might be most important. But, as degeneracy gets piled on degeneracy, there can be other unexpected accidents, and that is the case here. Some but not all of these are collected in Figure 7. Briefly, to understand this figure, one needs to know that

there is a classification of solutions of the twinning equation into Type I, Type II and Compound twins. (In microstructures with austenite and compound twins it is not known if one can eliminate the elastic energy.) The details can be found in [17]. Not pictured are zero elastic energy curved austenite/martensite interfaces, and zero elastic energy mechanisms for nucleation. There may well be other famlies of zero energy microstructures, and a compete understanding is lacking. It will be shown in Section 7 that unprecedented behavior is seen in the two known alloys that satisfy the cofactor conditions.

We should mention that there is potentially a completely different interpretation of the cofactor conditions. This concerns the relaxed energy [38]. The relaxed energy is the limiting energy of the lowest energy minimizing sequence having a given weak limit. From a materials science viewpoint: fix the average deformation, find a (possibly complex) microstructure that minimizes the energy and has this given average deformation. That is, assume $1 < p < \infty$ and Ω bounded and open, with a Lipschitz boundary. Let $y \in W^{1,p}$ be given, and for sequences

$$y^{(j)} \rightharpoonup y \text{ in } L^p \quad and \quad \nabla y^{(j)} \rightharpoonup \nabla y \text{ in } L^p \quad (29)$$

minimize the energy:

$$E_{macro}(y) = \inf_{\{y^{(j)}\}} \left\{ \liminf_{j \to \infty} \int_\Omega \varphi(\nabla y^{(j)}, \theta_c) \, dx \right\}. \quad (30)$$

24

It is known [38] (see also [27]) that if $\varphi(F, \theta_c) > c|F|^p$, $c > 0$, when $|F|$ is large, then E_{macro} is representable in terms of a macroscopic energy density $\tilde{\varphi}$ by

$$E_{macro}(y) = \int_\Omega \tilde{\varphi}(\nabla y(x)) \, dx. \qquad (31)$$

Properties of $\tilde{\varphi}$ are given by Kristensen [38]. Of interest here is its zero level set. Does it suddenly get larger when the cofactor conditions are satisfied? Naively, one would think "yes", since the new austenite/martensite interfaces, exhibited for example in Figure 7, should enlarge the flat region present on $\tilde{\varphi}$.

Our vague sense is rather that the presence of the zero elastic energy minimizers, which disrupts the delicate balance between elastic and interfacial energy, is perhaps most important consequence of the cofactor conditions. Therefore, by *supercompatibilty* we shall mean degeneracy conditions like $\lambda_2 = 1$ or the cofactor conditions that permit the phases to fit together with finite interfacial area and without stressed transition layers. Another interesting recent example is [20]. With a good collection of examples one can hope for a general theory of supercompatibility.

7 Reversibility

Two alloys have been found that accurately satisfy the cofactor conditions: $Zn_{45}Au_{30}Cu_{25}$ [56] and $Ti_{54.7}Ni_{30.7}Cu_{12.3}Co_{2.3}$ [18]. Briefly, they both exhibit exceptional reversibility of the phase transformation. They are quite different alloys both chemically and structurally: the ZnAuCu alloy undergoes a cubic to monoclinic transformation while the TiNiCuCo alloy undergoes a cubic to orthorhombic transformation. The ZnAuCu alloy was found by systematic alloy development: make a specimen starting from high purity elemental Zn, Au and Cu in the right proportions, check for changes of composition that may have occurred by loss to the environment during melting, measure accurately the lattice parameters of both phases by X-ray methods, calculate the quantities in (28) and repeat. After several specimens, one develops the relation between (28) and composition, from which satisfaction of (28) to high accuracy is possible. The TiNiCuCo alloy was made by thin[14] film sputtering methods, removal the substrate and polishing the boundary of the specimen. Both specimens were made to undergo stan-

---

[14]actually quite thick, so it would be considered more like bulk material

dard heat treatments after synthesis. Further information on synthesis and processing can be found in [18, 29]

To understand reversibility, a nice test to do is the *shape memory effect*: cool the material from austenite, deform the material in the martensite phase (which rearranges the variants) leaving the material with a large overall deformation, heat it up and it returns to its starting shape as the martensite transforms back to austenite. But this heating and cooling would take too long for thse highly reversible alloys, i.e., for these materials such tests would take years.
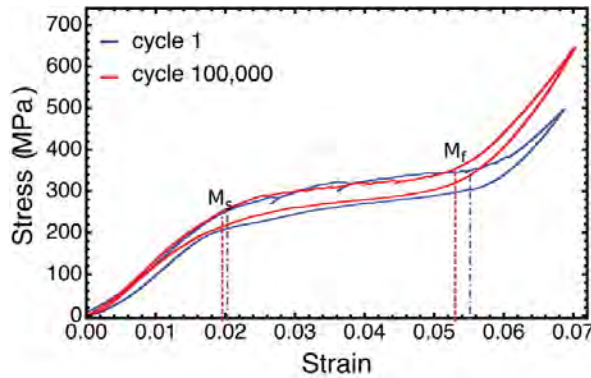
(a) $Zn_{45}Au_{30}Cu_{25}$ [46, 56]
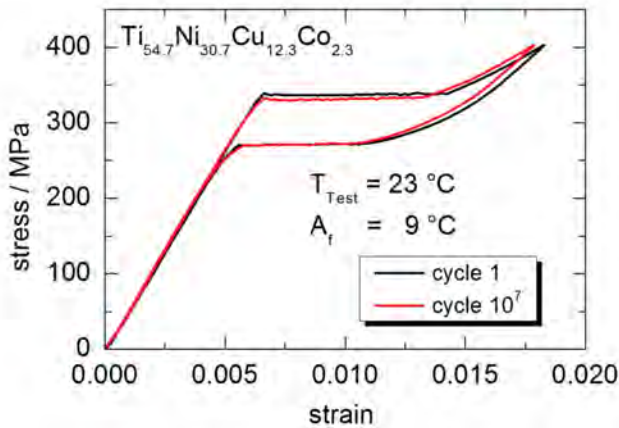
(b) $Ti_{54.7}Ni_{30.7}Cu_{12.3}Co_{2.3}$ [18]

Figure 8: Stress-induced transformation in ZnAuCu and TiNiCuCo. See text. (a) is reprinted with permission from [46], copyright 2016, American Chemical Society. (b) is from [18] and is reprinted with permission from AAAS.

Another test, which in many ways is even more demanding of the material but can be done in a matter of weeks, is *stress-induced trans-*

formation. In the simplest case – say the stress is a uniaxial tension in the direction $e, |e| = 1$, with stress $\sigma > 0$ – this corresponds mathematically to the study of the a modified energy minimization problem

$$\inf_{y \in \mathcal{A}} \int_{\Omega} (\varphi(\nabla y(x), \theta) - \sigma\, e \cdot \nabla y e)\, dx. \tag{32}$$

In the case of these two alloys the main effect of this particular stress is to raise the transformation temperature, which can be easily understood by solving[15] (32). Thus, one can do stress-induced transformation by fixing the temperature above θ_c and increasing the stress σ until the material transforms. This is seen as a flat region on the measured stress-strain curve.

Basic information on reversibility is seen in Figure 8. The ZnAuCu alloy shows nearly the same response after 100,000 cycles, under demanding conditions of almost 7% strain each cycle and peak (compressive) stresses of more than 500 MPa. The TiNiCuCo alloys is in many ways more impressive: even though the strain is lower (almost 2%) these tests were done in tension (more demanding) and the stress-strain curve at cycle 1 is extremely close to that at cycle 10 million. In both cases the material was made to undergo nearly complete transformation to martensite each cycle. For mathematicians not familiar with these units, a typical value of the yield stress of the steel that holds up a department of mathematics is 300 MPa. Several other measurements test reversibility in other ways. For example, $Zn_{45}Au_{30}Cu_{25}$ has a remarkably low thermal hysteresis (e.g., Figure 1), as low as 0.2°C.

In general martensites widely exhibit a high degree of repeatability of the pattern of microstructure on heating and cooling [54]. Transform a typical martensitic material by cooling and one sees a pattern of microstructure. Heat to austenite (which wipes out the microstructure) and cool again: the pattern is very nearly the same. Often, in a polycrystal, plates of martensite appear by growing out of a defect or triple junction in the same way each cycle.

Thus, it comes as a striking observation that, when $Zn_{45}Au_{30}Cu_{25}$ is heated and cooled back and forth through the transformation, the

---

[15]This is easily done under mild growth conditions on φ. To see the simplest asymptotic result quickly, let $K = \mathrm{SO}(3)U_1 \cup \cdots \cup \mathrm{SO}(3)U_n$, assume that $\varphi(\cdot, \theta)$ rises steeply from the energy wells, note that one can minimize the integral by minimizing the integrand, and therefore reduce the problem to $\min_{F \in \mathrm{SO}(3) \cup K}(f_a(\theta)\chi_{\mathrm{SO}(3)} + f_m(\theta)\chi_K - \sigma e \cdot Fe)$. The energy wells for the ZnCuAu alloy can be found in the supplement of [56].

microstructure is completely different each cycle[16]. See Figure 9 or, for the full video, the supplement of [56], or, at a bit lower resolution, `http://www.aem.umn.edu/~james/research/`. How can such a highly reversible alloy behave in such a highly nonrepeatable way? Is it that,

(a) cycle 1 (b) cycle 2

(c) cycle 3
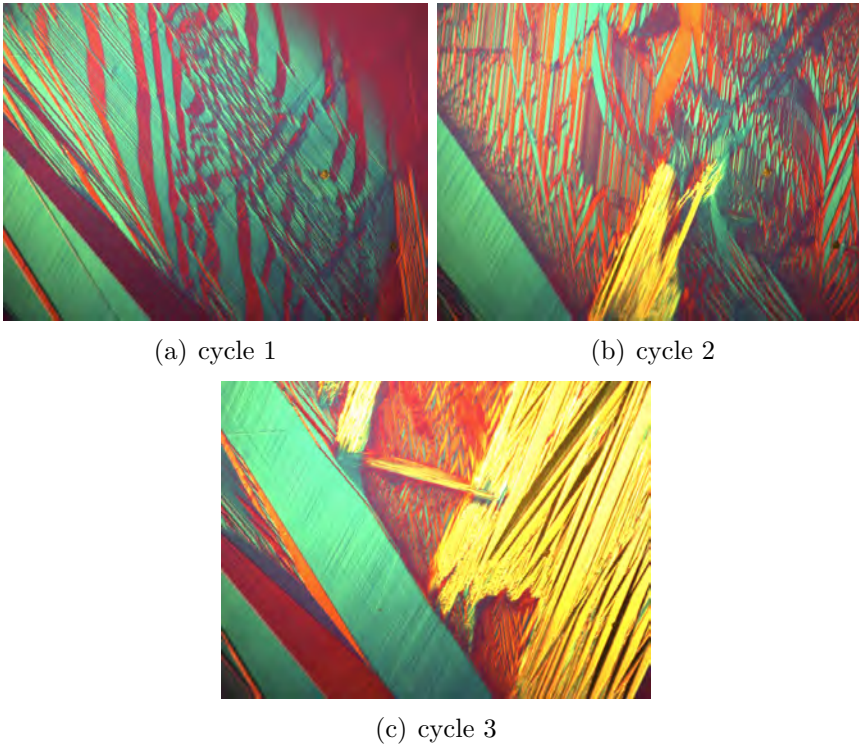
Figure 9: Three successive transformation cycles at the same location in a grain showing the non repeatability of microstructure in $Zn_{45}Au_{30}Cu_{25}$.

by satisfying the cofactor conditions, we have so flattened its infinite-dimensional energy landscape that the material can take any path? Sounds good at first, until one recalls that, in an ordinary martensite, the plates of martensite tend to emerge from triple junctions and defects. Shouldn't the defects then even more strongly bias the microstructure with an otherwise flat energy landscape?

These and related questions have inspired mathematicians to look more critically at this video. We collect two observations that are particularly interesting. One is due to J. M. Ball and F. Della Porta [22] and is illustrated by Figure 10. It is as if the microstructure is already set behind a blanket, and one moves the blanket aside to reveal it. That is, there is very little further relaxation once the microstructure

---

[16]It should be mentioned that every effort was made in the heating/cooling device to give a periodic temperature profile vs. time.

appears. Perhaps this also can be rationalized by a very flat energy landscape – or, anthropomorphically, wherever you are, there is little driving force to push you elsewhere. Upon reflection, it is clear that this is a very strong restriction on the microstructure: previously, it was compatible with the austenite across a (possibly irregular) interface [22] and, after that, it did not change. In general, this kind of restriction, which embodies the idea that there must have been a

(a) frame 1

(b) frame 2

(c) frame 3
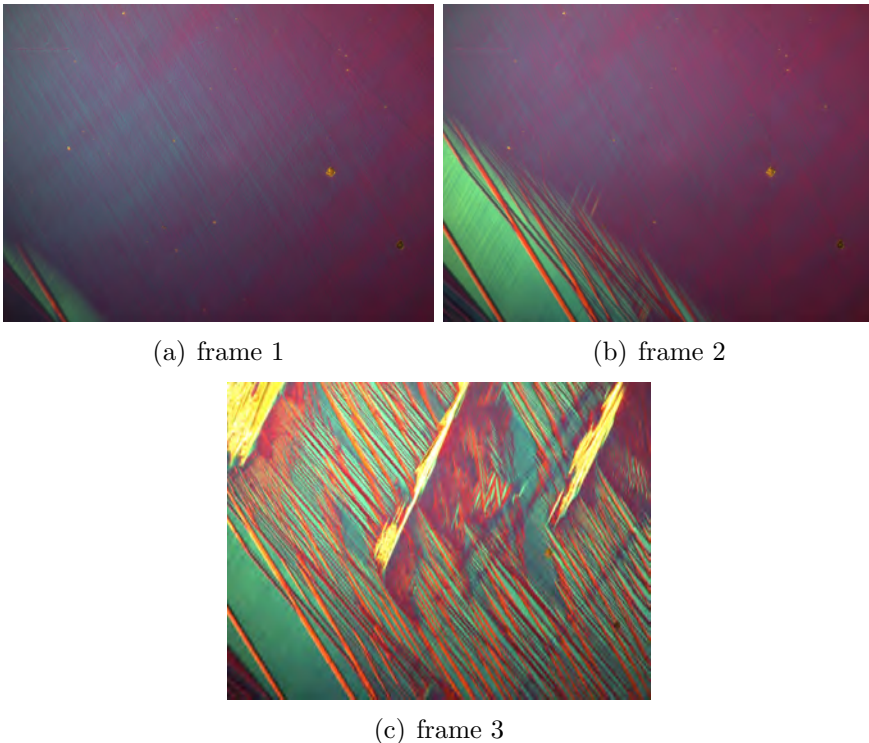
Figure 10: The process of transformation in $Zn_{45}Au_{30}Cu_{25}$ seen in three consecutive frames. Austenite is dark gray. Notice that once the microstructure appears there is very little further relaxation.

low energy pathway to an observed energy minimizing state, has not been studied much, and also relates to the study of barriers mentioned above. For another interesting barrier see the wonderful experiment of H. Seiner [6].

A fascinating observation on $Zn_{45}Au_{30}Cu_{25}$ due to Noemi Barrera and Giovanni Zanzotto [7] relates to power-law behavior [45, 19] and the theory of self-organized criticality [39]. It is known that the martensitic phase transformation often takes place through abrupt strain events ("avalanches") even when the temperature or loading is smoothly changing [48, 2]. Figure 11 shows a plot of the empirical probability of

avalanche sizes, where size S refers to the number of pixels in a connected domain that undergoes an austenite-to-martensite switch of color between successive frames. They note [7] that the empirical avalanche statistics of $Zn_{45}Au_{30}Cu_{25}$ (labelled CC1 + CC2) have an exceedingly good power-law character, to a degree that is rare in ma-
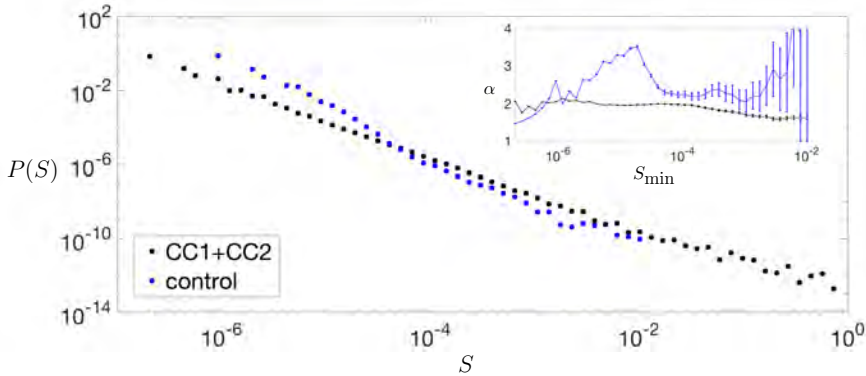
Figure 11: Empirical probability $P(S)$ for the size S of transformation avalanches during the video (supplement, [56]), where S is the number of pixels in connected domains where there is an austenite-martensite switch of color. The plot (blue) shows power-law behavior with exponent near 2 over a remarkable range of 6 orders of magnitude for $Zn_{45}Au_{30}Cu_{25}$ (CC1+CC2). The control plot (blue) is measured in the same way on a generic alloy. The inset shows the values of the power-law exponent α for CC1+CC2 (black) vs. the generic control (blue), determined by the maximum likelihood method [19] as a function of the lower cutoff imposed to the data.

terials science. Qualitatively, this lack of one or more characteristic scales indicates that $Zn_{45}Au_{30}Cu_{25}$ can perform a much wider and more efficient collection of adjustments of microstructure to environmental changes. It is a striking example of this type, that includes sand piles, earthquakes, stick-slip friction, the firing of neurons, and fluctuations in financial markets. Of course, it would be good to relate this behavior to the theory given above; see [5] for a model in this direction and [53] for recent diverse perspectives on the origins of behavior such as that shown in Figure 11.

8 The direct conversion of heat to electricity

We finish this article with a few brief remarks about one of the most interesting[17] applications of reversible transforming materials. It concerns the use of these materials for the direct conversion of heat to electricity in the small temperature difference regime. Here, "direct" means that the material itself creates the electricity without a separate electrical generator. The "small temperature difference regime" is the regime $10 - 200°$ C, for which there does not currently exist a reasonable energy conversion device. Sources in this regime are ubiquitous: concentrated solar-thermal sources, data centers (which now consume $\sim 3\%$ of the energy budget in the US), waste heat from industrial sources, desktop and laptop computers, air conditioning systems, power plants, and even hand-held electronic devices[18]

There are at several ways that transforming materials can be used for the direct conversion of heat to electricity, and we will mention two of them. They are enabled by the abrupt change of magnetoelectric properties that can occur in materials with big first order phase transformations, like those discussed here. The two cases are based on magnetism and ferroelectricity. In the first case one uses a material for which the low temperature phase is non-magnetic and the high temperature phase is strongly magnetic. For an example of such a material, which also has λ_2 quite close to 1, see [58].

One heats the material through the phase transformation. If left alone it would demagnetize itself[19] by forming domains, so we place it on top of a permanent magnet to bias it. This biasing can be well understood from the theory of micromagnetics. As it transforms to the strongly magnetic phase as we heat it up, it magnetizes. We can think of magnetization as a time-dependent vector field $M : \Omega \times [0, t_1) \to \mathbb{R}^3$ satisfying $|M(x, t)| = \chi_{\Omega_1(t)}(x)M_s$ on the region Ω occupied by the material. Here, $\Omega_1(t) \subset \Omega$, $\Omega_1(0) = \emptyset$, $\Omega_1(t_1) = \Omega$, much like the frames in Figure 10 run backwards. By heating we increase $\fint_\Omega M(x, t)dx$.

---

[17]in the opinion of the author

[18]for which, of course, the generated electricity would be used to help recharge the battery. In some of the computer examples mentioned, it is in fact a significant technological obstacle to *get rid of the heat.*

[19]since it is a soft magnet, which means, from the point of view of energy minimization, that its static behavior is well predicted by absolute energy minimization based on the theory of micromagnetics.

Ferromagnetism is governed by a well-known dipolar relation[20]

$$B = \mu_0(H + M) \tag{33}$$

as well as the Maxwell equations

$$\mathrm{div}\, B = 0, \quad \mathrm{curl}\, E = -\frac{\partial B}{\partial t}. \tag{34}$$

Interpreted physically, the latter means that, if the material is surrounded by a coil, a current will be generated, or, briefly, induction. In fact, one gets a current of opposite sign on cooling back through the phase transformation.

This is the barest explanation, but there are many subtle aspects. For example, the presence of the field of the permanent magnet changes the transformation temperature, since there is an effect of magnetic field on transformation temperature. Moreover, the current induced in the coil modifies this field in an important way. The effect on heating and cooling is different, and one gets two transformation temperatures. This splitting of the transformation temperature turns out to be terribly important to the efficiency and power output of such a device. Of course, minimizing the hysteresis is also critical. See [57] for a basic model.

As one can see from (33) and (34) the rapidly changing M is partitioned between B and H, but it is $\partial B/\partial t$ that creates the electricity. This partitioning is well-known to mathematicians who work in micromagnetics. A study of this partitioning reveals a deficiency of this method: the good shape of Ω for a favorable $\partial B/\partial t$ seems to be a bad shape for the also crucial aspect of heat transfer. This dilemma suggests that, in fact, the ferroelectric case (with capacitance, instead of induction) is preferred, as will be explained in forthcoming work.

References

[1] Edgar C. Bain and N. Y. Dunkirk. The nature of martensite. *Trans. AIME*, 70(1):25, 1924.

---

[20]For a mathematical perspective on this relation see [31].

[2] Xavier Balandraud, Noemi Barrera, Paolo Biscari, Michel Grédiac, and Giovanni Zanzotto. Strain intermittency in shape-memory alloys. *Physical Review B*, 91(17):174111, 2015.

[3] J. M. Ball and R. D. James. Fine phase mixtures as minimizers of energy. *Arch. Ration. Mech. Anal.*, 100(1):13, 1987.

[4] J. M. Ball and Richard D. James. Proposed experimental tests of a theory of fine microstructure and the two-well problem. *Phil. Trans.: Phys. Sci. Eng.*, 338(1650):389, 1992.

[5] John M Ball, Pierluigi Cesana, and Ben Hambly. A probabilistic model for martensitic avalanches. In *MATEC Web of Conferences*, volume 33. EDP Sciences, 2015.

[6] John M Ball, Konstantinos Koumatos, and Hanuš Seiner. Nucleation of austenite in mechanically stabilized martensite by localized heating. *Journal of Alloys and Compounds*, 577:S37–S42, 2013.

[7] Noemi Barrera and Giovanni Zanzotto. Private communication. 2017.

[8] Pavel Belik and Mitchell Luskin. Stability of microstructure for tetragonal to monoclinic martensitic transformations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 34(3):663–685, 2000.

[9] Kaushik Bhattacharya. *Microstructure of martensite: why it forms and how it gives rise to the shape-memory effect.* Oxford series on materials modeling. Oxford University Press, 2003.

[10] Kaushik Bhattacharya, Sergio Conti, Giovanni Zanzotto, and Johannes Zimmer. Crystal symmetry and the reversibility of martensitic transformations. *Nature*, 428(6978):55–59, 2004.

[11] Kaushik Bhattacharya and Robert V Kohn. Symmetry, texture and the recoverable strain of shape-memory polycrystals. *Acta Mater.*, 44(2):529–542, 1996.

[12] Kaushik Bhattacharya, Bo Li, and Mitchell Luskin. The simply laminated microstructure in martensitic crystals that undergo a cubic-to-orthorhombic phase transformation. *Archive for rational mechanics and analysis*, 149(2):123–154, 1999.

[13] Xavier Blanc, Claude Le Bris, and P-L Lions. Convergence de modèles moléculaires vers des modèles de mécanique des milieux continus. *C. R. Acad. Sci. Paris Sér. I Math.*, 332:949–956, 2001.

[14] Xavier Blanc, Claude Le Bris, and P-L Lions. From molecular models to continuum mechanics. *Archive for Rational Mechanics and Analysis*, 164(4):341–381, 2002.

[15] Antonio Capella and Felix Otto. A quantitative rigidity result for the cubic-to-tetragonal phase transition in the geometrically linear theory with interfacial energy. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 142(2):273–327, 2012.

[16] Xian Chen, Yintao Song, Nobumichi Tamura, and Richard D. James. Determination of the stretch tensor for structural transformations. *Journal of the Mechanics and Physics of Solids*, 93:34 – 43, 2016. Special Issue in honor of Michael Ortiz.

[17] Xian Chen, Vijay Srivastava, Vivekanand Dabade, and Richard D. James. Study of the cofactor conditions: conditions of supercompatibility between phases. *J. Mech. Phys. Solids*, 61(12):2566, 2013.

[18] Christoph Chluba, Wenwei Ge, Rodrigo Lima de Miranda, Julian Strobel, Lorenz Kienle, Eckhard Quandt, and Manfred Wuttig. Ultralow-fatigue shape memory alloy films. *Science*, 348(6238):1004–1007, 2015.

[19] Aaron Clauset, Cosma Rohilla Shalizi, and Mark E. J. Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

[20] S Conti, M Klar, and B Zwicknagl. Piecewise affine stress-free martensitic inclusions in planar nonlinear elasticity. *Proc. R. Soc. A*, 473(2203):20170235, 2017.

[21] J. Cui, Y. S. Chu, O. Famodu, Y. Furuya, J. Hattrick-Simpers, R. D. James, A. Ludwig, S. Thienhaus, M. Wuttig, Z. Zhang, and I. Takeuchi. Combinatorial search of thermoelastic shape memory alloys with extremely small hysteresis width. *Nature Materials*, pages 286–290, 2006.

[22] Francesco Della Porta. Modeling moving interfaces in reversible martensitic transformations. Preprint, 2017.

[23] R. Delville, S. Kasinathan, Z. Zhang, V. Humbeeck, R. D. James, and D. Schryvers. A transmission electron microscopy study of phase compatibility in low hysteresis shape memory alloys. *Philosophical Magazine*, pages 177–195, 2010.

[24] J. L. Ericksen. On the Cauchy-Born Rule. *Math. Mech. Solids*, 13:199, 2008.

[25] Leonhard Euler. Recherches physiques sur la nature des moindres parties de la matière. *Berlin Academy Memoires*, 1:28–32, 1745. (*Opera omnia*, III. 1: 615).

[26] LC Flatley and Florian Theil. Face-centered cubic crystallization of atomistic configurations. *Archive for Rational Mechanics and Analysis*, 218(1):363–416, 2015.

[27] Irene Fonseca and Jan Malỳ. Relaxation of multiple integrals below the growth exponent. In *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, volume 14, pages 309–338. Elsevier, 1997.

[28] Gero Friesecke and Florian Theil. Validity and failure of the cauchy-born hypothesis in a two-dimensional mass-spring lattice. *Journal of Nonlinear Science*, 12(5), 2002.

[29] Hanlin Gu, Lars Bumke, Christoph Chluba, Eckhard Quandt, and Richard D. James. Phase engineering and supercompatibility of shape memory alloys. *Materials Today*, 2017. in press.

[30] R. D. James and Z. Zhang. A way to search for multiferroic materials with unlikely combinations of physical properties. In A. Planes, L Manõsa, and A. Saxena, editors, *Magnetism and Structure in Functional Materials, Springer Series in Materials Science*, volume 9, pages 159–175. Springer-Verlag, Berlin, 2005.

[31] Richard D James and Stefan Müller. Internal variables and fine-scale oscillations in micromagnetics. *Continuum Mechanics and Thermodynamics*, 6(4):291–336, 1994.

[32] Hans Knüpfer and Robert V Kohn. Minimal energy for elastic inclusions. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, page rspa20100316. The Royal Society, 2010.

[33] Hans Knüpfer, Robert V Kohn, and Felix Otto. Nucleation barriers for the cubic-to-tetragonal phase transformation. *Communications on pure and applied mathematics*, 66(6):867–904, 2013.

[34] Robert V Kohn. Energy-driven pattern formation. In *International Congress of Mathematicians*, volume 1, pages 359–383, 2006.

[35] Robert V Kohn and Stefan Müller. Branching of twins near an austenitetwinned-martensite interface. *Philosophical Magazine A*, 66(5):697–715, 1992.

[36] Robert V Kohn and Stefan Müller. Surface energy and microstructure in coherent phase transitions. *Communications on Pure and Applied Mathematics*, 47(4):405–435, 1994.

[37] K. Koumatos and A. Muehlemann. Optimality of general lattice transformations with applications to the bain strain in steel. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 472(2188), 2016.

[38] Jan Kristensen. A necessary and sufficient condition for lower semicontinuity. *Nonlinear Analysis: Theory, Methods & Applications*, 120:43–56, 2015.

[39] Anna Levina, J Michael Herrmann, and Theo Geisel. Dynamical synapses causing self-organized criticality in neural networks. *Nature physics*, 3(12):857–860, 2007.

[40] Mitchell Luskin. On the computation of crystalline microstructure. *Acta Numerica*, 5:191–258, 1996.

[41] Robert D. MacPherson and David J. Srolovitz. The von Neumann relation generalized to coarsening of three-dimensional microstructures. *Nature*, 446(7139):1053–5, 2007.

[42] Isaak D Mayergoyz. *Mathematical models of hysteresis and their applications*. Academic Press, 2003.

[43] Stefan Müller. Variational models for microstructure and phase transitions. In *Calculus of variations and geometric evolution problems*, pages 85–210. Springer, 1999.

[44] W. W. Mullins. Two-dimensional motion of idealized grain boundaries. *Journal of Applied Physics*, 27:900–904, 1956.

[45] Mark E. J. Newman. Power laws, pareto distributions and zipf's law. *Contemporary physics*, 46(5):323–351, 2005.

[46] Xiaoyue Ni, Julia R Greer, Kaushik Bhattacharya, Richard D James, and Xian Chen. Exceptional resilience of small-scale $Au_{30}Cu_{25}Zn_{45}$ under cyclic stress-induced phase transformation. *Nano letters*, 16(12):7621–7625, 2016.

[47] Patent iNSIGHT Pro. Shape Memory Material Technology Insight Report. Technical report, Gridlogics Technologies Pvt. Ltd., 2015.

[48] Francisco-José Pérez-Reche, Marcelo Stipcich, Eduard Vives, Lluís Mañosa, Antoni Planes, and Michel Morin. Kinetics of martensitic transitions in cu-al-mn under thermal cycling: Analysis at multiple length scales. *Physical Review B*, 69(6):064101, 2004.

[49] M Pitteri. Reconciliation of local and global symmetries of crystals. *Journal of Elasticity*, 14(2):175–190, 1984.

[50] M. Pitteri and G. Zanzotto. *Continuum models for phase transitions and twinning in crystals*. Chapman & Hall/CRC, 2003.

[51] Guang-Rui Qian, Xiao Dong, Xiang-Feng Zhou, Yongjun Tian, Artem R Oganov, and Hui-Tian Wang. Variable cell nudged elastic band method for studying solid–solid structural phase transitions. *Computer Physics Communications*, 184(9):2111–2118, 2013.

[52] Angkana Rüland. The cubic-to-orthorhombic phase transition: rigidity and non-rigidity properties in the linear theory of elasticity. *Archive for Rational Mechanics and Analysis*, 221(1):23–106, 2016.

[53] Ekhard K. H. Salje, Avadh Saxena, and Antoni Planes. *Avalanches in Functional Materials and Geophysics*. Springer, 2017.

[54] James P Sethna, Karin Dahmen, Sivan Kartha, James A Krumhansl, Bruce W Roberts, and Joel D Shore. Hysteresis and hierarchies: Dynamics of disorder-driven first-order phase transformations. *Physical Review Letters*, 70(21):3347, 1993.

[55] Cyril Stanley Smith. Grain shapes and other metallurgical applications of topology. In *Metal interfaces: a seminar on metal interfaces held during the 33$^{\rm rd}$ National Metal Congress and Exposition, October 13-19, Detroit*, pages 65–108, Cleveland, 1951. American Society for Metals.

[56] Y. Song, X. Chen, V. Dabade, T. W. Shield, and R. D. James. Enhanced reversibility and unusual microstructure of a phase-transforming material. *Nature*, 502:85–88, 2013.

[57] Yintao Song, Chris Leighton, and Richard D James. Thermodynamics and energy conversion in heusler alloys. In *Heusler Alloys*, pages 269–291. Springer, 2016.

[58] Vijay Srivastava, Xian Chen, and Richard D. James. Hysteresis and unusual magnetic properties in the singular heusler alloy $Ni_{45}Co_5Mn_{40}Sn_{10}$. *Appl. Phys. Lett.*, 97(1):014101, 2010.

[59] Likun Tan and Kaushik Bhattacharya. Length scales and pinning of interfaces. *Phil. Trans. R. Soc. A*, 374(2066):20150167, 2016.

[60] Florian Theil. A proof of crystallization in two dimensions. *Communications in Mathematical Physics*, 262(1):209–236, 2006.

[61] Clarence Marvin Wayman. *Introduction to the crystallography of martensitic transformations*. Macmillan, 1964.

[62] E Weinan, Weiqing Ren, and Eric Vanden-Eijnden. String method for the study of rare events. *Physical Review B*, 66(5):052301, 2002.

[63] L. C. Young. Generalized curves and the existence of an attained absolute minimum in the calculus of variations. *Comptes Rendus de la Société des Sci. et des Lettres de Varsovie*, 30:212–234, 1937.

[64] L. C. Young. Generalized surfaces in the calculus of variations. *Annals of mathematics*, pages 84–103, 1942.

[65] L.C. Young. *Lectures on the Calculus of Variations and Optimal Control Theory*. Saunders, 1969. Reprinted by A.M.S. Chelsea 1969.

[66] Giovanni Zanzotto. The cauchy–born hypothesis, nonlinear elasticity and mechanical twinning in crystals. *Acta Crystallographica Section A: Foundations of Crystallography*, 52(6):839–849, 1996.

[67] Nikolai A Zarkevich and Duane D Johnson. Shape-memory transformations of NiTi: minimum-energy pathways between austenite, martensites, and kinetically limited intermediate states. *Physical review letters*, 113(26):265701, 2014.

