

What's new - 2007

Terence Tao

DEPARTMENT OF MATHEMATICS, UCLA, LOS ANGELES, CA
90095

E-mail address: tao@math.ucla.edu

The author is supported by NSF grant CCF-0649473 and a grant from the MacArthur foundation.

Contents

Preface	ix
Chapter 1. Open problems	1
§1.1. Best bounds for capsets	2
§1.2. Noncommutative Freiman theorem	4
§1.3. Mahler's conjecture for convex bodies	8
§1.4. Why global regularity for Navier-Stokes is hard	9
§1.5. Scarring for the Bunimovich stadium	10
§1.6. Triangle and diamond densities in large dense graphs	11
§1.7. What is a quantum honeycomb?	12
§1.8. Boundedness of the trilinear Hilbert transform	13
§1.9. Effective Skolem-Mahler-Lech theorem	14
§1.10. The parity problem in sieve theory	15
§1.11. Deterministic RIP matrices	16
§1.12. The nonlinear Carleson conjecture	17
Chapter 2. Expository articles	19
§2.1. Quantum mechanics and Tomb Raider	20
§2.2. Compressed sensing and single-pixel cameras	30
§2.3. Soft analysis, hard analysis, and the finite convergence principle	31

§2.4. The Lebesgue differentiation theorem and the Szemerédi regularity lemma	32
§2.5. Ultrafilters, nonstandard analysis, and epsilon management	33
§2.6. Dyadic models	34
§2.7. Nonfirstorderisability	35
§2.8. Amplification, arbitrage, and the tensor power trick	36
§2.9. The crossing number inequality	37
§2.10. Ratner’s theorems	38
§2.11. Unipotent elements of the Lorentz group, and conic sections	39
§2.12. John’s blowup theorem for the nonlinear wave equation	40
§2.13. Hilbert’s nullstellensatz	41
§2.14. The Hahn-Banach theorem, Mengers theorem, and Hellys theorem	42
Chapter 3. Lectures	43
§3.1. Simons Lecture Series: Structure and randomness	44
§3.2. Ostrowski lecture: The uniform uncertainty principle and compressed sensing	56
§3.3. Milliman lecture: Recent developments in arithmetic combinatorics	57
Bibliography	59

Preface

???

Chapter 1

Open problems

1.1. Best bounds for capsets

Perhaps my favourite open question is the problem on the maximal size of a *cap set* - a subset of \mathbf{F}_3^n (\mathbf{F}_3 being the finite field of three elements) which contains no lines, or equivalently no non-trivial arithmetic progressions of length three. As an upper bound, one can easily modify the proof of Roth's theorem [Ro1953] to show that cap sets must have size $O(3^n/n)$; see [Me1995]. This of course is better than the trivial bound of 3^n once n is large. In the converse direction, the trivial example $\{0, 1\}^n$ shows that cap sets can be as large as 2^n ; the current world record is $(2.2174\dots)^n$, held by Edel [Ed2004]. The gap between these two bounds is rather enormous; I would be very interested in either an improvement of the upper bound to $o(3^n/n)$, or an improvement of the lower bound to $(3 - o(1))^n$. (I believe both improvements are true, though a good friend of mine disagrees about the improvement to the lower bound.)

One reason why I find this question important is that it serves as an excellent model for the analogous question of finding large sets without progressions of length three in the interval $\{1, \dots, N\}$. Here, the best upper bound of $O(N\sqrt{\frac{\log \log N}{\log N}})$ is due to Bourgain [Bo1999] (with a recent improvement to $O(N\frac{(\log \log N)^2}{\log^{2/3} N})$ [Bo2008]), while the best lower bound of $Ne^{-C\sqrt{\log N}}$ is an ancient result of Behrend [Be1946]. Using the finite field heuristic (see Section 2.6) that \mathbf{F}_3^n "behaves like" $\{1, \dots, 3^n\}$, we see that the Bourgain bound should be improvable to $O(\frac{N}{\log N})$, whereas the Edel bound should be improvable to something like $3^n e^{-C\sqrt{n}}$. However, neither argument extends easily to the other setting. Note that a conjecture of Erdős asserts that any set of positive integers whose sum of reciprocals diverges contains arbitrarily long arithmetic progressions; even for progressions of length three, this conjecture is open, and is essentially equivalent (up to $\log \log$ factors) to the problem of improving the Bourgain bound to $o(\frac{N}{\log N})$.

The Roth bound of $O(3^n/n)$ appears to be the natural limit of the purely Fourier-analytic approach of Roth, and so any breakthrough would be extremely interesting, as it almost certainly would need a radically new idea. The lower bound might be improvable by some

sort of algebraic geometry construction, though it is not clear at all how to achieve this.

One can interpret this problem in terms of the wonderful game “Set”, in which case the problem is to find the largest number of cards one can put on the table for which nobody has a valid move. As far as I know, the best bounds on the cap set problem in small dimensions are the ones cited in [Ed2004].

There is a variant formulation of the problem which may be a little bit more tractable. Given any $0 < \delta \leq 1$, the fewest number of lines in a set of \mathbf{F}_3^n of density at least δ is known to be $(c(\delta) + o(1))3^{2n}$ for some $0 < c(\delta) \leq 1$; see [Cr2008]. The reformulated question is then to get as strong a bound on $c(\delta)$ as one can. For instance, the counterexample $0, 1^m \times \mathbf{F}_3^n$ shows that $c(\delta) \ll \delta^{\log_3/2 \cdot 9/2}$, while the Roth-Meshulam argument gives $c(\delta) \gg e^{-C/\delta}$.

1.1.1. Notes. This article was originally posted on Feb 23, 2007 at <http://terrytao.wordpress.com/2007/02/23/open-question-best-bounds-for-cap-sets/>

Thanks to Jordan Ellenberg for suggesting the density formulation of the problem.

1.2. Noncommutative Freiman theorem

This is another one of my favourite open problems, falling under the heading of *inverse theorems* in arithmetic combinatorics. “Direct” theorems in arithmetic combinatorics take a finite set A in a group or ring and study things like the size of its *sum set* $A + A := \{a + b : a, b \in A\}$ or *product set* $A \cdot A := \{ab : a, b \in A\}$. For example, a typical result in this area is the *sum-product theorem*, which asserts that whenever $A \subset \mathbf{F}_p$ is a subset of a finite field of prime order with $1 \leq |A| \leq p^{1-\delta}$, then

$$\max(|A + A|, |A \cdot A|) \geq |A|^{1+\varepsilon}$$

for some $\varepsilon = \varepsilon(\delta) > 0$. This particular theorem was first proven in [BoGIKo2006] with an earlier partial result in [BoKaTa2004]; more recent and elementary proofs with civilised bounds can be found in [TaVu2006], [GIKo2008], [Ga2008], [KaSh2008]. See Section 3.3.3 for further discussion.

In contrast, inverse theorems in this subject start with a hypothesis that, say, the sum set $A + A$ of an unknown set A is small, and try to deduce structural information about A . A typical goal is to completely classify all sets A for which $A+A$ has comparable size with A . In the case of finite subsets of integers, this is Freiman’s theorem [Fr1973], which roughly speaking asserts that if $|A + A| = O(|A|)$, if and only if A is a dense subset of a generalised arithmetic progression P of rank $O(1)$, where we say that A is a dense subset of B if $A \subset B$ and $|B| = O(|A|)$. (The “if and only if” has to be interpreted properly; in either the “if” or the “only if” direction, the implicit constants in the conclusion depends on the implicit constants in the hypothesis, but these dependencies are not inverses of each other.) In the case of finite subsets A of an arbitrary abelian group, we have the Freiman-Green-Ruzsa theorem [GrRu2007], which asserts that $|A + A| = O(|A|)$ if and only if A is a dense subset of a sum $P+H$ of a finite subgroup H and a generalised arithmetic progression P of rank $O(1)$.

