# The work of Terence Tao

Charles Fefferman

Mathematics at the highest level has several flavors. On seeing it, one might say:

(A) What amazing technical power!

(B) What a grand synthesis!

(C) How could anyone not have seen this before?

(D) Where on earth did this come from?

The work of Terence Tao encompasses all of the above. One cannot hope to capture its extraordinary range in a few pages. My goal here is simply to exhibit a few contributions by Tao and his collaborators, sufficient to produce all the reactions (A)...(D). I shall discuss the Kakeya problem, nonlinear Schrödinger equations and arithmetic progressions of primes.

Let me start with a vignette from Tao's work on the Kakeya problem, a beautiful and fundamental question at the intersection of geometry and combinatorics. I shall state the problem, comment briefly on its significance and history, and then single out my own personal favorite result, by Nets Katz and Tao.

The original Kakeya problem was to determine the least possible area of a plane region inside which a needle of length 1 can be turned a full 360 degrees. Besicovitch and Pál showed that the area can be taken arbitrarily small.

In its modern form, the Kakeya problem is to estimate the fractal dimension of a "Besicovitch set" $E \subset \mathbb{R}^n$, i.e., a set containing line segments of length 1 in all directions.

There are several relevant notions of "fractal dimension". Here, let us use the Minkowski dimension, defined in terms of coverings of $E$ by small balls of a fixed radius $\delta$. The Minkowski dimension is the infimum of all $\beta$ such that, for small $\delta$, $E$ can be covered by $\delta^{-\beta}$ balls of radius $\delta$. We want to prove that any Besicovitch set $E \subset \mathbb{R}^n$ has Minkowski dimension at least $\beta(n)$, with $\beta(n)$ as large as possible. (Perhaps $\beta(n) = n$.)

Regarding the central importance of this problem, perhaps it is enough to say that it is intimately connected with the multiplier problem for Fourier transforms, and with the restriction of Fourier transforms to hypersurfaces; these in turn are closely connected with non-linear PDE via Strichartz estimates and their variants. There are also connections with other hard, interesting problems in combinatorics.

Let me sketch some of the history of the problem over the last 30 years. The basic result of the 1970s is that $\beta(2) = 2$. (This is due to Davies, and is closely related

Proceedings of the International Congress
of Mathematicians, Madrid, Spain, 2006

to the early work of A. Córdoba. See [7], [8].) In the 1980s, Drury [9] showed that $\beta(n) \geq \frac{n+1}{2}$ for $n \geq 3$. (See also Christ et al [4].)

Then, about 1990, J. Bourgain and, shortly afterwards, T. Wolff discovered that Besicovitch sets of small fractal dimension have geometric structure (they contain "bouquets" and "hairbrushes"). During the 1990s, Bourgain also discovered a connection between the Kakeya problem and Gowers' work on the Balog–Szemerédi theorem from combinatorics. These insights led to small, hard-won improvements in the value of $\beta(n)$. The work looks deep and forbidding. See [1], [2], [24].

The connection with Gowers' work arises in the following result. (We write $\#(S)$ for the number of elements of a set $S$.)

**Deep Theorem** (Bourgain, using ideas from Gowers' improvement of Balog–Szemerédi). *Let $A$, $B$ be subsets of an abelian group, and let $G \subset A \times B$. Assume that $\#(A)$, $\#(B)$, and $\#\{a+b : (a,b) \in G\}$ are at most $N$. Then $\#\{a-b : (a,b) \in G\} \leq CN^{2-1/13}$, for a universal constant $C$.*

The point is that one improves on the trivial bound $N^2$. From the Deep Theorem, one quickly obtains a result on $\beta(n)$ by slicing the set $E$ with three parallel hyperplanes $H, H', H''$, with $H''$ halfway between $H$ and $H'$.

Enter Nets Katz and Terence Tao, who proved the following result in 1999.

**Little Lemma.** *Under the assumptions of the Deep Theorem, we have $\#\{a-b : (a,b) \in G\} \leq CN^{2-1/6}$, for a universal constant $C$.*

Note that the Little Lemma is strictly sharper than the Deep Theorem, Nevertheless, its proof takes only a few pages, and can be understood by a bright high-school student. After reading the proof, one has not the faintest clue where the idea came from (see (D)).

