Rectangles as Sums of Squares.

Mark Walters

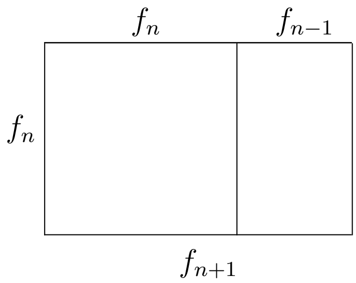
Peterhouse, Cambridge, CB2 1RD

Abstract

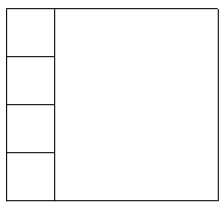
In this paper we examine generalisations of the following problem posed by Laczkovich: Given an $n \times m$ rectangle with n and m integers, it can be written as a disjoint union of squares; what is the smallest number of squares that can be used? He also asked the corresponding higher dimensional analogue. For the two dimensional case Kenyon proved a tight logarithmic bound but left open the higher dimensional case. Using completely different methods we prove good upper and lower bounds for this case as well as some other variants.

1 Introduction.

Laczkovich [4] asked the following question: Given an $n \times m$ rectangle with n and m integers, we can write it as a union of squares with disjoint interiors; for example as the union of nm 1×1 squares. We call such a decomposition a *tiling*. What is the minimum number of squares required? (We do not lose generality by restricting to rectangles with rational side ratios; as proved by Dehn [2] this is required for any such decomposition to exist.) It is clear that we require at least n/m squares; thus we restrict to the case where $n < 2m$ (again it is easy to see that the value two is arbitrary; any number greater than one will only increase the number of squares required by a constant). It is worth bearing two examples in mind. The first is when n and m are consecutive Fibonacci numbers. In this case we can remove a square from one end of the rectangle and reduce to an $m \times (n-m)$ rectangle; i.e., the preceding pair of Fibonacci numbers. It is easily seen that this takes $\theta(\log n)$ squares.



The second case is an $n \times (n+2)$ rectangle with n and $n+2$ both prime (which we, at least, cannot rule out). Then if we remove a square from one side (as above) we will need at least $n/2$ squares to tile the resulting $n \times 2$ rectangle. Since the two side lengths are prime we cannot remove a small strip of squares from either side to reduce to a smaller rectangle; i.e., the following picture cannot occur:



It is recommended that the reader consider this case for a moment to see the problems that can arise. This example appears to rule out any simple method.

Kenyon [3] proved a tight logarithmic bound for this problem. However, his proof does not extend to the generalisations we prove; in particular his method cannot be extended to higher dimensions. We briefly discuss his proof later in this section.

We look at some generalisations of this question which were also asked by Laczkovich. One is to work in higher dimensions with rectangle replaced by hypercuboid and square replaced by hypercube. For this problem we require the side lengths $n = n_1 \geq n_2 \geq \ldots \geq n_d$ (where d is the dimension of the space) to satisfy $n < 2n_d$. Again the 2 is arbitrary, but this time a different value will change the result by a constant factor (rather than by an additive constant). Indeed, we can split the hypercuboid into a (bounded) number of pieces satisfying the above restrictions.

Another generalisation is to view the above problem as that of writing the indicator function of an $n \times m$ rectangle as the sum of indicator functions of squares (with some slight technicalities on the boundaries); this leads to the question of how many squares we require to write the indicator function of the rectangle as a plus/minus sum of indicator functions of squares. In this case we do not need to put any restriction on m. This problem will be called the *plus/minus problem*; the previous problem will be called the *plus problem*.

A more extreme generalisation is to allow infinite sums of indicator functions of squares, allowing these to have real coefficients, and look at the sum of the modulus of the coefficients. (In this case we do not require that the ratio of side lengths is rational.) This is considered by Ruzsa [6]. He proves a uniform bound in the two-dimensional case. The higher-dimensional cases are still open.

Some motivation for this problem comes from discrepancy theory. Suppose that μ is a measure on $Q = [0,1]^d$, that λ is the Lebesgue measure on Q and that $\mathcal{H}$ is a collection of subsets of Q. Then the discrepancy of μ with respect to $\mathcal{H}$ is

$$\sup_{H \in \mathcal{H}} |\mu(H) - \lambda(H)|.$$

For example, suppose that we have N points in Q. For any subset A of Q let $\mu(A)$ be the proportion of the points that lie in A. Then the discrepancy of μ measures how evenly these points are distributed.