One can view these theorems as a “robust” or “rigid” analogue of the classification of finite abelian groups. It is well known that finite

abelian groups are direct sums of cyclic groups; the above results basically assert that finite sets that are “nearly groups” in that their sum set is not much larger than the set itself, are (dense subsets of) the direct sums of cyclic groups and a handful of arithmetic progressions.

The open question is to formulate an analogous conjectural classification in the non-abelian setting, thus to conjecture a reasonable classification of finite sets A in a multiplicative group G for which $|A \cdot A| = O(|A|)$. Actually for technical reasons it may be better to use $|A \cdot A \cdot A| = O(|A|)$; I refer to this condition by saying that A has *small tripling*. (Note for instance that if H is a subgroup and x is not in the normaliser of H , then $H \cup \{x\}$ has small doubling but not small tripling. On the other hand, small tripling is known to imply small quadrupling, etc., see e.g. [TaVu2006].) Note that I am not asking for a theorem here - just even stating the right conjecture would be major progress! An if and only if statement might be too ambitious initially: a first step would be to obtain a slightly looser equivalence, creating for each group G and fixed $\varepsilon > 0$ a class \mathcal{P} of sets (depending on some implied constants) for which the following two statements are true:

- (i) If A is a finite subset of G with small tripling, then A is a dense subset of $O(|A|^\varepsilon)$ left- or right- translates of a set P of the form \mathcal{P} .
- (ii) If P is a set of the form \mathcal{P} , then there exists a dense subset A of P with small tripling (possibly with a loss of $O(|A|^\varepsilon)$ in the tripling constant).

An obvious candidate for \mathcal{P} is the inverse image in $N(H)$ of a ball in a nilpotent subgroup of $N(H)/H$ of step $O(1)$, where H is a finite subgroup of G and $N(H)$ is the normaliser of H ; note that property (ii) is then easy to verify. Let us call this the *standard candidate*. I do not know if this candidate fully suffices, but it seems to be a reasonably good candidate nevertheless. In this direction, some partial results are known:

- For abelian groups G , from the Freiman-Green-Ruzsa theorem, we know that the standard candidate suffices.

- For $G = SL_2(\mathbf{C})$, we know from work of Elekes and Király[**EIKi2001**] and Chang[**Ch2008**] that the standard candidate suffices.
- For $G = SL_2(\mathbf{F}_p)$, there is a partial result of Helfgott [**He2008**], which (roughly speaking) asserts that if A has small tripling, then either A is a dense subset of all of G , or is contained in a proper subgroup of G . It is likely that by pushing this analysis further one would obtain a candidate for \mathcal{P} in this case.
- For $G = SL_3(\mathbf{Z})$, a result of Chang[**Ch2008**] shows that if A has small tripling, then it is contained in a nilpotent subgroup of G .
- For the lamplighter group $G = \mathbf{Z}/2\mathbf{Z} \wr \mathbf{Z}$, there is a partial result of Lindenstrauss[**Li2001**] which (roughly speaking) asserts that if A has small tripling, then A cannot be nearly invariant under a small number of shifts. It is also likely that by pushing the analysis further here one would get a good candidate for \mathcal{P} in this case.
- For a free non-abelian group, we know (since the free group embeds into $SL_2(\mathbf{C})$) that the standard candidate suffices; a much stronger estimate in this direction was recently obtained by Razborov [**Ra2008**].
- For a Heisenberg group G of step 2, there is a result of myself[**Ta2008**], which shows that sets of small tripling also have small tripling in the abelianisation of G , and are also essentially closed under the antisymmetric form that defines G . This, in conjunction with the Freiman-Green-Ruzsa theorem, gives a characterisation, at least in principle, but it is not particularly explicit, and it may be of interest to work it out further.
- For G torsion-free, there is a partial result of Hamidoune, Lladó, and Serra[**HaLiSe1998**], which asserts that $|A \cdot A| \geq 2|A| - 1$, and that if $|A \cdot A| \leq 2|A|$ then A is a geometric progression with at most one element removed; in particular, the standard candidate suffices in this case.

These examples do not seem to conclusively suggest what the full classification should be. Based on analogy with the classification of finite simple groups, one might expect the full classification to be complicated, and enormously difficult to prove; on the other hand, the fact that we are in a setting where we are allowed to lose factors of $O(1)$ may mean that the problem is in fact significantly less difficult than that classification. (For instance, all the sporadic simple groups have size $O(1)$ and so even the monster group is “negligible”.) Nevertheless, it seems possible to make progress on explicit groups, in particular refining the partial results already obtained for the specific groups mentioned above. An even closer analogy may be with Gromov’s theorem [Gr1981] on groups of polynomial growth; in particular, the recent effective proof of this theorem by Kleiner [Kl2008] may prove to be relevant for this problem.

1.2.1. Notes. This article was originally posted on Mar 2, 2007 at <http://terrytao.wordpress.com/2007/03/02/open-question-noncommutative-freiman-theorem/>

Thanks to Akshay Venkatesh and Elon Lindenstrauss to pointing out the analogy with Gromov’s theorem.

1.3. Mahler's conjecture for convex bodies

1.4. Why global regularity for Navier-Stokes is hard

1.5. Scarring for the Bunimovich stadium

1.6. Triangle and diamond densities in large dense graphs¹

**1.6. Triangle and diamond densities in large
dense graphs**

1.7. What is a quantum honeycomb?

1.8. Boundedness of the trilinear Hilbert transform

1.9. Effective Skolem-Mahler-Lech theorem

1.10. The parity problem in sieve theory

1.11. Deterministic RIP matrices

1.12. The nonlinear Carleson conjecture

Chapter 2

Expository articles

2.1. Quantum mechanics and Tomb Raider

Quantum mechanics has a number of weird consequences, but in this article I will focus on three (inter-related) ones:

- Objects can behave both like particles (with definite position and a continuum of states) and waves (with indefinite position and (in confined situations) quantised states);
- The equations that govern quantum mechanics are deterministic, but the standard interpretation of the solutions (the *Copenhagen interpretation*) of these equations is probabilistic; and
- If instead one applies the laws of quantum mechanics literally at the macroscopic scale (via the *relative state interpretation*, more popularly known as the *many worlds interpretation*), then the universe itself must split into the superposition of many distinct “worlds”.

What I will attempt to do here is to use the familiar concept of a computer game as a classical conceptual model with which to capture these non-classical phenomena. The exact choice of game is not terribly important, but let us pick *Tomb Raider* - a popular game from about ten years ago, in which the heroine, Lara Croft, explores various tombs and dungeons, solving puzzles and dodging traps, in order to achieve some objective. It is quite common for Lara to die in the game, for instance by failing to evade one of the traps. (I should warn that this analogy will be rather violent on certain computer-generated characters.)

The thing about such games is that there is an “internal universe”, in which Lara interacts with other game elements, and occasionally is killed by them, and an “external universe”, where the computer or console running the game, together with the human who is playing the game, resides. While the game is running, these two universes run more or less in parallel; but there are certain operations, notably the “save game” and “restore game” features, which disrupt this relationship. These operations are utterly mundane to people like us who reside in the external universe, but it is an interesting thought experiment to view them from the perspective of someone like Lara,

in the internal universe. (I will eventually try to connect this with quantum mechanics, but please be patient for now.) Of course, for this we will need to presume that the Tomb Raider game is so advanced that Lara has levels of self-awareness and artificial intelligence which are comparable to our own. In particular, we will imagine that Lara is independent enough to play the game without direct intervention from the player, whose role shall be largely confined to that of saving, loading, and observing the game.

