Structure, transference, and Hahn-Banach

W. T. Gowers

University of Cambridge

April 11, 2008

A few reminders from yesterday

Definition

Let $\|.\|$ be a norm on $\mathbb{R}^n$. Then $\|\phi\|^*$ is defined to be $\max\{\langle f, \phi\rangle : \|f\| \leq 1\}$.

Trivial Lemma

$\langle f, \phi\rangle \leq \|f\|\|\phi\|^*$ for every $f, \phi \in \mathbb{R}^n$.

Hahn-Banach Corollary

If $K_1, \ldots, K_r$ are closed convex sets that contain 0 and if $f \notin K_1 + \cdots + K_r$, then there is a function ϕ such that $\langle f, \phi\rangle > 1$ and $\langle g_i, \phi\rangle \leq 1$ for every i.

A rudimentary structure theorem

Theorem

Let $\|.\|$ be a norm on $\mathbb{R}^n$, let $c > 0$ and let f be a function with $\|f\|_2 \leq 1$. Then f can be written as $g + h$ in such a way that $\|g\|^* \leq c^{-1}$ and $\|h\| \leq c$.

Proof.

Suppose not. Let $K = \{g : \|g\|^* \leq c^{-1}\}$ and let $K' = \{h : \|h\| \leq c\}$. By the Hahn-Banach corollary, there exists ϕ such that $\langle f, \phi \rangle > 1$, $\|\phi\| \leq c$ and $\|\phi\|^* \leq c^{-1}$. It follows that $\|\phi\|_2^2 = \langle \phi, \phi \rangle \leq 1$. But then $\langle f, \phi \rangle \leq 1$ by Cauchy-Schwarz. This is a contradiction. $\square$

Drawbacks

Unfortunately, this structure theorem is too weak to be useful. There are two major reasons for this.

Drawbacks

Unfortunately, this structure theorem is too weak to be useful. There are two major reasons for this.

- We usually need $\|h\|$ to be much smaller than $\|g\|^*$ is large. That is, we want $\|h\| \leq \eta(\|g\|^*)$ for some function η that we can choose.

Drawbacks

Unfortunately, this structure theorem is too weak to be useful. There are two major reasons for this.

- We usually need $\|h\|$ to be much smaller than $\|g\|^*$ is large. That is, we want $\|h\| \leq \eta(\|g\|^*)$ for some function η that we can choose.
- We often need control on the ranges of g and h (given control on the range of f).

Drawbacks

Unfortunately, this structure theorem is too weak to be useful. There are two major reasons for this.

- We usually need $\|h\|$ to be much smaller than $\|g\|^*$ is large. That is, we want $\|h\| \leq \eta(\|g\|^*)$ for some function η that we can choose.
- We often need control on the ranges of g and h (given control on the range of f).

We shall see that, with the help of two additional ideas, we can get these properties from the Hahn-Banach method as well.

A less rudimentary structure theorem

Theorem

Let $\|.\|$ be a norm on $\mathbb{R}^n$, let $\epsilon > 0$, let f be a function with $\|f\|_2 \leq 1$ and let η be an arbitrary positive decreasing function from $\mathbb{R}$ to $\mathbb{R}$. Then there exists $M \leq M(\eta, \epsilon)$ such that we can write $f = f_1 + f_2 + f_3$ with $\|f_1\|^* \leq M$, $\|f_2\| \leq \eta(M)$ and $\|f_3\|_2 \leq \epsilon$.

In other words, we can arbitrarily strengthen the first structure theorem if we're prepared to allow a small L_2 error. This general phenomenon is far from new.

Sketch of proof

Sketch of proof

- We begin by applying Hahn-Banach in the same way as before: if we can't find a decomposition of the required kind for a particular choice of M, then we get a function ϕ such that $\langle f, \phi \rangle \geq 1$, $\|\phi\| \leq M^{-1}$ and $\|\phi\|^* \leq \eta(M)^{-1}$. But now we also have that $\|\phi\|_2 \leq \epsilon^{-1}$.