[68] R. Zarnetta, R. Takahashi, M. L. Young, A. Savan, Y. Furuya, S. Thienhaus, B. Maaß, M. Rahim, J. Frenzel, H. Brunken, Y. S. Chu, V. Srivastava, R. D. James, I. Takeuchi, G. Eggeler, and A. Ludwig. Identification of quaternary shape memory alloys with near zero thermal hysteresis and unprecedented functional stability. *Advanced Functional Materials*, pages 1917–1923, 2010.

[69] Zhiyong Zhang, R. D. James, and Stefan Müller. Energy barriers and hysteresis in martensitic phase transformations. *Acta Materialia (Invited Overview)*, 57:2332–4352, 2009.

[70] Barbara Zwicknagl. Microstructures in low-hysteresis shape memory alloys: Scaling regimes and optimal needle shapes. *Archive for Rational Mechanics and Analysis*, 213(2):355–421, 2014.

HOW COMPLICATED ARE POLYNOMIALS IN MANY VARIABLES?

CRAIG HUNEKE

ABSTRACT. The title question refers to systems of polynomial equations in many variables over a field. The question can be made precise in many ways, for example, through the complexity of detecting whether a given polynomial can be expressed as a linear combination (with polynomial coefficients) of other polynomials.

Another sense in which the question can be made precise is through comparisons of numerical data about the ideal generated by the polynomial equations, which generalize the numbers of generators and relations. Such additional numerical data was originally introduced in the 1890's by David Hilbert to count the number of polynomial invariants of the action of a group (this was the work that "killed" invariant theory for a brief time!). In the last two years, three long-standing problems about these numerical invariants have been solved.

These notes are the basis of my talk at the Current Events session, JMM, 2018, and will introduce the main themes in this story: Hilbert functions, free resolutions, projective dimension, Betti numbers, and regularity.

1. INTRODUCTION

What do the following have in common? Counting the number of k-colorings of a finite simple graph; finding the number of integer lattice points in kP as we dilate a polytope P by an integer k; finding the number of n by n magic squares with line sums k; counting the maximal number of unattacked squares in an n by n chess board after placing k queens on the board; determining the nature of the function that sends an integer k to the number of faces of dimension $k + 1$ in a given finite simplicial complex; counting the number of invariants of a fixed degree for a group acting linearly on a polynomial ring; and finding the lowest degree of a nonconstant morphism from a projective nondegenerate curve to the projective line?

The obvious answer is that all these problems involve finding numerical information. The non-obvious answer is that all of these problems can be analyzed and described by sets of polynomial equations. In practice this means that the numerical information being sought is given by the Hilbert function of a graded commutative algebra, or in some cases through the finer numerical information provided by graded Betti numbers. Our second section introduces these concepts.

The Hilbert function measures the vector space dimensions of the graded pieces of a graded commutative ring. It is the first in a series of finer and finer measurements one can make to analyze a wide variety of problems. This finer data, which is connected to free resolutions, is the main topic of this talk.

Three fundamental conjectures concerning the numerical invariants attached to the free resolutions of a quotient ring of a polynomial ring over a field have been resolved, one negatively, within the last two years. These notes will introduce the basic concepts and statements of the conjectures, and give at least a partial sense of the solution of one them.

All three of the conjectures deal with data associated to ideals in polynomial rings. Throughout these notes, $S = k[x_1, ..., x_n]$ denotes a polynomial ring over the complex numbers (or in general an algebraically closed field k of characteristic 0). Although some of the results hold in more generality, it is simpler to focus on this case. We say a polynomial $f(x_1, \ldots, x_n)$ is homogeneous of degree d if $f(tx_1, \ldots, tx_n) = t^d f(x_1, \ldots, x_n)$. For example, $x_1^3 + 4x_2x_3x_4$ is homogeneous of degree three, but $x_1^2 + x_2$ is not homogeneous. The set of all homogeneous polynomials of degree d, including the identically zero polynomial, is a k-vector space. In this way S decomposes as a sum of vector spaces $S = \oplus_{d=0}^{\infty} S_d$, where S_d consists of all polynomials of degree d, and such that $S_i \cdot S_j \subset S_{i+j}$. Evidently, S_d is finite-dimensional, having as a basis all monomials $x_1^{a_1} \cdots x_n^{a_n}$ where the a_i are nonnegative integers whose sum is d. A simple combinatorial argument shows that the number of such monomials in degree d is exactly $\binom{n+d-1}{n-1}$.

The generating function for the sequence of the dimensions of S_d, namely $H_S(z) = \sum_d dim_k(S_d)z^d$, can then be directly computed:

$$H_S(z) = \sum_d \binom{n+d-1}{n-1} z^d = \frac{1}{(1-z)^n}.$$

Observe that the generating function is a rational function.

Remark 1.1. It is not difficult to prove that the following two conditions concerning the generating function of a sequence of complex numbers $h(d)$, $d \geq 0$, are equivalent:

(1) $\sum_{d \geq 0} h(d)z^d = \frac{P(z)}{(1-z)^n}$, where $P(z) \in \mathbb{C}[z]$ is a polynomial of degree at most $n-1$; and
(2) $h(d)$ is a polynomial function of d of degree at most $n-1$.

We also observe that in these equivalent cases, $h(d)$ is a polynomial of degree exactly $n-1$ with leading coefficient $\frac{P(1)}{(n-1)!}$ if and only if $P(1) \neq 0$.

Counting polynomials of degree d in n variables is rather easy. But can we similarly count the number of polynomials of degree d with certain imposed restrictions or identifications, e.g., count the degree-by-degree dimensions of quotients of polynomials ring, or modules over such rings? These are the questions that pertain to our original set of counting problems.

Definition 1.2. A subset I of S which is closed under addition and arbitrary multiplication by polynomials in S is called an ideal. An ideal I is said to be homogeneous if $I = \oplus_d(I \cap S_d)$ as vector spaces. A generating set $\{f_j : j \in J\}$ for I is a set of polynomials $f_j \in I$ such that every polynomial $g \in I$ can be written $g = \sum_j r_j f_j$ for some (not necessarily unique) polynomials r_j. In this case, we

write $I = (f_j : j \in J)$. We say that I is finitely generated if it is generated by finitely many polynomials.

Whenever an ideal I is given, we can form a quotient ring of cosets $R = S/I$ in which the images of the elements of I become 0. Moreover, if I is homogeneous, then the elements of R inherit the degrees from S, and R becomes a graded ring itself: $R = \oplus R_d$, and $R_i \cdot R_j \subset R_{i+j}$. Most of the counting problems in our orginal list have an equivalent interpretation as finding either explicitly or implicitly the dimensions of the graded components R_d for suitably chosen I.

2. Hilbert's Method

Hilbert wrote two monumental papers [Hil90, Hil93a] in the early 1890s which in some sense answered many of the classical problems of invariant theory, but in a way which dissatisfied some mathematicians at the time. These papers established the foundations of commutative algebra by proving what are now known as the Hilbert Basis Theorem, Hilbert's Nullstellensatz, and the Hilbert Syzygy Theorem. The last-named theorem gave Hilbert a method to package the generating function of the degree dimensions of a ring of invariants, in other words to count the number of invariants of a given degree.

This section defines the basic objects whose study occupies the rest the paper.

The basis of Hilbert's method is the method of *syzygies*, and to get this method off of the ground one needs to know that polynomials rings are Noetherian, i.e., that every ideal is finitely generated. This is the famous Hilbert Basis Theorem:

Theorem 2.1. *Every ideal in the polynomial ring $S = k[x_1, \ldots, x_n]$ is finitely generated.*

To describe Hilbert's syzygy theorem, we need to expand the scope of our discussion to include modules over the polynomial ring S. In fact, we also need a stronger version of the Hilbert Basis Theorem, namely that submodules of finitely generated free S-modules are also finitely generated.

Definition 2.2 (Hilbert Function/Series)**.** Let M denote a finitely generated S-module, We call M graded if it can be written $M = \bigoplus_{i \in \mathbb{Z}} M_i$ where each M_i is a k-vector spaces with $S_i \cdot M_j \subseteq M_{i+j}$). The Hilbert function $h_M : \mathbb{Z} \to \mathbb{Z}_{\geq 0}$ is defined by

$$h_M(d) = \dim_k(M_d).$$

The Hilbert series $H_M(z)$ is the generating function associated to h_M:

$$H_M(z) = \sum_{d \in \mathbb{Z}} h_M(d) z^d.$$

We write $M(e)$ for the shifted module having $M(e)_i = M_{e+i}$. It follows that $h_{M(e)}(i) = h_M(i + e)$. The most important case of a shifted module is the module $S(-d)$. This is simply a copy of the ring S, but in which we set d to be the degree of the generator $1 \in S_0 = S(-d)_d$. The Hilbert series of $S(-d)$ is exactly $\frac{z^d}{(1-z)^n}$ since this module simply shifts the degrees by d. The utility of this notation is given by the example below.

Example 2.3. Let f be a non-zero homogeneous polynomial of degree d in S. We compute the Hilbert series of the graded ring $R = S/(f)$. There is a short exact sequence

$$0 \longrightarrow S \overset{f}{\longrightarrow} S \longrightarrow R \longrightarrow 0$$

in which the first map is multiplication by f. To keep track of the various graded components, it is more convenient to shift the degree of the generator of the first copy of S by d so that the map becomes homogeneous of degree 0. Thus, in our graded setting, the "correct" short exact sequence is

$$0 \longrightarrow S(-d) \overset{f}{\longrightarrow} S \longrightarrow R \longrightarrow 0$$

This short exact sequence upon restriction to an arbitrary fixed degree yields short exact sequences of vector spaces on which dimension is an additive function. Translating this additivity to the generating function gives that that $H_R(z) = H_S(z) - H_{S(-d)}(z) = \frac{1}{(1-z)^n} - \frac{z^d}{(1-z)^n} = \frac{1-z^d}{(1-z)^n}$, a rational function. It follows that

$$h_R(i) = h_S(i) - h_S(i-d) = \binom{i+n-1}{n-1} - \binom{i-d+n-1}{n-1},$$

which is a polynomial for large i (in fact for $i \geq d$), as predicted by Remark 1.1.

Our goal is to understand the Hilbert series of cyclic S-modules, namely, the Hilbert series $H_R(z)$ in the case $R = S/I$ with I a homogeneous ideal of R. An obvious approach is to try to lift this calculation back to the polynomial ring S. This can be done as follows: There is a short exact sequence of S-modules,

$$0 \longrightarrow I \longrightarrow S \longrightarrow R \longrightarrow 0$$

This restricts degree by degree to short exact sequences of vector spaces,

$$0 \longrightarrow I_d \longrightarrow S_d \longrightarrow R_d \longrightarrow 0$$

which gives $h_S(d) = h_I(d) + h_R(d)$, and therefore

$$H_R(z) = H_S(z) - H_I(z) = \frac{1}{(1-z)^n} - H_I(z).$$

On the face of it, we have only made a tautological change in the goal of computing $H_R(z)$ by replacing it by $H_I(z)$. However, as the example above illustrates, this seemingly tautological strategy can actually work in some cases if we know enough about the generators of the ideal I. We then at least have a chance to explicitly compute the polynomials in I of degree d. Hilbert's fundamental realization is that even without any information about the generators, we can "unwind" the structure of R one step at a time by repeating the same idea, and at least discover fundamental properties of the Hilbert function. This is the reason modules come into the story.

The ideal I is finitely generated by Hilbert's Basis Theorem. As a consequence, we can map a finitely generated free graded S-module onto it by sending the standard basis of the free module to the generators of I. Explicitly, let $\{f_1, \ldots, f_t\}$ generate I. Suppose their degrees are $d_1, \ldots, d_t$. We map a graded free S-module $F = S(-d_1) \oplus \cdots \oplus S(-d_t)$ onto I by sending the basis element $e_i = (0, \ldots, 1, \ldots 0)$, with 1 in the ith component, to f_i. The "shift" $-d_i$ gives e_i degree d_i so that the map is a graded map of graded modules. The kernel of this map is called a *syzygy*.

We repeat the process: we have a short exact sequence

$$0 \longrightarrow N \longrightarrow F \longrightarrow I \longrightarrow 0$$

where N is the kernel of the map defined above, giving that $H_I(z) = H_F(z) - H_N(z) = \sum_{i=0}^{t} \frac{z^{d_i}}{(1-z)^n} - H_N(z)$.

In general, consider a short exact sequence of finitely generated graded S-modules,

$$0 \longrightarrow K \stackrel{f}{\longrightarrow} M \stackrel{g}{\longrightarrow} N \longrightarrow 0$$

We assume that the homomorphisms all have degree zero, $f(K_i) \subset M_i$ for all $i \in \mathbb{Z}$, and similarly for the map g from M to N. It follows that we also obtain short exact sequences of finite dimensional vector spaces,

$$0 \longrightarrow N_i \longrightarrow M_i \longrightarrow K_i \longrightarrow 0,$$

yielding $h_N(i) = h_M(i) + h_K(i)$, and therefore

$$H_N(z) = H_M(z) + H_K(z).$$

Hilbert's discovery was that this process of moving back one syzygy at a time actually stops, allowing for computation:

Theorem 2.4 (Hilbert's Syzygy Theorem). *If M is a finitely generated graded S-module, then there exists a finite graded free resolution*

$$(2.1) \qquad 0 \longrightarrow F_n \longrightarrow F_{n-1} \longrightarrow \cdots \longrightarrow F_1 \longrightarrow F_0 \longrightarrow M \longrightarrow 0.$$

In his notes [Hil93b] he mentions that this theorem is "very difficult to prove".

We write $F_i = \bigoplus_{j \in \mathbb{Z}} S(-j)^{\beta_{i,j}}$. It is true that *a priori* these numbers $\beta_{i,j}$ are not uniquely determined by M, but the free resolution can be suitably pruned to make them unique. This pruned resolution is called minimal. After this pruning we call the multiplicities $\beta_{i,j} = \beta_{i,j}(M)$ the graded Betti numbers of M. Using a bit more homological algebra, one can prove the characterization that

$$\beta_{i,j} = \dim_k \operatorname{Tor}_i^S(M, k)_j.$$

The collection of all such Betti numbers is called the Betti table of M. Here we think of i as the homological degree, and of j as the internal degree of the syzygies. Observe that the length of this resolution is at most n, the number of variables.

The Hilbert function, and the Hilbert series of M can all be read off from the Betti table of M. We have

$$h_M(i) = \sum_{l=0}^{n} (-1)^l \cdot \sum_j \binom{i-j+n-1}{n-1} \beta_{l,j},$$

and

$$H_M(z) = \frac{\sum_j \left(\sum_l (-1)^l \beta_{l,j} \right) \cdot z^j}{(1-z)^n}.$$

An immediate consequence is the remarkable fact that the Hilbert series of any finitely generated graded S-module is a rational function. An equivalent statement is contained in the following corollary of Hilbert's Syzygy Theorem.

Theorem 2.5. *If M is a finitely generated graded S-module, then there exists a polynomial $p_M(t)$ with rational coefficients, such that $p_M(i) = h_M(i)$ for sufficiently large values of i.*

The polynomial p_M is called the Hilbert polynomial of M.

Example 2.6 (The Koszul complex). If we take R equal to $k = S/(x_1, \dots, x_n)$, the residue field, then its minimal graded free resolution takes the form:

$$0 \longrightarrow S(-n)^{\binom{n}{n}} \longrightarrow \cdots \longrightarrow S(-i)^{\binom{n}{i}} \longrightarrow \cdots \longrightarrow S(-1)^n \longrightarrow S \longrightarrow k \longrightarrow 0.$$

The graded Betti numbers are given by

$$\beta_{i,j}(k) = \begin{cases} 0 & j \neq i, \\ \binom{n}{i} & j = i. \end{cases}$$

This resolution of k can be thought of as a homological realization of the binomial expansion $(1-z)^n = 1 - nz + \binom{n}{2}z^2 - \dots + (-1)^n z^n$, since computing the Hilbert series of k (which just 1) from the resolution gives exactly that $\frac{1 - nz + \binom{n}{2}z^2 - \dots + (-1)^n z^n}{(1-z)^n} = H_k(t) = 1$.

The resolution of the residue field is a special case of a Koszul complex. More generally, one can define the Koszul complex $K(f_1, \dots, f_p; S)$ on p-elements $f_1, \dots, f_p$ to be the graded algebra $\wedge S^p = \oplus \wedge^i S^p$, whose ith graded piece is $F_i = \wedge^i S^p$. If we let $e_1, \dots, e_p$ be the standard basis of S^p, then as a basis of F_i we can choose the set of all $\{e_J\}$, where J runs through i-tuples $1 \leq j_1 < j_2 < \dots < j_i \leq p$, and e_J is the exterior product $e_{j_1} \wedge \cdots \wedge e_{j_i}$. This basis element sits in degree $\sum_{j \in J} \deg f_j$. The maps of $K(f_1, \dots, f_p; S)$ are described by taking $e_J \in F_i$ to the element $\sum_{1 \leq j \leq i} (-1)^{j+1} f_j e_{J-\{j\}} \in F_{i-1}$. With this notation, the Koszul complex resolving k is $K(x_1, \dots, x_n; S)$. However, the Koszul complex on arbitrary elements is not always exact. See Example 3.5 below for more on this point.

The graded Betti numbers of a module are recorded into the Betti table, where the entry in row i and column j is $\beta_{i,i+j}$:

	0	1	2	$\cdots$
0	$\beta_{0,0}$	$\beta_{1,1}$	$\beta_{2,2}$	$\cdots$
1	$\beta_{0,1}$	$\beta_{1,2}$	$\beta_{2,3}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

The Koszul complex in Example 2.6 has Betti table

	0	1	2	$\cdots$	i	$\cdots$	$n-1$	n
0	1	n	$\binom{n}{2}$	$\cdots$	$\binom{n}{i}$	$\cdots$	n	1

Important invariants of R, and more generally of S-modules, are encoded in the Betti table. Although each of these invariants has other definitions and meanings, the statements below can be taken as definitions.

(1) *Dimension*:

$$\dim(M) = \deg(p_M) + 1.$$

We will also need the codimension of an ideal I in S, namely

$$\operatorname{codim}(I) = \dim(S) - \dim(S/I).$$

(2) *Multiplicity or Degree*:

$$\deg(S/I) = (\dim(S/I) - 1)! \cdot (\text{leading coefficient of } p_{S/I}).$$

(3) *Projective dimension*:

$\mathrm{pd}(M) =$ length of the Betti table (the index of the last non-zero column).

(4) *Regularity.*

$\mathrm{reg}(M) =$ width of the Betti table (the index of the last non-zero row).

Example 2.7. We have seen one explicit example, namely the Koszul complex. To see these concepts in another example, let $T = \mathbb{C}[x, y]$ be the polynomial ring in two variables with complex coefficients. We let the cyclic group of third roots of unity act linearly on T by sending x to ωx and y to ωy, where ω is a primitive third root of unity. Let R be the ring of invariant polynomials under this action, namely the set of all polynomials $f(x, y)$ such that $f(x, y) = f(\omega x, \omega y)$. It is not difficult to prove that R is generated as an algebra over $\mathbb{C}$ by all the homogeneous polynomials of degree 3, namely $R = \mathbb{C}[x^3, x^2 y, xy^2, y^3]$. If we map a polynomial ring $S = \mathbb{C}[s, t, u, v]$ in four variables onto R by sending $s \mapsto x^3, t \mapsto x^2 y, u \mapsto xy^2, v \mapsto y^3$, then the relations on these invariants are exactly the ideal I in $S = \mathbb{C}[s, t, u, v]$ generated by three quadrics, namely $sv - tu, t^2 - su, u^2 - tv$. These three quadrics themselves have two relations, and the free resolution of R over S takes the form

$$0 \longrightarrow S(-3)^2 \longrightarrow S(-2)^3 \longrightarrow S \longrightarrow R \longrightarrow 0$$

We can read off important invariants from this data: the regularity of R over S is 1, and the projective dimension is 2. The Betti table is:

	0	1	2
0	1	.	.
1	.	3	2

The Hilbert series

$$H_R(z) = H_{S(-2)^3}(z) - H_{S(-3)^2}(z) - H_S(z) = \frac{2z^3 - 3z^2 + 1}{(1 - z)^4} = \frac{2z + 1}{(1 - z)}.$$

Expanding the Hilbert series gives $H_R(z) = \sum_t (3t + 1) z^t$ from which we deduce the Hilbert function is $h_R(t) = 3t + 1$.

One can think of a hierarchy of levels of information which enable one to study the ring $R = S/I$. At the first level is its Hilbert series giving the degrees of the various graded pieces. A second level is provided by the graded Betti numbers $\beta_{i,j}$ which give finer information; the Hilbert series can be recovered from these integers. A third level would be understanding in more detail the structure of the maps within the resolution, or relationships between the graded Betti numbers. See [BE77] for structural information on free resolutions, and [BS08], [ES09] for the development of Boij-Söderberg theory, which gives structure to the Betti table. See also the recent Bulletin article [FMP16] for many other problems about resolutions and their structure. A yet finer-grained analysis can be done by putting an algebra structure on the resolution itself compatible with the maps in the resolution. The scope of this article is the second level of trying to understand bounds on the actual Betti numbers.

Example 2.8 (Gonality). As an example of what finer information can be provided by the Betti table, as opposed to the Hilbert function, we consider the notion of *gonality*. Let C be a smooth complex projective curve of genus $g \geq 2$. The gonality of C, $\mathrm{gon}(C)$, is the least degree $C \longrightarrow \mathbf{P}^1$ of a branched covering. Let L be a very

ample line bundle of degree d on C defining an embedding $C \subset \mathbf{P}(H_0(C, L)) = \mathbf{P}^r$. Mark Green [Gre84] first studied connections between the geometry of C and L and their syzygies. Write $S = \mathrm{Sym}(H_0(C, L))$ for the homogeneous coordinate ring of projective r-space, and let $R = \oplus_m H_0(C, L^m)$ be the graded S-module, the ring of sections of L. If L is normally generated, then $R = S/I$ for the homogeneous ideal of C in $\mathbf{P}^r$. We let $\beta_{p,q}$ be the graded Betti numbers of R over S. A conjecture of Green, proved in [EL15] for line bundles of sufficiently high degree, states that the gonality can be read off of the resolution of I by looking for the longest linear strand of the resolution: explicitly, $\beta_{p,p+1} \neq 0$ if and only if $1 \leq p \leq r - \mathrm{gon}(C)$.

The three conjectures mentioned in the introduction are the Eisenbud-Goto conjecture, the Buchsbaum-Eisenbud-Horrocks conjecture, and Stillman's question. All three have been solved within the last two years, one negatively (although a case of great interest is still open). The Buchsbaum-Eisenbud-Horrocks conjecture was first in print in 1977, the Eisenbud-Goto conjecture was made in print in 1984, while Stillman's question was first made informally around 2000, and did not appear in print until a few years later. Combined, these conjectures resisted about 90 years of attempts to solve them. All three can be stated using data from Betti tables as they deal with regularity, total Betti numbers, and projective dimension respectively. In the next section we concentrate on the statement and solution of Stillman's conjecture. In a last section we give a short summary of the first two of these conjectures.

3. Stillman's Question

The Hilbert Syzygy Theorem proves that the number of variables, N, is a bound for the projective dimension of an arbitrary ideal in $S = k[x_1, ..., x_N]$. Around 2000, Michael Stillman had the vision to ask if Hilbert's bound was actually the best if one fixed data about the ideal. Stillman's conjecture asks if there exists a bound independent of the number of variables, provided the number and degrees of generators are fixed.

Conjecture 3.1 (Stillman's Question). *Fix positive integers $d_1, ..., d_n$. Consider all ideals J in an arbitrary polynomial ring $k[x_1, ..., x_N] = S$, with N varying, generated by n homogeneous polynomials of degrees $d_1, ..., d_n$. There a bound to the projective dimension, $pd_S(J)$, depending only on $d_1, ..., d_n$.*

In theory, unwinding the structure of the ideal J through its free S-resolution could become long and complicated, even have length $N - 1$. In theory, the resolution could also involve equations of larger and larger degree. The meaning of the conjecture is that this worst-case scenario cannot ever occur. Not even close. One can ask a similar question for regularity, but Giulio Caviglia proved that bounding the projective dimension independent of the number of variables is equivalent to bounding the regularity independent of the number of variables!

We let $b(d_1, ..., d_n)$ be the best possible bound for the projective dimension of S/J for all ideals J in an arbitrary polynomial ring $k[x_1, ..., x_N] = S$, with N varying, that are generated by n homogeneous polynomials of degrees $d_1, ..., d_n$. Through the work of Tigran Ananyan and Melvin Hochster [AH16], we now know that $b(d_1, ..., d_n) < \infty$ in general, i.e., that Stillman's conjecture is true. In this section we sketch some of the ideas in their proof.

Example 3.2. There are some cases in which Stillman's conjecture can be proved. If $n = 1$, then we only have one polynomial f of degree d, and we've already seen that $R := S/(f)$ has a free resolution,

$$0 \longrightarrow S(-d) \xrightarrow{f} S \longrightarrow R \longrightarrow 0$$

showing that its projective dimension is one regardless of the degree of f. If $n = 2$, then the fact that polynomial rings are unique factorization domains can be used to prove that the projective dimension is at most 2, which is sharp provided $N \geq 2$. It was shown by Diana Taylor in 1966 [Tay66] that the projective dimension of n monomials of arbitrary degrees can be at most n. If the ideal I is generated by linear polynomials, then a minimal generating set is given by linear polynomials in I which form a basis of I_1, and the resolution of these generators is given by their Koszul complex, which has length exactly the number of them.

These relatively elementary examples might encourage one to hypothesize that the projective dimension of S/I is always bounded by the number of generators of I. As more "evidence" for this naive guess, if $f_1, ..., f_p$ are chosen randomly with $p < N$, then they form a complete intersection (see Example 3.5 below). However, whenever $f_1, ..., f_p$ form a complete intersection, the Koszul complex $K(f_1, ..., f_p; S)$ gives a free resolution of length p.

The first case not covered by these elementary considerations is $n = 3$ and $d_1 = d_2 = d_3 = 2$, with the ideal having codimension 2. However, the ideal $I = (xy, uv, xu + yv)$ generated by three quadrics has the free resolution

$$0 \longrightarrow S(-6) \longrightarrow S(-5)^4 \longrightarrow S(-4)^5 \longrightarrow S(-2)^3 \longrightarrow S \longrightarrow S/I \longrightarrow 0$$

showing that its projective dimension is at least 4. In other words, $b(2, 2, 2) \geq 4$. In fact, $b(2, 2, 2) = 4$. In turns out that 3-generated ideals can be as complex as arbitrary ideals: Winfried Bruns [Bru76] proved that basically **every** possible free resolution is the resolution of a 3-generated ideal up to shifting degrees. However, the degrees of the generators must increase as one applies his construction to increase the length of the resolutions. One of the few sharp bounds known is that n-quadrics generating an ideal of codimension 2 has projective dimension at most $2n - 2$ [HMMS13].

Life quickly becomes hard. Even the value of $b(3, 3, 3, 3)$ is not known.

Example 3.3 (The Smooth Case)**.** An important case to consider geometrically is the case in which the ideal $I = (f_1, ..., f_n)$ is a homogeneous ideal in $S = k[x_1, ..., x_N]$ defining a smooth projective variety. This means that the ring S/I has an isolated singularity at the origin. The answer was provided in a 1981 paper of Gerd Faltings [Fal81] long before Stillman asked his question. Faltings proved that *either* I is a complete intersection (see Example 3.5 below) in which case the projective dimension of S/I is at most n, or $N \leq 3n$, in which case the projective dimension is at most $3n$. In either case, there is a bound depending only on the number of polynomials f_i, not even depending on their degrees. We will soon see that this dichotomy of either knowing an ideal is a complete intersection or bounding its singular locus is a crucial insight in the solution of Stillman's conjecture.

Remark 3.4 (A Gröbner Approach). For those familar with Gröbner bases, there is seemingly a "natural" line to approach Stillman's conjecture. In this discussion we assume familarity with the fundamentals of Gröbner bases.

A famous paper of Bayer and Stillman [BS88] in 1984 proved that if I is a homogeneous ideal of a polynomial ring S, the projective dimension and regularity of its generic initial ideal $\mathrm{Gin}(I)$ in the revlex order are the same as that of I. Thus Stillman's question is equivalent to asking about the projective dimension of the generic initial ideal of a homogeneous ideal generated by forms of degrees $d_1, ..., d_n$. Of course, the initial ideal is generated by monomials, and as mentioned above, the projective dimension of an ideal of monomials is bounded above by the number of monomials in their generating set. Thus the projective dimension of I will be bounded by the number of generators in $\mathrm{Gin}(I)$. The problem is that taking a Gröbner basis can vastly increase the number of generators. However, in revlex, the projective dimension of a generic initial ideal is simply the total number of variables which appear in any generator.

Buchberger's algorithm provides a method to compute the Gröbner basis of an ideal by repeated S-pairs. One might try to solve Stillman's conjecture by showing that from a fixed number of polynomials of fixed degrees in generic coordinates, one can only get initial terms involving a bounded number of variables. Proving this is equivalent to proving Stillman's conjecture. However, this idea seems to go nowhere, at least at the present time.

In 2012, Ananyan and Hochster [AH12] settled Stillman's conjecture for ideals generated by quadrics. Their upper bound for the projective dimension of n quadrics in that paper is approximately $2n^{2n}$. They proved this case by changing the question in a fundamental way. In 2016 [AH16], the same authors provided a full solution for arbitrary degrees. Their proof is a remarkable journey through six theorems nested by an induction. The new ideas in 2016 even provide a much better bound for quadrics. A full description of their method is beyond the scope of these notes, and we will instead concentrate on one key aspect of their proof. To describe some of the main ingredients, we first need to take a closer look at the all-important case of complete intersections.

Example 3.5 (Complete intersections). We say the homogeneous elements $f_1, ..., f_p$ in a polynomial ring $S = k[x_1, ..., x_n]$ which generate an ideal I form a *complete intersection* if the codimension of I is exactly p, where $\mathrm{codim}(I)$ is by definition $n - \dim(S/I)$. The sense of being a complete intersection is that each f_j cuts down the dimension by as much as possible, i.e. the vanishing loci of the polynomials meet transversally in some sense. For example, if $f_1 = x_1 x_2$ and $f_2 = x_3 x_4$, then they form a complete intersection, while $f_1 = x_1 x_2$ and $f_2 = x_2 x_3$ are not a complete intersection. An important point for our discussion is that whenever $f_1, ..., f_p$ is a complete intersection, the Koszul complex $K(f_1, ..., f_p; S)$ is acyclic, and is a free resolution of $S/(f_1, ..., f_p)$. In particular, as noted above, the projective dimension is just p, the number of f_i, or equivalently the projective dimension of I is $p-1$.

More critically, even more is true: if we let $A = k[f_1, ..., f_p]$ be the subring of S generated by elements $f_1, ..., f_p$ which form a complete intersection, A is isomorphic to a polynomial ring in p variables (in other words, any polynomial $g(Y_1, ..., Y_p)$ vanishing at $f_1, ..., f_p$ is identically zero), and the containment of A into S makes S into a free A-module! The freeness of S as an A-module has an important consequence for the solution of Stillman's conjecture. Namely, suppose that J is an ideal in A. By the Hilbert Syzygy Theorem, since A is itself isomorphic to a polynomial ring in p variables, the projective dimension of A/J is at most p. So there is a free A-resolution of A/J of length at most p. When one extends the

ring and this resolution to S (by tensoring with S over A), the A-free resolution of A/J becomes an S-free resolution of S/JS. Thus, if we are given polynomials in S which are in fact polynomials in a complete intersection of much smaller cardinality, we can bound their projective dimension by the number of polynomials in that complete intersection.

Remark 3.6. By itself, the fact that S is a free A-module is not quite enough for the proof. Ananyan and Hochster prove a variety of extension-type results in passing from A to S. One of the more surprising ones is that if the ideal generated by the complete intersection $f_1, ..., f_p$ in S is a prime ideal, then **every** prime ideal of A extends to a prime ideal of R. This fact, and others of similar nature, are crucial in their proof.

This all-important base-change principle is a simple generalization of the somewhat obvious remark that if $g_1, ...g_n$ are polynomials which only involve variables $x_1, ..., x_p$, then the projective dimension of the ideal they generate is at most $p-1$, regardless of what other variables are in the ambient polynomial ring. What the discussion in the above paragraph shows is that if $g_1, ..., g_n$ can be written as polynomials in auxiliary polynomials $f_1, ..., f_p$ which form a complete intersection, then it is still the case that the projective dimension is bounded by $p-1$.

The discussion above provides a strategy to attack Stillman's conjecture: given homogeneous polynomials $f_1, ..., f_p$ of degrees $d_1, ..., d_p$, suppose we can find a number of auxilliary forms, say $g_1, ..., g_m$, where m only depends on the number and degrees of the f_i, such that $f_1, ..., f_p \in k[g_1, ..., g_m]$, and such that $g_1, ..., g_m$ form a complete intersection in the original polynomial ring S. The discussion above then proves that the projective dimension of the $S/(f_1, ..., f_p)$ is at most m.

One of the fundamental insights of Ananyan and Hochster is that trying to write polynomials in terms of simpler polynomials, or polynomials of lower degree, is closely connected with the codimension of the singular locus of those polynomials. The following is a quote from their paper [AH16]:

"We define a nonzero homogeneous polynomial F of positive degree in S to have a k-collapse for k an integer, if F is in an ideal generated by k elements of strictly smaller positive degree, and we define F to have strength k if it has a $k+1$-collapse but no k-collapse. Nonzero linear forms have strength ∞, and a form has strength at least 1 if and only if it is irreducible. One of the main themes here is that F has a "small" collapse if and only if the singular locus of F has small codimension."

Remark 3.7. The "only if" direction of this quote is explained as follows: if $F = \sum_{i=1}^{k} G_i H_i$, the partial derivatives of F are in the ideal $(G_1, ..., G_k, H_1, ..., H_k)$. Since an ideal defined by m equations has at most codimension m, the partials of F, whose vanishing locus is the singular locus of F, have "small" codimension at most $2k$. The "if" part of their approach is extremely difficult.