Imagine first that Lara is about to navigate a tricky rolling boulder puzzle, when she hears a distant rumbling sound - the sound of her player saving her game to disk. From the perspective of the player, we suppose that what happens next is the following: Lara navigates the boulder puzzle but fails, being killed in the process; then the player restores the game from the save point and then Lara successfully makes it through the boulder puzzle.

Now, how does the situation look from Lara's point of view? At the save point, Lara's reality diverges into a superposition of two non-interacting paths, one in which she dies in the boulder puzzle, and one in which she lives. (Yes, just like that cat.) Her future becomes indeterministic. If she had consulted with an infinitely prescient oracle before reaching the save point as to whether she would survive the boulder puzzle, the only truthful answer this oracle could give is "50% yes, and 50% no".

This simple example shows that the *internal* game universe can become indeterministic, even though the *external* one might be utterly deterministic. However, this example does not fully capture the weirdness of quantum mechanics, because in each one of the two alternate states Lara could find herself in (surviving the puzzle or being killed by it), she does not experience any effects from the other state at all, and could reasonably assume that she lives in a classical, deterministic universe.

So, let's make the game a bit more interesting. Let us assume that every time Lara dies, she leaves behind a corpse in that location for future incarnations of Lara to encounter. Then Lara will start noticing the following phenomenon (assuming she survives at all): whenever she navigates any particularly tricky puzzle, she usually encounters a

number of corpses which look uncannily like herself. This disturbing phenomenon is difficult to explain to Lara using a purely classical deterministic model of reality; the simplest (and truest) explanation that one can give her is a “many-worlds” interpretation of reality, and that the various possible states of Lara’s existence have some partial interaction with each other. Another valid (and largely equivalent) explanation would be that every time Lara passes a save point to navigate some tricky puzzle, Lara’s “particle-like” existence splits into a “wave-like” superposition of Lara-states, which then evolves in a complicated way until the puzzle is resolved one way or the other, at which point Lara’s wave function “collapses” in a non-deterministic fashion back to a particle-like state (which is either entirely alive or entirely dead).

Now, in the real world, it is only microscopic objects such as electrons which seem to exhibit this quantum behaviour; macroscopic objects, such as you and I, do not directly experience the kind of phenomena that Lara does, and we cannot interview individual electrons to find out their stories either. Nevertheless, by studying the statistical behaviour of large numbers of microscopic objects we can indirectly infer their quantum nature via experiment and theoretical reasoning. Let us again use the Tomb Raider analogy to illustrate this. Suppose now that Tomb Raider does not only have Lara as the main heroine, but in fact has a large number of playable characters, who explore a large number deadly tombs, often with fatal effect (and thus leading to multiple game restores). Let us suppose that inside this game universe there is also a scientist (let’s call her Jacqueline) who studies the behaviour of these adventurers going through the tombs. However, Jacqueline does not experience the tombs directly, nor does she actually communicate with any of these adventurers. Each tomb is explored by only one adventurer; regardless of whether she lives or dies, the tomb is considered “used up”.

Jacqueline observes several types of trapped tombs in her world, and gathers data as to how likely an adventurer is to survive any given type of tomb. She learns that each type of tomb has a fixed survival rate - e.g. she may observe that a tomb of type A has a 20% survival rate, whilst a tomb of type B has a 50% survival rate - but

that it seems impossible to predict with any certainty whether any given adventurer will survive any given type of tomb. So far, this is something which could be explained classically; each tomb may have a certain number of lethal traps in them, and whether an adventurer survives these traps or not may entirely be due to random chance or other “hidden variables”.

But then Jacqueline encounters a mysterious *quantisation* phenomenon: the survival rate for various tombs are always one of the numbers 100%, 50%, 33.3...%, 25%, 20%, ...; in other words, the “frequency” of success for a tomb is always of the form $1/n$ for some integer n . This phenomenon would be difficult to explain in a classical universe, since the effects of random chance should be able to produce a continuum of survival probabilities.

Here’s what is going on. In order for Lara (or any other adventurer) to survive a tomb of a given type, she needs to stack together a certain number of corpses together to reach a certain switch; if she cannot attain that level of “constructive interference” to reach that switch, she dies. The type of tomb determines exactly how many corpses are needed; for instance, a tomb of type A might requires four corpses to be stacked together. Then the player who is playing Lara will have to let her die four times before she can successfully get through the tomb; and so from her perspective, Lara’s chances of survival are only 20%. In each possible state of the game universe, there is only one Lara which goes into the tomb, who either lives or dies; but her survival rate here is what it is because of her interaction with other states of Lara (which Jacqueline cannot see directly, as she does not actually enter the tomb).

In our own reality, a familiar example of this type of quantum effect is the fact that each atom (e.g. sodium or neon) can only emit certain wavelengths of light (which end up being quantised somewhat analogously to the survival probabilities above); for instance, sodium only emits yellow light, neon emits blue, and so forth. The electrons in such atoms, in order to emit such light, are in some sense clambering over skeletons of themselves to do so; the more commonly given explanation is that the electron is behaving like a wave within the

confines of an atom, and thus can only oscillate at certain frequencies (similarly to how a plucked string of a musical instrument can only exhibit a certain set of wavelengths, which coincidentally are also proportional to $1/n$ for integer n). Mathematically, this “quantisation” of frequency can be computed using the bound states of a Schrödinger operator with potential. [I will not attempt to stretch the Tomb Raider analogy so far as to try to model the Schrödinger equation! In particular, the complex phase of the wave function - which is a fundamental feature of quantum mechanics - is not easy at all to motivate in a classical setting.]

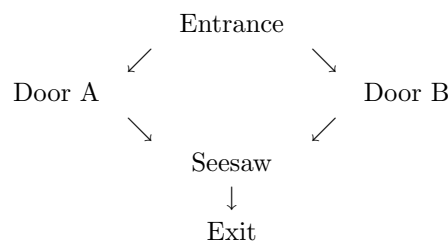
Now let’s use the Tomb Raider analogy to explain why microscopic objects (such as electrons) experience quantum effects, but macroscopic ones (or even mesoscopic ones, such as large molecules) seemingly do not. Let’s assume that Tomb Raider is now a two-player co-operative game, with two players playing two characters (let’s call them Lara and Indiana) as they simultaneously explore different parts of their world. The players can choose to save the entire game, and then restore back to that point; this resets both Lara and Indiana back to the state they were in at that save point.

Now, this game still has the strange feature of corpses of Lara and Indiana from previous games appearing in later ones. However, we assume that Lara and Indiana are *entangled* in the following way: if Lara is in tomb A and Indiana is in tomb B, then Lara and Indiana can each encounter corpses of their respective former selves, but only if *both* Lara *and* Indiana died in tombs A and B respectively in a single previous game. If in a previous game, Lara died in tomb A and Indiana died in tomb C, then this time round, Lara will not see any corpse (and of course, neither will Indiana). (This entanglement can be described a bit better by using *tensor products*; rather than saying that Lara died in A and Indiana died in B, one should instead think of Lara \otimes Indiana dying in $|A\rangle \otimes |B\rangle$, which is a state which is orthogonal to $|A\rangle \otimes |C\rangle$.) With this type of entanglement, one can see that there is going to be significantly less “quantum weirdness” going on; Lara and Indiana, adventuring separately but simultaneously, are going to encounter far fewer corpses of themselves than

Lara adventuring alone would. And if there were many many adventurers entangled together exploring simultaneously, the quantum effects drop to virtually nothing, and things now look classical unless the adventurers are somehow organised to “resonate” in a special way (much as *Bose-Einstein condensates* operate in our own world).