The Little Lemma and its refinements led to the estimate $\beta(n) \geq \frac{4n+3}{7}$ for the Kakeya problem, which at the time was the best result known for $n > 8$. In high dimensions, the high-school accessible paper [18] went further than all the deep, forbidding work that came before it. Since then, there has been further progress, with Nets Katz, Izabella Łaba, and Terence Tao playing a leading rôle. The subject still looks deep and forbidding. In particular, regarding (A), let me refer the reader to the tour-de-force [17] by these authors.

Unfortunately, the complete solution to the Kakeya problem still seems far away.

Next, I shall discuss "interaction Morawetz estimates". This simple idea, with profound consequences for PDE, was discovered by the "I-Team": J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and Terence Tao. Let me start with the 3D nonlinear Schrödinger equation (NLS):

$$i\partial_t u + \Delta_x u = \pm |u|^{p-1} u, \quad u(x,0) = u_0(x) \text{ given,} \tag{1}$$

where $u$ is a complex-valued function of $(x,t) \in \mathbb{R}^3 \times \mathbb{R}$, and $p > 1$ is given.

This equation is important in physics and engineering. For instance, it describes the propagation of light in a fiber-optic cable. The behavior of solutions of (1) depends strongly on the $\pm$ sign and on the value of $p$. In particular, the minus sign is "focussing", and we may expect solutions of (1) to develop singularities; while the plus sign is "defocussing", and we expect solutions of (1) to spread out over large regions of space, as $t \longrightarrow \pm\infty$. In the defocussing case, the non-linear term in (1) should eventually become negligibly small, and the solution of (1) ought to behave like a solution of the (linear) free Schrödinger equation $(i\partial_t + \Delta_x)u = 0$. From now on, we restrict attention to the defocussing case.

We note two obvious conserved quantities for (1): the "mass" $\int_{\mathbb{R}^3} |u(x,t)|^2 dx$, and the energy,

$$E = \frac{1}{2}\int_{\mathbb{R}^3} |\nabla_x u(x,t)|^2\, dx \;+\; \frac{1}{p+1}\int_{\mathbb{R}^3} |u(x,t)|^{p+1}\, dx.$$

How can we prove that solutions of (1) spread out for large time? A fundamental tool is the Morawetz estimate. C. Morawetz first discovered this wonderful, simple idea for the non-linear Klein–Gordon equation. Let me describe it here for cubic 3D NLS, i.e., for equation (1) with $p = 3$. There, the Morawetz estimate asserts that

$$\int_0^T \int_{\mathbb{R}^3} \frac{|u(x,t)|^4}{|x|}\, dx\, dt \le C \sup_{0\le t\le T} \|\, (-\Delta_x)^{1/4} u(\,\cdot\,,\, t)\, \|^2_{L^2(\mathbb{R}^3)}, \tag{2}$$

for any $T > 0$ and any solution of (1).

The good news is that (2) instantly shows that $u$ must eventually become small in any given bounded region of space. (If not, then the left-hand side of (2) grows linearly in $T$ as $T \longrightarrow \infty$, while the right-hand side of (2) remains bounded, thanks to conservation of mass and energy.)

The bad news is that (2) does not rule out a scenario in which $u(x,t)$ remains concentrated near a moving center $x = x_0(t)$. The trouble is that the weight function $\frac{1}{|x|}$ is concentrated near $x = 0$, whereas $u(\,\cdot, t)$ may be concentrated somewhere else.

The I-Team found an amazingly simple and straightforward way to overcome the bad news. Let me sketch the idea, starting with the classic proof of (2). To derive (2), we start with the quantity

$$M_0(t) = \mathrm{Im} \int_{\mathbb{R}^3} \bar{u}(x,t) \cdot \left[\frac{x}{|x|} \cdot \nabla_x u(x,t)\right] dx. \tag{3}$$

On one hand, $M_0(t)$ is controlled by the right-hand side of (2), when $0 \le t \le T$. On the other hand, a computation using (1) shows that

$$\frac{d}{dt} M_0(t) = 4\pi^2 |u(0,t)|^2 + 2\int_{\mathbb{R}^3} |\nabla_\Omega u(x,t)|^2 \frac{dx}{|x|} + \int_{\mathbb{R}^3} \frac{|u(x,t)|^4}{|x|}\, dx, \tag{4}$$

where $\nabla_\Omega$ denotes the angular part of the gradient.