They extend this notion of strength to include entire vector subspaces of polynomials: We say that a k-vector subspace of S is k-strong if all of the nonzero homogeneous polynomials in it are k-strong.

Their proof winds through a nested induction of six different theorems. In some respects the punchline for Stillman's question is provided by the following theorem that is Corollary B in their paper (the notation is changed slightly from their paper):

Theorem 3.8 (Ananyan-Hochster)**.** *There is a function $B(n,d)$, independent of k and N, such that for all polynomial rings $S = k[x_1, ..., x_N]$ over an algebraically closed field k and all graded vector subspaces V of S of dimension at most n whose homogeneous elements have positive degree at most d, the elements of V are contained in a subring $A = k[G_1, ..., G_B]$, where $B \leq B(n,d)$ and $G_1, ..., G_B$ form a complete intersection of forms of degree at most d.*

Evidently as described above, this theorem immediately implies that the projective dimension of S/I is at most $B(n,d)$, where I is the ideal generated by the forms in V, giving a positive answer to Stillman's conjecture.

To give a sense of their proof, the next subsection will outline their proof when the ideal is generated by quadrics. This case is considerably easier, though details must still be supressed.

3.1. **The Quadric Case.** We begin with a basic result about matrices of linear forms. This result will then be applied to the Jacobian matrix associated with given quadratic forms. One of the difficult problems in extending the proof from quadrics to the general case is that the Jacobian matrix no longer consists of linear forms, but rather higher degree terms.

Lemma 3.9. *Suppose that A is an $n \times N$ matrix of linear forms in a polynomial ring over an infinite field with $n \leq (<<)N$. Further assume that the codimension of the ideal generated by the entries of a general row is at least h. Then the ideal generated by the n by n minors of A has codimension at least $h - n + 1$.*

A consequence is the following fundamental principle:

Either the original set of quadrics having codimension c is a complete intersection, or some linear combination of any c of the quadrics is contained in a polynomial subring of S in $2c$-variables.

To see why, we apply Lemma 3.9 to the Jacobian matrix of an ideal I of codimension c minimally generated by n quadrics in a polynomial ring S with N-variables. We assume that $n > c$, else I is a complete intersection and has projective dimension $c - 1$. Choose c of the quadrics. Suppose that a generalized row of their Jacobian matrix has height at least $2c + 1$. Then the height of the actual Jacobian ideal of these c quadrics in the polynomial ring S is at least $2c + 1 - c + 1 = c + 2$, which forces the ideal of those c quadrics to be codimension c, in other words forces the quadrics to be a complete intersection. Even more, the Jacobian criterion of smoothness proves that the ideal generated by those c quadrics is a prime ideal! But then the entire set of our original quadrics is generated by those c-quadrics since otherwise the codimension would increase, a contradiction. This contradicts our assumption that $n > c$.

The above discussion proves that some linear combination of any c of the quadrics has its partial derivatives generating an ideal of codimension at most $2c$. But an ideal J generated by linear forms having codimension at most $2c$ must be contained in a polynomial ring of at most $2c$ variables, namely the linear forms that give a basis of the graded piece J_1. But even more is true: the quadric whose partials are those linear forms can be written as a polynomial in those variables. For example, if $f = xz + yz + xu + yu$, then the partials are generated by the linear forms $z + u$ and $x + y$, and letting $a = z + u, b = x + y$, we can change variables so that

$f = ab \in k[a, b]$. Thus, after a change of variables, we can assume that the last quadric, say f_n, can be written with only the variables $x_1, ..., x_{2c}$.

On the face on it, there is now an easy induction: without loss of generality we may assume that $f_n \in k[x_1, ..., x_{2c}]$. By induction we can then assume that the first $f_1, ..., f_{n-1} \in k[g_1, ..., g_m]$, where $g_1, ..., g_m$ are a complete intersection in S and m is bounded by a function only depending on $n - 1$. Then all of the quadrics are in a subring $k[x_1, ..., x_{2c}, g_1, ..., g_m]$. Since the x_i are simply variables, we presumably have that the sequence $x_1, ..., x_{2c}, g_1, ..., g_m$ form a complete intersection in S. Not so fast! Unfortunately, the last statement is not true.

This last difficulty then requires rather intricate manipulations which go beyond what we can do in this manuscript. The notion of strength comes into play. Ananyan and Hochster basically start the process over by including the notion of strength from the beginning. If our original n quadrics are $(n-1)$-strong then they must be a complete intersection. If not, one of them can be written in terms of $2n - 2$ linear forms and $n - 1$ quadrics. If the strength of the remaining quadrics is at least $3n - 4$, then they prove that the remaining forms must be a complete intersection modulo the linear forms. If not, one of the remaining quadrics can be written as a polynomial in $(2n-2) + 2(3n-4) = 8n - 10$ linear forms. They iterate this process until either all the quadrics are gone, or are sufficiently strong to be a complete intersection. Analyzing this process, in the special case of quadrics they prove the following theorem:

Theorem 3.10 ([AH16]). *A vector space of quadrics of dimension n in a polynomial ring S is contained in a polynomial subring generated by a sequence of polynomials of at most $2^{n+1}(n - 2) + 4$ linear and quadratic forms which form a complete intersection in S. Hence, the projective dimension of S/I, where I is the ideal generated by these forms, is at most $2^{n+1}(n - 2) + 4$.*

This result is probably far from sharp. The worst-case scenario now known for the projective dimension of n-quadrics is about $n^{\frac{3}{2}}$.

Remark 3.11. The work of Ananyan and Hochster is being used to prove many other results of similar type. In [ESS17], combining the techniques of Ananyan and Hochster with the ideas of Draisma on twisted commutative algebra [Dra17], Daniel Erman, Steven Sam and Andrew Snowden give large classes of numerical functions on graded ideals which are similarly bounded. They only require that the function have certain properties such as being weakly upper semi-continuous in families, and what they call *cone-stable*, meaning that the numerical value attached to an ideal I in S is the same as that of the extension of I in $S[t]$, where t is a new variable.

4. The Other Conjectures

In this section we introduce the other two conjectures. One deals with a conjectured bound on the regularity of prime ideals in polynomial rings, while the other gives a lower bound for the sum of all the Betti numbers.

4.1. **The Eisenbud-Goto Conjecture.** The two gross invariants which measure the complexity of a set of polynomials equations which generate an ideal I in a polynomial ring S are the length and width of a minimal "rectangle" enclosing the Betti table of S/I. The length is exactly the projective dimension of S/I, while the width is exactly the regularity of S/I. One might hope that ideals generated by polynomials of small degree might have small regularity. The following sequence of examples by Giulio Caviglia show that even simple examples can have high regularity.

Example 4.1. Let $S = k[x_1, x_2, x_3, x_4]$, and let I be the ideal generated by the three polynomials of degree d, $x_1^d, x_2^d, x_1 x_3^{d-1} - x_2 x_4^{d-1}$. Then the regularity of S/I is approximately d^2. For example the free resolution of S/I when $d = 5$ is quite short (by Hilbert's syzygy theorem), but the Betti table is very thick. The resolution without the shifts seems innocent enough:

$$0 \longrightarrow S^4 \longrightarrow S^{10} \longrightarrow S^8 \longrightarrow S^3 \longrightarrow S \longrightarrow S/I \longrightarrow 0$$

However, the degrees of the generators of the syzygy modules become large quickly. The regularity is 23. The Betti table is too big to print here.

In fact, famous examples of Mayr-Meyer [MM82], as reworked by Bayer and Stillman in [BS88], show that the regularity can become doubly exponential in the number and degrees of the generators. These papers show that there are ideals I in a polynomial ring with $10n + 1$ variables generated by quartics such that the regularity of the quotient ring defined by I is at least $2^{2^n} + 1$. It was known that there exists a doubly exponential bound on the regularity of any ideal [BM93], [CS05], so this type of growth is the worst possible. However these examples were unsatisfying from the point of view of algebraic geometry in that the ideals are very far from irreducible. These considerations led Eisenbud and Goto to the following conjecture:

Conjecture 4.2 (Eisenbud-Goto Conjecture). *Suppose that the field k is algebraically closed. If P is a homogeneous prime ideal of $S = k[x_1, \ldots, x_n]$ that does not contain a linear form, then*

$$reg(P) \leq deg(S/P) + n - dim(S/P) + 1.$$

A positive answer to the Eisenbud-Goto conjecture would yield a positive answer to the following weaker question: Suppose that the field k is algebraically closed. If P is a homogneous prime ideal of a polynomial ring S over k, that does not contain a linear form, then the largest degree of any minimal generator of P is bounded by the degree of S/P.

Jason McCullough and Irena Peeva recently used the original examples of Mayr-Meyer in a remarkable construction which yields homogeneous prime ideals which give a counterexample to the Eisenbud-Goto conjecture. Their methods also yield a counterexample to the weaker question. Their work makes it clear that there is huge gap in what we know about regularity and resolutions. In particular, their counterexamples prove that no polynomial bound for regularity exists. An exposition of their construction is given in [Swa17].

It is of great interest to find geometric conditions that would give the Eisenbud-Goto bound; for example the conjecture is not known for either smooth or toric varieties. It has been proved for irreducible curves by Gruson-Lazarsfeld-Peskine [GLP83], for smooth surfaces by Lazarsfeld [Laz87], and for most smooth 3-folds

by Ran [Ran90]. Finding optimal bounds on regularity for all prime ideals is a challenge. In particular, the question of whether there is a singly exponential bound was implicitly raised in an influential paper of Dave Bayer and David Mumford [BM93].

4.2. The Buchsbaum-Eisenbud-Horrocks Conjecture.

A famous technique in algebraic geometry is reduction to the diagonal. This technique can be used to study intersections of algebraic varieties, and rests on a process which converts information about intersections to the case in which one of the algebraic varieties is defined by linear polynomials and so has a Koszul complex as its free resolution. This process suggests that a typical resolution should be at least as complicated as the Koszul complex. In the 1970s, through completely different means, one avatar of this idea was formulated independently by Geoffrey Horrocks and by David Buchsbaum and Eisenbud. The Buchsbaum-Eisenbud-Horrocks rank conjecture says roughly that the Koszul complex is the smallest possible minimal free resolution. The conjecture was formulated by Buchsbaum and Eisenbud in [BE77, p. 453] and, independently, the conjecture is implicit in a question of Horrocks [Har79, Problem 24].

The strong form of this principle is still open:

Conjecture 4.3 (Strong Form). *Let $S = k[x_1, ..., x_n]$ be a polynomial ring over a field k, and let M be a non-zero S-module having finite dimension as a vector space over k. For any finite free resolution*

$$0 \longrightarrow F_n \longrightarrow \cdots \longrightarrow F_1 \longrightarrow F_0 \longrightarrow M \longrightarrow 0$$

of M we must have that the rank of F_i is at least $\binom{n}{i}$

The ranks of the free modules F_i in a minimal resolution are called the Betti numbers of M; in the graded case, these are the sums of the graded Betti numbers in homological degree i.

A weaker form of the orginal questions asks instead about the sum of the ranks of the free modules:

Conjecture 4.4 (Weak Form; the Total Rank Conjecture). *Let $S = k[x_1, \ldots, x_n]$ be a polynomial ring over a field, and let M be a graded S-module having finite dimension as a vector space over k. Then the sum of all the Betti numbers of M is at least 2^n.*

Notice that the free resolution of k itself is the Koszul complex, and the sum of all the Betti numbers is exactly $\sum \binom{n}{i} = 2^n$.

Although several special cases of this conjecture were settled (for example, see [AB93]), there was very little progress on it for many years. Thus, it was a great surprise when Mark Walker [Wal17] gave a marvelous proof of the weaker form of the conjecture (except in characteristic 2), and in far greater generality. His methods come from the use of Adams operations on complexes and algebraic K-theory, and are beyond the scope of these notes. Specifically, he proves the following theorem:

Theorem 4.5 (M. Walker). *Assume R is a local (Noetherian, commutative) ring of Krull dimension n and that M is a non-zero R-module of finite length and finite projective dimension. If either*

(1) *R is the quotient of a regular local ring by a sequence of elements forming a complete intersection and* 2 *is invertible in R, or*

(2) *R contains* $\mathbb{Z}/p\mathbb{Z}$ *as a subring for an odd prime p,*

then the sum of all the Betti numbers of M is at least 2^n.

Our title question was "how complicated are polynomials in many variables?" The message delivered by the solutions to these three conjectures is somewhat mixed. The proof by Ananyan and Hochster of Stillman's question shows that polynomials are perhaps not as complicated as one might think, while Walker's proof of the Buchsbaum-Eisenbud-Horrocks conjecture shows that there are really no special sets of polynomials that behave better than the variables. However, the negative solution to the Eisenbud-Goto conjecture by McCullough and Peeva teaches us that there is a long way to go in understanding exactly how complicated polynomials can be.

References

[AH12] Tigran Ananyan and Melvin Hochster, *Ideals Generated by Quadratic Polynomials*, Math. Res. Lett. **19** (2012), 233–244. ↑10

[AH16] ———, *Small Subalgebras of Polynomial Rings and Stillman's Conjecture*, arXiv **1610.09268** (2016). ↑8, 10, 11, 13

[AB93] Luchezar Avramov and Ragnar-Olaf Buchweitz, *Lower bounds for Betti numbers*, Compositio Math. **86** (1993), 147–158. ↑15

[BM93] David Bayer and David Mumford, *What can be computed in algebraic geometry?*, Computational algebraic geometry and commutative algebra, Sympos. Math. **XXXIV** (1993), 1–48. ↑14, 15

[BS88] David Bayer and Michael Stillman, *On the complexity of computing syzygies*, J. Symbolic Comput. **6** (1988), 135–147. ↑10, 14

[BS87] ———, *A criterion for detecting m-regularity*, Invent. Math. **87** (1987), 1–11. ↑

[BS08] Mats Boij and Jonas Söderberg, *Graded Betti numbers of Cohen-Macaulay modules and the multiplicity conjecture*, J. Lond. Math. Soc. **78** (2008), 85–106. ↑7

[Bru76] Winfried Bruns, *"Jede" endliche freie Auflösung ist freie Auflösung eines von drei Elementen erzeugten Ideals*, J. Algebra **39** (1976), 429–439. ↑9

[BE77] David A. Buchsbaum and David Eisenbud, *Algebra structures for finite free resolutions, and some structure theorems for ideals of codimension 3*, Amer. J. Math. **99** (1977), 447–485. ↑7

[CS05] Giulio Caviglia and Enrico Sbarra, *Characteristic-free bounds for the Castelnuovo-Mumford regularity*, Compositio Math. **141** (2005), 1365–1373. ↑14

[Dra17] Jan Draisma, *Topological noetherianity for polynomial functors*, arXiv: 1705.01419v1 (2017). ↑13

[EL15] Lawrence Ein and Robert Lazarsfeld, *The gonality conjecture on syzygies of algebraic curves of large degree*, Publ. Math. Inst. Hautes Études Sci. **122** (2015), 301–313. ↑8

[Eis95] David Eisenbud, *Commutative algebra*, Graduate Texts in Mathematics, vol. 150, Springer-Verlag, New York, 1995. With a view toward algebraic geometry. MR1322960 (97a:13001) ↑

[Eis05] ———, *The geometry of syzygies*, Graduate Texts in Mathematics, vol. 229, Springer-Verlag, New York, 2005. A second course in commutative algebra and algebraic geometry. MR2103875 (2005h:13021) ↑

[EG84] David Eisenbud and Shiro Goto, *Linear free resolutions and minimal multiplicity*, J. Algebra **88** (1984), 89–133. ↑

[ES09] David Eisenbud and Frank-Olaf Schreyer, *Betti numbers of graded modules and cohomology of vector bundles*, J. Amer. Math. Soc. **22** (2009), 859–888. ↑7

[Eng10] Bahman Engheta, *A bound on the projective dimension of three cubics*, J. Symbolic Comput. **45** (2010), 60–73. ↑

[ESS17] Daniel Erman, Steven V. Sam, and Andrew Snowden, *Generalizations of Stillman's conjecture via twisted commutative algebra*, preprint (2017). ↑13

[Fal81] Gerd Faltings, *Ein Kriterium für vollständige Durchschnitte*, Invent. Math. **62** (1981), 393–401. ↑9

[FMP16] Gunnar Floystad, Jason McCullough, and Irena Peeva, *Three themes of syzygies*, Bull. Amer. Math. Soc. **53** (2016), 415–435. ↑7

[Gre84] Mark Green, *Koszul cohomology and the geometry of projective varieties*, J. Diff. Geom. **19** (1984), 125–171. ↑8

[GLP83] L. Gruson, R. Lazarsfeld, and C. Peskine, *On a theorem of Castelnuovo and the equations defining projective varieties*, Invent. Math. **72** (1983), 491–506. ↑14

[Har79] Robin Hartshorne, *Algebraic vector bundles on projective spaces: a problem list*, Topology **18** (1979), 117–128. ↑

[Hil90] David Hilbert, *Über die Theorie der algebraischen Formen*, Math. Ann. **36** (1890,) 473–534. ↑3

[Hil93a] _____, *Über die vollen Invariantensysteme*, Math. Ann. **42** (1893), 313–373. ↑3

[Hil93b] _____, *Theory of algebraic invariants. Translated from the German and with a preface by Reinhard C. Laubenbacher. Edited and with an introduction by Bernd Sturmfels*, Cambridge University Press, Cambridge, 1993. ↑5

[HMMS13] Craig Huneke, Paolo Mantero, Jason McCullough, and Alexandra Seceleanu, *The projective dimension of codimension two algebras presented by quadrics*, J. Algebra **393** (2013), 170–186. ↑9

[Laz87] R. Lazarsfeld, Duke Math. J. **55** (1987), 423–438. ↑14

[MM82] E. Mayr and A. Meyer, *The complexity of the word problem for commutative semigroups and polynomial ideals*, Adv. Math. **46** (1982), 305–329. ↑14

[MP18] Jason McCullough and Irena Peeva, *Counterexamples to the Eisenbud-Goto regularity conjecture*, J. Amer. Math. Soc., to appear (2018). ↑

[Ran90] Z. Ran, *Local differential geometry and generic projections of threefolds*, J. Differential Geom. **32** (1990), 13–137. ↑15

[Sta83] Richard Stanley, *Combinatorics and commutative algebra*, Progress in Mathematics, vol. 41, Birkhäuser Boston, Inc., 1983. ↑

[Swa17] Irena Swanson, *Commutative Algebra Provides a Big Surprise for Craig Huneke's Birthday*, Notices AMS **64** (2017), 256–259. ↑14

[Tay66] Diana Kahn Taylor, *Ideals generated by monomials in an R-sequence*, Thesis, University of Chicago (1966). ↑9

[Wal17] Mark Walker, *Total Betti numbers of modules of finite projective dimension*, Ann. of Math. (2) **186** (2017), no. 2, 641–646. ↑15

UNIVERSITY OF VIRGINIA, DEPARTMENT OF MATHEMATICS, 141 CABELL DRIVE, KERCHOF HALL, CHARLOTTESVILLE, VA 22904

E-mail address: huneke@virginia.edu

FROM NEWTON TO NAVIER-STOKES, OR HOW TO CONNECT FLUID MECHANICS EQUATIONS FROM MICROSCOPIC TO MACROSCOPIC SCALES

ISABELLE GALLAGHER

ABSTRACT. In this survey we present an overview of some mathematical results concerning the passage from the microscopic description of fluids via Newton's laws to the macroscopic description via the Navier-Stokes equations.

1. INTRODUCTION AND PLAN OF THE SURVEY

The mathematical description of fluids is an intricate process which depends on the scale of observation:

- at a microscopic level one sees atoms, and the gas may be described by a classical mechanics picture via Newton's ordinary differential equations: this procedure is presented in Section 2.1;
- at a macroscopic scale one is interested in an average behaviour modeled by PDEs such as the Navier-Stokes equations, which are described in Section 2.2;
- an intermediate regime, called "mesoscopic scale" can also be used and the corresponding Boltzmann PDE is presented in Section 2.3.

Our goal in this survey is to explain how these apparently very different descriptions (ODEs vs PDEs, reversible vs irreversible dynamics...) can be related one to the other from a mathematical point of view: this question goes back to Hilbert (see Section 2.4) and has known quite a lot of progress in the recent years. In Section 3 we present some mathematical attempts to justify the passage from one scale to the other by a limiting process, and we show the limitations of those approaches which prevent from solving the full problem: justifying nonlinear fluid mechanics PDEs from the microscopic ODEs. Section 4 describes one situation where the full problem does have an answer, in a linear setting. Finally in Section 5 e few questions are presented.

2. MICROSCOPIC, MESOSCOPIC AND MACROSCOPIC SCALES IN FLUIDS

2.1. **Newton: a microscopic point of view.**

2.1.1. *The equations.* A gas is made of a very large number of particles evolving and interacting in a d-dimensional space domain. Throughout this survey we assume that the space dimension is $d \geq 2$ and we denote by $N \gg 1$ the number of particles. Typically N is larger than the Avogadro number $6.02 \cdot 10^{23}$.

A number of simplifying assumptions will be made throughout this survey, concerning the space domain, the types of particles considered, and their interactions. These are listed below, and comments on these assumptions can be found in the concluding section:

- The particles are all identical spheres of mass 1 and diameter $\varepsilon > 0$;
- The particles evolve in a periodic box of size 1 denoted $\mathbb{T}^d := [0,1]^d$;
- The particles interact elastically at each binary collision and there is no other type of interaction nor forcing.

We label by integers $i \in \{1, \ldots, N\}$ each particle – note that the particles are undistinguishable so this labelling is arbitrary and all the functions we shall consider will be symmetric with respect to permutations of the labels. We denote by $(x_i, v_i) \in \mathbb{T}^d \times \mathbb{R}^d$ the position and velocity of particle i for $1 \leq i \leq N$. Due to the fact that the particles are hard-spheres, the non-overlapping condition holds

$$|x_i - x_j| > \varepsilon\,.$$

We denote by $Z_N := (z_1, \ldots, z_N)$ the set of configurations of the particles, with for each particle $z_i := (x_i, v_i)$. We also denote in the following by $X_N := (x_1, \ldots, x_N)$ the set of positions and by $V_N := (v_1, \ldots, v_N)$ the set of velocities of the particles. The positions and velocities of the system of N particles obey the Newton laws, which are the following equations of motion

$$(2.1) \qquad \forall i \in [1, \ldots, N]\,, \quad \frac{dx_i(t)}{dt} = v_i(t)\,, \qquad \frac{dv_i(t)}{dt} = 0\,,$$

provided that the exclusion condition $|x_i(t) - x_j(t)| > \varepsilon$ is satisfied for all $j \neq i$. The flow takes therefore place in the domain

$$(2.2) \qquad \mathcal{D}_N^\varepsilon := \left\{ Z_N \in \mathbb{T}^{dN} \times \mathbb{R}^{dN} \,/\, \forall i \neq j,\, |x_i - x_j| > \varepsilon \right\}.$$

We further have to prescribe a reflection condition at the boundary of $\mathcal{D}_N^\varepsilon$: if there exist $j \neq i$ such that $|x_i - x_j| = \varepsilon$ then the incoming velocities v_i^{in}, v_j^{in} are related to the outgoing velocities v_i^{out}, v_j^{out} through the relations

$$(2.3) \qquad \begin{aligned} v_i^{in} &= v_i^{out} - \nu^{i,j} \cdot (v_i^{out} - v_j^{out})\,\nu^{i,j} \\ v_j^{in} &= v_j^{out} + \nu^{i,j} \cdot (v_i^{out} - v_j^{out})\,\nu^{i,j}\,, \end{aligned}$$

where

$$\nu^{i,j} := \frac{x_i - x_j}{|x_i - x_j|}\,.$$

Note that incoming velocities are defined by the fact that

$$\nu^{i,j} \cdot (v_i^{in} - v_j^{in}) < 0\,,$$

meaning that incoming velocities are precollisional, and similarly

$$\nu^{i,j} \cdot (v_i^{out} - v_j^{out}) > 0\,,$$

meaning that outgoing velocities are postcollisional.

2.1.2. *Solving the Newton equations.* It is not obvious to check that the Newton equations (2.1)-(2.3) define global dynamics. Indeed this is not a simple consequence of the Cauchy-Lipschitz theorem since the boundary condition is not smooth, and even not defined for all configurations. We call *pathological* a trajectory such that

- either there exists a collision involving more than two particles, or the collision is grazing (meaning that $\nu^{i,j} \cdot (v_i^{in} - v_j^{in}) = 0$) hence the boundary condition is not well defined;
- or there are an infinite number of collisions in finite time so the dynamics cannot be globally defined.

In [1, 2], it is proved that outside a negligible set of initial data there are no pathological trajectories.

Proposition 2.1. *Let N, ε be fixed. The set of initial configurations leading to a pathological trajectory is of measure zero in $\mathbb{T}^{dN} \times \mathbb{R}^{dN}$.*

Sketch of proof. Let us recall briefly the proof given in [34], following [1, 2]. For any integer $s \in \mathbb{N}^*$ and any $R > 0$, we denote

$$B_R^s := \{V_s \in \mathbb{R}^{ds}, |V_s| \leq R\}$$

where $|\cdot|$ is the euclidean norm. Now let us fix $R > 0$, $\delta < \varepsilon/2$ (recall that ε is the diameter of the particles) and $t > 0$ and assume t/δ is an integer. Then it is easy to see that the set

$$\left\{ Z_N \in B_R^N \times B_R^N \,/\, \text{one particle will collide with two others on the time } [0, \delta] \right\}$$

has measure smaller than $C(N, \varepsilon, R)\delta^2$. Moreover up to removing a measure zero set of initial data each collision on $[0, \delta]$ is non-grazing. We can repeat this argument starting again at time δ since the measure is invariant by the flow, so repeating the procedure t/δ times produces a subset $I_\delta(t, R)$ of $B_R^N \times B_R^N$, of measure

$$|I_\delta(t, R)| \leq C(N, R, t, \varepsilon)\delta,$$

such that for any initial configuration in $B_R^N \times B_R^N$ outside that set, the flow is well-defined up to time t. The intersection $I(t, R) := \bigcap_{\delta > 0} I_\delta(t, R)$ is of measure zero, and any initial configuration in $B_R^N \times B_R^N$ outside $I(t, R)$ generates a well-defined flow until time t. Finally any initial configuration in $\mathbb{T}^{dN} \times \mathbb{R}^{dN}$ outside $I := \bigcup_n I(t_n, R_n)$ where t_n and R_n go to infinity, generates a globally defined flow. The proposition is proved.   □

2.2. **Euler and Navier-Stokes: a macroscopic point of view.**

2.2.1. *The equations.* The history of the mathematical study of fluids goes back many centuries, but one can probably date to the mid-eighteenth century the first equations describing fluid flows. In 1748, the Academy of Sciences in Berlin announced a Mathematics Prize for 1750, regarding the resistance of fluids when a rigid body is immersed in the fluid. J. d'Alembert took part in the competition and submitted a manuscript [3] which was revolutionary in many aspects: in particular for the first time the movement of a fluid was described by Partial Differential Equations acting on the velocity fluid. This model was close to being correct, and

the final equations were coined by L. Euler in [32]: if u is the time and space dependent, three component vector field denoting the velocity of the fluid, then it solves the following system of PDEs:

$$(2.4) \qquad\qquad (E) \qquad \begin{cases} \partial_t u + u \cdot \nabla u = -\nabla p \\ \operatorname{div} u = 0 \,, \end{cases}$$

where p is the pressure of the fluid, also an unknown, guaranteeing that the fluid remains incompressible (meaning $\operatorname{div} u = 0$) for all times. The first equation in (E) translates the conservation of momentum while the second one stands for the conservation of mass.

However d'Alembert quickly realized that a solid body placed in a fluid whose velocity satisfies those equations can evolve without suffering any resistance, which is obviously contrary to intuition and physical experiments. This is known as "d'Alembert's paradox". To understand why a solid body is submitted in general to a force tending to slow it down, one needs to take into account friction phenomena, at a molecular level: as it evolves, the fluid will have a tendency to dissipate energy under the form of heat. This phenomenon is absent from the Euler equations. L. Navier [53] had the idea, in 1820, of introducing an additional term to the Euler equations, intended to represent this dissipation of energy. Followed among others by G. Stokes in 1845 ([62]), he suggested the following model to describe the evolution of a viscous fluid:

$$(NS) \qquad \begin{cases} \partial_t u + u \cdot \nabla u - \nu \Delta u = -\nabla p \\ \operatorname{div} u = 0 \,. \end{cases}$$

The parameter $\nu > 0$ is the fluid's viscosity, and measures the discrepancy between a viscous and a perfect fluid.

2.2.2. *Solving the Navier-Stokes equations.* The question of the resolution of (NS) is not the topic of this survey. It is however useful to go through rapidly its main properties and recall the main results concerning its solutions. The most important property of (NS) relates to the conservation of energy. Formally if one computes the scalar product of u with the momentum conservation equation in (NS), noticing that

$$(u \cdot \nabla u \mid u)_{L^2} = -\frac{1}{2} \int \operatorname{div} u \, |u|^2 \, dx = 0 \quad \text{and} \quad (\nabla p \mid u)_{L^2} = -(p \mid \operatorname{div} u)_{L^2} = 0$$

one finds that

$$\frac{1}{2} \frac{d}{dt} \|u(t)\|_{L^2}^2 + \nu \|\nabla u(t)\|_{L^2}^2 = 0 \,.$$

After integration in time this implies that for all $t \geq 0$, the solution $u(t)$ associated with the initial data u_0 satisfies (formally)

$$\frac{1}{2} \|u(t)\|_{L^2}^2 + \nu \int_0^t \|\nabla u(t')\|_{L^2}^2 \, dt' = \frac{1}{2} \|u_0\|_{L^2}^2 \,.$$

This implies in particular that the map $t \mapsto \|u(t)\|_{L^2}$ is decreasing. Moreover this equality shows the "smoothing effect" induced by viscosity since as soon as the initial data is of finite energy (meaning that it lies in L^2) then the solution is instantaneously smoother in the sense that ∇u belongs to $L^2(\mathbb{R}^+; L^2)$.

Another important property of (NS) is its scale invariance: if u solves (NS) on $[0, T] \times \mathbb{R}^d$, then for any $\lambda > 0$

$$u_\lambda(t, x) := \lambda u(\lambda^2 t, \lambda x)$$

solves (NS) on $[0, \lambda^{-2} T] \times \mathbb{R}^d$ (for the rescaled data $\lambda u_0(\lambda x)$).

The issue is now to use those properties to solve the equations: given an initial data u_0, is there a solution to (NS) associated with this initial data, is it unique, does it exist globally in time? Unfortunately to this day there is no complete, satisfactory answer to this question in general. If the flow has one invariant direction (which is often physically unrealistic) then it has been known since the fundamental work of J. Leray ([49]) in 1934 that the equations are wellposed in the sense that for any finite energy initial data there is a unique, global in time solution, which has decreasing energy. In three space dimensions on the other hand the situation is less clear. One can solve the equations (uniquely and globally in time) if the initial data is small enough, where the smallness is measured in a function space invariant under the scaling of the equations – see for instance [33, 41, 17, 57, 42]. If the data is large there are (possibly non unique) global solutions of finite energy [48], which solve the equation in the sense of distributions, but uniqueness and smoothness are only known to hold for a short time. We shall not describe more in detail the numerous contributions on the question of the resolution of the (NS) equations but rather refer to [5, 46, 47] for surveys on the Cauchy problem for (NS).

2.3. **Boltzmann: a mesocopic point of view.**

2.3.1. *The equations.* L. Boltzmann's equation goes back to 1872. It can be understood as an intermediate step in the analysis of fluid motion, between Newton's microscopic approach and the Navier-Stokes macroscopic description. The idea is to place the description at a more statistical level, describing the "number", or density, of microscopic particles which at a time t have position x and velocity v. Denoting this quantity by the probability density $f = f(t, x, v)$, Boltzmann's equation (introduced in [13, 14]) states that f evolves following

$$(B) \qquad \underbrace{\partial_t f + v \cdot \nabla_x f}_{\text{free transport}} = \underbrace{\alpha Q(f, f)}_{\text{localized binary collisions}},$$

where the parameter α is the inverse of the mean free path of the microscopic particles and keeps track of the collision rate. The Boltzmann collision operator present in the right-hand side of (B) is the quadratic form, acting on the velocity variable, associated with the bilinear operator

$$Q(f, f) := \iint [f' f_1' - f f_1] \left((v - v_1) \cdot \omega \right)_+ dv_1 d\omega$$

where we have used the standard abbreviations

$$(2.5) \qquad f = f(t, x, v), \quad f' = f(t, x, v'), \quad f_1' = f(t, x, v_1'), \quad f_1 = f(t, x, v_1),$$

with (v', v_1') given by

$$v' = v + \omega \cdot (v_1 - v) \, \omega, \quad v_1' = v_1 - \omega \cdot (v_1 - v) \, \omega.$$

The Boltzmann collision operator can be split into a gain term and a loss term (see [18, 66])

$$Q(f, f) = Q^+(f, f) - Q^-(f, f).$$

The loss term counts all collisions in which a given particle of velocity v will encounter another particle, of velocity v_1, and thus will change its velocity leading to a loss of particles of velocity v, whereas the gain term measures the number of particles of velocity v which are created due to a collision between particles of velocities v' and v'_1.

2.3.2. *Solving the Boltzmann equation.* We shall not detail all the mathematical literature concerning the resolution of (B) but merely recall its main properties. It is important to notice that formally for any test function φ, there holds with the notation (2.5),

$$\int Q(f, f)\varphi dv = \frac{1}{4} \int [f'f'_1 - ff_1](\varphi + \varphi_1 - \varphi' - \varphi'_1) \left((v - v_1) \cdot \omega\right)_+ dvdv_1d\omega\,.$$

In particular,

$$\int Q(f, f)\varphi dv = 0$$

for all regular enough f, if and only if $\varphi(v)$ is a collision invariant, i.e. $\varphi(v)$ is a linear combination of $\{1, v_1, \dots, v_d, |v|^2\}$. Thus, successively multiplying the Boltzmann equation (B) by the collision invariants and then integrating in velocity yields formally the local conservation laws

$$\partial_t \int_{\mathbb{R}^d} f \begin{pmatrix} 1 \\ v \\ \frac{|v|^2}{2} \end{pmatrix} dv + \nabla_x \cdot \int_{\mathbb{R}^d} f \begin{pmatrix} v \\ v \otimes v \\ \frac{|v|^2}{2}v \end{pmatrix} dv = 0\,.$$

This provides a link to a macroscopic description of the gas.