The Tomb Raider analogy is admittedly not a perfect model for quantum mechanics. In the latter, the various possible basis states of a system interfere with each other via linear superposition of their complex phases, whereas in the former, the basis states interfere in an ordered nonlinear fashion, with the states associated to earlier games influencing the states of later games, but not vice versa. Another very important feature of quantum mechanics - namely, the ability to change the set of basis states used to decompose the full state of the system - does not have a counterpart in the Tomb Raider model. Nevertheless, this model is still sufficiently non-classical (when viewed from the internal universe) to construct some partial analogues of well-known quantum phenomena. We illustrate this with two more examples.

2.1.1. A two-slit experiment. The famous *two-slit experiment* involves a particle, such as an electron, being sent through a barrier with two slits. It can turn out that the particle can reach a certain destination beyond the barrier if one of the slits is covered up, but that this destination becomes inaccessible if both slits are opened up. A somewhat similar phenomenon can be simulated in the Tomb Raider universe described above, using the following kind of tomb:



Suppose that Door A and Door B are one-way; on reaching the antechamber, Lara has to choose one of the two doors, and on doing so, is stuck on one end of the seesaw. Suppose that the seesaw

is lethally trapped in such a way that one has to keep the seesaw balanced for, say, five minutes, otherwise the trap is set off, killing anyone on either side of the seesaw. Classically, it would be impossible for Lara to reach the exit, as she can only be on one side of the seesaw and so cannot maintain that seesaw's balance. But if she goes through once, say on side A, and then dies, then when the game is restored, she can go in on side B and balance herself against the corpse from the previous game to defuse the trap. So she in fact has up to a 50% chance of survival here. (Actually, if she chooses a door randomly each time, and the player restores the game until she makes it through, the net chance of survival is only $2 \ln 2 - 1 = 38.6 \dots \%$ - why?) On the other hand, if either of the doors is locked in advance, then her survival rate drops to 0%.

This does not have an easy classical explanation within the game universe, even with hidden variables, at least if you make the locality assumption that Lara can only go through one of the two one-way doors, and if you assume that the locks have no effect other than to stop Lara from choosing one of the doors.

2.1.2. Bell's inequality violation. Before we begin this example, let us recall some inequalities from classical probability. If A and B are two events, then we have the inclusion-exclusion identity

$$\mathbb{P}(A \vee B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \wedge B)$$

where $A \vee B$ is the event that at least one of A and B occur, and $A \wedge B$ is the event that A and B both occur. Since $\mathbb{P}(A \vee B)$ clearly cannot exceed 1, we conclude that

$$(2.1) \quad \mathbb{P}(A \wedge B) \geq \mathbb{P}(A) + \mathbb{P}(B) - 1.$$

Note that this inequality holds regardless of whether A and B are independent or not.

Iterating (2.1), we conclude that for any three events A, B, C , we have

$$(2.2) \quad \mathbb{P}(A \wedge B \wedge C) \geq \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - 2.$$

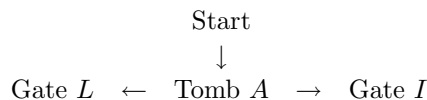
Now let $l_1, l_2, i_1, i_2 \in \{0, 1\}$ be four random variables (possibly dependent). Observe that if the event occurs that $l_1 = i_1$, $l_1 = i_2$, and

$l_2 = i_1$, then we necessarily have $l_2 = i_2$. We conclude that

$$(2.3) \quad \mathbb{P}(l_2 = i_2) \geq \mathbb{P}(l_1 = i_1) + \mathbb{P}(l_1 = i_2) + \mathbb{P}(l_2 = i_1) - 2.$$

Again, we emphasise that this inequality must hold regardless of whether l_1, l_2, i_1, i_2 are independent or not. This inequality is a variant of the famous *Bell inequality*, and is known as the CHSH inequality.

We will now create a Tomb Raider experiment that shows that the internal game reality cannot be modeled by classical probability, at least if one insists that only one instance of the game universe exists. We will need two game characters, Lara and Indiana, who are exploring this map:



Gate L and Gate I both have two up-down switches which either character can manipulate into any of the four positions before trying to open the gate: up-up, up-down, down-up, or down-down. However, the gates are trapped: only two of the positions allow the gate to be opened safely; the other two positions will ensure that the gate electrocutes whoever is trying to open it. Lara and Indiana know that the gates are anti-symmetric: if one flips both switches then that toggles whether the gate is safe or not (e.g. if down-up is safe, then up-down electrocutes). But they do not know exactly which combinations are safe.

Lara and Indiana (starting in the position “Start”) desperately need to open both gates before a certain time limit, but do not know which of the combinations are safe. They have just enough time for Lara to go to Gate L through Tomb A , and for Indiana to go to Gate I through Tomb A , but there is not enough time for Lara to communicate to Indiana what she sees at Gate L , or conversely.

They believe (inaccurately, as it turns out) that inside Tomb A , there is inscribed a combination (of one of the four positions) which will safely open both gates. Their plan is to jointly go to Tomb A , find the combination, write that combination down on two piece of paper (one for Lara, one for Indiana), and then Lara and Indiana will

travel separately to Gate L and Gate I to try that combination to unlock both gates. At this point, the player saves the game and play continues repeatedly from this restore point. We re-emphasise that the player actually has no control over Lara and Indianas actions; they are independent AIs, following the plan described above.

Unfortunately for Lara and Indiana, the combination in Tomb A is simply a random combination - up-up, up-down, down-up, and down-down are each 25% likely to be found in Tomb A . In truth, the combinations to Gate L and Gate I have been set by Jacqueline. Jacqueline has set Gate L to one of the following two settings:

- Setting L_1 : Gate L will open safely if the switches are up-up or up-down, but electrocutes if the switches are down-up or down-down
- Setting L_2 : Gate L will open safely if the switches are up-up or down-up, but electrocutes if the switches are up-down or down-down.

Similarly, Jacqueline has set Gate I to one of the following two settings:

- Setting I_1 : Gate I will open safely if the switches are up-up or up-down, but electrocutes if the switches are down-up or down-down.
- Setting I_2 : Gate I will open safely if the switches are up-down or down-down, but electrocutes if the switches are down-up or up-up.

Note that these settings obey the anti-symmetry property mentioned earlier.

Jacqueline sets Gate L to setting L_a for some $a = 1, 2$, and Gate I to setting I_b for some $b = 1, 2$, and measures the probability p_{ab} of the event that Lara and Indiana both survive, or both die, thus computing four numbers $p_{11}, p_{12}, p_{21}, p_{22}$. (To do this, one would have to assume that the experiment can be repeated a large number of times, for instance by assuming that a large number of copies of these tombs and gates exist across the game universe, with a different pair of adventurers exploring each such copy.)

Jacqueline does not know the contents (or “hidden variables”) of Tomb A, and does not know what Lara and Indiana’s strategy is to open the gates (in particular, the strategy could be randomly chosen rather than deterministic). However, if she assumes that communication between Lara and Indiana is local (thus Lara cannot transmit information about Gate L to Indiana at Gate I, or vice versa), and that the universe is classical (in particular, that no multiple copies of the universe exist), then she can deduce a certain theoretical inequality connecting the four numbers $p_{11}, p_{12}, p_{21}, p_{22}$. Indeed, she can write $p_{ab} = \mathbb{P}(l_a = i_b)$, where l_a is the random variable that equals 1 when Lara sets the switches of gate L to a position which is safe for L_a and 0 otherwise, and similarly i_b is the random variable that equals 1 when Indiana sets the switches of gate I to a position which is safe for I_b and 0 otherwise. Applying (2.3), we conclude that

$$(2.4) \quad p_{22} \geq p_{11} + p_{12} + p_{21} - 2$$

regardless of what goes on in Tomb A, and regardless of what strategy Indiana and Lara execute.