Sketch of proof

- We begin by applying Hahn-Banach in the same way as before: if we can't find a decomposition of the required kind for a particular choice of M, then we get a function ϕ such that $\langle f,\phi\rangle \geq 1$, $\|\phi\| \leq M^{-1}$ and $\|\phi\|^* \leq \eta(M)^{-1}$. But now we also have that $\|\phi\|_2 \leq \epsilon^{-1}$.
- Choose a sequence $M_1, M_2, \ldots, M_r$ such that $M_{i+1} >> \eta(M_i)^{-1}$ and for each M_i choose ϕ_i in this way.

Sketch of proof

- We begin by applying Hahn-Banach in the same way as before: if we can't find a decomposition of the required kind for a particular choice of M, then we get a function ϕ such that $\langle f, \phi \rangle \geq 1$, $\|\phi\| \leq M^{-1}$ and $\|\phi\|^* \leq \eta(M)^{-1}$. But now we also have that $\|\phi\|_2 \leq \epsilon^{-1}$.
- Choose a sequence $M_1, M_2, \ldots, M_r$ such that $M_{i+1} >> \eta(M_i)^{-1}$ and for each M_i choose ϕ_i in this way.
- Then $\|\phi_i\|_2 \leq \epsilon^{-1}$ for each i, and when $i < j$ we have

$$\langle \phi_i, \phi_j \rangle \leq \|\phi_i\|^* \|\phi_j\| \leq M_j^{-1} \eta(M_i)^{-1},$$

which is small.

Sketch of proof

- We begin by applying Hahn-Banach in the same way as before: if we can't find a decomposition of the required kind for a particular choice of M, then we get a function ϕ such that $\langle f, \phi \rangle \geq 1$, $\|\phi\| \leq M^{-1}$ and $\|\phi\|^* \leq \eta(M)^{-1}$. But now we also have that $\|\phi\|_2 \leq \epsilon^{-1}$.
- Choose a sequence $M_1, M_2, \ldots, M_r$ such that $M_{i+1} >> \eta(M_i)^{-1}$ and for each M_i choose ϕ_i in this way.
- Then $\|\phi_i\|_2 \leq \epsilon^{-1}$ for each i, and when $i < j$ we have

$$\langle \phi_i, \phi_j \rangle \leq \|\phi_i\|^* \|\phi_j\| \leq M_j^{-1} \eta(M_i)^{-1},$$

  which is small.
- This means that $\|\phi_1 + \cdots + \phi_r\|_2$ grows like $\epsilon^{-1}\sqrt{r}$, while $\langle f, \phi_1 + \cdots + \phi_r \rangle \geq r$. Since we assume that $\|f\|_2 \leq 1$, this contradicts Cauchy-Schwarz when $r \geq 2\epsilon^{-2}$.

Outline of proof of Green-Tao theorem

Outline of proof of Green-Tao theorem

- Basic idea: the primes are too sparse for the usual methods to apply; however, the expression $\mathbb{E}_{x,d} f(x)f(x+d)\dots f(x+(k-1)d)$ is robust under small U^{k-1} perturbations, so we'd be done if we could approximate the von Mangoldt function Λ in U^{k-1} by a bounded function.

Outline of proof of Green-Tao theorem

- Basic idea: the primes are too sparse for the usual methods to apply; however, the expression $\mathbb{E}_{x,d} f(x)f(x+d)\dots f(x+(k-1)d)$ is robust under small U^{k-1} perturbations, so we'd be done if we could approximate the von Mangoldt function Λ in U^{k-1} by a bounded function.
- This is not possible, but one can use the "W-trick" to pass to an arithmetic progression inside which Λ is dominated by a non-negative function ν that is close in U^{k-1} to a constant function (and has other good pseudorandomness properties). This was closely based on work of Goldston and Yildirim.