The other very important feature of the Boltzmann equation comes also from the symmetries of the collision operator. Disregarding integrability issues, we choose $\varphi = \log f$ and find

$$\begin{aligned} D(f) &:= -\int_{\mathbb{R}^d} Q(f, f) \log f dv \\ &= \frac{1}{4} \int_{\mathbb{T}^d \times \mathbb{R}^d \times \mathbb{S}_1^{d-1}} (f'f'_1 - ff_1) \log \frac{f'f'_1}{ff_1} \left((v - v_1) \cdot \omega\right)_+ dvdv_1d\omega \geq 0\,. \end{aligned}$$

The so-defined entropy dissipation is therefore a nonnegative functional.

This leads to Boltzmann's H-theorem, also known as second principle of thermodynamics, stating that the entropy is (at least formally) a Lyapunov functional for the Boltzmann equation.

(2.6) $$\partial_t \int_{\mathbb{R}^d} f \log f dv + \nabla_x \cdot \int_{\mathbb{R}^d} f \log f v dv \leq 0\,.$$

As to the equation $Q(f, f) = 0$, it is possible to show that it is only satisfied by the so-called Maxwellian distributions $M_{\rho, u, \theta}$, which are defined by

$$M_{\rho, u, \theta}(v) := \frac{\rho}{(2\pi\theta)^{\frac{d}{2}}} e^{-\frac{|v-u|^2}{2\theta}}\,,$$

where $\rho \in \mathbb{R}_+$, $u \in \mathbb{R}^d$ and $\theta \in \mathbb{R}_+$ are respectively the macroscopic density, bulk velocity and temperature, under some appropriate choice of units. In the following we set

$$M_\beta(v) := \left(\frac{\beta}{2\pi}\right)^{\frac{d}{2}} e^{-\frac{\beta|v|^2}{2}} \quad \text{and} \quad M(v) := M_1(v).$$

Concerning the Cauchy problem, the theory is from being complete, similarly to the three dimensional Navier-Stokes equations recalled in the previous section: global existence of weak (actually renormalized) solutions is known to hold [27] but uniqueness is only known in the case when the initial data is small enough (and decaying sufficiently fast at infinity in velocity space), see for instance [63, 64].

2.4. **Hilbert's Sixth Problem.** At the second International Congress of Mathematicians held in Paris in 1900, D. Hilbert presented his famous list of twenty-three open questions [39]. Some of those questions have since been solved, and some remain open to this day. Among these, we are interested here in the following question: *to develop "mathematically the limiting processes [...] which lead from the atomistic view to the laws of motion of continua"*.

Our aim in this survey is to present some some mathematical progress that has been made recently on this question. Note that an answer to this question contains in particular an explanation to the appearance of irreversibility when passing from one description to the other, since the system of hard spheres (2.1)-(2.3) is time-reversible while the Boltzmann equation (B) and the Navier-Stokes equations (NS) are not.

3. SOME MATHEMATICAL ATTEMPTS AT RECONCILING SCALES

3.1. **Introduction.** In the large N limit, individual trajectories become irrelevant, and our goal is to describe an average behaviour. This average is of course over particles which are indistiguishable. Because we have only a vague knowledge of the state of the system at initial time, we average over initial configurations. At time 0, we thus start with a distribution $f_N^0(Z_N)$, where we use the notation introduced in Paragraph 2.1, and we define a probability $f_N = f_N(t, Z_N)$, referred to as the N-*particle distribution function*. We assume that it satisfies for all permutations σ of $\{1, \ldots, N\}$,

$$(3.1) \qquad f_N(t, Z_{\sigma(N)}) = f_N(t, Z_N),$$

with $Z_{\sigma(N)} = (x_{\sigma(1)}, v_{\sigma(1)}, \ldots, x_{\sigma(N)}, v_{\sigma(N)})$. Since f_N is an invariant of the particle system, the *Liouville equation* relative to the particle system (2.1) is

$$(3.2) \qquad \partial_t f_N + \sum_{i=1}^{N} v_i \cdot \nabla_{x_i} f_N = 0$$

on the domain $\mathcal{D}_N^\varepsilon$ defined in (2.2), with the boundary condition

$$f_N(t, Z_N^{in}) = f_N(t, Z_N^{out}),$$

meaning that on the part of the boundary of $\mathcal{D}_N^\varepsilon$ such that $|x_i - x_j| = \varepsilon$, there holds

$$(3.3) \qquad f_N(t, \ldots, x_i, v_i^{out}, \ldots x_j, v_j^{out}, \ldots) = f_N(t, \ldots, x_i, v_i^{in}, \ldots x_j, v_j^{in}, \ldots)$$

where the ingoing and outgoing velocities are related by (2.3).

3.2. **A direct approach from microscopic to macroscopic scales.** A natural approach to derive fluid mechanics equations from particle systems is to start with the following empirical distributions acting on $\mathbb{R}^+ \times \mathbb{T}^d$:

$$\rho_N(t, x) := \frac{1}{N} \sum_{i=1}^{N} \delta_{x - x_i(t)} \,,$$

$$u_N(t, x) := \frac{1}{N} \sum_{i=1}^{N} v_i(t) \delta_{x - x_i(t)} \,,$$

$$e_N(t, x) := \frac{1}{2N} \sum_{i=1}^{N} |v_i(t)|^2 \delta_{x - x_i(t)} \,,$$

and to try to use laws of large numbers or large deviation principles to obtain their limiting behaviour as N goes to infinity. This has been achieved successfully in some asymptotic regimes, in the case when noise is added to the microscopic system. We refer for instance to [56] for a derivation of the Euler equations when the momenta of nearby particles are exchanged stochastically (with a noise of very small amplitude) or to [31, 59] for a derivation of the incompressible Navier-Stokes equations. We shall not give more details here as our goal is to derive fluid mechanics equations from deterministic particle systems, and to this day the direct approach starting from empirical distributions, with no additional randomness, seems out of reach.

3.3. **From mesoscopic to macroscopic scales.** Starting from the Boltzmann equation (B), it is possible to derive formally a number of (though not all classes of) fluid mechanics equations. The formal method goes back to Hilbert [39] and Chapman and Enskog [21] and consists in looking for asymptotic expansions in terms of α. More precisely expanding the solution f to (B) under the form

$$f(t, z) = \sum_{n \geq 0} \alpha^{-n} f_n(t, z)$$

recalling that $z = (x, v)$, plugging the expansion into (B) and identifying powers of α formally gives rise the Euler equations as well as the weakly viscous incompressible Navier-Stokes equations at the next order (but also at higher orders other equations such as the Burnett model). Using truncated asymptotic expansions, it has been possible [16, 43] to obtain a a rigorous justification of the compressible Euler limit up to the first singular time for the solution of the Euler system, and similarly for the Navier-Stokes system [22]. In [6, 7], Bardos, Golse and Levermore devised a program for deriving weak solutions of the Navier-Stokes equations (NS) from the DiPerna-Lions solutions of the Boltzmann equation (B). One of the fundamental ideas behind this program is that the proof should only require a priori estimates coming from physics (namely mass, energy and entropy bounds). The difficulty in the approach however lies in the very poor understanding of renormalized solutions. Nevertheless F. Golse and L. Saint-Raymond [36, 37] were able to achieve this program in the diffusive scaling limit. The precise statement, in three space dimensions, is the following.

Theorem 3.1 ([36, 37])**.** *Consider a family $(f_{\alpha,0})_{\alpha>1}$ of initial data, the relative entropy of which satisfies, uniformly in α,*

$$\frac{1}{\alpha^2} \int f_{\alpha,0} \log \left(\frac{f_{\alpha,0}}{M} \right) dxdv \leq C_0$$

and such that as α goes to infinity the following limits hold in the sense of distributions:

$$\frac{1}{\alpha} P \int f_{\alpha,0} \, vdv \rightharpoonup u_0 \quad and \quad \frac{1}{\alpha} \int (f_{\alpha,0} - M)(\frac{1}{5}|v|^2 - 1) \, dv \rightharpoonup \theta_0 \,,$$

where P denotes the projection onto divergence free vector fields.

If $f_\alpha = M(1 + \dfrac{1}{\alpha} g_\alpha)$ is an associate family of renormalized solutions to the scaled Boltzmann equation

$$\frac{1}{\alpha} \partial_t f + v \cdot \nabla_x f = \alpha Q(f, f) \,,$$

then g_α is relatively weakly compact in $L^1_{loc}(dtdx; L^1((1 + |v|^2)Mdv))$ as α goes to infinity, and any limit point g of g_α can be written as

$$g = \rho + u \cdot v + \theta \frac{|v|^2 - 3}{2}$$

where u satisfies the incompressible Navier-Stokes equations with data u_0 and ρ, θ are linked by the Fourier system

$$\partial_t \theta + u \cdot \nabla \theta - \kappa \Delta \theta = 0 \,, \quad \theta_{|t=0} = \theta_0 \,, \quad and \quad \nabla(\rho + \theta) = 0 \,.$$

The viscosities can be explicitly computed.

The proof of this result is difficult will not be described in these notes. Let us simply mention that it relies on an approach known as the "moment method" going back to [7] and [50], and one of the main difficulties to implement this method is to control large velocities while gaining some equi-integrability properties in the space variable. That is the main achievement of [36, 37], following an idea of [60].

In Paragraph 4 we describe the proof of a considerably simpler result, going from a linear scaled Boltzmann equation to the heat equation: this is by no means intended as an explation of the proof of Theorem 3.1, but will give an idea of the reason why a diffusive equation appears as an asymptotic regime for the Boltzmann equation.

3.4. **From microscopic to mesoscopic scales.** The previous paragraph showed that it is possible to derive the incompressible Navier-Stokes equations from the nonlinear Boltzmann equation as α goes to infinity, in diffusive times (see Theorem 3.1). The question is now to derive the Boltzmann equation from particle systems, as putting both arguments together should provide a complete derivation of the incompressible Navier-Stokes equations from particle systems. As we shall see in this paragraph, this part of the program remains largely unsolved.

We shall present the strategy of Lanford [44], which is essentially the only one known to this day (we refer the interested reader to a variant introduced in [52] via a semi-group approach to the study of the probability of trees). It consists in studying the asymptotics of the first marginal $f_N^{(1)}$ of the distribution function f_N, defined by

$$f_N^{(1)}(t, z_1) := \int f_N(t, Z_N) \, dz_2 \ldots dz_N \,.$$

More generally we define the marginal of order $s \in [1, N]$ by

$$f_N^{(s)}(t, Z_s) := \int f_N(t, Z_N) \, dz_{s+1} \dots dz_N \, .$$

Lanford's theorem is the following.

Theorem 3.2 ([44]). *Consider a system of N hard spheres of diameter ε on the d-dimensional periodic box $\mathbb{T}^d = [0, 1]^d$ (with $d \geq 2$), initially "independent" in the sense that*

$$(3.4) \qquad f_{N,0}(Z_N) = \frac{1}{\mathcal{Z}_N} \prod_{i=1}^N f_0(z_i) \prod_{k \neq j} \mathbf{1}_{|x_k - x_j| > \varepsilon} \, ,$$

where f_0 is a continuous density such that

$$(3.5) \qquad \big\| f_0 \exp(\mu + \frac{\beta}{2} |v|^2) \big\|_{L^\infty(\mathbb{T}_x^d \times \mathbb{R}_v^d)} \leq 1 \, ,$$

for some $\beta > 0, \mu \in \mathbb{R}$. We have denoted by $\mathcal{Z}_N$ the partition function, that is the normalizing constant for $f_{N,0}$ to be a probability.

In the limit $N \to \infty$ with $N\varepsilon^{d-1} = \alpha$, the one particle distribution $f_N^{(1)}$ converges almost everywhere to the solution of the Boltzmann equation (B) *with initial data f_0, on a time interval $[0, t^*/\alpha]$ where t^* depends only on the parameters β, μ.*

Remark 3.3.   • The statement and main steps of the proof of Theorem 3.2 go back to [44]. We refer also to [19, 20] and [34] for details of the proof.
 • The limit $N \to \infty$ with $N\varepsilon^{d-1} = \alpha$ is known as the Boltzmann-Grad, or low density, limit (see [38]). It corresponds to a dilute gas since the volume occupied by the gas $N\varepsilon^d$ goes to zero as N goes to infinity. On average a particle of given speed has $O(\alpha)$ collisions in a given time.
 • The main drawback of the statement lies of course on the time interval on which the convergence is proved. Recall indeed that our aim is to take the limit $\alpha \to \infty$ to recover fluid mechanics equations, and that is impossible with Theorem 3.2 since the life span shrinks to zero in that limit.

Sketch of proof of Theorem 3.2. Let us explain the strategy of the proof, which is due to [44]. The details are rather long and technical and we refer the interested reader to [34] for instance.

It is not difficult to check that $f_N^{(1)} = f_N^{(1)}(t, x_1, v_1)$ satisfies the equation

$$(3.6) \qquad \partial_t f_N^{(1)} + v_1 \cdot \nabla_{x_1} f_N^{(1)} = \mathcal{C}_{1,2} f_N^{(2)} \, ,$$

where

$$\mathcal{C}_{1,2} = \mathcal{C}_{1,2}^+ - \mathcal{C}_{1,2}^-$$

and

$$(3.7) \qquad \big(\mathcal{C}_{1,2}^\pm f_N^{(2)}\big)(z_1) := (N-1)\varepsilon^{d-1} \int_{\mathbb{S}_1^{d-1} \times \mathbb{R}^d} f_N^{(2)}(x_1, v_1, x_1 + \varepsilon\omega, v_2)$$
$$\times \, (\omega \cdot (v_2 - v_1))_\pm \, d\omega dv_2 \, ,$$

the index $+$ corresponding to post-collisional configurations and the index $-$ to pre-collisional configurations. The boundary condition (3.3) imposes that

$$(3.8) \qquad \big(\mathcal{C}_{1,2}^+ f_N^{(2)}\big)(z_1) = (N-1)\varepsilon^{d-1} \int_{\mathbb{S}_1^{d-1} \times \mathbb{R}^d} f_N^{(2)}(x_1, v_1', x_1 + \varepsilon\omega, v_2')$$
$$\times \, (\omega \cdot (v_2 - v_1))_+ \, d\omega dv_2 \, ,$$

where
$$v_1' := v_1 - \omega \cdot (v_1 - v)\,\omega\,, \quad v_2' := v_2 - \omega \cdot (v_2 - v)\,\omega\,.$$

On the other hand after a change of variables $\omega \mapsto -\omega$ one finds
$$\left(\mathcal{C}_{1,2}^{-} f_N^{(2)}(Z_2) := (N-1)\varepsilon^{d-1} \int_{\mathbb{S}_1^{d-1} \times \mathbb{R}^d} f_N^{(2)}(x_1, v_1, x_1 - \varepsilon\omega, v_2)\right.$$
$$\times \left(\omega \cdot (v_2 - v_1)\right)_+ d\omega dv_2\,,$$

so that
$$(3.9) \quad \left(\mathcal{C}_{1,2} f_N^{(2)}\right)(z_1) = (N-1)\varepsilon^{d-1} \int_{\mathbb{S}_1^{d-1} \times \mathbb{R}^d} \left(f_N^{(2)}(x_1, v_1', x_1 + \varepsilon\omega, v_{s+1}')\right.$$
$$\left. - f_N^{(2)}(x_1, v_1, x_1 - \varepsilon\omega, v_1)\right)\left(\omega \cdot (v_2 - v_1)\right)_+ d\omega dv_2\,.$$

Notice that the process of transforming (3.7) into (3.9) may seem arbitrary but it is actually not since the value of $f_N^{(2)}$ at outgoing configurations is prescribed by the boundary condition (3.3) so the transformation (3.8) is actually not optional.

Assuming that $f_N^{(1)}$ has a limit f when $N \to \infty$ under the Boltzmann-Grad scaling $N\varepsilon^{d-1} = \alpha$, and similarly for $f_N^{(2)}$ (which we denote $f^{(2)}$), we find that formally
$$\partial_t f + v_1 \cdot \nabla_{x_1} f = \mathcal{C}_{1,2}^0 f^{(2)}$$

where
$$\mathcal{C}_{1,2}^0 = \mathcal{C}_{1,2}^{0+} - \mathcal{C}_{1,2}^{0-}$$

with
$$\left(\mathcal{C}_{1,2}^{0,\pm} f^{(2)}\right)(z_1) := \int_{\mathbb{S}_1^{d-1} \times \mathbb{R}^d} \left(f^{(2)}(x_1, v_1', x_1, v_2') - f^{(2)}(x_1, v_1, x_1, v_2)\right)$$
$$\times \left(\omega \cdot (v_2 - v_1)\right)_+ d\omega dv_2\,.$$

To conclude one now assumes that
$$(3.10) \qquad f^{(2)}(x_1, v_1, x_2, v_2) = f(x_1, v_1) f(x_2, v_2)$$

which is known as the *propagation of chaos assumption*, and the Boltzmann equation (B) appears immediately.

The difficulty in transforming the above argument into a rigorous proof lies in the justification of the different limits taken above, as well as the propagation of chaos (3.10). Let us explain Lanford's main ideas (which were subsequently developed and precised by, among others, [19, 20, 23, 34, ?, 61]). First since the equation (3.6) on $f_N^{(1)}$ involves $f_N^{(2)}$, one needs to write the whole hierarchy of equations known as the BBGKY hierarchy
$$(3.11) \qquad \partial_t f_N^{(s)} + \sum_{1 \le i \le s} v_i \cdot \nabla_{x_i} f_N^{(s)} = \mathcal{C}_{s,s+1} f_N^{(s+1)}\,,$$

where as above one can write
$$\mathcal{C}_{s,s+1} = \sum_{i=1}^{s} \mathcal{C}_{s,s+1}^i$$

the index i referring to the index of the interaction particle among the s "fixed" particles, with the notation

$$\left(\mathcal{C}^i_{s,s+1}f^{(s+1)}_N\right)(Z_s) := (N-s)\varepsilon^{d-1}\int_{\mathbb{S}^{d-1}_1\times\mathbb{R}^d}\int\left(\omega\cdot(v_{s+1}-v_i)\right)_+$$

$$\times\left(f^{(s+1)}_N(z_1,\ldots,x_i,v'_i,\ldots,z_s,x_i+\varepsilon\omega,v'_{s+1}) - f^{(s+1)}_N(Z_s,x_i-\varepsilon\omega,v_{s+1})\right)d\omega dv_{s+1}\,.$$

Denote by $\boldsymbol{\Psi}_s(t)$ the s-particle flow associated with the hard-spheres system, and by $\mathbf{S}_s$ the associated solution operator:

$$\mathbf{S}_s(t): \qquad f\in L^\infty(\mathcal{D}^\varepsilon_s;\mathbb{R})\mapsto f(\boldsymbol{\Psi}_s(-t,\cdot))\in L^\infty(\mathcal{D}^\varepsilon_s;\mathbb{R})\,.$$

The time-integrated form of equation (3.11) is

$$f^{(s)}_N(t,Z_s) = \mathbf{S}_s(t)f^{(s)}_N(0,Z_s) + \int_0^t \mathbf{S}_s(t-\tau)\mathcal{C}_{s,s+1}f^{(s+1)}_N(\tau,Z_s)\,d\tau\,.$$

Notice that actually the only way to make sense of the collision operators is to use the above Duhamel formulation consisting in applying a transport operator to lift the singularity of the collision integral (where a trace on a hypersurface is taken); we refer to [34] for details. The *total flow* and *total collision* operators $\mathbf{S}$ and $\mathbf{C}_N$ are defined on finite sequences $G_N = (g_s)_{1\le s\le N}$ as follows:

$$(3.12) \qquad \begin{cases} \forall s\le N\,, \ \left(\mathbf{S}(t)G_N\right)_s := \mathbf{S}_s(t)g_s\,, \\ \forall s\le N-1\,, \ \left(\mathbf{C}_N G_N\right)_s := \mathcal{C}_{s,s+1}g_{s+1}\,, \quad \left(\mathbf{C}_N G_N\right)_N := 0\,. \end{cases}$$

We finally define solutions to the BBGKY hierarchy to be solutions of

$$F_N(t) = \mathbf{S}(t)F_N(0) + \int_0^t \mathbf{S}(t-\tau)\mathbf{C}_N F_N(\tau)\,d\tau\,, \qquad F_N = (f^{(s)}_N)_{1\le s\le N}\,.$$

The main idea is then to define a limit hierarchy by formally taking the limit $N\to\infty$ under the Boltzmann-Grad scaling $N\varepsilon^{d-1} = \alpha$. We thus define the limiting collision operators

$$\mathcal{C}^0_{s,s+1} = \sum_{i=1}^s \mathcal{C}^{0,i}_{s,s+1}$$

with

$$\mathcal{C}^{0,i}_{s,s+1}f^{(s+1)}(t,Z_s) := \alpha\int\left(\omega\cdot(v_{s+1}-v_i)\right)_+$$

$$\times\left(f^{(s+1)}(t,x_1,v_1,\ldots,x_i,v'_i,\ldots,x_s,v_s,x_i,v'_{s+1}) - f^{(s+1)}(t,Z_s,x_i,v_{s+1})\right)d\omega dv_{s+1}\,.$$

Similarly to (3.12), we can then define the total Boltzmann flow and collision operators $\mathbf{S}^0$ and $\mathbf{C}^0$ as follows:

$$(3.13) \qquad \begin{cases} \forall s\ge 1\,, \ \left(\mathbf{S}^0(t)G\right)_s := \mathbf{S}^0_s(t)g_s\,, \\ \forall s\ge 1\,, \ \left(\mathbf{C^0}G\right)_s := \mathcal{C}^0_{s,s+1}g_{s+1}\,, \end{cases}$$

so that solutions to the Boltzmann hierarchy solve

$$(3.14) \qquad F(t) = \mathbf{S}^0(t)F(0) + \int_0^t \mathbf{S}^0(t-\tau)\mathbf{C^0}F(\tau)\,d\tau\,, \qquad F = (f^{(s)})_{s\ge 1}\,.$$

The crucial point is to notice that if

$$f^{(s)}(t,Z_s) = \prod_{i=1}^s f(t,z_i)$$

(meaning $f^{(s)}(t)$ is *tensorized*) then f satisfies the Boltzmann equation (B). It follows that the chaos property (3.10) will be automatically satisfied if one proves the convergence of one hierarchy to the other, as well as uniqueness for the limit hierarchy.

It turns out that the restriction on the time interval on which Theorem 3.2 holds is precisely due to the proof of the wellposedness of the hierarchy. Indeed the proof relies on a Cauchy-Kowalevskaya type argument (in the spirit of [54, 55, 65]) which completely misses the structure of the collision operators, which are dealt with as if the nonlinear term in the Boltzmann equation was f^2 instead of $Q(f, f)$. We refer to [34] for details.

The main difficulty in the proof of the convergence of one hierarchy to the other lies in the possibility of recollisions in the BBGKY flow, meaning that two particles that have collided in the past (directly or indirectly via collisions in chain with other particles) may collide again in the future and be deflected one by the other. This type of situation is impossible in the Boltzmann hierarchy, where each $\mathcal{C}^{0,i}_{s,s+1}$ term in the Duhamel formulation corresponds to a collision of a particle labeled i with a "fresh" particle labeled $s+1$, and in between each collision there is free flow – particles finding themselves at the same place at some moment simply continue their trajectory without being deflected. Eliminating recollisions is possible thanks to geometric arguments which are valid as long as

- there are not too many particles at play (of the order of $\log N$ at most): it is therefore necessary to truncate the Duhamel sum, which expresses the number of particles that have actually interacted at time t, directly or indirectly. This is done thanks to the wellposedness of the hierarchy, which provides the necessary a priori bounds on the Duhamel series;
- the velocities of the interacting particles are under control (at most of size $O(\log N)$): truncating velocities is possible thanks to the a priori bounds;
- collision times are not too close, namely in the iterated Duhamel formula $|t_i - t_{i+1}| \geq \delta$ where δ scales like a power of ε: this is possible thanks to a Lebesgue dominated convergence argument.

We refer to [34] fo more details. □

4. Both limits reconciled: linear models of fluids

Summarizing the two previous paragraphs, the limit from the mesoscopic to the macroscopic description of fluids corresponds to taking $\alpha \to \infty$ in the Boltzmann equation (B) (and rescaling time), and is known in some situations – namely near equilibrium or for weak solutions. On the other hand the limit from the microscopic to the mesoscopic description corresponds to taking $N \to \infty$ with $N\varepsilon^{d-1} = \alpha$, and is known for small times only, of the order of α^{-1}. This prevents from combining both limits to go from particles to fluids.

There are however some (linear) cases where much progress on the Hilbert program has been made, we refer for instance to the works [15, 25, 26, 35]. The full program we are after has been achieved recently in two linear contexts [10, 11], namely in deriving the linear heat and Stokes-Fourier equations. We describe briefly the case studied in [10] in this paragraph: in the case of a tagged particle in a background at equilibrium, it is proved in [10] that its distribution satisfies

a heat equation at the limit $N \to \infty$, using the linear Boltzmann equation as an intermediate step (with a parameter α going slowly to infinity with N). The precise result is the following.

Theorem 4.1. [10] *Consider the initial distribution*

$$f_N^0(Z_N) := \frac{1}{\mathcal{Z}_N} \rho^0(x_1) M_\beta^{\otimes N}(V_N) \prod_{k \neq j} \mathbf{1}_{|x_k - x_j| > \varepsilon},$$

with $\mathcal{Z}_N$ the normalizing constant. Assume that ρ^0 belongs to $C^0(\mathbb{T}^d)$. Then for all $T > 0$ and all $\tau \in [0, T]$, the distribution $f_N^{(1)}(\alpha\tau, x, v)$ remains close for the L^∞-norm to $\rho(\tau, x) M_\beta(v)$ where $\rho(\tau, x)$ is the solution of the linear heat equation

$$\partial_\tau \rho - \kappa_\beta \Delta_x \rho = 0 \quad in \quad \mathbb{T}^d, \quad \rho_{|\tau=0} = \rho^0,$$

and the diffusion coefficient κ_β is given by

$$(4.1) \qquad\qquad \kappa_\beta := \frac{1}{d} \int_{\mathbb{R}^d} v \mathcal{L}^{-1} v \, M_\beta(v) dv,$$

where $\mathcal{L}$ is the linear Boltzmann operator (4.3) and $\mathcal{L}^{-1}$ is its pseudo-inverse defined on $(Ker\mathcal{L})^\perp$. More precisely,

$$(4.2) \qquad\qquad \left\| f_N^{(1)}(\alpha\tau, x, v) - \rho(\tau, x) M_\beta(v) \right\|_{L^\infty([0,T] \times \mathbb{T}^d \times \mathbb{R}^d)} \to 0$$

in the limit $N \to \infty$, with $\alpha = N\varepsilon^{d-1}$ going to infinity much slower than $\sqrt{\log \log N}$. In the same asymptotic regime, the process $\Xi(\tau) = x_1(\alpha\tau)$ associated with the tagged particle converges in law towards a Brownian motion of variance κ_β, initially distributed under the measure ρ^0.

This is an extension of the works [9, 45] where the linear Boltzmann equation was derived for long times. We shall not describe further those results here, but simply mention that the main achievement consists in deriving the linear Boltzmann equation for an arbitrarily long time (contrary to the Lanford theorem which only holds for times of the order! α^{-1}) thanks to the use of the maximum principle associated with this very special type of initial data. To conclude this paragraph we shall merely explain why the linear Boltzmann equation does have the heat equation as an asymptotic regime. Compared with Theorem 3.1 this can be considered as an exercise, but we feel it has some interest as it at least gives a flavor of the reason why a transport-type equation like (B) can lead asymptotically to a diffusive equation.

The linear Boltzmann equation is defined by linearizing the Boltzmann equation (B) around a Maxwellian M_β and in dropping two of the four terms appearing in the linearization of the collision integral: factoring out the Maxwellian leads to the equation

$$\partial_t \varphi_\alpha + v \cdot \nabla_x \varphi_\alpha = -\alpha \, \mathcal{L}\varphi_\alpha$$

$$(4.3) \qquad \mathcal{L}\varphi_\alpha(v) := \iint [\varphi_\alpha(v) - \varphi_\alpha(v')] \, M_\beta(v_1) \big((v - v_1) \cdot \omega\big)_+ \, dv_1\omega.$$

It is not difficult to prove that as soon as the initial data belongs to L^∞, then there is a unique global solution to (4.3), which remains uniformly bounded in α, for all times. The precise result describing the limit $\alpha \to \infty$ of φ_α in diffusive times is the following.

Proposition 4.2. *Consider ρ^0 a continuous density of probability on $\mathbb{T}^d$ and let φ_α be the associate solution of (4.3) with initial data ρ^0. There holds for all $T \in [0,T]$*

(4.4)
$$\sup_{\tau \in [0,T]} \sup_{(x,v) \in \mathbb{T}^d \times \mathbb{R}^d} \left| M_\beta(v) \big(\varphi_\alpha(\alpha\tau, x, v) - \rho(\tau, x) \big) \right| \to 0, \quad \alpha \to \infty,$$

where ρ solves

$$\partial_\tau \rho - \kappa_\beta \Delta_x \rho = 0 \quad in \quad \mathbb{T}^d, \quad \rho_{|\tau=0} = \rho^0,$$

and the diffusion coefficient κ_β is given by (4.1).

Main steps of the proof. Let us define

$$\widetilde{\varphi}_\alpha(\tau, x, v) := \varphi_\alpha(\alpha\tau, x, v),$$

which satisfies

(4.5)
$$\partial_\tau \widetilde{\varphi}_\alpha + \alpha \, v \cdot \nabla_x \widetilde{\varphi}_\alpha + \alpha^2 \mathcal{L} \widetilde{\varphi}_\alpha = 0.$$

Notice that by the maximum principle on the heat equation, we may assume without loss of generality that ρ^0 is smooth. As recalled above, the formal Hilbert expansion consists in writing an asymptotic expansion of $\widetilde{\varphi}_\alpha$ in terms of powers of α^{-1}

$$\widetilde{\varphi}_\alpha(\tau, x, v) = \widetilde{\rho}_0(\tau, x, v) + \frac{1}{\alpha} \widetilde{\rho}_1(\tau, x, v) + \frac{1}{\alpha^2} \widetilde{\rho}_2(\tau, x, v) + \dots,$$

in plugging that expansion in Equation (4.5), and in canceling successively all the powers of α. This gives formally the following set of equations, keeping only the $O(1)$, $O(\alpha)$ and $O(\alpha^2)$ terms

(4.6)
$$\mathcal{L}\widetilde{\rho}_0 = 0,$$
$$v \cdot \nabla_x \widetilde{\rho}_0 + \mathcal{L}\widetilde{\rho}_1 = 0,$$
$$\partial_\tau \widetilde{\rho}_0 + v \cdot \nabla_x \widetilde{\rho}_1 + \mathcal{L}\widetilde{\rho}_2 = 0.$$

In order to find the expressions for $\widetilde{\rho}_1$ and $\widetilde{\rho}_2$, as well as the equation on $\widetilde{\rho}_0$ (which we expect to be the heat equation), it is necessary to be able to invert the operator $\mathcal{L}$. It is known (see [39]) that $\mathcal{L}$ is invertible on the set of functions

$$\left\{ g \in L^2(\mathbb{R}^d, a_\beta M_\beta dv), \int_{\mathbb{R}^d} g(v) \, M_\beta(v) dv = 0 \right\},$$

where

$$a_\beta(v) := \int_{\mathbb{S}^{d-1} \times \mathbb{R}^d} M_\beta(v_1) \big((v - v_1) \cdot \omega \big)_+ d\omega dv_1.$$

The first equation in (4.6) therefore reflects the fact that $\widetilde{\rho}_0$ does not depend on v. We next define the vector $b(v) = \big(b_k(v) \big)_{k \leq d}$ with $\int_{\mathbb{R}^d} b(v) \, M_\beta(v) dv = 0$, by

(4.7)
$$\mathcal{L}b(v) := v.$$

Returning to (4.6), we have

$$\widetilde{\rho}_1(\tau, x, v) = \rho_1(\tau, x, v) + \overline{\rho}_1(\tau, x),$$

with

$$\rho_1(\tau, x, v) := -b(v) \cdot \nabla_x \widetilde{\rho}_0(\tau, x) \quad \text{and} \quad \overline{\rho}_1 \in \operatorname{Ker} \mathcal{L}.$$

Next we consider the last equation in (4.6) and we notice that for $\widetilde{\rho}_2$ to exist it is necessary for $\partial_\tau \widetilde{\rho}_0 + v \cdot \nabla_x \widetilde{\rho}_1$ to belong to the range of $\mathcal{L}$. Since $\widetilde{\rho}_0$ does not depend on v, this means that

$$(4.8) \qquad \partial_\tau \widetilde{\rho}_0 + \int_{\mathbb{R}^d} v \cdot \nabla_x \widetilde{\rho}_1(\tau, x, v) M_\beta(v)\, dv = 0 \,.$$

We then define the diffusion matrix $D(v) = \big(D_{k,\ell}(v)\big)_{k,\ell \leq d}$ by

$$(4.9) \qquad \mathcal{L}D(v) := v \otimes b(v) - \int_{\mathbb{R}^d} v \otimes b(v)\, M_\beta(v) dv \,.$$

From the symmetry of the model, one can check (see [24] for instance) that there is a function γ such that

$$b(v) = \gamma(|v|)v \,.$$

The end of the formal proof is an easy computation, noticing that by symmetry of b

$$\frac{1}{d} \int_{\mathbb{R}^d} v \mathcal{L}^{-1} v\, M_\beta(v) dv = \frac{1}{d} \int_{\mathbb{R}^d} \gamma(|v|)|v|^2\, M_\beta(v) dv \,.$$

Turning these formal arguments into a proof of convergence is not difficult and follows from the maximum principle. We refer for instance to [8, 10] for details. $\square$

5. Conclusion and open problems

In this survey we have presented some recent mathematical results concerning the derivation of fluid mechanics equations from the fundamental laws of mechanics and shown that the full derivation of the Navier-Stokes equations from Newton's laws is still widely open, due in particular to our lack of understanding of the derivation of the Boltzmann equation for large times. One of the reasons that this derivation fails to hold for large times (or large α) is that to this day we are unable to use the fundamental property of the Boltzmann equation – namely the entropy dissipation – at the level of particles.