We now show that in the actual Tomb Raider universe, the inequality (2.4) is violated - which proves to Jacqueline that her universe must either be non-local (with instantaneous information transmission) or non-classical (with the true state of the game universe being described as a superposition of more than one classical state).

2.1.3. Notes. This article was originally posted on Feb 26, 2007 at <http://terrytao.wordpress.com/2007/02/26/quantum-mechanics-and-tomb-raider/>

It was derived from an interesting conversation I had several years ago with my friend Jason Newquist, on trying to find some intuitive analogies for the non-classical nature of quantum mechanics.

2.2. Compressed sensing and single-pixel cameras

2.3. Soft analysis, hard analysis, and the finite convergence principle

2.3. Soft analysis, hard analysis, and the finite convergence principle

2.4. The Lebesgue differentiation theorem and the Szemerédi regularity lemma

2.5. Ultrafilters, nonstandard analysis, and epsilon management

**2.5. Ultrafilters, nonstandard analysis, and
epsilon management**

2.6. Dyadic models

2.7. Nonfirstorderisability

2.8. Amplification, arbitrage, and the tensor power trick

2.9. The crossing number inequality

2.10. Ratner's theorems

2.11. Unipotent elements of the Lorentz group, and conic sections

**2.11. Unipotent elements of the Lorentz group,
and conic sections**

**2.12. John's blowup theorem for the nonlinear
wave equation**

2.13. Hilbert's nullstellensatz

2.14. The Hahn-Banach theorem, Mengers theorem, and Hellys theorem

Chapter 3

Lectures

3.1. Simons Lecture Series: Structure and randomness

On Apr 5-7, 2007, I gave one of the Simons Lecture Series at MIT (the other lecture series was given by David Donoho). I gave three lectures, each expounding on some aspects of the theme “the dichotomy between structure and randomness” (see also my ICM talk [Ta2006], [Ta2006a] on this topic). This theme seems to pervade many of the areas of mathematics that I work in, and my lectures aim to explore how this theme manifests itself in several of these. In the first lecture, I describe the dichotomy as it appears in Fourier analysis and in number theory. In the second, I discuss the dichotomy in ergodic theory and graph theory, while in the third, I discuss PDE.)

3.1.1. Structure and randomness in Fourier analysis and number theory. The “dichotomy between structure and randomness” seems to apply in circumstances in which one is considering a “high-dimensional” class of objects (e.g. sets of integers, functions on a space, dynamical systems, graphs, solutions to PDE, etc.). For sake of concreteness, let us focus today on sets of integers (later lectures will focus on other classes of objects). There are many different types of objects in these classes, however one can broadly classify them into three categories:

- *Structured* objects - objects with a high degree of predictability and algebraic structure. A typical example are the odd integers $A := \{\dots, -3, -1, 1, 3, 5, \dots\}$. Note that if some large number n is known to lie in A , this reveals a lot of information about whether $n + 1$, $n + 2$, etc. will also lie in A . Structured objects are best studied using the tools of *algebra* and *geometry*.
- *Pseudorandom* objects - the opposite of structured; these are highly unpredictable and totally lack any algebraic structure. A good example is a randomly chosen set B of integers, in which each element n lies in B with an independent probability of $1/2$. (One can imagine flipping a coin for each integer n , and defining B to be the set of n for which the coin flip resulted in heads.) Note that if some integer n is

known to lie in B , this conveys no information whatsoever about the relationship of $n + 1$, $n + 2$, etc. with respect to B . Pseudorandom objects are best studied using the tools of analysis and probability.

- *Hybrid* sets - sets which exhibit some features of structure and some features of pseudorandomness. A good example is the primes $P := 2, 3, 5, 7, \dots$. The primes have some obvious structure in them: for instance, the prime numbers are all positive, they are all odd (with one exception), they are all adjacent to a multiple of six (with two exceptions), and their last digit is always 1, 3, 7, or 9 (with two exceptions). On the other hand, there is evidence that the primes, despite being a deterministic set, behave in a very “pseudorandom” or “uniformly distributed” manner. For instance, from the prime number theorem in arithmetic progressions we know that the last digits of large prime numbers are uniformly distributed in the set $\{1, 3, 7, 9\}$; thus, if N is a large integer, the number of primes less than N ending in (say) 3, divided by the total number of primes less than N , is known to converge to $1/4$ in the limit as N goes to infinity. In order to study hybrid objects, one needs a large variety of tools: one needs tools such as algebra and geometry to understand the structured component, one needs tools such as analysis and probability to understand the pseudorandom component, and one needs tools such as decompositions, algorithms, and evolution equations to separate the structure from the pseudorandomness.

A recurring question in many areas of analysis is the following: given a specific object (such as the prime numbers), can one determine precisely what the structured components are within the object, and how pseudorandom the remaining components of the object are? One reason for asking this question is that it often helps one compute various *statistics* (averages, sums, integrals, correlations, norms, etc.) of the object being studied. For instance, one can ask for how many twin pairs $\{n, n + 2\}$, with n between 1 and N , one can find within a given set. In the structured set A given above, the answer is roughly

$N/2$. For the random set B given above, the answer is roughly $N/4$; thus one sees that while A and B have exactly the same density (namely, $1/2$), their statistics are rather different due to the fact that one is structured and one is random. As for the prime numbers, nobody knows for certain what the answer is (although the Hardy-Littlewood prime tuples conjecture [HaLi1923] predicts the answer to be roughly $1.32 \frac{N}{\log^2 N}$), because we do not know enough yet about the pseudorandomness of the primes. On the other hand, the parity structure of the prime numbers is enough to show that the number of *adjacent* pairs $\{n, n + 1\}$ in the primes is exactly one: $\{2, 3\}$.

The problem of determining exactly what the structured and pseudorandom components are of any given object is still largely intractable. However, what we have learnt in many cases is that we can at least show that an arbitrary object can be *decomposed* into some structured component and some pseudorandom component. Also there is often an *orthogonality* property (or *dichotomy*): if an object is orthogonal (or has small correlation with) all structured objects, then it is necessarily pseudorandom, and vice versa. Finally, we are sometimes lucky enough to be able to *classify* all the structured objects which are relevant for any given problem (e.g. computing a particular statistic). In such cases, one merely needs (in principle) to compute how the given object correlates with each member in one's list of structured objects in order to determine what the desired statistic is. This is often simpler (though still non-trivial) than computing the statistic directly.

To illustrate these general principles, let us focus now on a specific area in analytic number theory, namely that of finding additive patterns in the prime numbers $\{2, 3, 5, 7, \dots\}$. Despite centuries of progress on these problems, many questions are still unsolved, for instance:

- (Twin prime conjecture) There are infinitely many positive integers n such that $n, n + 2$ are both prime.
- (Sophie Germain prime conjecture) There are infinitely many positive integers n such that $n, 2n + 1$ are both prime.

- (Even Goldbach conjecture) For every even number $N \geq 4$, there is a natural number n such that $n, N - n$ are both prime.

On the other hand, we do have some positive results:

- (Vinogradov’s theorem)[Vi1937] For every sufficiently large odd number N , there are positive integers n, n' such that $n, n', N - n - n'$ are all prime. (The best explicit bound currently known for “sufficiently large” is $N \geq 10^{1346}$ [LiWa2002]; the result has also been verified for $7 \leq N \leq 10^{20}$ [Sa1998].)
- (van der Corput’s theorem)[vdC1939] There are infinitely many positive integers n, n' such that $n, n + n', n + 2n'$ are all prime.
- (Green-Tao theorem)[GrTa2008] For any positive integer k , there are infinitely many positive integers n, n' such that $n, n + n', \dots, n + (k - 1)n'$ are all prime.
- (A polynomial generalisation) For any integer-valued polynomials $P_1(n), \dots, P_k(n)$ with $P_1(0) = \dots = P_k(0) = 0$, there are infinitely many positive integers n, n' such that $n + P_1(n'), \dots, n + P_k(n')$ are all prime.