Outline of proof of Green-Tao theorem

- Basic idea: the primes are too sparse for the usual methods to apply; however, the expression $\mathbb{E}_{x,d} f(x)f(x+d)\dots f(x+(k-1)d)$ is robust under small U^{k-1} perturbations, so we'd be done if we could approximate the von Mangoldt function Λ in U^{k-1} by a bounded function.
- This is not possible, but one can use the "W-trick" to pass to an arithmetic progression inside which Λ is dominated by a non-negative function ν that is close in U^{k-1} to a constant function (and has other good pseudorandomness properties). This was closely based on work of Goldston and Yildirim.
- This reduced the problem to a purely arithmetico-combinatorial one.

The Green-Tao transference theorem

They were left needing to prove the following result.

Theorem

Let ν be a pseudorandom measure, and let $0 \leq f \leq \nu$. Then f can be approximated in U^{k-1} by a function g that takes values in $[0,1]$.

The Green-Tao transference theorem

They were left needing to prove the following result.

Theorem

Let ν be a pseudorandom measure, and let $0 \leq f \leq \nu$. Then f can be approximated in U^{k-1} by a function g that takes values in $[0,1]$.

This was enough, since Szemerédi's theorem (in its functional form) shows that $\mathbb{E}_{x,d} g(x)g(x+d)\dots g(x+(k-1)d) \geq c(\delta)$, where δ is an absolute constant (arising from the construction of ν). Since f is a U^{k-1} perturbation of g, one expects the same of f.

The Green-Tao transference theorem

They were left needing to prove the following result.

Theorem

Let ν be a pseudorandom measure, and let $0 \leq f \leq \nu$. Then f can be approximated in U^{k-1} by a function g that takes values in $[0, 1]$.

This was enough, since Szemerédi's theorem (in its functional form) shows that $\mathbb{E}_{x,d} g(x)g(x+d)\dots g(x+(k-1)d) \geq c(\delta)$, where δ is an absolute constant (arising from the construction of ν). Since f is a U^{k-1} perturbation of g, one expects the same of f.

That is not immediate because f is not bounded, but Green and Tao could prove it under the weaker assumption that they did have: that f is bounded above by a pseudorandom measure. (Similar to other counting lemmas "relative to a random set.")

We shall consider the following general statement, and try to prove it, adding extra assumptions when we need them.

Theorem

Let μ and ν be positive functions with $\mathbb{E}\mu(x) = \mathbb{E}\nu(x) = 1$, and suppose that $\|\mu - \nu\|$ is very small. Let $0 \leq f \leq \mu$. Then there exists a function g such that $0 \leq g \leq \nu$ and $\|f - g\|$ is pretty small.

Step 1: Apply Hahn-Banach

Theorem

Let μ and ν be positive functions with $\mathbb{E}\mu(x) = \mathbb{E}\nu(x) = 1$, and suppose that $\|\mu - \nu\|$ is very small. Let $0 \leq f \leq \mu$. Then there exists a function g such that $0 \leq g \leq \nu$ and $\|f - g\|$ is pretty small.

Step 1: Apply Hahn-Banach

Theorem

Let μ and ν be positive functions with $\mathbb{E}\mu(x) = \mathbb{E}\nu(x) = 1$, and suppose that $\|\mu - \nu\|$ is very small. Let $0 \leq f \leq \mu$. Then there exists a function g such that $0 \leq g \leq \nu$ and $\|f - g\|$ is pretty small.

Let B be the unit ball of $\|.\|$ and let K be the set of all functions g such that $0 \leq g \leq \nu$. Then we would like to prove that $(1+\delta)^{-1} f \in \epsilon B + K$ for some small ϵ.

Step 1: Apply Hahn-Banach

Theorem

Let μ and ν be positive functions with $\mathbb{E}\mu(x) = \mathbb{E}\nu(x) = 1$, and suppose that $\|\mu - \nu\|$ is very small. Let $0 \leq f \leq \mu$. Then there exists a function g such that $0 \leq g \leq \nu$ and $\|f - g\|$ is pretty small.