This problem set aside, there are still many other open problems even in the short-time derivation of the Boltzmann equation.

- We have assumed that the particles are all identical spheres of mass 1 and diameter $\varepsilon > 0$. This assumption could be slightly relaxed to masses of comparable size, however if the masses, shapes and sizes of the particles differ substantially then some different phenomena may appear (see [12, 28, 29, 30, 40] for instance).
- We have assumed that the particles evolve in a periodic box of size 1 denoted $\mathbb{T}^d = [0, 1]^d$. Other situations could be considered, like the whole space $\mathbb{R}^d$ – the absence of boundaries simplifies substantially the analysis. In particular although the passage from Boltzmann to Navier-Stokes does hold in the presence of boundaries [51], essentially nothing is known in the derivation of the Boltzmann equation in a domain with boundaries – except of course if the problem can be reduced to the whole space or to a periodic box by symmetry.
- We have assumed that the particles interact elastically at each binary collision and there is no other type of interaction nor forcing. The case when more than two particles collide at the same time can be neglected rather

easily (see Proposition 2.1). The case of more general interaction potentials is of course very interesting and physically relevant but to this day this has only been considered under very stringent conditions on the potential [4, 34, 58].

Finally we have focussed on the derivation of the incompressible Navier-Stokes equations, but almost all remains to be done concerning the first limit of Boltzmann when $\alpha \to \infty$ (without rescaling in time), namely the compressible Euler equation – we refer to Section 3.3 for some comments.

REFERENCES

1. R. Alexander, The Infinite Hard Sphere System, *Ph.D. dissertation,* Dept. Mathematics, Univ. California, Berkeley, 1975.
2. R. Alexander, Time Evolution for Infinitely Many Hard Spheres, *Commun. in Math. Phys.* 49 (1976), p.217–232.
3. J. d'Alembert, *Essai d'une nouvelle théorie de la résistance des fluides* (1752), Paris.
4. N. Ayi, From Newton's law to the linear Boltzmann equation without cut-off, *Communications in Mathematical Physics*, 350 (3), 1219–1274.
5. H. Bahouri, J.-Y. Chemin and R. Danchin, Fourier analysis and nonlinear partial differential equations, Grundlehren der Mathematischen Wissenschaften, Springer, Heidelberg, 343, 2011.
6. C. Bardos, F. Golse, C.D. Levermore, Fluid Dynamic Limits of the Boltzmann Equation I, *J. Stat, Phys.* 63 (1991), 323–344.
7. C. Bardos, F. Golse, C.D. Levermore, Fluid Dynamic Limits of Kinetic Equations II: Convergence Proofs for the Boltzmann Equation, *Comm. Pure Appl. Math* 46 46 (1993), 667–753.
8. C. Bardos, R. Santos, R. Sentis, Diffusion approximation and computation of the critical size. *Trans. Amer. Math. Soc*, **284** (2), p. 617-649 (1984).
9. H. van Beijeren, O.E. III Lanford, J.L Lebowitz, H. Spohn, Equilibrium Time Correlation Functions in the Low Density Limit. *Jour. Stat. Phys.* 22, (1980), 237-257.
10. T. Bodineau, I. Gallagher, L. Saint-Raymond, The Brownian motion as the limit of a deterministic system of hard-spheres, *Inventiones mathematicae* 203 (2016), 493–553.
11. T. Bodineau, I. Gallagher, L. Saint-Raymond, *From hard sphere dynamics to the Stokes-Fourier equations: an L^2 analysis of the Boltzmann-Grad limit, Annals of PDEs* 3 (2017), no. 1, Art. 3:2, 118 pp.
12. T. Bodineau, I. Gallagher and L. Saint-Raymond, Derivation of an Ornstein-Uhlenbeck process for a massive particle in a rarified gas of particles, *submitted*.
13. L. Boltzmann, Weitere Studien uber das Warme gleichgenicht unfer Gasmolakular. Sitzungsberichte der Akademie der Wissenschaften 66 (1872), 275-370. Translation : Further studies on the thermal equilibrium of gas molecules, in Kinetic Theory 2, 88-174, Ed. S.G. Brush, Pergamon, Oxford (1966).
14. L. Boltzmann, *Leçons sur la théorie des gaz*, Gauthier-Villars (Paris, 1902-1905). Ré-édition Jacques Gabay, 1987.
15. L.A. Bunimovich, Y. G. Sinai . Statistical properties of Lorentz gas with periodic configuration of scatterers. *Comm. Math. Phys.* 78 (1980/81), 479-497.
16. R.E. Caflisch,The Boltzmann equation with a soft potential. I. Linear, spatially homogeneous, *Commun. Math. Phys.* 74 (1980), 71–95.
17. M. Cannone, A generalization of a theorem by Kato on Navier-Stokes equations, *Rev. Mat. Iberoamericana*, 13(3):515–541, 1997.
18. C. Cercignani, *The Boltzmann equation and its application*, Springer-Verlag, 1988.
19. C. Cercignani, V. I. Gerasimenko, D. I. Petrina, *Many-Particle Dynamics and Kinetic Equations*, Kluwer Academic Publishers, Netherlands, 1997.
20. C. Cercignani, R. Illner, M. Pulvirenti, *The Mathematical Theory of Dilute Gases*, Springer Verlag, New York NY (1994).
21. S. Chapman, T.G. Cowling, *The mathematical theory of non-uniform gases: An account of the kinetic theory of viscosity, thermal conduction, and diffusion in gases*, Cambridge University Press, New York, 1960.
22. A. DeMasi, R. Esposito, J. Lebowitz, Incompressible Navier-Stokes and Euler Limits of the Boltzmann Equation, *Comm Pure Appl. Math.* 42 (1990), 1189–1214.

23. R. Denlinger, *The propagation of chaos for a rarefied gas of hard spheres in the whole space*, https://arxiv.org/abs/1605.00589.

24. L. Desvillettes, F. Golse, *A remark concerning the Chapman-Enskog asymptotics*, in "Advances in kinetic theory and computing", B. Perthame ed., 191–203, Ser. Adv. Math. Appl. Sci., 22, World Sci. Publishing, River Edge, NJ, 1994.

25. L. Desvillettes, M. Pulvirenti. The linear Boltzmann equation for long-range forces: a derivation from particles. *Math. Meth. Mod. Appl. Sci.* 9 (1999), no. 8, 11231145.

26. L. Desvillettes, V. Ricci, A rigorous derivation of a linear kinetic equation of Fokker-Planck type in the limit of grazing collisions. *J. Statist. Phys.* 104 (2001), no. 5-6, p. 1173-1189.

27. R. DiPerna and P.-L. Lions, On the Cauchy problem for Boltzmann equations: global existence and weak stability, *Ann. of Math.* (2) 130 (1989), no. 2, 321–366.

28. M. Dobson, F. Legoll, T. Lelièvre, G. Stoltz, Derivation of Langevin dynamics in a nonzero background flow field, *ESAIM: M2AN*, 47, 6 (2013), 1583–1626.

29. D. Dürr, S. Goldstein, J. L. Lebowitz, A mechanical model of Brownian motion, *Comm. Math. Phys.*, 78, 4 (1981), 507–530.

30. D. Dürr, S. Goldstein, J. L. Lebowitz, *A mechanical model for the Brownian motion of a convex body*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 62, 4 (1983), 427–448.

31. R. Esposito, R. Marra, H.T. Yau, Navier-Stokes equations for stochastic particle systems on the lattice, *Comm. Math. Phys.* 182 (1996), 395–456.

32. L. Euler, Principles of the motion of fluids, *Physica D* 237 (2008), 18401854. (English translation of the original 1756-1757 article).

33. H. Fujita et T. Kato, On the Navier-Stokes initial value problem I, *Archive for Rational Mechanics and Analysis*, 16 (1964), 269–315.

34. I. Gallagher, L. Saint-Raymond, B. Texier, From Newton to Boltzmann : the case of hard-spheres and short-range potentials, *ZLAM* (2014).

35. G. Gallavotti. Statistical mechanics. A short treatise. *Texts and Monographs in Physics. Springer-Verlag, Berlin*, 1999.

36. F. Golse, L. Saint-Raymond, The Navier-Stokes limit of the Boltzmann equation for bounded collision kernels, *Invent. Math.* 155 (2004), no. 1, 81–161.

37. F. Golse, L. Saint-Raymond, The incompressible Navier-Stokes limit of the Boltzmann equation for hard cutoff potentials, *J. Math. Pures Appl.* 91 (2009), no. 5, 508–552.

38. H. Grad, On the kinetic theory of rarefied gases, *Comm. Pure Appl. Math.* 2 (1949), p. 331-407.

39. D. Hilbert. Begründung der kinetischen Gastheorie, *Math. Ann.* 72 (1912), no. 4, 562–577.

40. R. Holley, The motion of a heavy particle in an infinite one dimensional gas of hard spheres, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 17 (1971), 181–219.

41. T. Kato, Strong L^p-solutions of the Navier-Stokes equation in $\mathbb{R}^m$ avec applications to weak solutions, *Mathematische Zeitschrift*, 187 (1984), 471-480 .

42. H. Koch et D. Tataru, Well–posedness for the Navier–Stokes equations, *Advances in Mathematics*, 157 (2001), 22–35.

43. M. Lachowicz, On the initial layer and the existence theorem for the nonlinear Boltzmann equation, *Math. Methods Appl. Sci.* 9 (1987), 342–366.

44. O.E. Lanford, *Time evolution of large classical systems*, Lect. Notes in Physics 38, J. Moser ed., 1–111, Springer Verlag (1975).

45. J. Lebowitz, H. Spohn, Steady state self-diffusion at low density. *J. Statist. Phys.* 29 (1982), 39-55.

46. P.-G. Lemarié-Rieusset, Recent developments in the Navier-Stokes problem. Chapman and Hall/CRC Research Notes in Mathematics, 43, 2002.

47. P.-G. Lemarié-Rieusset, The Navier-Stokes problem in the 21st century, Chapman and Hall/CRC Research Notes in Mathematics, 2016.

48. J. Leray, Essai sur le mouvement d'un liquide visqueux emplissant l'espace, *Acta Matematica*, 63 (1933), 193–248.

49. J. Leray, Étude de diverses équations intégrales non linéaires et de quelques problèmes que pose l'hydrodynamique. *Journal de Mathématiques Pures et Appliquées*, 12 (1933), 1–82.

50. P.-L. Lions, N. Masmoudi, From Boltzmann Equation to the Navier-Stokes and Euler Equations II, *Archive Rat. Mech. Anal.* 158 (2001), 195–211.

51. N. Masmoudi, L. Saint-Raymond, From the Boltzmann equation to the Stokes- Fourier system in a bounded domain, Comm. Pure Appl. Math., 56 (2003), 1263 1293.

52. K. Matthies and F. Theil, A semi-group approach to the justification of kinetic theory, *SIAM Journal of Math. Anal* **44** (2012), 4345–4379.
53. C. Navier, Mémoire sur les lois du mouvement des fluides, *Mémoire de l'Académie des Sciences de l'Institut de France*, 6 (1822), 375–394.
54. L. Nirenberg, An abstract form of the nonlinear Cauchy-Kowalewski theorem, *Journal of Differential Geometry*, 6 (1972), 561–576.
55. T. Nishida, A note on a theorem by Nirenberg, *Journal of Differential Geometry*, 12 (1977), 629–633.
56. S. Olla, S. Varadhan, H. Yau, Hydrodynamical limit for a Hamiltonian system with weak noise, *Commun. Math. Phys.* 155 (1993), 523–560.
57. F. Planchon, Asymptotic behavior of global solutions to the Navier-Stokes equations in R^3, *Rev. Mat. Iberoamericana*, 14(1):71–93, 1998.
58. M. Pulvirenti, C. Saffirio, S. Simonella, *On the validity of the Boltzmann equation for short range potentials,* Rev. Math. Phys. **26** (2014), no. 2.
59. J. Quastel and H.-T. Yau, Lattice gases, large deviations, and the incompressible Navier-Stokes equations, *Ann. of Math.* 155 (1998), 51–108.
60. L. Saint-Raymond, From the BGK model to the Navier-Stokes equations, *Ann. Sci. Ecole Norm. Sup.* (4) 36 (2003), 271?317.
61. H. Spohn, Boltzmann hierarchy and Boltzmann equation, in *Kinetic theories and the Boltzmann equation* (Montecatini, 1981), p. 207-220.
62. H. Spohn, Large scale dynamics of interacting particles, *Springer-Verlag* **174** (1991).
63. G. Stokes, On the Theories of Internal Friction of Fluids in Motion and of the Equilibrium and Motion of Elastic Solids, *Trans. Camb. Phil. Soc.*, 8 (1845), 287–319.
64. S. Ukai, On the existence of global solutions of mixed problem for non-linear Boltzmann equation, *Proc. Japan Acad.,* 50 (1974), p. 179–184.
65. S. Ukai, Les solutions globales de l'équation de Boltzmann dans l'espace tout entier et dans le demi-espace, *C.R. Acad. Sci. Paris,* 282(6) (1976), p. 317–320.
66. S. Ukai, The Boltzmann-Grad Limit and Cauchy-Kovalevskaya Theorem, *Japan J. Indust. Appl. Math.,* 18 (2001), 383–392.
67. C. Villani, *A review of mathematical topics in collisional kinetic theory*, Handbook of mathematical fluid dynamics, Vol. I, 71-305, North-Holland, Amsterdam, 2002.

Current address: Université Paris-Diderot and École Normale Supérieure, Paris, France
E-mail address: gallagher@math.ens.fr

THE CAP SET CONJECTURE, THE POLYNOMIAL METHOD, AND APPLICATIONS (AFTER CROOT–LEV–PACH, ELLENBERG–GIJSWIJT, AND OTHERS)

JOSHUA A. GROCHOW

ABSTRACT. The cap set problem is the question of how large a subset of $(\mathbb{Z}/3\mathbb{Z})^n$ can be and contain no lines. This problem was motivated by deep questions about structure in the prime numbers, the geometry of lattice points, and the design of statistical experiments. In 2016, Croot, Lev, and Pach solved the analogous problem in $(\mathbb{Z}/4\mathbb{Z})^n$, showing that the largest set without lines had size at most c^n for some $c < 4$. Their proof was as elegant as it was unexpected, being a departure from the tried and true methods of Fourier analysis that had dominated the field for half a century. Shortly thereafter Ellenberg and Gijswijt leveraged their method to resolve the original cap set problem. This expository article covers the history and motivation for the cap set problem and some of the many applications of the technique: to relations between polynomials, to rigidity of matrices, and to algorithms for matrix multiplication. The latter application turns out to give back to the original problem, sharpening our understanding of the techniques involved and of what's needed to show that the current bounds are tight. Most of our exposition assumes only familiarity with basic linear algebra, polynomials, and the integers modulo N.

1. INTRODUCTION

The published proof [32] of the Cap Set Conjecture is so elegant, elementary, and short—and others have already provided expositions of it [51, 112, 120]—that we can hardly do better here. Although we'll include the quick proof in Section 2, our main purpose here is to provide motivation for the conjecture, put it in its proper historical context, and discuss some of the consequences of the new technique.

The Cap Set Conjecture is that the largest subset of $(\mathbb{Z}/3\mathbb{Z})^n$ which contains no lines—that is, no three points x, y, z such that $x + y = 2z$, or equivalently (mod 3), $x + y + z = 0$—has size at most c^n for some c strictly less than 3.

If you are an (additive) combinatorialist, you may find this problem intrinsically interesting and immediately fall in love with it. For the rest of us, however, it is natural to wonder how one arrives at this conjecture. Why were people studying such questions in the first place? What connections does it have to other areas of mathematics? As is often the case in mathematics, if we take the time to get to know the problem a bit better—take it to dinner, ask about its history, its family, what kind of recreational activities it enjoys—we find that these questions have good answers, and we come to appreciate a problem who's upbringing is perhaps not so much like our own. My main goal in this exposition is to share some of the

answers to these questions. And, okay, sure, I'll show you the proof, too—how can I resist?

One final note before diving in: I have attempted to make this exposition accessible to as wide an audience as possible, say starting from the early undergraduate level. As this may well include students who haven't seen group theory before, or have little exposure to combinatorics, I spell out (nearly) all the details for expository purposes, including some combinatorial arguments that are easy exercises once you've seen a bit of combinatorics. We ask for a little patience from our readers for whom such arguments are too standard to be worth writing down. In the few cases where I wanted to make some remark that was difficult to make under these conditions, footnotes are there to help the reader along; if such remarks confuse you, you ought to be able to skip over them and still understand the rest of the article. Conversely, if such footnotes annoy you, you ought to be able to skip over them without losing much. But my hope is that the footnotes will help entice readers with less background to learn new and exciting things!

1.1. **A frivolous and fun motivation.** The popular imagination is perhaps drawn to this problem because of its connection with the card game SET®. The SET® deck consists of 81 cards, which look like this:

FIGURE 1. A Set in the popular card game.

Each card has four attributes, each of which can take three values: color (values: red, green, purple), shape (diamond, oval, squiggle), number (1,2,3), and fill (solid, shaded, open). A "Set" is a collection of three cards such that, in each attribute, either all cards have the same value, or all cards have distinct values. Figure 1 shows a Set in which, in each attribute, all the cards are distinct (they have three different colors—or, if you're reading this in black and white—are all the same color—three different shapes, three different numbers, and three different fills). As another example, the cards (red,diamond,1,solid), (red,diamond,2,shaded), (red,diamond,3,open) also form a Set. Twelve cards are laid face up, and players compete to find Sets as fast as possible. When the players agree there is no Set on the table, three more cards are laid face up—or just one, if you're feeling lucky—until someone finds a

Set. This raises the natural question: How many cards can be on the table with no Set?

If we identify the values of each attribute with the elements of the integers mod 3, then each card corresponds to a point in $(\mathbb{Z}/3\mathbb{Z})^4$, and a Set is precisely a collection of three points $\vec{x}, \vec{y}, \vec{z} \in (\mathbb{Z}/3\mathbb{Z})^4$ such that $\vec{x} + \vec{y} + \vec{z} \equiv \vec{0} \pmod 3$ (exercise for the reader). Since $2 \equiv -1 \pmod 3$, this is the same as saying $\vec{y} - \vec{x} \equiv \vec{z} - \vec{y} \pmod 3$, in other words, that the points $\vec{x}, \vec{y}, \vec{z}$ lie on a line. Our question thus becomes: How large can a subset of $(\mathbb{Z}/3\mathbb{Z})^4$ be and still contain no line? If we generalize this from dimension 4 to arbitrary dimension n, we get the cap set problem.

In dimension four, the answer turns out to be 20 [86]. Note that this answer was found three years before Marsha Jean Falco invented the game—to help her visualize the combinatorics of genes related to epilepsy in German Shepherds [100]—and almost 20 years before the game was made public. Why were people looking at such questions? Our next motivation is older and quite a bit deeper.

1.2. Motivating additive combinatorics with the primes.

Additive combinatorics is the study of the additive structure of sets (duh!—wait, what? Wasn't that circular?). To give some meaning to the notion of "additive structure" and why it might be interesting, let's start with the example of the natural numbers $\mathbb{N} = \{1, 2, 3, \dots\}$. In terms of just addition, it seems pretty simple: Start from 1, and just keep adding 1. It looks like a (discrete) line, heading off in one direction. In terms of just multiplication, it looks a bit more complicated, but not so much more: Every number can be written uniquely as a product of primes. This implies that the multiplicative structure of $\mathbb{N}$ looks like the additive structure of the set of sequences of elements of $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, only finitely many of which aren't zero (which we denote $\mathbb{N}_0^{<\infty}$): To any such sequence of natural numbers $(a_1, a_2, \dots, a_d, 0, 0, 0, \dots)$, we may associate the number $2^{a_1} 3^{a_2} 5^{a_3} \cdots p_d^{a_d}$ where p_d is the d-th prime number. This identification shows that the multiplicative structure of $\mathbb{N}$ is isomorphic to the additive structure of $\mathbb{N}_0^{<\infty}$, an infinite-dimensional grid, as regular as can be.

When we consider both the multiplicative and additive structure of $\mathbb{N}$ together though, something remarkable happens. Note that our notion of "size" in $\mathbb{N}$ is essentially additive: how many times you need to add 1 to get to a given number. Thus, the two simplest questions we can ask that mix the additive and multiplicative structures on $\mathbb{N}$ are:

(1) How big is the n-th prime p_n? (Additive structure of a multiplicatively-defined sequence)
(2) What is the factorization of n? (Multiplicative structure of an additively-defined sequence)

These also turn out to be some of the deepest questions about the interaction between additive and multiplicative structure. The first will be traditionally recognized as deep, leading quickly to the Prime Number Theorem and the Riemann Hypothesis. The depth of the second can already be glimpsed in the techniques used in algorithms for factoring numbers [82, 88, 75, 16, 102] (see, e. g., [74, 89, 118] for surveys).

The algorithmic viewpoint adds evidence for the idea that it is the *mixture* of the additive and multiplicative structure that leads to complexity. When we represent numbers in, say, base 10, adding them is relatively easy but factoring them seems to be difficult, or at least a much deeper problem. But when we represent

numbers as their prime factorizations, multiplying them becomes as easy as adding them was before (just add up the exponents), but adding them becomes equivalent to factoring numbers written in base 10 [23]. The lesson is that, regardless of whether we view the multiplicative structure of $\mathbb{N}$ through an additive lens or its additive structure through a multiplicative lens, we run into the same, much deeper complexity than if we only considered one structure at a time.

We may also ask the next questions along these lines:

(1') What is the additive gap $p_{n+1} - p_n$ between successive primes?
(2') What is the relationship between the factorization of n and that of $n + 1$?

Only in 2013 was it shown that there is a universal constant C such that $p_{n+1} - p_n < C$ for infinitely many n [121]; C was eventually improved to 246 [79, 87], and the Twin Prime Conjecture is that C can be lowered all the way to 2. The relationship between the factorization of n and $n+1$ leads to notoriously difficult problems like the Collatz Conjecture [83], about which Erdős famously said, "Mathematics may not be ready for such problems" [59, p. 330] (see also [81]). Guy's chapter [59, Problem E16] and Lagarias's annotated bibliography [70] are excellent sources of references on this difficult problem.

Now, a zeroth-order heuristic for questions (1') and (2') is that the answer to both is essentially random: The factorizations of n and $n + 1$ are "independent" of one another, and the prime gap $p_{n+1} - p_n$ jumps around "randomly." There is some truth in this heuristic. But along with the randomness, there is also significant structure present, as evidenced already by the aforementioned results.

We may thus ask, for example, what further additive structure is there to the prime numbers? A natural generalization of (1') is to ask for structure in the differences between several primes (not necessarily consecutive). Here, perhaps the simplest structure to ask for would be for a set of primes with a common distance between them, that is, which form an *arithmetic progression* $p, p + r, p + 2r, p + 3r, \ldots, p + dr$. A classical folklore conjecture, going back perhaps a century or more, is that the primes contain arithmetic progressions of every length—quite a lot of additive structure for a multiplicatively defined set!

Green and Tao [56] proved this theorem in 2008, but we will see that the history of this theorem provides motivation and impetus for many topics in additive combinatorics, including our main topic, the Cap Set Conjecture.

1.3. **Additive combinatorics more generally.** As a young boy, Erdős (re)proved that $\sum_{p \text{ prime}} 1/p = \infty$, and it has been postulated by several authors that this early exciting mathematical experience, in combination with the long-standing conjecture about arithmetic progressions in the primes, led to:

Conjecture 1.1 (Erdős, 1940s or 1950s?, see [33, 34, 35, 103][1]). *If $A \subseteq \mathbb{N}$ satisfies $\sum_{n \in A} \frac{1}{n} = \infty$, then A contains arbitrarily long arithmetic progressions.*

In his 1976 talk, Erdős [34] offered \$3000 USD for its resolution—the highest prize he had ever offered at that point—and in 1996 he upped the prize to \$5000 USD [36], which I believe was his third-largest ever.[2] As pointed out by Soifer [103, p. 354], the high prize and the frequency with which he raised this conjecture in his talks and writings suggests it was one of his favorite.

Gowers [50] points out that Erdős's conjecture is "morally" about sets A such that the density of A in $\{1, \ldots, N\}$ is around $1/\log N$. For if the density is $1/\log N$ then the sum diverges, while if the density is, say, $1/(\log N (\log \log N)^2)$, then it converges. Now, of course, arbitrary subsets need not have a density function that varies so smoothly as a function of N. But Erdős's conjecture is sandwiched between two statements about sets of a certain density. For if $\sum_{n \in A} 1/n$ diverges, then there are infinitely many N such that $|A \cap \{1, \ldots, N\}| \geq 1/(\log N (\log \log N)^2)$, so to prove the conjecture it suffices to show that subsets of density $1/(\log N (\log \log N)^2)$ contain arbitrary long arithmetic progressions; to disprove the conjecture, it suffices to find a set of density $1/\log N$ that does not contain arithmetic progressions. Nonetheless, the conjecture is what it is. It remains open even to prove that a set A satisfying the hypothesis contains 3-term arithmetic progressions.

However, before Erdős put forth this conjecture, he was indeed thinking about the density of sets. In his 1936 paper with Turán [39], they conjectured:

Conjecture 1.2 (Erdős and Turán [39]). *If $A \subseteq \mathbb{N}$ has positive upper density—that is, $\limsup_{N \to \infty} \frac{|A \cap \{1, \ldots, N\}|}{N} > 0$—then A contains infinitely many k-term arithmetic progressions, for every k.*

While the primes do not have positive upper density—indeed the Prime Number Theorem states that their density up to N is $\sim 1/\log N$, so their upper density is

---

[1]Regarding the date of this conjecture, the earliest written reference I could find for the case of 3-term arithmetic progressions was a 1973 seminar report [33, Conjecture 1.2], and for general arithmetic progressions was a talk from 1976 [34]. This conjecture is often attributed to Erdős and Turán's 1936 paper [39], but the conjecture does not appear there in print—even as a question— and in his later writings, including a touching tribute to Turán [34, p. 40], although Erdős raises the problem in connection with his work on the primes with Turán, he refers to it as an "old conjecture *of mine*" (emph. added). Soifer [103, p. 355] found references in a 1982 talk in which Erdős says it was more than 40 years old, and a 1986 talk in which Erdős said it was "about 30 years ago," so we conclude with Soifer that the conjecture was made sometime between the early 1940s and mid 1950s.

Regarding its motivation from arithmetic progressions in the primes, we have at least the following evidence. In [33], Erdős relates it with Goldbach's Conjecture—which, in particular, implies that for any prime p, $2p = p_1 + p_2$ for some other primes p_1, p_2, and hence that the primes have infinitely many 3-term arithmetic progressions—and Chowla's unconditional result [18] that there are infinitely many 3-term arithmetic progressions in the primes. (The earlier paper of van der Corput proving the same [116] was apparently forgotten until later.) In his 1977 paper [34] he writes down the conjecture as stated here, that is, for arbitrarily long arithmetic progressions, and points out that, in particular, it would imply arbitrarily long arithmetic progressions in the primes, and therefore resolving the conjecture should be quite hard. In 1981, he restates the conjecture yet again [35, p. 28], this time explicitly "in connection" with the problem of showing arbitrarily long arithmetic progressions in the primes.

[2]The only larger Erdős prizes I'm aware of are \$10,000 USD to show that $p_{n+1} - p_n$ is "large" infinitely often, and \$25,000 USD to show there are only finitely many consecutive pairs of primes p_n such that $p_n \leq \frac{p_{n+1} + p_{n-1}}{2}$, though he offered only \$100 USD for a disproof; see [69] for details on these large prizes and a list of other Erdős prizes.

zero—this conjecture turned out to be a crucial step towards proving that there are arbitrarily long arithmetic progressions in the primes. Szemerédi [107] proved the Erdős–Turán conjecture, and Szemerédi's proof was a crucial ingredient in Green & Tao's proof [56] that the primes contain arbitrarily long arithmetic progressions.

When encountering a difficult conjecture, two natural tactics are to consider special subcases or to consider analogous conjectures in slightly different settings. First, instead of arbitrarily long arithmetic progressions, let's ask for any nontrivial arithmetic progressions at all. By "nontrivial" we mean consisting of at least three distinct points. Okay, great: We'll focus on 3-term arithmetic progressions for a bit. Historically, this has indeed been a good place to start; for example, two decades before Szemerédi's Theorem was proved, in 1953 Roth [90] proved the analogous result for 3-term arithmetic progressions, by a significantly easier argument. (Szemerédi himself proved the $k = 4$ case in 1969 [106], before proving the general case.) In terms of arithmetic progressions in the primes, it's trivial to show existence of a 3-term arithmetic progression (3,5,7), and almost 70 years before Green and Tao, van der Corput [116] showed there were infinitely many 3-term arithmetic progressions in the primes. But even the existence of infinitely many 4-term arithmetic progressions in the primes remained open until Green & Tao's result. Moreover, while 3-term arithmetic progressions have a nice formulation in terms of convolutions of Fourier transforms, 4-term arithmetic progressions do not; a conundrum which eventually led to higher-order Fourier analysis (see, e. g., [119] for a nice discussion of this difficulty and how it was overcome). Okay, fine, 4 is a lot harder than 3, so let's stick with 3-term arithmetic progressions.

And now, instead of only considering subsets of $\mathbb{N}$ (or $\mathbb{Z}$), let's consider subsets of arbitrary abelian groups.[3] These are, arguably, the most natural setting in which the notion of "additive structure," makes sense, since these are precisely the sets which have a notion of addition.[4] As a start, it should be clear that if we have $A \subseteq \mathbb{N}$ and we consider $A \cap \{1, \ldots, N\}$, then from the point of view of arithmetic progressions, this is essentially equivalent to considering A as a subset of $\mathbb{Z}/N\mathbb{Z}$, the integers mod N. If we're considering 3-term arithmetic progressions, maybe we should instead consider A as a subset of $\mathbb{Z}/3N\mathbb{Z}$, just to make sure there's no accidental wrapping around, but philosophically, and even mathematically, this turns out to make little difference. And note that we can indeed rephrase, for example, Szemerédi's Theorem in terms of $A \cap \{1, \ldots, N\}$: For all $k \in \mathbb{N}, \varepsilon > 0$, there exists N_0 such that if $N > N_0$ and $|A \cap \{1, \ldots, N\}| > \varepsilon N$, then A contains a k-term arithmetic progression. Okay, great, so other abelian groups might be good models for the phenomena we're interested in, but considering only cyclic groups (the integer mod m) seems to really be considering the *same* phenomena, rather than phenomena in analogous settings.

At the "opposite end" of some sort of natural spectrum, we might consider n-dimensional vectors over the integers modulo m, for small m and $n \to \infty$, viz. $(\mathbb{Z}/2\mathbb{Z})^n, (\mathbb{Z}/3\mathbb{Z})^n, (\mathbb{Z}/4\mathbb{Z})^n$. To what extent can we use results about arithmetic

---

[3]If you don't know what an abelian group is, don't be scared! The integers $\mathbb{Z}$, the integers mod a number $\mathbb{Z}/N\mathbb{Z}$, and vectors of such $(\mathbb{Z}/m\mathbb{Z})^n$ are all abelian groups. In fact, these are almost the general case, so just keep these in mind as your examples and you should have smooth sailing.

[4]Okay, technically maybe we should consider abelian *semi*groups. But many semigroups can naturally be embedded into groups, and those which can't have an addition operation which differs quite substantially from our intuition for addition: for example, if $x + x + x = x + x$ but $x \neq 0$, then our addition operation seems to be somewhat far from our main interest, namely $\mathbb{N}$.

progressions in these sets, which look more like vector spaces (and indeed, when m is prime, *are* vector spaces), to understand arithmetic progressions in $\mathbb{N}$?

There are two answers to this question, one historical and one formal. Historically, it's been useful to first consider the vector spaces $(\mathbb{Z}/p\mathbb{Z})^n$ (p prime), where we have lots of substructures to play with and induct on. To transfer results from this setting to $\mathbb{Z}/N\mathbb{Z}$, one fruitful approach is to use so-called Bohr sets, which are kind of an "approximate subspace" (in some Fourier-analytic sense). For any given result, however, this transference remains somewhat of an art, but has been very effective in the past. The papers [52, 119] are entirely devoted to the analogy between vector spaces over finite fields and $\mathbb{Z}$.

Formally, there is also a tool for comparing the additive structure of subsets in one abelian group with subsets of another abelian group which I can't resist mentioning: Freiman homomorphisms [43] (for a textbook treatment, see, e. g., [114, Section 5.3]). For me, when I first learned of this notion, it helped clarify what we mean by "additive structure" in general; I hope it has the same effect for you. Let's define an *additive set* to be a pair (A, Z) where Z is an abelian group and $A \subseteq Z$, and let's say that it's "k-additive structure" is completely determined by all equalities of the form $a_1 + a_2 + \cdots + a_k = a_1' + a_2' + \cdots + a_k'$ where $a_i, a_i' \in A$. For example, the 2-additive structure captures any k-term arithmetic progressions: $a_1, a_2, a_3, \ldots, a_k$ form an arithmetic progression if and only if $a_{i+1} - a_i = a_{i+2} - a_{i+1}$ for all $i = 1, \ldots, k - 2$, which we can rewrite entirely additively as $a_i + a_{i+2} = a_{i+1} + a_{i+1}$.

Definition 1.3 (Freiman homomorphism)**.** A *Freiman k-homomorphism* between additive sets $(A, Z), (B, W)$ is a function $f \colon A \to B$ such that, for all $a_i, a_i' \in A$

$$a_1 + \cdots + a_k = a_1' + \cdots + a_k' \implies f(a_1) + \cdots + f(a_k) = f(a_1') + \cdots + f(a_k').$$

A *Freiman k-isomorphism* is a bijective Freiman k-homomorphism whose inverse is also a k-homomorphism; equivalently, we require f to be a bijection such that the one-way implication above becomes two-way: $a_1 + \cdots + a_k = a_1' + \cdots + a_k'$ if and only if $f(a_1) + \cdots + f(a_k) = f(a_1') + \cdots + f(a_k')$.