As a general rule, it appears that it is feasible (after non-trivial effort) to find patterns in the primes involving two or more degrees of freedom (as described by the parameters n, n' in above examples), but we still do not have the proper technology for finding patterns in the primes involving only one degree of freedom n . (This is of course an oversimplification; for instance, the pattern $n, n + 2, n', n' + 2$ has two degrees of freedom, but finding infinitely many of these patterns in the primes is equivalent to the twin prime conjecture, and thus presumably beyond current technology. If however one makes a non-degeneracy assumption, one can make the above claim more precise; see [GrTa2008b].)

One useful tool for establishing some (but not all) of the above positive results is Fourier analysis (which in this context is also known as the *Hardy-Littlewood circle method*). Rather than give the textbook presentation of that method here, let us try to motivate why Fourier analysis is an essential feature of many of these problems from

the perspective of the dichotomy between structure and randomness, and in particular viewing structure as an obstruction to computing statistics which needs to be understood before the statistic can be accurately computed.

To treat many of the above questions concerning the primes in a unified manner, let us consider the following general setting. We consider k affine-linear forms $\psi_1, \dots, \psi_k : \mathbb{Z}^r \rightarrow \mathbb{Z}$ on r integer unknowns, and ask

Question 3.1. *Does there exist infinitely many r -tuples $\vec{n} = (n_1, \dots, n_r) \in \mathbb{Z}_+^r$ of positive integers such that $\psi_1(\vec{n}), \dots, \psi_k(\vec{n})$ are simultaneously prime?*

For instance, the twin prime conjecture is the case when $k = 2$, $r = 1$, $\psi_1(n) = n$, and $\psi_2(n) = n + 2$; van der Corput's theorem is the case when $k = 3$, $r = 2$, and $\psi_j(n, n') = n + (j - 1)n'$ for $j = 0, 1, 2$; and so forth.

Because of the “obvious” structures in the primes, the answer to the above question can be “no”. For instance, since all but one of the primes are odd, we know that there are not infinitely many patterns of the form $n, n + 1$ in the primes, because it is not possible for $n, n + 1$ to both be odd. More generally, given any prime q , we know that all but one of the primes is coprime to q . Hence, if it is not possible for $\psi_1(\vec{n}), \dots, \psi_k(\vec{n})$ to all be coprime to q , the answer to the above question is basically no (modulo some technicalities which I wish to gloss over) and we say that there is an *obstruction at q* . For instance, the pattern $n, n + 1$ has an obstruction at 2. The pattern $n, n + 2, n + 4$ has no obstruction at 2, but has an obstruction at 3, because it is not possible for $n, n + 2, n + 4$ to all be coprime to 3. And so forth.

Another obstruction comes from the trivial observation that the primes are all positive. Hence, if it is not possible for $\psi_1(\vec{n}), \dots, \psi_k(\vec{n})$ to all be positive for infinitely many values of \vec{n} , then we say that there is an *obstruction at infinity*, and the answer to the question is again “no” in this case. For instance, for any fixed N , the pattern $n, N - n$ can only occur finitely often in the primes, because there are only finitely many n for which $n, N - n$ are both positive.

It is conjectured that these “local” obstructions are the only *obstructions* to solvability of the above question. More precisely, we have

Conjecture 3.2. (*Dickson’s conjecture*)[Di1904] *If there are no obstructions at any prime q , and there are no obstructions at infinity, then the answer to the above question is “yes”.*

This conjecture would imply the twin prime and Sophie Germain conjectures, as well as the Green-Tao theorem; it also implies the Hardy-Littlewood prime tuples conjecture[HaLi1923] as a special case. There is a quantitative version of this conjecture which predicts a more precise count as to how many solutions there are in a given range, and which would then also imply Vinogradov’s theorem, as well as Goldbach’s conjecture (for sufficiently large N); see [GrTa2008b] for further discussion. As one can imagine, this conjecture is still largely unsolved, however there are many important special cases that have now been established - several of which were achieved via the Hardy-Littlewood circle method.

One can view Dickson’s conjecture as an *impossibility statement*: that it is impossible to find any other obstructions to solvability for linear patterns in the primes than the obvious local obstructions at primes q and at infinity. (It is also a good example of a *local-to-global principle*, that local solvability implies global solvability.) Impossibility statements have always been very difficult to prove - one has to locate all possible obstructions to solvability, and eliminate each one of them in turn. In particular, one has to exclude various exotic “conspiracies” between the primes to behave in an unusually structured manner that somehow manages to always avoid all the patterns that one is seeking within the primes. How can one disprove a conspiracy?

To give an example of what such a “conspiracy” might look like, consider the twin prime conjecture, that of finding infinitely many pairs $n, n + 2$ which are both prime. This pattern encounters no obstructions at primes q or at infinity and so Dickson’s conjecture predicts that there should be infinitely many such patterns. In particular, there are no obstructions at 3 because prime numbers can equal 1 or 2 mod 3, and it is possible to find pairs $n, n + 2$ which also have this property. But suppose that it transpired that all but

finitely many of the primes ended up being $2 \pmod 3$. From looking at tables of primes this seems to be unlikely, but it is not immediately obvious how to disprove it; it could well be that once one reaches, say, 10^{100} , there are no more primes equal to $1 \pmod 3$. If this unlikely "conspiracy" in the primes was true, then there would be only finitely many twin primes. Fortunately, we have *Dirichlet's theorem*, which guarantees infinitely many primes equal to $a \pmod q$ whenever a, q are coprime, and so we can rule out this particular type of conspiracy. (This does strongly suggest, though, that knowledge of Dirichlet's theorem is a necessary but not sufficient condition in order to solve the twin prime conjecture.) But perhaps there are other conspiracies that one needs to rule out also?

To look for other conspiracies that one needs to eliminate, let us rewrite the conspiracy "all but finitely many of the primes are $2 \pmod 3$ " in the more convoluted format

$$0.6 < \left\{ \frac{1}{3}p \right\} < 0.7 \text{ for all but finitely many primes } p$$

where $\{x\}$ is the fractional part of x . This type of conspiracy can now be generalised; for instance consider the statement

$$(3.1) \quad 0 < \{\sqrt{2}p\} < 0.01 \text{ for all but finitely many primes } p$$

Again, such a conspiracy seems very unlikely - one would expect these fractional parts to be uniformly distributed between 0 and 1, rather than concentrate all in the interval $[0, 0.01]$ - but it is hard to rule this conspiracy out *a priori*. And if this conspiracy (3.1) was in fact true, then the twin prime conjecture would be false, as can be quickly seen by considering the identity

$$\{\sqrt{2}(n+2)\} - \{\sqrt{2}n\} = 2\sqrt{2} \pmod 1,$$

which forbids the two fractional parts on the left-hand side to simultaneously fall in the interval $[0, 0.01]$. Thus, in order to solve the twin prime conjecture, one must rule out (3.1). Fortunately, it has been known since the work of Vinogradov [Vi1937] that $\{\sqrt{2}p\}$ is in fact uniformly distributed in the interval $[0, 1]$, and more generally that $\{\alpha p\}$ is uniformly distributed in $[0, 1]$ whenever α is irrational. Indeed, by Weyl's famous equidistribution theorem (see e.g. [KuNe1974]),

this uniform distribution, this is equivalent to the exponential sum estimate

$$\sum_{p < N} e^{2\pi i \alpha p} = o\left(\sum_{p < N} 1\right),$$

and we now see the appearance of Fourier analysis in this subject.