The Morawetz estimate (2) follows at once.

The I-Team simply replaced $M_0(t)$ by a weighted average of translates,

$$M(t) = \int_{\mathbb{R}^3} M_y(t)\, |u(y,t)|^2\, dy,$$

where

$$M_y(t) = \operatorname{Im} \int_{\mathbb{R}^3} \bar{u}(x,t) \cdot \left[ \frac{x-y}{|x-y|} \cdot \nabla_x u(x,t) \right] dx.$$

This puts the greatest weight on those $y \in \mathbb{R}^3$ where $u(y,t)$ lives – an eminently sensible idea.

Starting with $M(t)$ and proceeding more or less as in the proof of the Morawetz estimate, one obtains easily the

**Interaction Morawetz Estimate.**

$$\int_0^T \int_{\mathbb{R}^3} |u(x,t)|^4\, dx\, dt \le C \,\| u(\cdot,0) \|^2_{L^2(\mathbb{R}^3)} \cdot \sup_{0 \le t \le T} \| (-\Delta_x)^{1/4} u(\cdot,t) \|^2_{L^2(\mathbb{R}^3)}. \tag{5}$$

Again, the right-hand side is bounded for large $T$, thanks to conservation of mass and energy. This time however, the left-hand side grows linearly in $T$, even if our solution is concentrated in a moving ball. The I-Team has overcome the bad news. They made it look effortless. Why did no one think of it before? (See (C).)

Observe that the right-hand side of (5) is much weaker than the energy; we need only half an $x$-derivative, as opposed to a full gradient. The original purpose of the interaction Morawetz estimate was to derive global existence for cubic defocussing 3D NLS in Sobolev spaces in which the energy may be infinite. That is a big achievement (see [6]), but I will not discuss it further here, except to point out that the proof involves additional ideas and formidable work.

Instead, let me say a few words about the defocussing quintic 3D NLS, i.e., the case $p = 5$ of equation (1). This equation is particularly natural and deep, because it is critical for the energy. One knows that finite-energy initial data lead to solutions for a short time, and that small-energy initial data lead to global solutions. The challenge is to prove global existence for initial data with large, finite energy.

To appreciate the difficulty of the problem, we have only to turn to the tour-de-force [3] by J. Bourgain, solving the problem in the radially symmetric case. The general case is an order of magnitude harder; a singularity can form only at the origin in the radial case, but it may form anywhere in the general case. The I-Team settled the general case using a version of the interaction Morawetz estimate for quintic NLS (with cutoffs, which unfortunately greatly complicate the analysis). This is natural, since in a sense one must overcome the same bad news as before. Their result [5] is as follows.

**Theorem.** *Take $p = 5$ in the defocussing case in* (1). *Then, for any finite-energy initial data $u_0$, there is a global solution $u(x,t)$ of NLS. If $u_0$ belongs to $H^s$ with $s > 1$, then $u(\cdot, t)$ also belongs to $H^s$ for all $t$. Moreover, there exist solutions $u_\pm$ of the free Schrödinger equation, such that*

$$\int_{\mathbb{R}^3} |\nabla_x (u(x,t) - u_\pm(x,t))|^2 \, dx$$

*tends to zero as $t$ tends to $\pm\infty$.*

I will not try to describe their proof, except to say that they use an interaction Morawetz estimate with cutoffs, along with ideas from Bourgain [3], especially the "induction on energy", as well as other ideas that I cannot begin to describe here. The details are highly formidable; see (A).

We come now to Tao's great joint paper ([16]; see also Green [15]) with Ben Green, in which they prove the following result. Here again, $\#(S)$ denotes the number of elements of a set $S$.

**Theorem GT.** *There exist arbitrarily long arithmetic progressions of primes. More precisely, given $k \geq 3$, there exist constants $c(k) > 0$ and $N_0(k) \geq 1$, such that for any $N > N_0(k)$, we have $\#\{k$-term arithmetic progressions among the primes less than $N\} > \frac{c(k)N^2}{(\log N)^k}$.*

The lower bound here agrees in order of magnitude with a natural guess. (Green and Tao are currently working on a more precise result, with an optimal $c(k)$.)