Let B be the unit ball of $\|.\|$ and let K be the set of all functions g such that $0 \leq g \leq \nu$. Then we would like to prove that $(1+\delta)^{-1} f \in \epsilon B + K$ for some small ϵ.

The sets ϵB and K are convex, so we can apply Hahn-Banach. If we can't decompose $(1+\delta)^{-1} f$ as we want, then we can find ϕ such that $\langle f, \phi \rangle \geq 1 + \delta$, $\|\phi\|^* \leq \epsilon^{-1}$, and $\langle g, \phi \rangle \leq 1$ whenever $0 \leq g \leq \nu$.

Step 1: Apply Hahn-Banach

Theorem

Let μ and ν be positive functions with $\mathbb{E}\mu(x) = \mathbb{E}\nu(x) = 1$, and suppose that $\|\mu - \nu\|$ is very small. Let $0 \leq f \leq \mu$. Then there exists a function g such that $0 \leq g \leq \nu$ and $\|f - g\|$ is pretty small.

Let B be the unit ball of $\|.\|$ and let K be the set of all functions g such that $0 \leq g \leq \nu$. Then we would like to prove that $(1+\delta)^{-1} f \in \epsilon B + K$ for some small ϵ.

The sets ϵB and K are convex, so we can apply Hahn-Banach. If we can't decompose $(1+\delta)^{-1} f$ as we want, then we can find ϕ such that $\langle f, \phi \rangle \geq 1 + \delta$, $\|\phi\|^* \leq \epsilon^{-1}$, and $\langle g, \phi \rangle \leq 1$ whenever $0 \leq g \leq \nu$.

The last condition is equivalent to the assertion that $\langle \nu, \phi_+ \rangle \leq 1$.

Conditional completion of proof

We are given: $\|\mu - \nu\|$ tiny, $0 \leq f \leq \mu$. We have obtained: $\langle f, \phi \rangle > 1 + \delta$, $\|\phi\|^* \leq \epsilon^{-1}$, $\langle \nu, \phi_+ \rangle \leq 1$.

Conditional completion of proof

We are given: $\|\mu - \nu\|$ tiny, $0 \le f \le \mu$. We have obtained: $\langle f, \phi \rangle > 1 + \delta$, $\|\phi\|^* \le \epsilon^{-1}$, $\langle \nu, \phi_+ \rangle \le 1$.

If only we knew that $\|\phi_+\|^*$ was not too large, then we'd be done, since it would follow that $\langle \mu - \nu, \phi_+ \rangle$ was tiny, so $\langle \mu, \phi_+ \rangle \le 1 + \delta/2$, so $\langle f, \phi_+ \rangle \le 1 + \delta/2$, so $\langle f, \phi \rangle \le 1 + \delta/2$, a contradiction.

Conditional completion of proof

We are given: $\|\mu - \nu\|$ tiny, $0 \leq f \leq \mu$. We have obtained: $\langle f, \phi\rangle > 1 + \delta$, $\|\phi\|^* \leq \epsilon^{-1}$, $\langle \nu, \phi_+\rangle \leq 1$.

If only we knew that $\|\phi_+\|^*$ was not too large, then we'd be done, since it would follow that $\langle \mu - \nu, \phi_+\rangle$ was tiny, so $\langle \mu, \phi_+\rangle \leq 1 + \delta/2$, so $\langle f, \phi_+\rangle \leq 1 + \delta/2$, so $\langle f, \phi\rangle \leq 1 + \delta/2$, a contradiction.

It is not in general true that $\|\phi_+\|^*$ can be bounded in terms of $\|\phi\|^*$, so we need to make extra assumptions about the norm $\|.\|^*$.