Although this is not how many of the results were transferred from finite fields to the integers, in principle this notion could let us transfer (some) results about arithmetic progressions in one abelian group, such as $(\mathbb{Z}/m\mathbb{Z})^n$, to another abelian group, such as $\mathbb{Z}$. At any rate, as I said, I think this notion helps clarify what we mean by "additive structure."

1.4. **The Cap Set Conjecture.** Now, for arithmetic progressions, $(\mathbb{Z}/2\mathbb{Z})^n$ isn't interesting, as it only has trivial arithmetic progressions (of length two): if x, y, z form an arithmetic progression, then $z = y + (y - x) = x$, because we're working modulo 2. So the simplest interesting "toy model" to consider in our family $(\mathbb{Z}/m\mathbb{Z})^n$ is $(\mathbb{Z}/3\mathbb{Z})^n$. By a similar argument, because $3 = 0$ in $(\mathbb{Z}/3\mathbb{Z})^n$, it has no arithmetic progressions of length larger than 3 (which are proper, that is, consisting of all distinct elements). But that's alright, because 3-term arithmetic progressions were where we wanted to start anyways. So $(\mathbb{Z}/3\mathbb{Z})^n$ has the simultaneous virtues of (1) being analogous to but not the same as our original question(s) in $\mathbb{N}$, (2) among such analogous structures, being the simplest one which is still interesting, and (3) forcing us to focus our attention on the smallest case, namely that of 3-term arithmetic progressions, without having to worry about any "higher" additive structure.

Before getting into the Cap Set Conjecture itself, let's return to history to motivate it from two other angles. The first angle is geometric. Harborth [60] introduced the function $s(m, n)$ to be the smallest number s such that any s distinct points in $\mathbb{Z}^n$ contains a subset of size m whose centroid also has integer coordinates. It is not hard to see that this is equivalent to the smallest s such that any sequence of s elements in $(\mathbb{Z}/m\mathbb{Z})^n$ contains a subsequence of length m whose sum is zero. For $m = 3$, note that this is nearly the same as the largest subset of $(\mathbb{Z}/3\mathbb{Z})^n$ which contains no 3-term arithmetic progression, since modulo 3 we have that x, y, z form an arithmetic progression if and only if $x + y = 2z$ if and only if $x + y + z = 0$ (the only difference is that $s(m, n)$ allows sequences with repeated elements). Alon and Dubiner [3, 4] were perhaps the first to raise this question for $m = 3$ explicitly, but their primary interest in those papers was the case of small n and large m, which is closer to our original motivation of $\mathbb{Z}/N\mathbb{Z}$ than to our new toy model of $(\mathbb{Z}/3\mathbb{Z})^n$. For general m, n, the current best bounds are still those due to Harborth [60]:

$$(m - 1)2^n + 1 \leq s(m, n) \leq (m - 1)m^n + 1.$$

(Note that the upper bound is larger than the size of $(\mathbb{Z}/m\mathbb{Z})^n$: The question is about sequences of elements which need not be distinct.) Alon and Dubiner [3, 4] asked whether there was some $c < 3$ such that $s(3, n) \leq c^n$.

Finally, a third motivation. When studying the design of statistical experiments, Bose in 1947 [12] was led to study subsets of the projective geometry $\mathbb{P}(\mathbb{F}_q^n)$ over a finite field $\mathbb{F}_q$ that contained no 3 collinear points, which are called *caps*. In particular, he was interested in the size of the largest caps. This question was taken up by Segre, who provided upper and lower bounds on the maximum size of a cap [98, 99], and has been well-studied since (e. g., a quick search for on MathSciNet reveals at least a dozen papers in the last five years alone). Over $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$, for whatever reasons caps are called "cap sets," and a cap set is essentially the same as a set of points in $(\mathbb{Z}/3\mathbb{Z})^n$ containing no 3-term arithmetic progression; indeed, the extremal sizes of these two objects have identical asymptotic behavior as $n \to \infty$.[5]

With these motivations in mind, let us recall what's known about 3-term arithmetic progressions in $\mathbb{Z}/N\mathbb{Z}$ and in $(\mathbb{Z}/3\mathbb{Z})^n$. For any abelian group Z, let $r_3(Z)$ be the size of the largest subset of Z without 3-term arithmetic progressions. If we think of 3-term arithmetic progressions as some of the simplest nontrivial additive structure a set could have, $r_3(Z)$ is an upper bound on the size of (mostly) "unstructured" subsets of Z.

Figure 2 shows the history. Since we want to compare $\mathbb{Z}/N\mathbb{Z}$ against $(\mathbb{Z}/3\mathbb{Z})^n$, we will always use N to denote $|Z|$. Lower bounds on $r_3(Z)$ come from constructions of sets without 3-term arithmetic progressions. In $\mathbb{Z}/N\mathbb{Z}$, the current best bound is due to Elkin [30], who constructed such a set of size $\geq N(\log N)^{1/4}/e^{c\sqrt{\log N}}$ for some $c > 0$ (improving on Behrend's classic bound [8] by $\sqrt{\log N}$; see [58] for a shorter, albeit less constructive, proof). In $(\mathbb{Z}/3\mathbb{Z})^n$, the current best bound is due to Edel [28], who constructed a cap set of size $\geq N^{0.724851\cdots}$. In case you're a little rusty on your asymptotics,[6] $N/e^{O(\sqrt{\log N})}$ grows faster than $N^{1-\varepsilon}$ for arbitrarily small $\varepsilon > 0$, but slower than $N/(\log N)^c$ for arbitrarily large c. In contrast, Edel's lower bound is of the form $N^{0.72\cdots}$.

———————

[5]If you aren't familiar with how to compare the asymptotic behavior of two functions $f, g \colon \mathbb{N} \to \mathbb{N}$, it's not hard, but now would be a good time to consult Appendix A.

[6]What's the sound of an analytic number theorist drowning? "Log log log log log log..."

It is thus natural to ask where the truth lies for $(\mathbb{Z}/3\mathbb{Z})^n$: Does $r_3((\mathbb{Z}/3\mathbb{Z})^n)$ grow more quickly than $(3^n)^{1-\varepsilon}$ for all $\varepsilon > 0$, or is there some $c < 3$ such that $r_3((\mathbb{Z}/3\mathbb{Z})^n) \leq c^n$? While this was raised as a question by several authors [3, 4, 17, 28], and it seems plausible that other experts may have believed the upper bound to be c^n for some $c < 3$, the earliest reference we can find in which someone explicitly expressed the belief that $r_3((\mathbb{Z}/3\mathbb{Z})^n) \leq c^n$ for some $c < 3$ is from 2004 [52]; other authors expressed the opposite belief [110].
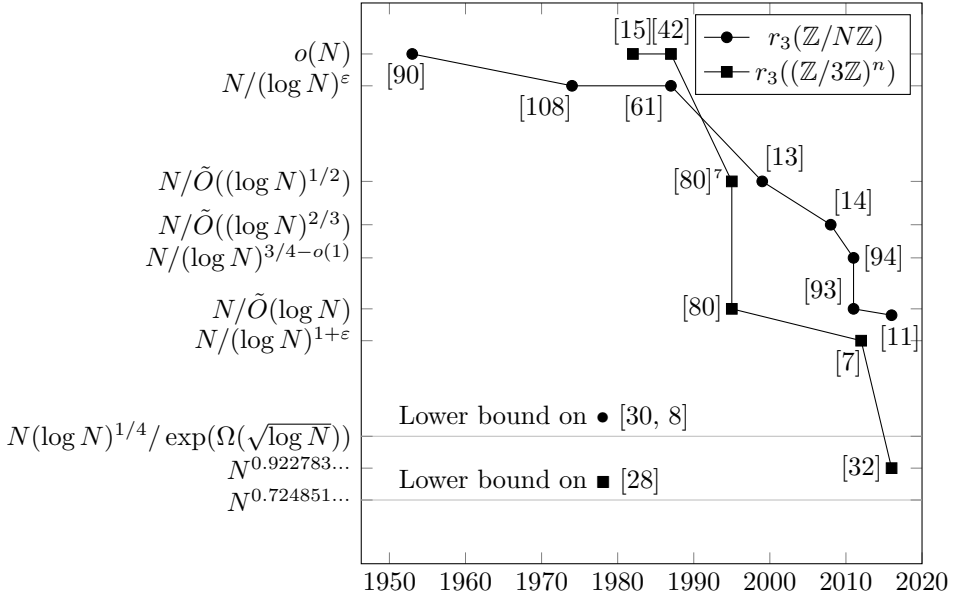

FIGURE 2.   History of upper bounds on $r_3(\mathbb{Z}/N\mathbb{Z})$ and $r_3((\mathbb{Z}/3\mathbb{Z})^n)$. In all cases, ε denotes some constant strictly between 0 and 1/2, but each use of ε denotes a different constant. The $\tilde{O}$ hides terms of the form $(\log \log N)^c$ or smaller. Note that Behrend's [8] and Elkin's [30] lower bounds on $r_3(\mathbb{Z}/N\mathbb{Z})$ grow faster than $N^{1-\varepsilon}$ for any $\varepsilon > 0$. If Bloom's $N(\log \log N)^4/\log N$ [11] could be improved to $N \log \log N/\log N$ (see, e. g., [50, p. 273]), it would imply a theorem of Green [53], known as "Roth's theorem in the primes."

Conjecture 1.4 (Cap Set Conjecture [52] (cf. [3, 4, 17, 28]). $r_3((\mathbb{Z}/3\mathbb{Z})^n) \leq c^n$ *for some $c < 3$.*

Before we come to its resolution, let's pause to discuss the lower bounds known on the cap set problem. Within $\mathbb{Z}/3\mathbb{Z}$, we have that the set $\{0, 1\}$ is a cap set; it follows that $\{0, 1\}^n$ is a cap set of size 2^n in $(\mathbb{Z}/3\mathbb{Z})^n$, for $x + y + z = 0$ in $(\mathbb{Z}/3\mathbb{Z})^n$ if and only if $x_i + y_i + z_i = 0$ for each coordinate $1 \leq i \leq n$. Similarly, if we have a cap set C of size s in $(\mathbb{Z}/3\mathbb{Z})^d$, then we get a cap set of size $s^{n/d}$ in $(\mathbb{Z}/3\mathbb{Z})^n$ (when d divides n) by partitioning the n coordinates into n/d groups of d, and considering the cap set $C^{n/d}$. Table 1 shows the bounds achieved using this idea.

Note that this technique always produces lower bounds of the form c^n for some $c < 3$. The only hope to disprove the conjecture this way would be to find an infinite

d	$r_3((\mathbb{Z}/3\mathbb{Z})^d) \geq$	$r_3((\mathbb{Z}/3\mathbb{Z})^n) \geq$
1	2	2^n
3	9 [12]	$9^{n/3} \approx 2.08^n$
4	20 [86]	$20^{n/4} \approx 2.11^n$
5	45 [27]	$45^{n/5} \approx 2.14^n$
6	112 [17, 29]	$112^{n/6} \approx 2.19^n$
62	$2(2 \cdot 112^{10} + 2 \cdot 10 \cdot 112^9 \cdot 12)$ [28]	$\approx 2.21^n$

TABLE 1.   Lower bounds on $r_3((\mathbb{Z}/3\mathbb{Z})^n)$. The bounds for $d \leq 5$ are tight. The current best bound, due to Edel [28], uses a recursive construction to build a cap set of size $\approx 2.57 \times 10^{21}$ in $(\mathbb{Z}/3\mathbb{Z})^{62}$ ($N \approx 3.18 \times 10^{29}$).

family of better and better such constructions, and given the level of complexity of Edel's construction, finding such an infinite family (of course, before we knew it was impossible) seemed like a tall order.

Given the lack of consensus on which way this conjecture should have been resolved, it was then quite a surprise to see it resolved (positively) in 2016. All the previous upper bounds on $r_3((\mathbb{Z}/3\mathbb{Z})^n)$ had used Fourier analytic techniques, and several people had speculated on ways to extend these techniques to get better bounds (e.g., [110, 49]). Then in 2016, Croot, Lev, and Pach [22] left Fourier analysis behind and introduced a beautiful new use of the polynomial method to show that $r_3((\mathbb{Z}/4\mathbb{Z})^n) \leq (4^n)^{0.926\cdots}$. Shortly thereafter, and nearly simultaneously with one another, Ellenberg and Gijswijt leveraged the Croot–Lev–Pach technique to give a positive resolution to the Cap Set Conjecture [32]. This exponentially small upper bound on $r_3((\mathbb{Z}/3\mathbb{Z})^n)$ was all the more surprising because it shows a striking asymptotic difference between arithmetic progressions in $\mathbb{Z}/N\mathbb{Z}$ and those in $(\mathbb{Z}/3\mathbb{Z})^n$. Moreover, while Fourier methods work in both settings, the Croot–Lev–Pach use of the polynomial method only yields trivial bounds for $\mathbb{Z}/N\mathbb{Z}$.

Such a strong upper bound is also tantalizing in its connection with the integers. Since the density of primes in $\{1, \ldots, N\}$ is $\sim 1/\log N$ (the Prime Number Theorem), if the upper bound on $r_3(\mathbb{Z}/N\mathbb{Z})$ could be improved from its current record of $N(\log \log N)^4/\log N$ [11] to, say, $N \log \log N/\log N$, it would immediately imply Roth's theorem in the primes (a theorem due to Green [53], with improved bounds by Helfgott and de Roton [62]). With the upper bound on cap sets being so small, it perhaps provides renewed hope that the upper bound on $r_3(\mathbb{Z}/N\mathbb{Z})$ could at least be improved this seemingly tiny amount to give a pure density proof of Roth's theorem in the primes.

Even for cap sets, there is still an exponential gap between the upper bound of 2.756^n [32] and the lower bound of 2.217^n [28], and closing this gap is an interesting problem. Improving the lower bound "just" requires a new construction; the smallest dimension in which a finite construction could get within .01 of 2.756 is $d = 42$.[8] We'll see in Section 3.2, through its connection with algorithms for

_____

[8]Calculated by assuming the bound [9, Theorem 4] is tight for relatively small values of d, namely that $r_3((\mathbb{Z}/3\mathbb{Z})^d) \leq \frac{d+1}{d^2} 3^d$. We know asymptotically that this bound isn't tight, since it was already beat out for sufficiently large d by Bateman and Katz, and now by Ellenberg and Gijswijt, but we haven't calculated exactly when those asymptotics kick in. If the Bierbrauer–Edel bound isn't tight for d up to 42, then the smallest useful dimension would only be larger.

matrix multiplication (of all things!), that the Croot–Lev–Pach–Ellenberg–Gijswijt technique extends from sets without arithmetic progressions to so-called "tricolored sum-free sets," and that in the tricolored setting the upper bound of $\approx 2.756^n$ is indeed tight [66]. Thus any attempt to improve the upper bound must differ substantially enough to *not* extend to the tricolored setting. Alternatively, to improve the lower bound one might try to turn the Kleinberg–Sawin–Speyer construction of a tricolored sum-free set [66] into an ordinary cap set of the same size, but this too seems difficult. See the end of Section 3 for a more detailed discussion of the difficulties in improving the upper bound.

2. Proof of the Cap Set Conjecture

The method introduced by Croot, Lev, and Pach [22], which ultimately led to the resolution of the Cap Set Conjecture [32], is an application of the polynomial method. In general, the polynomial method is the application of algebraic geometry to combinatorics (and other fields that, at first blush, seem unrelated to algebraic geometry). In particular, natural combinatorial structures can often be defined in terms of polynomial equations, and then by reasoning about these systems of polynomial equations (the domain of algebraic geometry) we can often learn about the original combinatorial structures. In addition to the Cap Set Conjecture, the polynomial method was also instrumental in the recent solution of another long-standing combinatorial problem, the finite field Kakeya problem [24] (see also [26] for nearly tight bounds, again using the polynomial method). For general introductions to the polynomial method see [63, Chapter 16] (in the context of extremal combinatorics), [114, Chapter 9] (in the context of additive combinatorics), and [111] for a recent tutorial and survey.

The version of the Ellenberg–Gijswijt proof we will follow here is the "symmetric" version due to Tao [112]. The idea is essentially the same as Ellenberg–Gijswijt, only the syntax is different. In the original proof [32], when studying solutions to the equation $x + y + z = 0$—which is clearly symmetric in all three of x, y, and z—they single out two of the variables and consider the rank of a matrix of the form $M_{x,y} = f(x + y)$ for some polynomial(s) f. What makes Tao's version more symmetric is that each of the three variables are put on an equal footing.

2.1. **Tensors.** Since we want to consider solutions to the equation $x+y+z = 0$ with $x, y, z \in (\mathbb{Z}/3\mathbb{Z})^n$, we will consider three-variable functions $F \colon X \times Y \times Z \to \mathbb{Z}/3\mathbb{Z}$, where $X, Y, Z \subseteq (\mathbb{Z}/3\mathbb{Z})^n$. Just as we could visualize a two-variable function on a finite domain $F \colon X \times Y \to \mathbb{Z}/3\mathbb{Z}$ as a $|X| \times |Y|$ matrix with entries from $\mathbb{Z}/3\mathbb{Z}$, when X, Y, Z are finite sets we may visualize a three-variable function as a three-dimensional array of numbers (see Figure 3)—sometimes called a 3-tensor—where the rows are indexed by the elements of X, the columns by the elements of Y, and the "depths" (the row-like thing, but in the third dimension) by the elements of Z:

We thus refer to $|X|$, $|Y|$, and $|Z|$ as the *side lengths* of the 3-tensor F.

The first observation is that if $A \subseteq (\mathbb{Z}/3\mathbb{Z})^n$ is a cap set, then the function

$$F(x, y, z) = \delta_0(x + y + z) = \begin{cases} 1 & x + y + z = 0 \\ 0 & \text{otherwise} \end{cases}$$

when restricted to $A \times A \times A$, looks like a three-dimensional version of the identity matrix. $F|_{A \times A \times A}(x, y, z) = 1$ if and only if $x = y = z$, and is zero otherwise. If
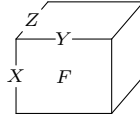
FIGURE 3.   Visualizing a three-variable function $F\colon X \times Y \times Z \to \mathbb{Z}/3\mathbb{Z}$ as a three-dimensional array or "3-tensor."

the only nonzero entries in a tensor F are those for which $x = y = z$, we call F a *diagonal* tensor (see Figure 4).
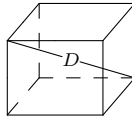

FIGURE 4.   Cap sets correspond to diagonal tensors; in this picture, the only nonzero entries are along the body diagonal D.

Continuing the analogy with matrices, what we would like is some notion of "rank" for 3-tensors such that

(1) diagonal tensors have rank equal to their side length (just as diagonal matrices do), but
(2) the function δ_0 (on all of $(\mathbb{Z}/3\mathbb{Z})^n \times (\mathbb{Z}/3\mathbb{Z})^n \times (\mathbb{Z}/3\mathbb{Z})^n$) has rank exponentially smaller than its side length 3^n, and
(3) the rank of a tensor is always at least the rank of any of its sub-tensors, gotten by restricting $F\colon X \times Y \times Z \to \mathbb{Z}/3\mathbb{Z}$ to $X' \times Y' \times Z'$ for $X' \subseteq X, Y' \subseteq Y, Z' \subseteq Z$.

Given such a notion of "rank," we quickly prove the Cap Set Conjecture:

Proof of the Cap Set Conjecture, assuming a notion of "rank" satisfying (1)–(3).
Let $F(x, y, z) = \delta_0(x + y + z)$ be the tensor above. Then for any cap set A:

$$\begin{aligned}
|A| &= \text{"rank"}(F|_{A \times A \times A}) &&\text{by property (1), since } F|_{A \times A \times A} \text{ is diagonal} \\
&\leq \text{"rank"}(F) &&\text{by property (3)} \\
&\leq c^n &&\text{for some } c < 3, \text{ by property (2).}
\end{aligned}$$

And that's it! $\square$

The following notion is a direct generalization of the rank of matrices, which we'll see has the desired properties. Showing that it satisfies properties (1)–(3) above will then turn the above "proof" of the Cap Set Conjecture into a Proof.

Definition 2.1 (Tao [112])**.** Given a 3-tensor $F\colon X \times Y \times Z \to \mathbb{Z}/3\mathbb{Z}$, its *slice rank* is the least r such that F can be written as the following sum of r terms:

$$F(x, y, z) = \sum_{i=1}^{a} f_i(x) g_i(y, z) + \sum_{i=a+1}^{b} f_i(y) g_i(x, z) + \sum_{i=b+1}^{r} f_i(z) g_i(x, y).$$

Any such expression for F, even with r not minimal, is called a *slice decomposition*.

This notion was introduced by Tao [112] in developing this symmetric version of the Ellenberg–Gijswijt proof, and following [10] we call it "slice rank." Further properties of slice rank were elaborated in [10, 113].

A 3-tensor of slice-rank one thus has the form $f(x)g(y,z)$ (or any of its symmetric versions, gotten by permuting the variables). In terms of our three-dimensional array visualization, we may think of $g(y,z)$ as a matrix placed on a horizontal slab coming out of the page, and then the function $f(x)g(y,z)$ consists of stacking up a bunch of scalar multiples of this slab on top of one another.

Property (3), that slice rank cannot increase when passing to sub-tensors, is the easiest to see:

Observation 2.2. *For any tensor* $F\colon X \times Y \times Z \to \mathbb{Z}/3\mathbb{Z}$, *and any* $X' \subseteq X, Y' \subseteq Y, Z' \subseteq Z$:
$$\text{slice-rank}(F|_{X' \times Y' \times Z'}) \leq \text{slice-rank}(F).$$

Proof. Given a slice decomposition for F with functions $f_i(\bullet), g_i(\bullet, \bullet)$, we get a slice decomposition for $F|_{X' \times Y' \times Z'}$ by restricting each f_i and g_i to the primed subsets. That is, for $f_i\colon X \to \mathbb{Z}/3\mathbb{Z}$, we restrict it to $f_i' = f_i|_{X'}\colon X' \to \mathbb{Z}/3\mathbb{Z}$, and restrict $g_i\colon Y \times Z \to \mathbb{Z}/3\mathbb{Z}$ to $g_i|_{Y' \times Z'}\colon Y' \times Z' \to \mathbb{Z}/3\mathbb{Z}$. □

Property (1), that the slice rank of diagonal tensors is maximal, can be shown by induction from the 2-variable (i. e., 2-tensor, i. e., matrix) case, which we leave as an exercise:

Lemma 2.3 (Tao [112])**.** *The slice rank of a diagonal tensor is equal to its number of non-zero entries.*

Finally, Property (2), the exponential upper bound, brings us to the key idea of the proof, which is an application of the polynomial method.

2.2. **Key idea of the proof.** The following lemma is the **key idea** from Croot–Lev–Pach and Ellenberg–Gijswijt, which unlocks the whole proof. Since a polynomial $F(\vec{x}, \vec{y}, \vec{z})$ on $((\mathbb{Z}/3\mathbb{Z})^n)^3$, is, in particular, a function $F\colon (\mathbb{Z}/3\mathbb{Z})^n \times (\mathbb{Z}/3\mathbb{Z})^n \times (\mathbb{Z}/3\mathbb{Z})^n \to \mathbb{Z}/3\mathbb{Z}$, we may view it as a 3-tensor of side length 3^n. In the language of slice rank we have:

Lemma 2.4 (Croot–Lev–Pach [22], slightly generalized by Ellenberg–Gijswijt [32])**.** *If* $F(x_1, \ldots, x_n, y_1, \ldots, y_n, z_1, \ldots, z_n)$ *is a polynomial over* $\mathbb{Z}/3\mathbb{Z}$ *of degree* $\leq d$, *then, when viewed as a 3-tensor of side length* 3^n *as above, we have:*
$$\text{slice-rank}(F) \leq 3 \sum_{\substack{a,b,c \geq 0 \\ a+b+c=n \\ b+2c \leq d/3}} \frac{n!}{a!b!c!}.$$

Proof. Each monomial m in F has the form $m = x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n} y_1^{e_1'} \cdots y_n^{e_n'} z_1^{e_1''} \cdots z_n^{e_n''}$, where the exponents satisfy $\sum_i (e_i + e_i' + e_i'') \leq d$. If we consider the degrees of this monomial in the x's, the y's, and the z's separately, namely $d_x(m) := \sum_i e_i$, $d_y(m) = \sum_i e_i'$, $d_z(m) = \sum_i e_i''$, then for each monomial m, at least one of $d_x(m)$, $d_y(m)$, and $d_z(m)$ must be $\leq d/3$.

Now, let M_x be the set of monomials for which $d_x(m) \leq d/3$, let M_y be the set of monomials for which $d_y(m) \leq d/3$, and define M_z similarly. Although not strictly necessary, it will make things simpler if M_x, M_y, M_z are disjoint, so let's

remove from M_y anything in M_x, and then remove from M_z anything in M_x or in M_y. Then we can write F as:

$$F(x, y, z) = \sum_{m \in M_x} m + \sum_{m \in M_y} m + \sum_{m \in M_z} m.$$

The key trick here is to rewrite each of these three sums by factoring out the relevant variables, viz. factor out the x variables as much as possible from $\sum_{m \in M_x} m$:

$$\sum_{m \in M_x} m = \sum_{\substack{e_1, \ldots, e_n \geq 0 \\ \sum_i e_i \leq d/3}} x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n} \overline{g}_{e_1, \ldots, e_n}(\vec{y}, \vec{z})$$

where the $\overline{g}_{e_1, \ldots, e_n}$ are precisely what they need to be to make this equality hold; but the only fact we need about the $\overline{g}$'s is that they only depend on $\vec{y}, \vec{z}$, and not on $\vec{x}$, as then the right-hand side here is a slice decomposition of the left-hand side. (The $\overline{g}$'s here are overlined because we're about to replace them.)

The next thing to note is that we can also restrict the e_i so that they are all at most 2. For any $\alpha \in \mathbb{Z}/3\mathbb{Z}$, note that $\alpha^3 = \alpha$, and thus the polynomial x^3, as a function on $\mathbb{Z}/3\mathbb{Z}$, computes the same function as the polynomial x. This lets us reduce the degree of each variable in each monomial until it is strictly less than 3. We are then left with:

$$\sum_{m \in M_x} m = \sum_{\substack{e_1, \ldots, e_n \in \{0,1,2\} \\ \sum_i e_i \leq d/3}} x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n} g_{e_1, \ldots, e_n}(\vec{y}, \vec{z})$$

The g's here may be combinations of some of the $\overline{g}$'s from before, but again, all we care about is that they do not depend on $\vec{x}$. Thus

$$\text{slice-rank}\left(\sum_{m \in M_x} m\right) \leq |\{(e_1, \ldots, e_n) \in \{0,1,2\}^n : \sum_i e_i \leq d/3\}|.$$

By swapping the role of $\vec{x}$, $\vec{y}$, and $\vec{z}$, we get the same bound on $\sum_{m \in M_y} m$ and $\sum_{m \in M_z} m$, and thus have

(2.1) $$\text{slice-rank}(F) \leq 3|\{(e_1, \ldots, e_n) \in \{0,1,2\}^n : \sum_i e_i \leq d/3\}|.$$

All that remains is to show that the set on the right-hand side has the size claimed in the statement of the lemma. Given $(e_1, \ldots, e_n) \in \{0,1,2\}^n$, let a be the number of e_i that are 0, b be the number of e_i that are 1, and c be the number of e_i that are 2. Then

(2.2) $$a + b + c = n,$$

for each e_i takes exactly one of these three values. Also, we have that $\sum_i e_i = a \cdot 0 + b \cdot 1 + c \cdot 2$, so $\sum_i e_i \leq d/3$ if and only if

(2.3) $$b + 2c \leq d/3.$$

Thus we can rewrite (2.1) as:

(2.4) $$\text{slice-rank}(F) \leq 3 \sum_{\substack{a,b,c \geq 0 \\ a+b+c=n \\ b+2c \leq d/3}} |\{(e_1, \ldots, e_n) \text{ with } a \text{ 0s}, b \text{ 1s, and } c \text{ 2s}\}|.$$

Finally, given values of a, b, c satisfying the constraints (2.2) and (2.3), we need to know how many vectors $(e_1, \ldots, e_n)$ have a 0s, b 1s, and c 2s. Every such vector comes from permuting the coordinates of the vector $(0, 0, \ldots, 0, 1, 1, \ldots, 1, 2, 2, \ldots, 2)$ (with a 0s, b 1s, and c 2s). There are $n!$ such permutations. However, this is significantly overcounting, since if we permute only those coordinates with the same value, we get back the same vector. Thus we have overcounted by a factor of $a!b!c!$, so our final count is:

$$(2.5) \qquad |\{(e_1, \ldots, e_n) \in \{0, 1, 2\}^n : \text{ there are } a \text{ 0s, } b \text{ 1s, and } c \text{ 2s}\}| = \frac{n!}{a!b!c!}.$$

Combining (2.4) with (2.5) yields the lemma. $\qquad\qquad\qquad\qquad\square$

2.3. **Finishing it off.** We now use the key Lemma 2.4 to prove Property (2) for slice rank, which will thus complete the proof of the Cap Set Conjecture.

Observation 2.5. *Let $F_0 \colon (\mathbb{Z}/3\mathbb{Z})^n \to \mathbb{Z}/3\mathbb{Z}$ be any function. Then the 3-tensor $F \colon (\mathbb{Z}/3\mathbb{Z})^n \times (\mathbb{Z}/3\mathbb{Z})^n \times (\mathbb{Z}/3\mathbb{Z})^n \to \mathbb{Z}/3\mathbb{Z}$ defined by $F(x, y, z) = F_0(x + y + z)$ has degree at most $2n$.*

Proof. The idea is to use interpolation to write the function F_0 as a polynomial. More formally, for $\alpha \in \mathbb{Z}/3\mathbb{Z}$, we can write the indicator function $\delta_\alpha(x)$, which is 1 if and only if $x = \alpha$, and 0 otherwise, as

$$\delta_\alpha(x) = 1 - (x - \alpha)^2.$$

Then for any $\vec{\alpha} \in (\mathbb{Z}/3\mathbb{Z})^n$, we can write the indicator function $\delta_{\vec{\alpha}}(\vec{x})$ as

$$\delta_{\vec{\alpha}}(\vec{x}) = \prod_{i=1}^n \delta_\alpha(x) = \prod_{i=1}^n \left(1 - (x - \alpha)^2\right).$$

Note that $\deg \delta_{\vec{\alpha}} \leq 2n$. Thus we can write any function F_0 as a polynomial of degree at most $2n$:

$$F_0(\vec{x}) = \sum_{\vec{\alpha} \in (\mathbb{Z}/3\mathbb{Z})^n} \delta_{\vec{\alpha}}(\vec{x}) F_0(\vec{\alpha}),$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

In particular, since the 3-tensor we care about, $F(x, y, z) = \delta_0(x + y + z)$, has the form in the preceding observation, we may apply the key Lemma 2.4 to a function of degree $\leq 2n$. And now, we come to the crucial estimate.

Lemma 2.6 (Key numerical estimate)**.**

$$\sum_{\substack{a,b,c \geq 0 \\ a+b+c=n \\ b+2c \leq 2n/3}} \frac{n!}{a!b!c!} \leq 2.756\ldots^{n(1+o(1))}.$$

Proof. Stirling's Formula says that, for large n, $n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$. Thus, if $a+b+c = n$, then

$$
\begin{aligned}
\frac{n!}{a!b!c!} &\sim \left(\frac{n}{e}\right)^n \left(\frac{e}{a}\right)^a \left(\frac{e}{b}\right)^b \left(\frac{e}{c}\right)^c \sqrt{\frac{2\pi n}{8\pi^3 abc}} \\
&= \frac{n^n}{a^a b^b c^c} \sqrt{\frac{n}{4\pi^2 abc}} \\
&\sim \frac{1}{2\pi} \left(\frac{n^n}{a^a b^b c^c}\right)^{1+o(1)}
\end{aligned}
$$

(If you're not familiar with the $o(1)$ notation in the exponent, now might be a good time to quickly check Appendix A.)

Now, since $a + b + c = n$, and we want to consider the behavior as $n \to \infty$, it's useful to rescale these three to get a probability distribution that essentially doesn't depend on n: For $\alpha = a/n, \beta = b/n, \gamma = c/n$, we have $\alpha, \beta, \gamma \geq 0$ and $\alpha + \beta + \gamma = 1$. With this notation, we then have

$$
\begin{aligned}
\frac{n!}{a!b!c!} &\sim \frac{1}{2\pi} \left(\frac{1}{\alpha^a \beta^b \gamma^c}\right)^{1+o(1)} \\
&= \frac{1}{2\pi} \left(\exp\left(-(a\log\alpha + b\log\beta + c\log\gamma)\right)^{1+o(1)}\right. \\
&= \frac{1}{2\pi} \left(\exp\left(-n(\alpha\log\alpha + \beta\log\beta + \gamma\log\gamma)\right)(1 + o(1))\right. \\
&= \frac{1}{2\pi} \exp\left(nh(\alpha,\beta,\gamma)(1+o(1))\right)
\end{aligned}
$$

where $h(\alpha, \beta, \gamma)$ can be recognized as the Shannon entropy of a probability distribution.

Our sum can thus be rewritten asymptotically as

$$
\sim \frac{1}{2\pi} \sum_{\substack{a,b,c \geq 0 \\ a+b+c=n \\ b+2c \leq 2n/3}} \exp\left(nh(a/n, b/n, c/n)(1 + o(1))\right).
$$

Since this is a sum of exponentials, and the sum only has $O(n^2)$ terms, as n gets large this will be dominated by the single largest exponential: For if the largest is e^{Cn} and the next largest is $e^{C'n}$, with $C' < C$, then even if all the remaining terms had magnitude $e^{C'n}$, they would still only add up to $O(n^2 e^{C'n}) \leq e^{C'n(1+o(1))}$, which is still $o(e^{Cn})$.