To convey something of the range and depth of the ideas in the proof, let me start with the classic theorem of Szemerédi on sets of positive density. Here, $\mathbb{Z}_N$ denotes the cyclic group of order $N$.

**Theorem Sz 1.** *Given $k$ and $\delta$, we have for large enough $N$ that any subset $E \subset \mathbb{Z}_N$ with $\#(E) > \delta \cdot N$ contains an arithmetic progression of length $k$.*

Szemerédi's theorem also gives a lower bound for the number of $k$-term progressions in $E$. (See [23].) It is convenient to speak of functions $f$ rather than sets $E$. (One obtains Theorem Sz 1 from Theorem Sz 2 below, simply by taking $f$ to be the indicator function of $E$.) Thus, Szemerédi's theorem may be rephrased as follows.

**Theorem Sz 2.** *Given $k$, $\delta$, the following holds for large enough $N$. Let $f : \mathbb{Z}_N \to \mathbb{R}$, with $0 \leq f(x) \leq 1$ for all $x$, and with*

(1) $Av_{x \in \mathbb{Z}_N} f(x) > \delta$.

*Then*

(2) $Av_{x,r \in \mathbb{Z}_N} \{f(x) \cdot f(x+r) \ldots f(x+(k-1)r)\} \geq c(k,\delta) > 0$, *where $c(k,\delta)$ depends only on $k$, $\delta$ (and not on $N$ or $f$).*

(In (1),(2) and similar formulas, "$Av$" denotes the mean.)

In Theorems Sz 1 and 2, $\delta$ stays fixed as $N$ grows. If instead we could take $\delta \sim 1/\log N$, then the Green–Tao theorem would follow. However, such an improvement of Theorems Sz 1, 2 seems utterly out of reach, and may be false.

There are three very different proofs of Theorems Sz 1, 2; they are due to Szemerédi [21], Furstenberg [10], and Gowers [14]. Without doing justice to the remarkable ideas in these arguments, let me just say that Szemerédi used combinatorics, Furstenberg used ergodic theory, and Gowers used (non-linear) Fourier analysis. It is hard to see anything in common in these three proofs. In a sense, the Green–Tao paper synthesizes them all, by quoting Theorem Sz 2 and using ideas that go back to the proofs of Furstenberg and Gowers. See (B).

Green and Tao prove a powerful extension of Theorem Sz 2, in which the hypothesis $0 \leq f(x) \leq 1$ is replaced by $0 \leq f(x) \leq \nu(x)$ for a suitable non-negative weight function $\nu(x)$. The function $\nu(x)$ is assumed to satisfy three conditions, which we describe crudely here.

- $Av_{x \in \mathbb{Z}_N}\, \nu(x) = 1.$

- We assume an upper bound on the quantity

$$Av_{\vec{x}=(x_1,\dots,x_t)\,\in\,(\mathbb{Z}_N)^t} \left\{ \prod_{i=1}^{m} \nu\left(\lambda_i(\vec{x})\right) \right\}$$

  for certain affine functions $\lambda_1, \dots, \lambda_m : (\mathbb{Z}_N)^t \longrightarrow \mathbb{Z}_N$.

- For any $h_1, \dots, h_m \in \mathbb{Z}_N$, we assume that

$$Av_{x \in \mathbb{Z}_N} \{\nu(x+h_1) \dots \nu(x+h_m)\} \leq \sum_{i \neq j} \tau_m(h_i - h_j),$$

  for a function $\tau_m : \mathbb{Z}_N \longrightarrow \mathbb{R}$ that satisfies

$$Av_{h \in \mathbb{Z}_N} \left\{ (\tau_m(h)^q \right\} \leq C(m, q)$$

  for any $q$.

Such a function $\nu(x)$ is called a "pseudo-random measure" by Green and Tao. Their extension of Szemerédi's theorem is as follows.

**Theorem GTS** (Green–Tao–Szemerédi). *Let $k, \delta > 0$, suppose $N$ is large enough and let $\nu$ be a pseudo-random measure. Let $f : \mathbb{Z}_N \longrightarrow \mathbb{R}$, with $0 \leq f(x) \leq \nu(x)$, and with $Av_{x \in \mathbb{Z}_N} f(x) \geq \delta$.*

*Then $Av_{x,r \in \mathbb{Z}_N} \{f(x) \cdot f(x+r) \dots f(x+(k-1)r)\} \geq c(k,\delta) > 0$, where $c(k,\delta)$ depends on $k$, $\delta$, but not on $N$ or $f$.*

The point is that there are pseudo-random measures $\nu(x)$ that are large on sparse subsets of $\mathbb{Z}_N$ (e.g., the primes up to $N$). We will return to this point.