Polynomial approximation

Suppose we have a norm such that an upper bound for $\|\phi\|^*$ implies control on both $\|\phi\|_\infty$ and on $\|\phi^m\|^*$ for every positive integer m. Then it also implies that ϕ_+ can be uniformly approximated by a function with not too large $\|.\|^*$ norm.

Proof.

Let $\delta > 0$ and suppose that $\|\phi\|_\infty \leq K$. Let $J(x) = 0$ if $x \leq 0$ and x if $x \geq 0$, and let P be a polynomial that approximates J to within δ on $[-K, K]$. Then $J\phi = \phi_+$ and $\|P\phi - J\phi\|_\infty \leq \delta$. Since $P\phi$ is built out of powers of ϕ, we have control on $\|P\phi\|^*$. □

A soft inverse theorem, or approximate duality

For the Green-Tao theorem we need $\|.\|$ to be $\|.\|_{U^k}$ for an appropriate k. Unfortunately, one cannot control $\|\phi^m\|_{U^k}^*$ in terms of $\|\phi\|_{U^k}^*$. However, for certain functions ϕ one can.

Green and Tao define a class of functions called *basic anti-uniform functions* and prove the following three facts about them, under the assumption that ν is a pseudorandom measure.

- If $|f| \leq 1 + \nu$ and $\|f\|_{U^k} \geq c$ then there is a basic anti-uniform function ϕ such that $\langle f, \phi \rangle \geq c^{2^k}$.
- If ϕ is a basic anti-uniform function then $\|\phi\|_\infty \leq 2^{2^{k-1}}$.
- A product of m basic anti-uniform functions has $(U^k)^*$ norm bounded above by a function of k and m.

Using an *approximate* dual norm

- Let $\|\phi\|^*_{BAU}$ be the norm whose unit ball is the (symmetric) convex hull of all basic anti-uniform functions. Explicitly, $\|\phi\|^*_{BAU}$ is the minimum of $\sum_i |\lambda_i|$ over all ways of writing $\phi = \sum_i \lambda_i \beta_i$ with each β_i a basic anti-uniform function.

Using an *approximate* dual norm

- Let $\|\phi\|^*_{BAU}$ be the norm whose unit ball is the (symmetric) convex hull of all basic anti-uniform functions. Explicitly, $\|\phi\|^*_{BAU}$ is the minimum of $\sum_i |\lambda_i|$ over all ways of writing $\phi = \sum_i \lambda_i\beta_i$ with each β_i a basic anti-uniform function.
- Now suppose that $\|h\|_{U^k}$ is small. Then $\langle h, \psi\rangle$ is small for every basic anti-uniform function ψ, since we have a uniform bound for $\|\psi\|^*_{U^k}$. It follows that $\|h\|_{BAU}$ is small too.

Using an *approximate* dual norm

- Let $\|\phi\|^*_{BAU}$ be the norm whose unit ball is the (symmetric) convex hull of all basic anti-uniform functions. Explicitly, $\|\phi\|^*_{BAU}$ is the minimum of $\sum_i |\lambda_i|$ over all ways of writing $\phi = \sum_i \lambda_i \beta_i$ with each β_i a basic anti-uniform function.
- Now suppose that $\|h\|_{U^k}$ is small. Then $\langle h, \psi \rangle$ is small for every basic anti-uniform function ψ, since we have a uniform bound for $\|\psi\|^*_{U^k}$. It follows that $\|h\|_{BAU}$ is small too.
- Applying this to $h = \mu - \nu$ we deduce that $\|\mu - \nu\|_{BAU}$ is small.