To find the single largest term, we need to maximize the entropy $h(\alpha, \beta, \gamma)$ subject to the constraints that α, β, γ form a probability distribution satisfying $\beta + 2\gamma \leq 2/3$. A routine Lagrange multiplier calculation will find the exact values of α, β, γ (see [112]), but since entropy is convex, simple numerical hill-climbing will also yield the correct maximum value. The largest value of $h(\alpha, \beta, \gamma)$ subject to these constraints is ≈ 1.013455, resulting in an upper bound of $\sim \exp(1.013455\ldots n(1 + o(1))) = 2.756\ldots^{n(1+o(1))}$. □

This completes the proof of the Cap Set Conjecture. Here are the bread crumbs of the proof (this how I remember it):

(1) Observe that a cap set corresponds to a diagonal 3-tensor, which is a sub-tensor of $F(x, y, z)$, the indicator function of $x + y + z = 0$. Since diagonal 3-tensors are like diagonal matrices, the size of the cap set is the slice rank of the diagonal 3-subtensor, which is therefore upper bounded by the slice rank of F itself.

(2) Write the indicator function F as a polynomial. Note that it has degree $\leq 2n$ (it's a product of n indicator functions, each of which has degree 2).

(3) Pigeonhole the monomials by degree, $d = 2n$ pigeons into 3 holes (one for each of x, y, z).

(4) Group together the monomials which have x-degree $\leq d/3$, (resp., y-degree, resp., z-degree)

(5) Count monomials and estimate the growth rate using entropy maximization.

3. Tricolored sum-free sets and the question of tight bounds

For both $r_3(\mathbb{Z}/N\mathbb{Z})$ and $r_3((\mathbb{Z}/3\mathbb{Z})^n)$ there is still a gap between the best upper and lower bounds known: In the former case, between $N/(\log N)^{1+\varepsilon}$ [11] and $N/e^{c\sqrt{\log N}-(1/4)\log\log N}$ [30, 8], and in the latter between 2.756^n [32] and 2.21^n [28]. In both cases, however, if we broaden our scope to slightly more general objects, we find essentially tight bounds.

3.1. **Tight bounds in the integers?** An arithmetic progression is a sequence of integers satisfying $x + z = 2y$. In the case of the integers (or $\mathbb{Z}/N\mathbb{Z}$), if we generalize to other translation-invariant linear equations—that is, of the form $\sum_{i=1}^{k} a_i x_i = 0$ where $\sum_i a_i = 0$—we find essentially tight bounds. Translation-invariance is a natural condition here, as we want to consider subsets of $\mathbb{Z}$ satisfying some condition on their relative differences $x_i - x_j$, which are unaffected by translating the entire subset by an additive constant.

Just last year Schoen and Sisask [97] (following [96], who showed the $k = 6$ case) showed that any subset $A \subseteq \{1, \ldots, N\}$ of size $\geq N/\exp(c(\log N)^{1/7})$ contains distinct elements $x_1, \ldots, x_4$ such that

$$x_1 + x_2 + x_3 = 3x_4.$$

Behrend's construction for arithmetic progressions adapts easily to this setting, resulting in essentially tight bounds for the preceding linear equation: The difference is only between $1/2$ and $1/7$ in the exponent of the exponent (sic!).

In fact, their argument works for any translation-invariant with at least four terms. As $x + y = 2z$ is of this form with only three terms, four here cannot be improved without showing that Behrend's construction is essentially tight for $r_3(\mathbb{Z}/N\mathbb{Z})$. It's worth noting, given our discussion above, that their techniques also give similar bounds for such equations over finite fields (but over finite fields no construction as large as Behrend's is known for $k = 4$, even though such a construction is not ruled out by the Cap Set Conjecture). In fact, they first present the argument over finite fields as it's simpler, and then use Bohr sets to extend the argument to the integers, as discussed above.

This result has the following interesting implication. Either their result extends to the $k = 3$ case, in which case Behrend's bound is essentially tight for $r_3(\mathbb{Z}/N\mathbb{Z})$, or showing a significantly better upper bound—say, $N/(\log N)^c$ for some $c > 1$, which would show that Bloom's upper bound is nearly tight—must use techniques

that are sensitive to the difference between 3-term linear equations and 4-term linear equations. We note that, although the difference here is between 3-term and 4-term linear equations, this situation actually seems quite different than the syntactically-similar difference between 3-term and 4-term arithmetic progressions. In particular, since 3-term arithmetic progressions can be captured by a single equation, they are relatively easy to analyze using Fourier analysis; since 4-term arithmetic progressions require two equations ($x_1 + x_3 = 2x_2$ and $x_2 + x_4 = 2x_3$), there is no single Fourier expression that captures them, thus necessitating the "higher-order Fourier analysis" suggested by Gowers [47, 48] and developed by Green, Tao, and Ziegler [57]. In contrast, for both the $k = 3$ and $k = 4$ cases considered in this section, they are still just a single equation and so are—at least in principle—amenable to standard Fourier-analytic techniques. The difference, if any, between 3-term and 4-term linear equations is apparently more subtle.

3.2. **Tight bounds in vector spaces over finite fields?** In the case of $(\mathbb{Z}/3\mathbb{Z})^n$, a slightly different generalization yields 2.756^n as a tight bound, matching the Ellenberg–Gijswijt upper bound. This generalization was motivated by algorithms for matrix multiplication (see Section 4.2), but our starting point here will be a simple observation about slice rank, which generalizes the fact that the rank of a matrix is invariant under change of basis. Given a tensor $F\colon X \times Y \times Z \to \mathbb{F}$ and an invertible $|X| \times |X|$ matrix S, $|Y| \times |Y|$ matrix T, and $|Z| \times |Z|$ matrix U, we may use these as change of basis matrices on $\mathbb{F}^X$, $\mathbb{F}^Y$, and $\mathbb{F}^Z$, resulting in a new tensor $F'(x', y', z') = \sum_{x,y,z} S(x, x')T(y, y')U(z, z')F(x, y, z)$.

Observation 3.1. *If $F, F'\colon X \times Y \times Z \to \mathbb{F}$ are two 3-tensors that differ by a change of basis (as above), then* slice-rank$(F) =$ slice-rank(F').

Proof. Suppose slice-rank$(F) = r$; then there is a slice decomposition

$$F(x, y, z) = \sum_{i=1}^{a} f_i(x)g_i(y, z) + \sum_{i=a+1}^{b} f_i(y)g_i(x, z) + \sum_{i=b+1}^{r} f_i(z)g_i(x, y).$$

Given a change of basis (S, T, U), let us apply it to the preceding decomposition. Now, for simplicity, let's just consider the first summation $\sum_i f_i(x)g_i(y, z)$ of the slice decomposition of F, and how it appears in the expression for F'. The other two summations will be handled similarly (one advantage of the symmetry of the notion of slice rank). We have

$$\sum_{x,y,z} S(x, x')T(y, y')U(z, z') \sum_{i=1}^{a} f_i(x)g_i(y, z)$$

$$= \sum_{i=1}^{a} \sum_{x} S(x, x')f_i(x) \sum_{y,z} T(y, y')U(z, z')g_i(y, z)$$

$$= \sum_{i=1}^{a} f_i'(x')g_i'(y', z'),$$

where $f_i'(x') = \sum_x S(x, x')f_i(x)$, which indeed only depends on x', since we sum over all values of x. Similarly, $g_i'(y', z') = \sum_{y,z} T(y, y')U(z, z')g_i(y, z)$ only depends on y', z'. Thus our first sum can be written after the change of basis using exactly as many slice-rank-one terms as before. Similarly for the second and third sums. Thus slice-rank$(F) =$ slice-rank(F'). $\qquad\square$

Although this may seem a rather trivial consequence of the definition, note that when we considered a cap set $A \subseteq (\mathbb{Z}/3\mathbb{Z})^n$, it at least *felt* important that we were using values from the *same* set A for all three variables x, y, z. But the above change-of-basis observation says that we can change basis in X independent from Y independent from Z. In particular, if we change bases using permutation matrices, this corresponds to simply re-ordering the elements of $(\mathbb{Z}/3\mathbb{Z})^n$ in each of X, Y, Z. What does a diagonal sub-tensor of our favorite tensor, $F(x, y, z) = \delta_0(x + y + z)$, look like after permuting basis elements? It's a restriction of F to $A \times B \times C$ with $A, B, C \subseteq (\mathbb{Z}/3\mathbb{Z})^n$ such that $F(a_i, b_j, c_k) = 1$ if and only if $i = j = k$, where $A = \{a_1, \ldots, a_{|A|}\}$ and similarly for B, C. This leads to the following notion:

Definition 3.2 (Tricolored sum-free set [5, 10]). A *tricolored sum-free set* in an abelian group Z consists of three subsets $(a_1, \ldots, a_k), (b_1, \ldots, b_k), (c_1, \ldots, c_k) \in Z^k$ such that
$$(\forall i, j, k)[a_i + b_j + c_k = 0 \Longleftrightarrow i = j = k].$$
(The indexing is only relevant for identifying the matching between A, B, and C.)

Cap sets are examples of tricolored sum-free sets, but they are far from the only ones.

From our observations above, we thus have:

Lemma 3.3. *In any abelian group Z, the slice-rank of $\delta_0(x + y + z)$ $(x, y, z \in Z)$ is an upper bound on the size of any tricolored sum-free set in Z.*

Proof. From the discussion above, tricolored sum-free sets yield diagonal subtensors in some basis. Apply Observation 3.1 and Lemma 2.3. $\square$

Finally, we see that for tricolored sum-free sets, the Ellenberg–Gijswijt bound is essentially exactly tight:

Theorem 3.4 (Kleinberg–Sawin–Speyer [66], with a lemma from Norin [84] and Pebody [85]). *Let θ be the base of the exponent in the Ellenberg–Gijswijt bound, that is, $r_3(\mathbb{Z}/N\mathbb{Z}) \leq \theta^{n(1+o(1))}$, $\theta \approx 2.756$. There is a tricolored sum-free set in $(\mathbb{Z}/3\mathbb{Z})^n$ of size $\geq \theta^{n(1-o(1))}$.*

To me, one of the really cool things here is not just that they achieved a tight bound, but the method of proof: They use a "pull-back" of Behrend's construction in the integers! Namely, they essentially choose three random mappings $h_1, h_2, h_3 \colon \mathbb{Z}^n \to \mathbb{Z}/p\mathbb{Z}$ for a prime $p \sim \exp(cn)$ for some c, use the Behrend (/Elkin) construction to get a large set $S \subseteq \mathbb{Z}/p\mathbb{Z}$ without arithmetic progressions, and then build their tricolored sum-free set as a large subset of $\{(a, b, c) \in (\mathbb{Z}/3\mathbb{Z})^n : h_1(a) = h_2(b) = h_3(c) \in S\}$. The $o(1)$ in the exponent of the preceding statement hides a factor which is nearly exactly the density of the Behrend/Elkin construction relative to the prime p.

Finally, let us return to the question of tight bounds on $r_3((\mathbb{Z}/3\mathbb{Z})^n)$. Of course, we may take Theorem 3.4 as some indication that the Ellenberg–Gijswijt bound is already tight for cap sets. At the end of Section 1.4 we began discussing what is needed to improve the lower bound. But with a little bit about the proof of Theorem 3.4, we may take some inspiration: Perhaps by taking pull-backs of Behrend's construction, we can indeed get an infinite family of better and better cap sets in $(\mathbb{Z}/3\mathbb{Z})^d$ for $d \to \infty$ that would meet the Ellenberg–Gijswijt bound.

To improve the upper bound, the key barrier to be avoided (at the moment) is the use of slice rank itself. (This is not particular to Tao's symmetric formulation; rather, any technique, such as Croot–Lev–Pach or Ellenberg–Gijswijt, which yields a slice rank upper bound falls prey to this limitation.) For slice rank upper bounds the size of tricolored sum-free sets, and Theorem 3.4 says that these bounds cannot be further improved. Thus, to improve the upper bound what is needed is a method that is somehow sensitive to the difference between a tricolored sum-free set and a cap set, or equivalently, between a diagonal tensor in arbitrary bases versus a diagonal tensor in three identical bases. Said another way, one needs a property of 3-tensors that is invariant under change of bases of the form (S, S, S), but *not* invariant under change of bases of the form (S, T, U).

4. Applications and extensions

Next we come to the question of the relationship between the Cap Set Conjecture and other problems, or even other areas of mathematics. In this section we'll cover several applications of the Croot–Lev–Pach polynomial method, as well as extensions of the Cap Set Conjecture motivated by other questions. Since Croot, Lev, and Pach first posted their preprint, progress on these applications happened very rapidly, and there are more than we can possibly cover in this short space. I will cover the ones with which I am most familiar. Here are a few I know of that are left out (I cannot hope to be exhaustive): relations between polynomials, viz. a polynomial Sárközy's Theorem [55], sum-sets as unions of sum-sets of subsets [31], subsets containing no right angles [45], and ordered tricolored sum-free sets [65]. As our purpose here is just to highlight a few of the many connections the Cap Set Conjecture has with other areas of mathematics, we won't be quite as expository in this section as we've been so far, but will point the reader to the relevant literature for further details. We will, however, give at least some motivation for each of the problems considered.

4.1. **Sunflowers.** If you thought addition and lines were pretty basic mathematical objects, let's leave them behind for a moment to get even more basic: we'll just consider sets and their intersections. A *sunflower* is a collection of sets $A_1, \ldots, A_k$ such that their pairwise intersections are the same as their k-wise intersection: $A_i \cap A_j = A_1 \cap A_2 \cap \cdots \cap A_k$ for all $i \neq j$. (If you draw the Venn diagram of such a collection of sets, you'll see where the name comes from.) This notion was introduced by Erdős and Rado [37] in 1960 as a generalization of Dirichlet's box argument, and has since found many uses in combinatorics, number theory, and computer science (see the introduction to [5] for many excellent references).

Dirichlet's box argument says that for any finite list $x_0, \ldots, x_{a^2}$ of $a^2 + 1$ elements, there is a sublist of size $(a + 1)$ such that either all the elements of the sublist are equal, or all are distinct from one another. This should sound a little familiar, as $x + y + z = 0$ in $\mathbb{Z}/3\mathbb{Z}$ if and only if $x = y = z$ or x, y, z are all distinct from one another. We'll see that the connection between sunflowers and cap sets is very tight indeed.

Theorem 4.1 (Erdős and Rado [37])**.** *Let $\mathcal{F}$ be a family of sets each of size s. If $|\mathcal{F}| \geq (k - 1)^s s!$ then $\mathcal{F}$ contains a k-sunflower.*

Conjecture 4.2 (The Sunflower Conjecture [37])**.** *For every $k > 0$, there is a constant c_k such that "$(k - 1)^s s!$" in the above theorem can be replaced by "c_k^s."*

This conjecture itself has also had many applications in extremal graph theory, the construction of Ramsey graphs, and circuit complexity (again, see [5] for references).

A slight variant of the Sunflower Conjecture, which will bring us even closer to cap sets, is:

Conjecture 4.3 (The Erdős–Szemerédi Sunflower Conjecture [38]). *There is a constant $c < 2$ such that any family $\mathcal{F}$ of subsets of $[n] = \{1, \ldots, n\}$ of size $|\mathcal{F}| \geq c^n$ contains a 3-sunflower.*

The difference between this conjecture and the preceding one is that this one doesn't require every set in $\mathcal{F}$ to have the same size, it *does* depend on the size of the ambient set from which $\mathcal{F}$ is built, and it only posits the existence of a 3-sunflower (instead of k-sunflowers for arbitrary k).

In connection with the complexity of matrix multiplication (see Section 4.2), Alon, Shpilka, and Umans [5] studied the Sunflower Conjecture and several of its variants, showing implications and equivalences between them. They introduced the following notion, which draws out the connection with cap sets.

Definition 4.4 (Sunflowers in $(\mathbb{Z}/m\mathbb{Z})^n$ [5, Definition 2.7]). A *k-sunflower in* $(\mathbb{Z}/m\mathbb{Z})^n$ is a collection of k vectors $v_1, \ldots, v_k \in (\mathbb{Z}/m\mathbb{Z})^n$ such that for every coordinate $i \in [n]$, either all the v_j have the same value in their i-th coordinate or these values are all distinct. Equivalently, for each i, $|\{(v_1)_i, (v_2)_i, \ldots, (v_k)_i\}|$ must be either 1 or k.

This is equivalent to a k-sunflower of sets (the usual notion) if the ambient set is partitioned into n pairwise disjoint blocks of size m, and every set in $\mathcal{F}$ contains exactly one element from each block.

Observation 4.5. *A 3-sunflower in $(\mathbb{Z}/3\mathbb{Z})^n$ is the same as a cap set.*

Conjecture 4.6 (Sunflower conjecture in $(\mathbb{Z}/m\mathbb{Z})^n$ [5, Conjecture 2.8]). *For every k, there is a constant b_k such that for all m, n, any set of $\geq b_k^n$ vectors in $(\mathbb{Z}/m\mathbb{Z})^n$ contains k vectors forming a k-sunflower.*

While this sounds different from the original Sunflower Conjecture, Alon, Shpilka, and Umans showed that the two are actually equivalent [5, Theorem 2.9].

By Observation 4.5, the Cap Set Conjecture thus resolves the $k = 3$ case of a weak form of Conjecture 4.6, in which we also restrict m to be 3. We'll see in the next section that the same method used to resolve the Cap Set Conjecture can be applied to resolve the full Sunflower Conjecture.

4.2. **Algorithms for matrix multiplication, and tricolored sum-free sets in other abelian groups.** Multiplying matrices—and its computationally equivalent sibling, solving linear systems of equations—is a fundamental linear algebra primitive used throughout the algorithmic world. Understanding its complexity is a central question in algebraic complexity theory that has led to new insights and conjectures in the representation theory of finite groups and algebraic geometry (see, e.g., Landsberg [71] and references therein).

The naive method of multiplying two $n \times n$ matrices takes $O(n^3)$ steps, which was thought to be optimal until Strassen showed [105] that this could be done in only $O(n^{2.81\cdots})$ steps. This led to the introduction of the *exponent ω of matrix*

multiplication, namely

$$\omega = \inf\{w : n \times n \text{ matrices can be multiplied in } O(n^{w+\varepsilon}) \text{ steps } \forall \varepsilon > 0\}.$$

The best lower bound known [72] is only a constant multiple of the obvious $\Omega(n^2)$: Any algorithm must at least read all $2n^2$ entries of the input matrices. Currently the best algorithm known takes $O(n^{2.3729\cdots})$ steps [73], and it is a folklore conjecture that $\omega = 2$. Closing this gap is a major open problem in algebraic complexity theory.

Starting in 1969 with Strassen's result, there was a relatively steady stream of improvements to the best upper bound for ω. This culminated in 1990 when Coppersmith and Winograd [21] used the Salem–Spencer construction [92] of arithmetic-progression-free sets to develop an infinite family of matrix multiplication algorithms, whose exponent limited to $2.375477\ldots$. This was the first hint of a relationship between matrix multiplication and arithmetic progressions. Then progress on ω hit a standstill for 20 years.

Although improvements in the upper bound on ω would wait until 2010, in 2003, Cohn and Umans [20] introduced a new approach to algorithms for matrix multiplication, which will draw out just how deep the connection is between such algorithms and arithmetic progressions. Briefly, their approach requires finding finite groups with only low-dimensional irreducible representations and containing three subsets satisfying a certain condition (see Definition 4.7 below). In 2005, with Kleinberg and Szegedy [19], they showed how to use this approach to develop new algorithms, and to capture the Coppersmith–Winograd algorithm as a Cohn–Umans-style construction in abelian groups of bounded exponent (that is, an infinite family of finite abelian groups such that every element of every group in the family had order $\leq b$). Starting in 2010, by analyzing higher tensor powers of the basic object used by Coppersmith and Winograd, Stothers [104], then Vassilevska Williams [117], and finally Le Gall [73] made improvements, resulting in the current world record $\omega < 2.3728639\ldots$. However, it was then shown [6] that this particular technique—analyzing higher tensor powers of Coppersmith–Winograd—could get an exponent no better than $2.3078\ldots$.

Shortly after the resolution of the Cap Set Conjecture [32], Blasiak–Church–Cohn–Grochow–Nasland–Sawin–Umans [10], and independently N. Alon, showed that Cohn–Umans-style constructions in an abelian group yielded not cap sets, but tricolored sum-free sets (first introduced in connection with matrix multiplication in [5], elaborated in [10]). To get a sense for where these come from, let's see how Cohn and Umans proposed using a finite group to multiply matrices. The idea is that the group algebra $\mathbb{C}[G]$ is a direct sum of matrix algebras $\mathbb{C}[G] \cong M_{d_1}(\mathbb{C}) \oplus \cdots M_{d_c}(\mathbb{C})$, where the d_i are the dimensions of the irreducible representations of G. If we could somehow embed $n \times n$ matrix multiplication into $\mathbb{C}[G]$ with $n > \max\{d_i\}$, then we could recursively multiply the smaller $d_i \times d_i$ matrices in order to multiply $n \times n$ matrices, thereby getting a nontrivial algorithm. One then gets a nontrivial upper bound on ω as the infimum of w falsifying the inequality $n^w \leq \sum_i d_i^w$ [20].

To embed a matrix product larger than any of the d_i into a group algebra, Cohn and Umans proposed the following construction. If we want to multiply $A \cdot B = C$, we'll use three subsets $S, T, U \subseteq G$, such that A is an $|S| \times |T|$ matrix and B is a $|T| \times |U|$ matrix. We embed A into the group algebra as $\iota_1(A) = \sum_{i,j} a_{ij} s_i t_j^{-1}$ and B into the group algebra as $\iota_2(B) = \sum_{j,k} b_{j,k} t_j u_k^{-1}$. We would like to be able

to read off the entries of C as the coefficients of the group elements $s_i u_k^{-1}$ in the product $\iota_1(A)\iota_2(B)$. When we perform this multiplication, however, we end up with $\sum_{i,j,j',k} a_{i,j} b_{j',k} s_i t_j^{-1} t_{j'} u_k^{-1}$. If the only way that a group element $s_i t_j^{-1} t_{j'} u_k^{-1}$ can be of the form $s_{i'} u_{k'}^{-1}$ is with $i = i'$, $j = j'$, and $k = k'$, then indeed we get that the coefficient of $s_i u_k^{-1}$ in the product is precisely $\sum_j a_{i,j} b_{j,k} = c_{i,k}$, as desired. Rewriting this condition we have:

Definition 4.7 (Triple product property (TPP) [20])**.** Given a group G, three subsets $S, T, U \subseteq G$ satisfy the *triple product property* if

$$s_{i'}^{-1} s_i t_j^{-1} t_{j'} u_k^{-1} u_{k'} = 1 \iff i = i' \text{ and } j = j' \text{ and } k = k'.$$

In fact, the constructions of [19, 21, 104, 117, 73] are all instances of a generalization of this called the *simultaneous* triple product property—in which one embeds several independent copies of matrix multiplication simultaneously—but the preceding definition is already enough to give us the flavor of the connection with tricolored sum-free sets. For if we write $Q(S) = S^{-1}S = \{s^{-1}s' : s, s' \in S\}$, then the TPP can be rewritten as: for all $q_1 \in Q(S)$, $q_2 \in Q(T)$, $q_3 \in Q(U)$,

$$q_1 q_2 q_3 = 1 \iff q_1 = q_2 = q_3 = 1.$$

This condition is precisely the nonabelian generalization of the defining condition of a tricolored sum-free set (the nonabelian version is sometimes called a "multiplicative matching" [1, 95]). And here, we finally see where this notion of "tricolored" came from: It's because we wanted three different sets to index the rows of A, the rows of B, and the columns of C.

Blasiak *et al.* [10] extended the Ellenberg–Gijswijt bound from vector spaces $(\mathbb{Z}/p\mathbb{Z})^n$ to $(\mathbb{Z}/m\mathbb{Z})^n$ for arbitrary m, and even more generally to abelian groups of bounded exponent. This generalization implies the $(\mathbb{Z}/m\mathbb{Z})^n$ Sunflower Conjecture 4.6, which was previously shown equivalent to the original Sunflower Conjecture 4.2 [5]. Additionally, using the connection between TPP constructions and tricolored sum-free sets, they showed:

Theorem 4.8 (Blasiak, Church, Cohn, Grochow, Nasland, Sawin, & Umans [10])**.** *One cannot show that $\omega = 2$ using simultaneous TPP constructions in families of abelian groups of bounded exponent.*

This includes and goes significantly beyond the class of Coppersmith–Winograd-style algorithms [21, 104, 117, 73]. Thus, if $\omega = 2$, proving so requires significantly new techniques.

4.3. **Triangle removal.** Szemerédi's Regularity Lemma is a powerful tool in graph theory, essentially giving the structure of an arbitrary graph. A well-known consequence of the regularity lemma is the following (whose origin is a bit murky, but from Green [54] this should be attributed to some combination of Szemerédi & Ruzsa [109, 91]):

Theorem 4.9 (Triangle Removal Lemma, Szemerédi & Ruzsa [109, 91][9])**.** *If a graph G on n vertices contains only $o(n^3)$ triangles, then by removing only $o(n^2)$ edges one can make the resulting graph triangle-free. More precisely, if G has $\leq \delta n^3$ triangles, then one can remove $\varepsilon(\delta)n^2$ edges, where $\varepsilon(\delta) \to 0$ as $\delta \to 0$.*

---

[9]It's not exactly clear where this now-well-known version of the lemma was first stated; see [54] for a discussion of the history.

While on the surface this seems to have little to do with arithmetic progressions and cap sets, we note that the Triangle Removal Lemma can be used to give a very simple proof [91] of Roth's result [90] that $r_3(\mathbb{Z}/N\mathbb{Z}) \leq o(N)$. To see the connection, we give the brief proof here:

Proof of Roth's Theorem from the Triangle Removal Lemma [91]. Suppose $A \subseteq [N]$ has size $|A| \geq \varepsilon N$. We build a graph G as follows: Its vertex set V will be the disjoint union of three sets V_1, V_2, V_3, each of size $3N$, which we identify with $[3N]$ (so $|V| = 9N$). The edges are as follows: $(i, j) \in V_1 \times V_2$ is an edge if and only if $j - i \in A$; $(j, k) \in V_2 \times V_3$ is an edge if and only if $k - j \in A$, and $(k, i) \in V_3 \times V_1$ is an edge if and only if $\frac{k-i}{2} \in A$. There are no other edges. Then $(i, j, k) \in V_1 \times V_2 \times V_3$ form a triangle if and only if $j - i = a_1 \in A$ and $k - j = a_3 \in A$ and $\frac{k-i}{2} = a_2 \in A$ if and only if a_1, a_2, a_3 is an arithmetic progression in A, for we have $a_2 - a_1 = \frac{k+i}{2} - j = a_3 - a_2$. Note that this also allows the trivial arithmetic progression (a, a, a), as nothing here forces the difference $\frac{k+i}{2} - j$ to be nonzero. For each $i \in [N]$ and each $a \in A$ we get a triangle corresponding to the trivial arithmetic progression (a, a, a), namely the triangle with vertices $i \in V_1$, $i + a \in V_2$, and $i + 2a \in V_3$. Since $|A| \geq \varepsilon N$, we thus have at least $|A||V_1| \geq 3\varepsilon N^2$ triangles in G. Furthermore, these triangles are all disjoint from one another, so to make G triangle-free would require removing at least $3\varepsilon N^2$ edges (one for each such triangle). As this is *not* $o(|V|^2) = o(81N^2) = o(N^2)$, it must *not* be the case that G has only $o(N^3)$ triangles, by the Triangle Removal Lemma. In other words, there is some $\delta > 0$ that depends only on ε (but not on A nor N) such that G contains at least $\delta|V|^3 = 729\delta N^3$ triangles. However, the total number of triangles corresponding to the trivial arithmetic progressions is $|A||V_1| \leq 3N^2$, so at least $729\delta N^3 - 3N^2$ of the triangles correspond to proper arithmetic progressions in A. In particular, since δ is independent of N, for sufficiently large N it must be the case that A contains at least one proper arithmetic progression of length 3. □

With this connection in mind, it is natural to define a *triangle* in an abelian group Z to be three elements $x, y, z \in Z$ such that $x + y + z = 0$.

Theorem 4.10 (Green [54]). *Let Z be an abelian group of order N. If $A \subseteq Z$ has only $o(N^2)$ triangles, then by removing only $o(N)$ elements from A one can make the resulting set triangle-free.*

As with the original Triangle Removal Lemma, we can rephrase this in terms of δ, ε. Unfortunately, the best quantitative upper bound was that $1/\delta$ was of the form

$2^{2^{2^{\cdot^{\cdot^{\cdot}}}}}$, where the height of this tower was $\log(1/\varepsilon)$ [40]. But with the techniques used to resolve the Cap Set Conjecture (and the result of [10]), Fox and Lovász showed a tight bound, dropping this from an exponential tower all the way down to a polynomial! They also generalized it from a single set A to a tricolored version: given $A, B, C \subseteq Z$, we say $(a, b, c) \in A \times B \times C$ form a triangle if $a + b + c = 0$.

Theorem 4.11 (Fox and Lovász [41]). *For each prime p there is a constant C_p such that the following holds. If $A, B, C \subseteq (\mathbb{Z}/p\mathbb{Z})^n$ have only δN^2 tricolored triangles where $\delta = (\varepsilon/3)^{C_p}$ (and $N = p^n$, as usual), then by removing only εN elements from $A \cup B \cup C$ one can make the resulting set triangle-free. Furthermore, this is essentially tight, in that it only holds with $\delta \leq \varepsilon^{C_p - o(1)}$.*

Continuing the recurring theme of the relationship between $(\mathbb{Z}/p\mathbb{Z})^n$ and $\mathbb{Z}/N\mathbb{Z}$, Aaronson [1] extended this connection between tricolored sum-free sets and triangles to $\mathbb{Z}/N\mathbb{Z}$.

4.4. **Matrix rigidity.** A natural question in computational complexity is, for a fixed matrix A, how hard is it to compute the function $x \mapsto Ax$? The naive approach, for $n \times n$ matrices, takes $O(n^2)$ arithmetic operations. There are several famous matrices for which this number is reduced, most notably Fourier matrices, which can be applied in only $O(n \log n)$ operations, nearly linear in the size of the vector space. Aside from a few other highly structured classes of matrices, very little is known about this question in general. (And if we can't even answer this question with modern techniques, what hope do we have of proving $\mathsf{P} \neq \mathsf{NP}$?) Two natural properties of a matrix A that make the corresponding linear function easy to compute are: (1) sparsity, that is, if A has only very few nonzero entries, or (2) low rank. And any two such "easy" cases can be added together. That is, if an $n \times n$ matrix A is the sum of a matrix A' with only s nonzero entries and a matrix A'' of rank r, then $x \mapsto Ax$ can be computed in $O(s + rn)$ arithmetic operations. For a given matrix A, this raises the question of how few entries you need to change in order to make the difference have low rank.

Definition 4.12 (Matrix rigidity [115])**.** The *rank-r rigidity* of a matrix A, denoted $R_A(r)$, is the least number s such that A is the sum of a matrix A' with $\leq s$ nonzero entries and a matrix A'' of rank $\leq r$.

Perhaps the most natural way to express the computation of an n-dimensional linear function $x \mapsto Ax$ is with a *linear circuit*: a sequence of instructions $g_1, \ldots, g_\ell$ of the form $g_i = x_j$ for some coordinate x_j of the input, or a linear combination $g_j = \alpha_{j1} g_{j_1} + \alpha_{j2} g_{j_2}$ for constants α_{jk} and previous instructions $j_1, j_2 < j$. The "output" of such a sequence is its last n values. The *size* of the linear circuit is the number ℓ of instructions. The $O(n \log n)$-step algorithm for the Fourier transform, for example, translates into a linear circuit of size $O(n \log n)$. To any such linear circuit we may naturally associate a directed acyclic graph on vertex set $[\ell]$, with arrows from $g_i \to g_j$ if g_i appears as a summand in the instruction g_j. The *depth* of a linear circuit is the length of the longest directed path in this graph.

Theorem 4.13 (Valiant [115])**.** *For every n, let A_n be an $n \times n$ matrix over a field. If $R_{A_n}(n/\log\log n) \geq \Omega(n^{1+\varepsilon})$ for some $\varepsilon > 0$, then for sufficiently large n, the linear function A_n cannot be computed by linear circuits of size $O(n)$ and depth $O(\log n)$.*

It is not hard to see that $R_A(r) \leq (n-r)^2$ for all r, and Valiant proved that almost all matrices are at least this rigid. However, to date, the best lower bound on any *explicit* matrix A is $R_A(r) \geq \Omega(\frac{n^2}{r} \log \frac{n}{r})$ [44, 101]. Other techniques that have been used to study rigid matrices include elimination theory [68], degree bounds [77, 76], spectral methods [64], and algebraic geometry [46]; for a mostly up-to-date survey, that also includes relations to other areas, see [78]. For a long time it was believed that the Hadamard matrices were sufficiently rigid to apply Theorem 4.13, but this was recently disproved [2].

And now, we can add to this list of techniques the Croot–Lev–Pach polynomial method:

Theorem 4.14 (Dvir and Edelman [25])**.** *Let p be a fixed prime, and $\varepsilon > 0$. For any function $f\colon (\mathbb{Z}/p\mathbb{Z})^n \to \mathbb{Z}/p\mathbb{Z}$, let $N = p^n$, and let M_f be the $N \times N$ matrix $M_f(x, y) = f(x + y)$. Then there is a $\delta > 0$ such that for all sufficiently large n, $R_{M_f}(N^{1-\delta}) \leq N^{1+\varepsilon}$. In particular, such matrices are not rigid enough to apply Theorem 4.13.*

5. Conclusion and outlook

Originally motivated by trying to find structure in the prime numbers, we were led to study arithmetic progressions in vector spaces over $\mathbb{Z}/p\mathbb{Z}$ as a model for arithmetic progression in the primes or in $\mathbb{Z}$. This turns out to be quite a fruitful toy model, and the Cap Set Conjecture was developed as a keystone problem, whose solution was expected to unlock the mysteries of many other problems in combinatorics and number theory. And indeed, as evidenced by the long list of applications already, the technique used to resolve the Cap Set Conjecture had precisely the desired effect! (It may be worth noting that almost none of these applications follow as corollaries of the result itself, they only followed by using the Croot–Lev–Pach *technique*.)