Let me say a few words about the proof of Theorem GTS, and then afterwards describe how it applies to the primes.

It is in the proof of Theorem GTS that Szemerédi's theorem is combined with ideas from Furstenberg and Gowers.

Green and Tao break up the function $f$ into a "uniform" and an "anti-uniform" part, $f = f_U + f_{U^\perp}$.

They expand out $Av_{x,r\in\mathbb{Z}_N}\{f(x)\cdot f(x+r)\dots f(x+(k-1)r)\}$ into a sum of terms

(3) $Av_{x,r\in\mathbb{Z}_N}\{f_0(x)\cdot f_1(x+r)\dots f_{k-1}(x+(k-1)r)\}$, where each $f_i$ is either $f_U$ or $f_{U^\perp}$.

The terms (3) that contain any factor $f_U$ are $o(1)$, thanks to ideas that go back to Gowers' proof.

This leaves us with the term (3) in which each $f_i$ is $f_{U^\perp}$. Let us call this the "critical term".

To control that term, Green and Tao partition $\mathbb{Z}_N$ into subsets $E_1, E_2, \dots, E_A$, and then define a function $\bar{f}_{U^\perp}$ on $\mathbb{Z}_N$ by averaging $f_{U^\perp}$ over each $E_\alpha$. Green and Tao then prove that

(4) replacing $f_{U^\perp}$ by $\bar{f}_{U^\perp}$ makes a difference $o(1)$ in the critical term,

and moreover,

(5) $0 \le \bar{f}_{U^\perp} \le 1$ and $Av_{x\in\mathbb{Z}_N}\,\bar{f}_{U^\perp}(x) \ge \delta$.

Consequently, the classic Szemerédi theorem (Theorem Sz 2) applies to $\bar{f}_{U^\perp}$, completing the proof of the Green–Tao–Szemerédi theorem.

The proof of (4) and (5) is based on ideas that go back to Furstenberg's proof of Szemerédi's theorem.

Once the Green–Tao–Szemerédi theorem is established, one can take $f(x) = \log x$ for $x$ prime, $f(x) = 0$ otherwise. If we can find a pseudo-random measure $\nu$ such that

(6) $0 \le f(x) \le \nu(x)$ for all $x$,

then Theorem GTS applies, and it yields arbitrarily long arithmetic progressions of primes as in Theorem GT. A first guess for $\nu(x)$ is the standard Von Mangoldt function $\Lambda(x) = \log p$ for $x = p^k$, $p$ prime; $\Lambda(x) = 0$ otherwise. $\Lambda$ may indeed be a pseudo-random measure, but that would be very hard to prove. Fortunately, another function $\nu$ can be seen to be a pseudo-random measure satisfying (6), thanks to important work of Goldston–Yıldırım [11], [12], [13], using not-so-hard analytic number theory.

Thus, in the end, a great theorem on the prime numbers is proven without hard analytic number theory. The difficulty lies elsewhere.

I have repeatedly used the phrase "tour-de-force"; I promise that I am not exaggerating.

There are additional first-rate achievements by Tao that I have not mentioned at all. For instance, he has set forth a program [22] for proving the global existence and regularity of wave maps, by using the heat flow for harmonic maps. This has an excellent chance to work, and it may well have important applications in general relativity. I should also mention Tao's joint work with Knutson [19] on the saturation conjecture in representation theory. It is most unusual for an analyst to solve an outstanding problem in algebra.

Tao seems to be getting stronger year by year. It is hard to imagine what can top the work he has already done, but we await Tao's future contributions with eager anticipation.

## References


[1] Bourgain, J., Besicovitch-type maximal operators and applications to Fourier analysis. *Geom. Funct. Anal.* **1** (2) (1991), 147–187.

[2] Bourgain, J., On the dimension of Kakeya sets and related maximal inequalities. *Geom. Funct. Anal.* **9** (2) (1999), 256–282.