Using an *approximate* dual norm

- Let $\|\phi\|^*_{BAU}$ be the norm whose unit ball is the (symmetric) convex hull of all basic anti-uniform functions. Explicitly, $\|\phi\|^*_{BAU}$ is the minimum of $\sum_i |\lambda_i|$ over all ways of writing $\phi = \sum_i \lambda_i \beta_i$ with each β_i a basic anti-uniform function.
- Now suppose that $\|h\|_{U^k}$ is small. Then $\langle h, \psi \rangle$ is small for every basic anti-uniform function ψ, since we have a uniform bound for $\|\psi\|^*_{U^k}$. It follows that $\|h\|_{BAU}$ is small too.
- Applying this to $h = \mu - \nu$ we deduce that $\|\mu - \nu\|_{BAU}$ is small.
- But now everything works, because we have the properties we want of $\|.\|^*_{BAU}$.

Using an *approximate* dual norm

- Let $\|\phi\|_{BAU}^*$ be the norm whose unit ball is the (symmetric) convex hull of all basic anti-uniform functions. Explicitly, $\|\phi\|_{BAU}^*$ is the minimum of $\sum_i |\lambda_i|$ over all ways of writing $\phi = \sum_i \lambda_i \beta_i$ with each β_i a basic anti-uniform function.
- Now suppose that $\|h\|_{U^k}$ is small. Then $\langle h, \psi \rangle$ is small for every basic anti-uniform function ψ, since we have a uniform bound for $\|\psi\|_{U^k}^*$. It follows that $\|h\|_{BAU}$ is small too.
- Applying this to $h = \mu - \nu$ we deduce that $\|\mu - \nu\|_{BAU}$ is small.
- But now everything works, because we have the properties we want of $\|.\|_{BAU}^*$.
- It follows that we can find g with $\|f - g\|_{BAU}$ small. But this implies that $\|f - g\|_{U^k}$ is small (or else we could find a basic anti-uniform function that correlated with it).

A comparison with the approach of Green and Tao

Green and Tao prove the transference theorem using an energy incrementation argument.

A comparison with the approach of Green and Tao

Green and Tao prove the transference theorem using an energy incrementation argument.

- Many of the same ingredients: e.g., polynomial approximations, same lemmas about basic anti-uniform functions, etc.

A comparison with the approach of Green and Tao

Green and Tao prove the transference theorem using an energy incrementation argument.

- Many of the same ingredients: e.g., polynomial approximations, same lemmas about basic anti-uniform functions, etc.
- The energy approach builds an approximating function g in stages; the Hahn-Banach approach merely derives a contradiction if such a function does not exist.

A comparison with the approach of Green and Tao

Green and Tao prove the transference theorem using an energy incrementation argument.

- Many of the same ingredients: e.g., polynomial approximations, same lemmas about basic anti-uniform functions, etc.
- The energy approach builds an approximating function g in stages; the Hahn-Banach approach merely derives a contradiction if such a function does not exist.
- The Hahn-Banach approach is a lot shorter than the energy approach and avoids some awkward technicalities.

A comparison with the approach of Green and Tao

Green and Tao prove the transference theorem using an energy incrementation argument.

- Many of the same ingredients: e.g., polynomial approximations, same lemmas about basic anti-uniform functions, etc.
- The energy approach builds an approximating function g in stages; the Hahn-Banach approach merely derives a contradiction if such a function does not exist.
- The Hahn-Banach approach is a lot shorter than the energy approach and avoids some awkward technicalities.
- The Hahn-Banach approach uses the Weierstrass approximation theorem in a much simpler way. It just uses the fact that one can approximate $|x|$ on a bounded interval, and it uses it once. Green and Tao use a multidimensional theorem several times at each iteration, and moreover they are approximating level sets, so they need to worry about cutoffs.

The same method works with only minor changes for the transference results used by Green and Tao in their paper on linear equations in primes, and by Tao and Ziegler in their paper on polynomial progressions in the primes.

The same method works with only minor changes for the transference results used by Green and Tao in their paper on linear equations in primes, and by Tao and Ziegler in their paper on polynomial progressions in the primes.

One possible moral: the connection between transference theorems and ergodic theory is less significant than it looks.