Of course, Erdős's Conjecture on arithmetic progression (Conjecture 1.1) stands out as one of the most significant open problems we've discussed. Closing the gap between the best known upper and lower bounds for $r_3(\mathbb{Z}/N\mathbb{Z})$ and $r_3((\mathbb{Z}/p\mathbb{Z})^n)$ is also an interesting open problem. On the one hand, closing the gap for $r_3((\mathbb{Z}/p\mathbb{Z})^n)$ may seem like somewhat of a "clean-up operation," given how close the bounds now are. On the other hand, the discussion in Section 3 reveals that closing these gaps seems to require really novel methods, and one might hope that such new methods would have other applications.

Along the way, we've discussed many constructions of additive sets with various properties [8, 30, 58, 12, 86, 27, 17, 29, 28, 66]. Are these constructions native to the groups they were designed for? Or can they be used in other groups as well? It may be fruitful to study this question from the following angle. Along with the notion of Freiman isomorphism, for any additive set (A, Z) there is a notion of a *universal ambient group* for that additive set:

Definition 5.1 (Universal ambient group (see, e. g., [114, Section 5.5]))**.** Let (A, Z) be an additive set and $k \in \mathbb{N}$. An abelian group U is a *universal ambient group* (of order k) for A if there is a Freiman k-isomorphism $(A, Z) \cong (A', U)$, and every Freiman k-homomorphism $(A', U) \to (B, W)$ extends to a unique group homomorphism $U \to W$.

Lev and Konyagin [67] showed that universal ambient groups always exist. As a way of getting at the preceding questions, I think it would be interesting to develop a generalization of Freiman homomorphisms for "multi-colored additive sets" (whatever that ought to mean), and to determine the universal ambient groups for the constructions mentioned in this exposition.

Finally, given all the applications of the Croot–Lev–Pach technique after just one year, what other applications of the polynomial method are waiting to be explored?

Appendix A. Asymptotic growth

We will be considering many quantities as a function of some (usually integer) parameter N, as $N \to \infty$. This allows us to get at the essence of certain constructions and bounds—and to compare different constructions with one another—without getting caught up in the details of their *exact* sizes (which can often be hard to compute) or how large N must be before one sees a difference between two techniques.

The most common notations we will be using are:

- $f(N) \sim g(N)$: $\lim_{N \to \infty} \frac{f(N)}{g(N)} = 1$, and we say that f and g are *asymptotically equal*. The advantage of this notation is that it lets us focus on the highest-order terms only.
- $f(N) \le O(g(N))$: there is a constant $c > 0$ such that for all sufficiently large N (that is, "there is an N_0 such that for all $N > N_0$") $f(N) \le cg(N)$. For most purposes, it is equivalent to say that $\lim_{N \to \infty} \frac{f(N)}{g(N)}$ is finite (these are not entirely equivalent as there are functions f, g such that this limit doesn't exist, yet nonetheless $f(N) \le O(g(N))$, but I don't think we encounter any such pathologies here).

  Similarly, we may use the notation $O(g(N))$ in a formula to denote an unspecified function f such that $f \le O(g)$, e.g. $O(n^2)e^n$. The advantage of this notation is that it lets us focus on the highest-order terms *and* not worry about multiplicative constants, that are independent of N. For example, N and $100N$ are both $O(N)$, even though $N \not\sim 100N$.
- $f(N) \le o(g(N))$: for *all* $c > 0$, for all sufficiently large N, $f(N) \le cg(N)$. Equivalently, $\lim_{n \to \infty} \frac{f(N)}{g(N)} = 0$. If you think of O as the asymptotic version of $\le$, then o is the asymptotic version of $<$. Again, we may use $o(g(N))$ in a formula to denote an unspecified f such that $f \le o(g)$.
- $o(1)$: A particular case of the preceding that shows up frequently, namely an unspecified function of N that goes to zero as N goes to infinity. This is especially useful when it appears in exponents, such as $c^{n(1+o(1))}$. For example, $c^n n^2 = c^{n + \log_c 2} = c^{n(1 + (1/n) \log_c 2)} = c^{n(1+o(1))}$. The advantage of this notation is that it lets us focus on the exponential growth rate without worrying about lower-order *multiplicative* terms (even when they depend on N). For we have that $c < d$ if and only if $c^{n(1+o(1))} = o(d^{n(1+o(1))})$, even if the two $o(1)$ terms are different: let $f(N), g(N) \le o(1)$, then

$$\lim_{n \to \infty} \frac{c^{n(1+f(n))}}{d^{n(1+g(n))}} = \lim_{n \to \infty} \left(\frac{c^{1+f(n)}}{d^{1+g(n)}} \right)^n$$

$$= \lim_{n \to \infty} \left(\frac{c}{d} \right)^n = 0$$

  The jump from the first to second line here is allowed because if $c < d$, then there is some $\varepsilon_0 > 0$ such that if $0 < \varepsilon < \varepsilon_0$, we have $c^{1+\varepsilon} < d^{1+\varepsilon}$, and there is some n_0 such that for all $n > n_0$, we have $f(n), g(n) < \varepsilon_0$.

Here is a list of the most frequent growth rates we'll be considering and the relations between them. If you haven't seen these before, working out the relations for yourself is a nice but not terribly difficult exercise, that helps acquaint you with these growth rates. They are listed in strictly increasing order, so that if you see $f(N), g(N)$ in this list, it means that $f(N) \le o(g(N))$.

$$(\log N)^{1/2}, \log N, (\log N)^2, (\log N)^3, \ldots, \sqrt{N}, N^{0.75}, N, N^2, N^3, \ldots, 1.01^N, 2^N, e^N, 3^N$$

Indeed, $(\log N)^c \leq o((\log N)^d)$ for constants $c, d > 0$ if and only if $c < d$, and similarly $N^c \leq o(N^d)$ and $\exp(cn) \leq o(\exp(dn))$ if and only if $c < d$. In particular, for exponentials of the form c^N $(c > 1)$, the base of the exponent matters: $c^N \leq d^N$ if and only if $1 < c < d$.

Furthermore, for growth rates that are exponentially separated, altering constant exponents never changes this, viz. $(\log \log N)^c \leq o((\log N)^\varepsilon)$ for any constants $c > 0, \varepsilon > 0$, no matter how large c is and how small ε is. Similarly $(\log N)^c \leq o(N^\varepsilon)$, and $N^c \leq o((1 + \varepsilon)^n)$ for all $c > 0, \varepsilon > 0$.

References

1. James Aaronson, *A connection between matchings and removal in abelian groups*, arXiv:1612.05172 [math.CO], 2016.

2. Josh Alman and Ryan Williams, *Probabilistic rank and matrix rigidity*, STOC '17: 49th Annual ACM Symposium on Theory of Computing, ACM, New York, 2017, arXiv:1611.05558 [cs.CC], pp. 641–652. MR 3678217

3. Noga Alon and Moshe Dubiner, *Zero-sum sets of prescribed size*, Combinatorics, Paul Erdős is eighty, Vol. 1, Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest, 1993, pp. 33–50. MR 1249703

4. ———, *A lattice point problem and additive number theory*, Combinatorica **15** (1995), no. 3, 301–309. MR 1357277

5. Noga Alon, Amir Shpilka, and Christopher Umans, *On sunflowers and matrix multiplication*, Comput. Complexity **22** (2013), no. 2, 219–243. MR 3055780

6. Andris Ambainis, Yuval Filmus, and François Le Gall, *Fast matrix multiplication: Limitations of the Coppersmith–Winograd method*, STOC '15: 47th Annual ACM Symposium on Theory of Computing, 2015, arXiv:1411.5414 [cs.CC], pp. 585–593.

7. Michael Bateman and Nets Hawk Katz, *New bounds on cap sets*, J. Amer. Math. Soc. **25** (2012), no. 2, 585–613. MR 2869028

8. F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U. S. A. **32** (1946), 331–332. MR 0018694

9. Jürgen Bierbrauer and Yves Edel, *Bounds on affine caps*, J. Combin. Des. **10** (2002), no. 2, 111–115. MR 1883512

10. Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A. Grochow, Eric Naslund, William F. Sawin, and Chris Umans, *On cap sets and the group-theoretic approach to matrix multiplication*, Discrete Anal. (2017), Paper No. 3, 27. MR 3631613

11. T. F. Bloom, *A quantitative improvement for Roth's theorem on arithmetic progressions*, J. Lond. Math. Soc. (2) **93** (2016), no. 3, 643–663. MR 3509957

12. R. C. Bose, *Mathematical theory of the symmetrical factorial design*, Sankhyā **8** (1947), no. 2, 107–166.

13. Jean Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. **9** (1999), no. 5, 968–984. MR 1726234

14. ———, *Roth's theorem on progressions revisited*, J. Anal. Math. **104** (2008), 155–192. MR 2403433

15. T. C. Brown and J. P. Buhler, *A density version of a geometric Ramsey theorem*, J. Combin. Theory Ser. A **32** (1982), no. 1, 20–34. MR 640624

16. J. P. Buhler, H. W. Lenstra, Jr., and C. Pomerance, *Factoring integers with the number field sieve*, The development of the number field sieve, Lecture Notes in Math., vol. 1554, Springer, Berlin, 1993, pp. 50–94. MR 1321221

17. A. R. Calderbank and P. C. Fishburn, *Maximal three-independent subsets of* $\{0, 1, 2\}^n$, Des. Codes Cryptogr. **4** (1994), no. 3, 203–211. MR 1277940

18. S. Chowla, *There exists an infinity of 3—combinations of primes in A. P.*, Proc. Lahore Philos. Soc. **6** (1944), no. 2, 15–16. MR 0014125

19. Henry Cohn, Robert Kleinberg, Balazs Szegedy, and Christopher Umans, *Group-theoretic algorithms for matrix multiplication*, FOCS '05: 46th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, 2005, arXiv:math/0511460 [math.GR], pp. 379–388.

20. Henry Cohn and Christopher Umans, *A group-theoretic approach to fast matrix multiplication*, FOCS '03: 44th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, 2003, arXiv:math/0307321 [math.GR], pp. 438–449.

21. Don Coppersmith and Shmuel Winograd, *Matrix multiplication via arithmetic progressions*, J. Symbolic Comput. **9** (1990), no. 3, 251–280. MR 1056627 (91i:68058)

22. Ernie Croot, Vsevolod F. Lev, and Péter Pál Pach, *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small*, Ann. of Math. (2) **185** (2017), no. 1, 331–337. MR 3583357

23. M. S. Dousti and Kaveh Ghasemloo, *Answers to "Adding integers represented by their factorization is as hard as factoring?"*, https://cstheory.stackexchange.com/q/7491/129, 2011.

24. Zeev Dvir, *On the size of Kakeya sets in finite fields*, J. Amer. Math. Soc. **22** (2009), no. 4, 1093–1097. MR 2525780

25. Zeev Dvir and Ben Edelman, *Matrix rigidity and the Croot–Lev–Pach lemma*, arXiv:1708.01646 [cs.CC], 2017.

26. Zeev Dvir, Swastik Kopparty, Shubhangi Saraf, and Madhu Sudan, *Extensions to the method of multiplicities, with applications to Kakeya sets and mergers*, SIAM J. Comput. **42** (2013), no. 6, 2305–2328. MR 3143848

27. Y. Edel, S. Ferret, I. Landjev, and L. Storme, *The classification of the largest caps in* AG(5, 3), J. Combin. Theory Ser. A **99** (2002), no. 1, 95–110. MR 1911459

28. Yves Edel, *Extensions of generalized product caps*, Des. Codes Cryptogr. **31** (2004), no. 1, 5–14. MR 2031694

29. Yves Edel and Jürgen Bierbrauer, *Large caps in small spaces*, Des. Codes Cryptogr. **23** (2001), no. 2, 197–212. MR 1830941

30. Michael Elkin, *An improved construction of progression-free sets*, Israel J. Math. **184** (2011), 93–128. MR 2823971

31. Jordan S. Ellenberg, *Sumsets as unions of sumsets of subsets*, arXiv:1612.01929 [math.CO], 2016.

32. Jordan S. Ellenberg and Dion Gijswijt, *On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression*, Ann. of Math. (2) **185** (2017), no. 1, 339–343. MR 3583358

33. Pál Erdős, *Résultats et problèmes en théorie des nombres*, Séminaire Delange–Pisot–Poitou (14e année: 1972/73), Théorie des nombres, Fasc. 2, Exp. No. 24, Secrétariat Mathématique, Paris, 1973, p. 7. MR 0396376

34. ———, *Problems in number theory and combinatorics*, Proceedings of the Sixth Manitoba Conference on Numerical Mathematics (Univ. Manitoba, Winnipeg, Man., 1976), Congress. Numer., XVIII, Utilitas Math., Winnipeg, Man., 1977, pp. 35–58. MR 532690

35. ———, *On the combinatorial problems which I would most like to see solved*, Combinatorica **1** (1981), no. 1, 25–42.

36. ———, *Some of my favorite problems and results*, The mathematics of Paul Erdős, I, Algorithms Combin., vol. 13, Springer, Berlin, 1997, pp. 47–67. MR 1425174

37. Pál Erdős and R. Rado, *Intersection theorems for systems of sets*, J. London Math. Soc. **35** (1960), 85–90. MR 0111692

38. Pál Erdős and E. Szemerédi, *Combinatorial properties of systems of sets*, J. Combinatorial Theory Ser. A **24** (1978), no. 3, 308–313. MR 0491202

39. Pál Erdős and Paul Turán, *On some sequences of integers*, J. London Math. Soc. **S1-11** (1936), no. 4, 261–264. MR 1574918

40. Jacob Fox, *A new proof of the graph removal lemma*, Ann. of Math. (2) **174** (2011), no. 1, 561–579. MR 2811609

41. Jacob Fox and László Miklós Lovász, *A tight bound for Green's arithmetic triangle removal lemma in vector spaces*, Adv. Math. **321** (2017), 287–297, arXiv:1606.01230 [math.CO]. MR 3715712

42. P. Frankl, R. L. Graham, and V. Rödl, *On subsets of abelian groups with no 3-term arithmetic progression*, J. Combin. Theory Ser. A **45** (1987), no. 1, 157–161. MR 883900

43. G. A. Freĭman, *Foundations of a structural theory of set addition*, American Mathematical Society, Providence, R. I., 1973, Translated from the Russian, Translations of Mathematical Monographs, Vol 37. MR 0360496

44. Joel Friedman, *A note on matrix rigidity*, Combinatorica **13** (1993), no. 2, 235–239. MR 1237045

45. Gennian Ge and Chong Shangguan, *Rank counting and maximum subsets of $\mathbb{F}_q^n$ containing no right angles*, arXiv:1612.08255 [math.CO], 2016.

46. Fulvio Gesmundo, Jonathan D. Hauenstein, Christian Ikenmeyer, and J. M. Landsberg, *Complexity of linear circuits and geometry*, Found. Comput. Math. **16** (2016), no. 3, 599–635. MR 3494506

47. W. Timothy Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. **8** (1998), no. 3, 529–551. MR 1631259

48. _____, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588. MR 1844079

49. _____, *What is difficult about the cap-set problem?*, Gowers's Weblog, https://gowers.wordpress.com/2011/01/11/what-is-difficult-about-the-cap-set-problem/, 2011.

50. _____, *Erdős and arithmetic progressions*, Erdős Centennial, Bolyai Soc. Math. Stud., vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 265–287. MR 3203599

51. _____, *Reflections on the recent solution of the cap-set problem I*, Gowers's Weblog, https://gowers.wordpress.com/2016/05/19/reflections-on-the-recent-solution-of-the-cap-set-problem-i/, 2016.

52. Ben Green, *Finite field models in additive combinatorics*, Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser., vol. 327, Cambridge Univ. Press, Cambridge, 2005, pp. 1–27. MR 2187732

53. _____, *Roth's theorem in the primes*, Ann. of Math. (2) **161** (2005), no. 3, 1609–1636. MR 2180408

54. _____, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), no. 2, 340–376. MR 2153903

55. _____, *Sárközy's theorem in function fields*, Q. J. Math. **68** (2017), no. 1, 237–242. MR 3658291

56. Ben Green and Terence Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), no. 2, 481–547. MR 2415379

57. Ben Green, Terence Tao, and Tamar Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$-norm*, Ann. of Math. (2) **176** (2012), no. 2, 1231–1372. MR 2950773

58. Ben Green and Julia Wolf, *A note on Elkin's improvement of Behrend's construction*, Additive number theory, Springer, New York, 2010, pp. 141–144. MR 2744752

59. Richard K. Guy, *Unsolved problems in number theory. 3rd ed.*, 3rd ed. ed., New York, NY: Springer-Verlag, 2004 (English).

60. Heiko Harborth, *Ein Extremalproblem für Gitterpunkte*, J. Reine Angew. Math. **262/263** (1973), 356–360, Collection of articles dedicated to Helmut Hasse on his seventy-fifth birthday. MR 0327666

61. D. R. Heath-Brown, *Integer sets containing no arithmetic progressions*, J. London Math. Soc. (2) **35** (1987), no. 3, 385–394. MR 889362

62. Harald Andrés Helfgott and Anne de Roton, *Improving Roth's theorem in the primes*, Int. Math. Res. Not. IMRN (2011), no. 4, 767–783. MR 2773330

63. Stasys Jukna, *Extremal combinatorics*, second ed., Texts in Theoretical Computer Science. An EATCS Series, Springer, Heidelberg, 2011, With applications in computer science. MR 2865719

64. B. S. Kashin and A. A. Razborov, *New lower bounds for the stability of Hadamard matrices*, Mat. Zametki **63** (1998), no. 4, 535–540. MR 1680943

65. Taegyun Kim and Sang il Oum, *An upper bound on tricolored ordered sum-free sets*, arXiv:1708.07263 [math.CO], 2017.

66. Robert Kleinberg, William F. Sawin, and David E. Speyer, *The growth rate of tri-colored sum-free sets*, arXiv:1607.00047 [math.CO], 2016.

67. Sergei V. Konyagin and Vsevolod F. Lev, *Combinatorics and linear algebra of Freiman's isomorphism*, Mathematika **47** (2000), no. 1-2, 39–51 (2002). MR 1924486

68. Abhinav Kumar, Satyanarayana V. Lokam, Vijay M. Patankar, and M. N. Jayalal Sarma, *Using elimination theory to construct rigid matrices*, Comput. Complexity **23** (2014), no. 4, 531–563. MR 3274825

69. Greg Kuperberg, *Answer to question "Open problems with monetary rewards" on Math-Overflow*, https://mathoverflow.net/a/66219/38434, 2011.

70. Jeffrey C. Lagarias, *The $3x + 1$ problem: an annotated bibliography (1963–1999)*, The ultimate challenge: the $3x + 1$ problem, Amer. Math. Soc., Providence, RI, 2010, arXiv:math/0309224, pp. 267–341. MR 2560718

71. J. M. Landsberg, *Geometry and the complexity of matrix multiplication*, Bull. Amer. Math. Soc. (N.S.) **45** (2008), no. 2, 247–284. MR 2383305 (2009b:68055)

72. ———, *New lower bounds for the rank of matrix multiplication*, SIAM J. Comput. **43** (2014), no. 1, 144–149. MR 3162411

73. François Le Gall, *Powers of tensors and fast matrix multiplication*, ISSAC '14: Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation, ACM, 2014, arXiv:1401.7714 [cs.DS], pp. 296–303.

74. Arjen K. Lenstra, *Integer factoring*, Des. Codes Cryptogr. **19** (2000), no. 2-3, 101–128, Towards a quarter-century of public key cryptography. MR 1758972

75. H. W. Lenstra, Jr., *Factoring integers with elliptic curves*, Ann. of Math. (2) **126** (1987), no. 3, 649–673. MR 916721

76. Satyanarayana V. Lokam, *On the rigidity of Vandermonde matrices*, Theoret. Comput. Sci. **237** (2000), no. 1-2, 477–483. MR 1756225

77. ———, *Quadratic lower bounds on matrix rigidity*, Theory and applications of models of computation, Lecture Notes in Comput. Sci., vol. 3959, Springer, Berlin, 2006, pp. 295–307. MR 2277251

78. ———, *Complexity lower bounds using linear algebra*, Found. Trends Theor. Comput. Sci. **4** (2008), no. 1-2, front matter, 1–155 (2009). MR 2539154

79. James Maynard, *Small gaps between primes*, Ann. of Math. (2) **181** (2015), no. 1, 383–413. MR 3272929

80. Roy Meshulam, *On subsets of finite abelian groups with no 3-term arithmetic progressions*, J. Combin. Theory Ser. A **71** (1995), no. 1, 168–172. MR 1335785

81. Pascal Michel, *Busy beaver competition and Collatz-like problems*, Arch. Math. Logic **32** (1993), no. 5, 351–367. MR 1223396

82. Michael A. Morrison and John Brillhart, *A method of factoring and the factorization of F_7*, Math. Comp. **29** (1975), 183–205, Collection of articles dedicated to Derrick Henry Lehmer on the occasion of his seventieth birthday. MR 0371800

83. Greg Muller, *Answer to "What is the importance of the Collatz conjecture?"*, https://math.stackexchange.com/a/10608/224688, 2010.

84. Sergey Norin, *A distribution on triples with maximum entropy marginal*, arXiv:1608.00243 [math.CO], 2016.

85. Luke Pebody, *Proof of a conjecture of Kleinberg–Sawin–Speyer*, arXiv:1608.05740 [math.CO], 2016.

86. Giuseppe Pellegrino, *Sul massimo ordine delle calotte in $S_{4,3}$*, Matematiche (Catania) **25** (1970), 149–157 (1971). MR 0363952

87. D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. Sci. **1** (2014), Art. 12, 83, arXiv:1407.4897 [math.NT]. MR 3373710

88. C. Pomerance, *Analysis and comparison of some integer factoring algorithms*, Computational methods in number theory, Part I, Math. Centre Tracts, vol. 154, Math. Centrum, Amsterdam, 1982, pp. 89–139. MR 700260

89. Hans Riesel, *Prime numbers and computer methods for factorization*, second ed., Progress in Mathematics, vol. 126, Birkhäuser Boston, Inc., Boston, MA, 1994. MR 1292250

90. K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109. MR 0051853

91. I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II, Colloq. Math. Soc. János Bolyai, vol. 18, North-Holland, Amsterdam-New York, 1978, pp. 939–945. MR 519318

92. R. Salem and D. C. Spencer, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U. S. A. **28** (1942), 561–563. MR 0007405

93. Tom Sanders, *On Roth's theorem on progressions*, Ann. of Math. (2) **174** (2011), no. 1, 619–636. MR 2811612

94. ———, *On certain other sets of integers*, Journal d'Analyse Mathématique **116** (2012), no. 1, 53–82.

95. William F. Sawin, *Bounds for matchings in nonabelian groups*, arXiv:1702.00905 [math.CO], 2017.

96. Tomasz Schoen and Ilya D. Shkredov, *Roth's theorem in many variables*, Israel J. Math. **199** (2014), no. 1, 287–308. MR 3219538

97. Tomasz Schoen and Olof Sisask, *Roth's theorem for four variables and additive structures in sums of sparse sets*, Forum Math. Sigma **4** (2016), e5, 28. MR 3482282

98. Beniamino Segre, *Le geometrie di Galois*, Ann. Mat. Pura Appl. (4) **48** (1959), 1–96. MR 0116259

99. ———, *Introduction to Galois geometries*, Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez. I (8) **8** (1967), 133–236. MR 0238846

100. Set Enterprises, Inc., *Marsha Jean Falco—the creative genius behind SET®*, https://puzzles.setgame.com/set/history.htm.

101. M. A. Shokrollahi, D. A. Spielman, and V. Stemann, *A remark on matrix rigidity*, Inform. Process. Lett. **64** (1997), no. 6, 283–285. MR 1608240

102. Peter W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Comput. **26** (1997), no. 5, 1484–1509. MR 1471990

103. Alexander Soifer, *The mathematical coloring book*, Springer, New York, 2009, Mathematics of coloring and the colorful life of its creators, With forewords by Branko Grünbaum, Peter D. Johnson, Jr. and Cecil Rousseau. MR 2458293

104. Andrew Stothers, *On the complexity of matrix multiplication*, Ph.D. thesis, University of Edinburgh, Edinburgh, UK, 2010.

105. Volker Strassen, *Gaussian elimination is not optimal*, Numer. Math. **13** (1969), 354–356. MR 0248973 (40 #2223)

106. E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Acta Math. Acad. Sci. Hungar. **20** (1969), 89–104. MR 0245555

107. ———, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245, Collection of articles in memory of Juriĭ Vladimirovič Linnik. MR 0369312

108. ———, *Integer sets containing no arithmetic progressions*, Acta Math. Hungar. **56** (1990), no. 1-2, 155–158. MR 1100788

109. Endre Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401. MR 540024

110. Terence Tao, *Open question: best bounds for cap sets*, What's new, https://terrytao.wordpress.com/2007/02/23/open-question-best-bounds-for-cap-sets/, 2007.

111. ———, *Algebraic combinatorial geometry: the polynomial method in arithmetic combinatorics, incidence combinatorics, and number theory*, EMS Surv. Math. Sci. (2014), no. 1, 1–46, arXiv:1310.6482 [math.CO].

112. ———, *A symmetric formulation of the Croot–Lev–Pach–Ellenberg–Gijswijt capset bound*, What's new, https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound/, 2016.

113. Terence Tao and William F. Sawin, *Notes on the "slice rank" of tensors*, What's new, https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/, 2016.

114. Terence Tao and Van Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2006. MR 2289012

115. Leslie G. Valiant, *Graph-theoretic arguments in low-level complexity*, Mathematical foundations of computer science (Proc. Sixth Sympos., Tatranská Lomnica, 1977), Springer, Berlin, 1977, pp. 162–176. Lecture Notes in Comput. Sci., Vol. 53. MR 0660702

116. J. G. van der Corput, *über Summen von Primzahlen und Primzahlquadraten*, Math. Ann. **116** (1939), no. 1, 1–50. MR 1513216

117. Virginia Vassilevska Williams, *Multiplying matrices faster than Coppersmith–Winograd*, STOC '12: 44th Annual ACM Symposium on Theory of Computing, ACM, 2012, pp. 887–898.

118. Samuel S. Wagstaff, Jr., *The joy of factoring*, Student Mathematical Library, vol. 68, American Mathematical Society, Providence, RI, 2013. MR 3135977

119. J. Wolf, *Finite field models in arithmetic combinatorics—ten years on*, Finite Fields Appl. **32** (2015), 233–274. MR 3293412

120. Doron Zeilberger, *A motivated rendition of the Ellenberg–Gijswijt gorgeous proof that the largest subset of F_3^n with no three-term arithmetic progression is $O(c^n)$, with $c = \sqrt[3]{(5589 + 891\sqrt{33})/8} = 2.75510461302363300022127\ldots$*, arXiv:1607.01804 [math.CO], 2016.
121. Yitang Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), no. 3, 1121–1174, arXiv:1311.4600 [math.NT]. MR 3171761

E-mail address: `jgrochow@colorado.edu`

CURRENT EVENTS BULLETIN

Previous speakers and titles

For PDF files of talks, and links to Bulletin of the AMS articles,
 see http://www.ams.org/ams/current-events-bulletin.html.

January 6, 2017 (Atlanta, GA)

Lydia Bieri, University of Michigan
Black hole formation and stability: a mathematical investigation.

Matt Baker, Georgia Tech
Hodge Theory in Combinatorics.

Kannan Soundararajan, Stanford University
Tao's work on the Erdos Discrepancy Problem.

Susan Holmes, Stanford University
Statistical proof and the problem of irreproducibility.

January 8, 2016 (Seattle, WA)

Carina Curto, Pennsylvania State University
What can topology tell us about the neural code?

Lionel Levine, Cornell University and *Yuval Peres, Microsoft Research
and University of California, Berkeley
Laplacian growth, sandpiles and scaling limits.

Timothy Gowers, Cambridge University
Probabilistic combinatorics and the recent work of Peter Keevash.

Amie Wilkinson, University of Chicago
What are Lyapunov exponents, and why are they interesting?

January 12, 2015 (San Antonio, TX)

Jared S. Weinstein, Boston University
Exploring the Galois group of the rational numbers: Recent breakthroughs.

Andrea R. Nahmod, University of Massachusetts, Amherst
The nonlinear Schrödinger equation on tori: Integrating harmonic analysis, geometry, and probability.

Mina Aganagic, University of California, Berkeley
String theory and math: Why this marriage may last.

Alex Wright, Stanford University
From rational billiards to dynamics on moduli spaces.

January 17, 2014 (Baltimore, MD)

Daniel Rothman, Massachusetts Institute of Technology
Earth's Carbon Cycle: A Mathematical Perspective

Karen Vogtmann, Cornell University
The geometry of Outer space

Yakov Eliashberg, Stanford University
Recent advances in symplectic flexibility

Andrew Granville, Université de Montréal
Infinitely many pairs of primes differ by no more than 70 million (and the bound's getting smaller every day)

January 11, 2013 (San Diego, CA)

Wei Ho, Columbia University
How many rational points does a random curve have?

Sam Payne, Yale University
Topology of nonarchimedean analytic spaces

Mladen Bestvina, University of Utah
Geometric group theory and 3-manifolds hand in hand: the fulfillment of Thurston's vision for three-manifolds

Lauren Williams, University of California, Berkeley
Cluster algebras

January 6, 2012 (Boston, MA)

Jeffrey Brock, Brown University
*Assembling surfaces from random pants: the surface-subgroup
and Ehrenpreis conjectures*

Daniel Freed, University of Texas at Austin
*The cobordism hypothesis: quantum field theory + homotopy
invariance = higher algebra*

Gigliola Staffilani, Massachusetts Institute of Technology
Dispersive equations and their role beyond PDE

Umesh Vazirani, University of California, Berkeley
How does quantum mechanics scale?

January 6, 2011 (New Orleans, LA)

Luca Trevisan, Stanford University
*Khot's unique games conjecture: its consequences and the evidence
for and against it*

Thomas Scanlon, University of California, Berkeley
Counting special points: logic, Diophantine geometry and transcendence theory

Ulrike Tillmann, Oxford University
Spaces of graphs and surfaces

David Nadler, Northwestern University
The geometric nature of the Fundamental Lemma

January 15, 2010 (San Francisco, CA)

Ben Green, University of Cambridge
*Approximate groups and their applications: work
of Bourgain, Gamburd, Helfgott and Sarnak*

David Wagner, University of Waterloo
Multivariate stable polynomials: theory and applications

Laura DeMarco, University of Illinois at Chicago
The conformal geometry of billiards

Michael Hopkins, Harvard University
-On the Kervaire Invariant Problem

January 7, 2009 (Washington, DC)

Matthew James Emerton, Northwestern University
Topology, representation theory and arithmetic: Three-manifolds and the Langlands program

Olga Holtz, University of California, Berkeley
Compressive sensing: A paradigm shift in signal processing

Michael Hutchings, University of California, Berkeley
From Seiberg-Witten theory to closed orbits of vector fields: Taubes's proof of the Weinstein conjecture

Frank Sottile, Texas A & M University
Frontiers of reality in Schubert calculus

January 8, 2008 (San Diego, California)

Günther Uhlmann, University of Washington
Invisibility

Antonella Grassi, University of Pennsylvania
Birational Geometry: Old and New

Gregory F. Lawler, University of Chicago
Conformal Invariance and 2-d Statistical Physics

Terence C. Tao, University of California, Los Angeles
Why are Solitons Stable?

January 7, 2007 (New Orleans, Louisiana)

Robert Ghrist, University of Illinois, Urbana-Champaign
Barcodes: The persistent topology of data

Akshay Venkatesh, Courant Institute, New York University
Flows on the space of lattices: work of Einsiedler, Katok and Lindenstrauss

Izabella Laba, University of British Columbia
From harmonic analysis to arithmetic combinatorics

Barry Mazur, Harvard University
The structure of error terms in number theory and an introduction to the Sato-Tate Conjecture

January 14, 2006 (San Antonio, Texas)

Lauren Ancel Myers, University of Texas at Austin
Contact network epidemiology: Bond percolation applied to infectious disease prediction and control

Kannan Soundararajan, University of Michigan, Ann Arbor
Small gaps between prime numbers

Madhu Sudan, MIT
Probabilistically checkable proofs

Martin Golubitsky, University of Houston
Symmetry in neuroscience

January 7, 2005 (Atlanta, Georgia)

Bryna Kra, Northwestern University
The Green-Tao Theorem on primes in arithmetic progression: A dynamical point of view

Robert McEliece, California Institute of Technology
Achieving the Shannon Limit: A progress report

Dusa McDuff, SUNY at Stony Brook
Floer theory and low dimensional topology

Jerrold Marsden, Shane Ross, California Institute of Technology
New methods in celestial mechanics and mission design

László Lovász, Microsoft Corporation
Graph minors and the proof of Wagner's Conjecture

January 9, 2004 (Phoenix, Arizona)

Margaret H. Wright, Courant Institute of Mathematical Sciences,
New York University
*The interior-point revolution in optimization: History, recent
developments and lasting consequences*

Thomas C. Hales, University of Pittsburgh
What is motivic integration?

Andrew Granville, Université de Montréal
It is easy to determine whether or not a given integer is prime

John W. Morgan, Columbia University
Perelman's recent work on the classification of 3-manifolds

January 17, 2003 (Baltimore, Maryland)

Michael J. Hopkins, MIT
Homotopy theory of schemes

Ingrid Daubechies, Princeton University
Sublinear algorithms for sparse approximations with excellent odds

Edward Frenkel, University of California, Berkeley
Recent advances in the Langlands Program

Daniel Tataru, University of California, Berkeley
The wave maps equation

2018 CURRENT EVENTS BULLETIN

Committee

The back cover graphic is reprinted courtesy of Andrei Okounkov.

Cover graphic associated with Richard D. James' talk courtesy of Xian Chen and Richard D. James.

Cover graphic associated with Craig L. Huneke's talk courtesy of Clebsch cubic image, © Greg Egan 2016.

Cover graphic associated with Isabelle Gallagher's talk courtesy of PLAINVIEW/GettyImages.

Cover graphic associated with Joshua A. Grochow's talk courtesy of Conan Chadbourne (www.conanchadbourne.com).