[3] Bourgain, J., Global well-posedness of defocusing 3D critical NLS in the radial case. *J. Amer. Math. Soc.* **12** (1999), 145–171.

[4] Christ, M., Duoandikoetxea, J., Rubio de Francia, J. L., Maximal operators associated to the Radon transform and the Calderón-Zygmund method of rotations. *Duke Math. J.* **53** (1986), 189–209.

[5] Colliander, J., Keel, M., Staffilani, G., Takaoka, H., Tao, T., Global well-posedness and scattering for the energy-critical nonlinear Schrödinger equation in $\mathbb{R}^3$. *Ann. of Math.*, to appear.

[6] Colliander, J., Keel, M., Staffilani, G., Takaoka, H., Tao, T., Global existence and scattering for rough solutions of a nonlinear Schrödinger equation in $\mathbb{R}^3$. *Comm. Pure Appl. Math.* **57** (8) (2004), 987–1014.

[7] Córdoba, A., The Kakeya maximal function and the spherical summation multipliers. *Amer. J. Math.* **99** (1977), 1–22.

[8] Davies, R., Some remarks on the Kakeya problem. *Proc. Cambridge Philos. Soc.* **69** (1971), 417–421.

[9] Drury, S., $L^p$ estimates for the $x$-ray transform. *Illinois J. Math.* **27** (1983), 125–129.

[10] Furstenberg, H., Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Analyse Math.* **31** (1977), 204–256.

[11] Goldston, D., and C. Y. Yıldırım, Higher correlations of divisor sums related to primes, I: Triple correlations. *Integers* **3** (2003), A5 (electronic).

[12] Goldston, D., and C. Y. Yıldırım, Higher correlations of divisor sums related to primes, III: $k$-correlations. Preprint, September 2002; arXiv: math.NT/0209102.

[13] Goldston, D., and C. Y. Yıldırım, Small gaps between primes, I. Preprint, April 2005; arXiv:math.NT/0504336.

[14] Gowers, T., A new proof of Szemerédi's theorem. *Geom. Funct. Anal* **11** (2001), 465–588.

[15] Green, B., Roth's theorem in the primes. *Ann. of Math.* **161** (2005), 1609–1636.

[16] Green, B. and Tao, T., The primes contain arbitrarily long arithmetic progressions. *Ann. of Math.*, to appear.

[17] Katz, N. H., Łaba, I., Tao, T., An improved bound on the Minkowski dimension of Besicovitch sets in $\mathbb{R}^3$. *Ann. of Math.* **152** (2000), 383–446.

[18] Katz, N. H., and Tao, T., Bounds on arithmetic projections, and applications to the Kakeya conjecture. *Math. Res. Lett.* **6** (1999), 625–630.

[19] Knutson, A. and Tao, T., The honeycomb model of $\mathrm{GL}_n(\mathbb{C})$ tensor products I: Proof of the saturation conjecture. *J. Amer. Math. Soc.* **12** (4) (1999), 1055–1090.

[20] Morawetz, C., Time decay for the nonliner Klein-Gordon equation. *Proc. Roy. Soc. Ser. A* **306** (1968), 291–296.

[21] Szemerédi, E., On sets of integers containing no $k$ elements in arithmetic progression. *Acta Arith.* **27** (1975), 299–345.

[22] Tao, T., Geometric renormalization of large energy wave maps. Preprint, November 2004; arXiv:math.AP/0411354.

[23] Varnavides, P., On certain sets of positive density. *J. London Math. Soc.* **34** (1959), 358–360.

[24] Wolff, T., Recent work connected with the Kakeya problem. In *Prospects in mathematics* (Princeton, NJ, 1996), Amer. Math. Soc., Providence, RI, 1999, 129–162.

Department of Mathematics, Fine Hall, Washington Road, Princeton, NJ 08544-1000, U.S.A.
E-mail: cf@math.princeton.edu

**Wendelin Werner**

Ph.D. Université Paris-Sud, 1993

Positions until today

| | |
|---|---|
| 1991–1993 and 1995–1997 | Charge de recherche CNRS |
| 1993–1995 | European postdoctoral fellowship, University of Cambridge |
| Since 1997 | Professor at the Université Paris-Sud (Orsay) |