A generalization of Tao's structure theorem

Let us return to structure theorems. Earlier we saw how to decompose f as $f_1 + f_2 + f_3$ with $\|f_1\|^* \leq M$, $\|f_2\| \leq \eta(M)$ and $\|f_3\|_2 \leq \epsilon$. However, in applications we nearly always want more. E.g., if f takes values in $[0,1]$ then we want f_1 to do so too. In fact, even that is not enough.

Theorem

If $\|.\|^*$ is an algebra norm and f takes values in $[0,1]$, then we can decompose f as $f_1 + f_2 + f_3$ in such a way that $\|f_1\|^* \leq M$, $\|f_2\| \leq \eta(M)$ and $\|f_3\| \leq \epsilon$. Moreover, we can do so in such a way that the functions f_1 and $f_1 + f_3$ both take values in $[0,1]$.

An algebra norm is one that satisfies $\|fg\|^* \leq \|f\|^* \|g\|^*$, which implies that $\|f\|^* \geq \|f\|_\infty$.

Outline of proof

Outline of proof

- The first step is to apply the previous structure theorem and obtain a decomposition $f_1 + f_2 + f_3$ with $\|f_1\|^* \leq M$, $\|f_2\| \leq \eta(M)$ and $\|f_3\|_2 \leq \epsilon$. We shall then make adjustments to the functions in order to obtain the extra information about their ranges.

Outline of proof

- The first step is to apply the previous structure theorem and obtain a decomposition $f_1 + f_2 + f_3$ with $\|f_1\|^* \leq M$, $\|f_2\| \leq \eta(M)$ and $\|f_3\|_2 \leq \epsilon$. We shall then make adjustments to the functions in order to obtain the extra information about their ranges.
- Let $J(x) = 0$ if $x \leq 0$, x if $0 \leq x \leq 1$ and 1 if $x \geq 1$. Then J can be uniformly approximated by a polynomial P on the interval $[-M, M]$, in which f_1 takes its values. By the algebra property, $\|Pf_1\|^*$ is not too big, so we can uniformly approximate Jf_1 by a function with not too large $\|.\|^*$ norm.

Outline of proof

- The first step is to apply the previous structure theorem and obtain a decomposition $f_1 + f_2 + f_3$ with $\|f_1\|^* \leq M$, $\|f_2\| \leq \eta(M)$ and $\|f_3\|_2 \leq \epsilon$. We shall then make adjustments to the functions in order to obtain the extra information about their ranges.
- Let $J(x) = 0$ if $x \leq 0$, x if $0 \leq x \leq 1$ and 1 if $x \geq 1$. Then J can be uniformly approximated by a polynomial P on the interval $[-M, M]$, in which f_1 takes its values. By the algebra property, $\|Pf_1\|^*$ is not too big, so we can uniformly approximate Jf_1 by a function with not too large $\|.\|^*$ norm.
- A slightly fiddly lemma shows that $\|f_1 - Pf_1\|_2$ is small, so we can replace f_1 by Pf_1, at a small cost to the norm. Now we have the right range for our new f_1 (after a very small shrinking).

Outline of proof, continued

We now have $f = f_1 + f_2 + f_3$ and we know that f_1 (and f) take values in $[0, 1]$. Next, we turn attention to f_2.

Outline of proof, continued

We now have $f = f_1 + f_2 + f_3$ and we know that f_1 (and f) take values in $[0, 1]$. Next, we turn attention to f_2.

- We want $0 \leq f_1 + f_3 \leq 1$. This is equivalent to $f - 1 \leq f_2 \leq f$.

Outline of proof, continued

We now have $f = f_1 + f_2 + f_3$ and we know that f_1 (and f) take values in $[0, 1]$. Next, we turn attention to f_2.

- We want $0 \leq f_1 + f_3 \leq 1$. This is equivalent to $f - 1 \leq f_2 \leq f$.
- We shall go about this in a disastrously simple-minded way and then mitigate the disaster. So to begin with, replace f_2 by $g_2 = \min\{f_2, f\}$.