One can rather easily concoct an endless stream of further conspiracies, each of which could contradict the twin prime conjecture; this is one of the reasons why this conjecture is considered so difficult. Let us thus leave this conjecture for now and consider some two-parameter problems. Consider for instance the problem of finding infinitely many patterns of the form $n, n + n', n + 2n' + 2$ (i.e. arithmetic progressions of length 3, but with the last element shifted by 2). Once again, the conspiracy (3.1), if true, would obstruct solvability for this pattern, due to the easily verified identity

$$\{\sqrt{2}n\} - 2\{\sqrt{2}(n + n')\} + \{\sqrt{2}(n + 2n' + 2)\} = 2\sqrt{2} \pmod{1}$$

which is related to the fact that the function $\sqrt{2}n$ has a vanishing second derivative. (Note however that the same conspiracy does not obstruct solvability of an unmodified arithmetic progression $n, n + n', n + 2n'$. This highlights a special property of arithmetic progressions, which most other patterns do not have, namely that arithmetic progressions tend to exist both in structured objects and in pseudorandom objects (and also in hybrids of the two). This is why results about arithmetic progressions have tended to be easier to establish than those about more general patterns, as one does not need to know as much about the structured and random components of the set in which one is looking for progressions.)

More generally, we can see that if the primes correlate in some unusual way with a linear character $e^{2\pi i \alpha p}$, then this is likely to bias or distort the number of patterns $\{n, n + n', n + 2n' + 2\}$ in a significant manner. However, thanks to Fourier analysis, we can show that these “Fourier conspiracies” are in fact the *only* obstructions to counting this type of pattern. Very roughly, one can sketch the reason for this as follows. Firstly, it is helpful to create a counting function for the primes, namely the *von Mangoldt function* $\Lambda(n)$, defined as $\log p$ whenever n is a power of a prime p , and 0 otherwise. This rather strange-looking function is actually rather natural, because of the

identity

$$\sum_{d|n} \Lambda(d) = \log n$$

for all positive integers n , where the sum is over all positive integers d which divide n ; this identity is a restatement of the fundamental theorem of arithmetic, and in fact defines the von Mangoldt function uniquely. The problem of counting patterns $\{n, n + n', n + 2n' + 2\}$ is then roughly equivalent to the task of computing sums such as

$$(3.2) \quad \sum_n \sum_{n'} \Lambda(n) \Lambda(n + n') \Lambda(n + 2n' + 2)$$

where we shall be intentionally vague as to what range the variables n, n' will be summed over. We have the *Fourier inversion formula*

$$\Lambda(n) = \int_0^1 e^{2\pi i n \theta} \hat{\Lambda}(\theta) d\theta$$

where

$$\hat{\Lambda}(\theta) := \sum_n \Lambda(n) e^{-2\pi i n \theta}$$

is a sum very similar in nature to the sums $\sum_{p < N} e^{2\pi i p \alpha}$ mentioned earlier. Substituting this formula into (3.2), we essentially get an expression of the form

$$\int_0^1 \hat{\Lambda}(\theta)^2 \hat{\Lambda}(-2\theta) e^{4\pi i \theta} d\theta$$

(again ignoring issues related to the ranges that n, n' are being summed over). Thus, if one gets good enough control on the Fourier coefficients $\hat{\Lambda}(\theta)$, which can be viewed as a measure of how much the primes “conspire” with a linear phase oscillation with frequency θ , then one can (in principle, at least) count the solutions to the pattern $\{n, n + n', n + 2n' + 2\}$ in the primes. This is the Hardy-Littlewood circle method in a nutshell, and this is for instance how van der Corput’s theorem and Vinogradov theorem were first proven.

I have glossed over the question of how one actually *computes* the Fourier coefficients $\hat{\Lambda}(\theta)$. It turns out that there are two cases. In the “major arc” case when θ is rational, or close to rational (with a reasonably small denominator), then the problem turns out to be essentially equivalent to counting primes in arithmetic progressions, and so one uses tools related to Dirichlet’s theorem (i.e. L -functions,

the Siegel-Walfisz theorem [Wa1936], etc.). In the “minor arc” case when θ is far from rational, one instead uses identities such as

$$\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d},$$

where μ is the *Möbius function* (i.e. $\mu(n) := (-1)^k$ when n is the product of k distinct prime factors for some $k \geq 0$, and $\mu(n) = 0$ otherwise), to split the Fourier coefficient as

$$\hat{\Lambda}(\theta) = \sum_d \sum_m \mu(d) \log(m) e^{2\pi i \alpha d m}$$

and then one uses the irrationality of α to exhibit some significant oscillation in the phase $e^{2\pi i \alpha d m}$, which cannot be fully canceled out by the oscillation in the $\mu(d)$ factor. (In practice, the above strategy does not work directly, and one has to work with various truncated or smoothed out versions of the above identities; this is technical and will not be discussed here.)

Now suppose we look at progressions of length 4: $n, n+n', n+2n', n+3n'$. As with progressions of length 3, “linear” or “Fourier” conspiracies such as (3.1) will bias or distort the total count of such progressions in the primes less than a given number N . But, in contrast to the length 3 case where these are the only conspiracies that actually influence things, for length 4 progressions there are now “quadratic” conspiracies which can cause trouble. Consider for instance the conspiracy

$$(3.3) \quad 0 < \{\sqrt{2}p^2\} < 0.01 \text{ for all but finitely many primes } p.$$

This conspiracy, which can exist even when all linear conspiracies are eliminated, will significantly bias the number of progressions of length 4, due to the identity

$$\{\sqrt{2}n^2\} - 3\{\sqrt{2}(n+n')^2\} + 3\{\sqrt{2}(n+2n')^2\} - \{\sqrt{2}(n+3n')^2\} = 0 \pmod{1}$$

which is related to the fact that the function $\sqrt{2}n^2$ has a vanishing third derivative. In this case, the conspiracy works in one’s favour, increasing the total number of progressions of length 4 beyond what one would have naively expected; as mentioned before, this is related to a remarkable “indestructability” property of progressions, which can be used to establish things like the Green-Tao theorem without

having to deal directly with these obstructions. Thus, in order to count progressions of length 4 in the primes accurately (and not just to establish the qualitative result that there are infinitely many of them), one needs to eliminate conspiracies such as (3.3), which necessitates understanding exponential sums such as $\sum_{p < N} e^{2\pi i \alpha p^2}$ for various rational or irrational numbers α . What's worse, there are several further "generalised quadratic" conspiracies which can also bias this count, for instance the conspiracy

$$0 < \{ \lfloor \sqrt{2p} \rfloor \sqrt{3p} \} < 0.01 \text{ for all but finitely many primes } p,$$

where $x \mapsto \lfloor x \rfloor$ is the greatest integer function. The point here is that the function $\lfloor \sqrt{2x} \rfloor \sqrt{3x}$ has a third divided difference which does not entirely vanish (as with the genuine quadratic $\sqrt{2x^2}$), but does vanish a significant portion of the time (because the greatest integer function obeys the linearity property $\lfloor x+y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor$ a significant fraction of the time), which does lead ultimately to a non-trivial bias effect. Because of this, one is also faced with estimating exponential sums such as $\sum_{p < N} e^{2\pi i \lfloor \sqrt{2p} \rfloor \sqrt{3p}}$. It turns out that the correct way to phrase all of these obstructions is via the machinery of *2-step nilsequences*: details can be found in [GrTa2008b, GrTa2008c, GrTa2008d]. As a consequence, we can in fact give a precise count as to how many arithmetic progressions of primes of length 4 with all primes less than N ; it turns out to be

$$\left(\frac{3}{4} \prod_{p \geq 5} \left(1 - \frac{3p-1}{(p-1)^3} \right) + o(1) \right) \frac{N^2}{\log^4 N} \approx 0.4764 \frac{N^2}{\log^4 N}.$$

The method also works for other linear patterns of comparable "complexity" to progressions of length 4. We are currently working on the problem of longer progressions, in which cubic and higher order obstructions appear (which should be modeled by 3-step and higher nilsequences); some work related to this should appear here shortly.

3.1.2. Structure and randomness in ergodic theory and graph theory.

3.1.3. Structure and randomness in PDE.

3.1. Simons Lecture Series: Structure and randomness 55

3.1.4. Notes. These articles were originally posted on Apr 5-8, 2007
at

[http://terrytao.wordpress.com/2007/04/05/
simons-lecture-i-structure-and-randomness-in-fourier-analysis-and-number-theory/](http://terrytao.wordpress.com/2007/04/05/simons-lecture-i-structure-and-randomness-in-fourier-analysis-and-number-theory/)

[http://terrytao.wordpress.com/2007/04/07/
simons-lecture-ii-structure-and-randomness-in-ergodic-theory-and-graph-theory/](http://terrytao.wordpress.com/2007/04/07/simons-lecture-ii-structure-and-randomness-in-ergodic-theory-and-graph-theory/)

[http://terrytao.wordpress.com/2007/04/08/
simons-lecture-iii-structure-and-randomness-in-pde/](http://terrytao.wordpress.com/2007/04/08/simons-lecture-iii-structure-and-randomness-in-pde/)

3.2. Ostrowski lecture: The uniform uncertainty principle and compressed sensing

3.3. Milliman lecture: Recent developments in arithmetic combinatorics

3.3. Milliman lecture: Recent developments in arithmetic combinatorics

3.3.1. Additive combinatorics and the primes.

3.3.2. Additive combinatorics and random matrices.

3.3.3. Sum-product estimates, expanders, and exponential sums.

Bibliography

- [Be1946] F. A. Behrend, *On sets of integers which contain no three terms in arithmetic progression*, Proc. Nat. Acad. Sci., **32** (1946), 331–332.
- [Bo1999] J. Bourgain, *On triples in arithmetic progression*, Geom. Func. Anal. **9** (1999), 968–984.
- [Bo2008] J. Bourgain, *Roth’s theorem on arithmetic progressions revisited*, preprint.
- [BoGIKo2006] J. Bourgain, A. Glibichuk, S. Konyagin, *Estimates for the number of sums and products and for exponential sums in fields of prime order*, J. London Math. Soc. (2) **73** (2006), no. 2, 380–398.
- [BoKaTa2004] J. Bourgain, N. Katz, T. Tao, *A sum-product estimate in finite fields, and applications*, Geom. Funct. Anal. **14** (2004), no. 1, 27–57.
- [Ch2008] M-C. Chang, *Product theorems in SL_2 and SL_3* , preprint.
- [Cr2008] E. Croot, *The Minimal Number of Three-Term Arithmetic Progressions Modulo a Prime Converges to a Limit*, preprint.
- [Di1904] L. E. Dickson, *A new extension of Dirichlet’s theorem on prime numbers*, Messenger of Math. **33** (1904), 155–161.
- [Ed2004] Y. Edel, *Extensions of generalized product caps*, Designs, Codes, and Cryptography, **31** (2004), 5–14.
- [ElKi2001] G. Elekes, Z. Király, *On the combinatorics of projective mappings*, J. Algebraic Combin. **14** (2001), no. 3, 183–197.
- [Fr1973] G. Freiman, *Foundations of a structural theory of set addition*. Translated from the Russian. Translations of Mathematical Monographs, Vol 37. American Mathematical Society, Providence, R. I., 1973. vii+108 pp.

- [Ga2008] M. Garaev, *An explicit sum-product estimate in \mathbb{F}_p* , preprint.
- [GIKo2008] A. Glibichuk, S. Konyagin, *Additive properties of product sets in fields of prime order*, preprint.
- [GrRu2007] B. Green, I. Ruzsa, *Freiman's theorem in an arbitrary abelian group*, J. Lond. Math. Soc. (2) **75** (2007), no. 1, 163–175.
- [GrTa2008] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, to appear in Annals of Math.
- [GrTa2008b] B. Green, T. Tao, *Linear equations in primes*, to appear, Annals of Math.
- [GrTa2008c] B. Green, T. Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*, preprint.
- [GrTa2008d] B. Green, T. Tao, *Quadratic uniformity of the Möbius function*, preprint.
- [Gr1981] M. Gromov, *Groups of polynomial growth and expanding maps*, Inst. Hautes Études Sci. Publ. Math. No. 53 (1981), 53–73.
- [HaLiSe1998] Y. Hamiduone, A. Lladó, O. Serra, *On subsets with small product in torsion-free groups*, Combinatorica **18** (1998), 529–540.
- [HaLi1923] G.H. Hardy and J.E. Littlewood *Some problems of “partitio numerorum”; III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.
- [He2008] H. Helfgott, *Growth and generation in $SL_2(\mathbf{Z}/p\mathbf{Z})$* , preprint.
- [KaSh2008] N. Katz, C-Y. Shen, *A Slight Improvement to Garaev's Sum Product Estimate*, preprint.
- [Kl2008] B. Kleiner, *A new proof of Gromov's theorem on groups of polynomial growth*, preprint.
- [KuNe1974] L. Kuipers and H. Neiderreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.
- [Li2001] E. Lindenstrauss, *Pointwise theorems for amenable groups*, Invent. Math. **146** (2001), 259–295.
- [LiWa2002] M.-C. Liu, T. Wang, *On the Vinogradov bound in the three primes Goldbach conjecture*, Acta Arith. **105** (2002), no. 2, 133–175.
- [Me1995] R. Meshulam, *On subsets of finite abelian groups with no 3-term arithmetic progressions*, J. Combin. Theory Ser. A., **71** (1995), 168–172.
- [Ra2008] A. Razborov, *A Product Theorem in Free Groups*, preprint.
- [Ro1953] K.F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 245–252.
- [Sa1998] Y. Saouter, *Checking the odd Goldbach conjecture up to 10^{20}* , Math. Comp. **67** (1998), no. 222, 863–866.

-
- [Ta2006] T. Tao, *The dichotomy between structure and randomness, arithmetic progressions, and the primes*, 2006 ICM proceedings, Vol. I., 581–608.
- [Ta2006a] T. Tao, *The dichotomy between structure and randomness*, unpublished slides, available at <http://www.math.ucla.edu/~tao/preprints/Slides/icmslides2.pdf>
- [Ta2008] T. Tao, *Product set estimates in noncommutative groups*, preprint.
- [TaVu2006] T. C. Tao and V. H. Vu, *Additive Combinatorics*, Cambridge University Press, 2006.
- [TaZi] T. Tao, T. Ziegler, *The primes contain arbitrarily long polynomial progressions*, to appear, Acta Math.
- [vdC1939] J.G. van der Corput, *Über Summen von Primzahlen und Primzahlquadraten*, Math. Ann. **116** (1939), 1–50.
- [Vi1937] I. M. Vinogradov, *Some theorems concerning the primes*, Mat. Sbornik. N.S. **2** (1937), 179–195.
- [Wa1936] A. Walfisz, *Zur additiven Zahlentheorie. II*, (German) Math. Z. **40** (1936), no. 1, 592–607.