Outline of proof, continued

We now have $f = f_1 + f_2 + f_3$ and we know that f_1 (and f) take values in $[0, 1]$. Next, we turn attention to f_2.

- We want $0 \leq f_1 + f_3 \leq 1$. This is equivalent to $f - 1 \leq f_2 \leq f$.
- We shall go about this in a disastrously simple-minded way and then mitigate the disaster. So to begin with, replace f_2 by $g_2 = \min\{f_2, f\}$.
- Another fiddly lemma shows that the difference has small L_2 norm. The disaster is that there is no reason for $\|g_2\|$ to be small.

Outline of proof, continued

We now have $f = f_1 + f_2 + f_3$ and we know that f_1 (and f) take values in $[0, 1]$. Next, we turn attention to f_2.

- We want $0 \leq f_1 + f_3 \leq 1$. This is equivalent to $f - 1 \leq f_2 \leq f$.
- We shall go about this in a disastrously simple-minded way and then mitigate the disaster. So to begin with, replace f_2 by $g_2 = \min\{f_2, f\}$.
- Another fiddly lemma shows that the difference has small L_2 norm. The disaster is that there is no reason for $\|g_2\|$ to be small.
- To deal with this, we use transference! Let $\mu = (f_2)_+$ and $\nu = (f_2)_-$. Then $\|\mu - \nu\| = \|f_2\|$ is very small. Also, $0 \leq (g_2)_+ \leq \mu$. So by the transference principle earlier (in the easy case where $\|.\|^*$ is an algebra norm) we can find a function h such that $0 \leq h \leq \nu$ and $\|(g_2)_+ - h\|$ is small. Take as our new f_2 the function $(g_2)_+ - h$.

Outline of proof, continued

We now have $f = f_1 + f_2 + f_3$ and we know that f_1 (and f) take values in $[0, 1]$. Next, we turn attention to f_2.

- We want $0 \leq f_1 + f_3 \leq 1$. This is equivalent to $f - 1 \leq f_2 \leq f$.
- We shall go about this in a disastrously simple-minded way and then mitigate the disaster. So to begin with, replace f_2 by $g_2 = \min\{f_2, f\}$.
- Another fiddly lemma shows that the difference has small L_2 norm. The disaster is that there is no reason for $\|g_2\|$ to be small.
- To deal with this, we use transference! Let $\mu = (f_2)_+$ and $\nu = (f_2)_-$. Then $\|\mu - \nu\| = \|f_2\|$ is very small. Also, $0 \leq (g_2)_+ \leq \mu$. So by the transference principle earlier (in the easy case where $\|.\|^*$ is an algebra norm) we can find a function h such that $0 \leq h \leq \nu$ and $\|(g_2)_+ - h\|$ is small. Take as our new f_2 the function $(g_2)_+ - h$.
- Another fiddly lemma: h is close to $(g_2)_- = (f_2)_-$ in L_2.

Outline of proof, continued

We now have $f = f_1 + f_2 + f_3$ and we know that f and f_1 take values in $[0, 1]$ and that $f_2 \leq f$. It remains to get $f_2 \geq f - 1$.

To do this we just turn everything upside down and repeat the argument of the previous slide. So the structure theorem is proved.

Wild speculation

Tao proved the structure theorem (for some particular algebras that he constructed, but the method is a general one) using energy arguments. Can one complete the following square?

Energy arguments used to prove a structure theorem for bounded functions.	Inductive construction of characteristic factors.
The Hahn-Banach theorem used to prove a structure theorem for bounded functions.	A softer approach to characteristic factors.

More wild speculation

More generally, can the Hahn-Banach theorem do everything that energy arguments can do?

There certainly are several examples, e.g. connected with regularity lemmas.

What about the rest of Tao's quantitative ergodic proof of Szemerédi's theorem?

Might it even be possible to simplify the ergodic-theory proof itself?