A bound for any of the above problems allows us to bound the discrepancy with respect to aligned rectangles (hypercuboids) in terms of the discrepancy with respect to aligned squares (hypercubes). A bound for the more general problem that Ruzsa considered would allow this as well. In fact, it would be equivalent in the sense that if such a decomposition does not exist then there exist measures on $[0,1]^d$ with the discrepancy with respect to hypercubes arbitrarily much smaller than the discrepancy with respect to hypercuboids. In particular, Ruzsa's uniform bound in two dimensions implies that the discrepancy with respect to squares is, to within a constant factor, the same as the discrepancy with respect to rectangles.

We prove polylogarithmic ($C(\log n)^d$) upper and lower bounds for the number of hypercubes needed to tile a hypercuboid in either of the above senses. In more than two dimensions this is a huge improvement on anything previously known. It was not even known whether an $n \times 1 \times 1$ cuboid could be tiled with fewer than n cubes in the plus/minus sense.

We do not aim to get the best constants at any point. In fact we do not aim to get the best exponent for the polylogarithm except when we prove true logarithmic bounds.

As mentioned earlier Kenyon [3] proved a tight logarithmic bound in the two-dimensional plus case (which trivially proves a logarithmic bound for the two-dimensional plus/minus problem). His method is to split the rectangle into a bounded number of pieces, each of which has a ratio of side lengths that has a bounded continued fraction expansion. It is easy to show that the greedy algorithm (remove the largest possible square at each stage) takes a logarithmic number of steps on each of these rectangles. In more than two dimensions there is no sensible notion of a greedy algorithm, removing a cube from one corner will leave an unpleasant shape (whereas in two dimensions it leaves a rectangle). Thus, his method will not generalise to higher dimensions. Indeed, he explicitly states that this is the case.

He gave the same (easy) proof of the lower bound in the two-dimensional plus case; we include the proof here because we generalise it to the plus/minus case and to higher dimensions.

2 The Upper Bounds.

In general we will let $f(n)$ denote the maximum number of hypercubes required to tile a hypercuboid with longest side n satisfying the constraints on side length (e.g. the shortest side being at least $n/2$ for the plus problem). All our upper bounds will be proved inductively. Since there is no reason to suppose that $f(n)$ is monotonic it is convenient to define $F(n) = \max_{m \leq n} f(m)$. The recursive constructions will relate $f(n)$ and $F(m)$ for some $m < n$.

2.1 The Polylogarithmic Bound in Arbitrary Dimension.

First, we have a simple lemma which turns out to be very useful for proving polylogarithmic (upper) bounds.

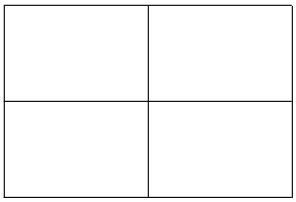
Lemma 1 *Let f be a function $\mathbb{N} \to \mathbb{R}$ and let $F(n) = \max_{m \leq n} f(m)$. Suppose that, for some k and $\alpha < 1$, and all sufficiently large n,*

$$f(n) \leq kF(n^{\alpha}). \qquad (*)$$

Then f has polylogarithmic growth.

The proof is standard and is omitted.

The idea is to split the d-dimensional hypercuboid into 2^d sub-hypercuboids by splitting it in each direction (see below for illustration in two dimensions) such that each sub-hypercuboid has almost the same shape as the original hypercuboid, and each sub-hypercuboid has side lengths with a "high" common factor (n^{ε}). Since tiling an $n_1 h \times n_2 h \times \ldots \times n_d h$ hypercuboid clearly requires the same number of hypercubes as tiling an $n_1 \times n_2 \times \ldots \times n_d$ hypercuboid, this leads to an immediate application of Lemma 1 (with $\alpha = 1 - \varepsilon$).



Lemma 2 *Let $p_1, p_2, \ldots, p_r$ be the first r primes. Then, for any k, the numbers $\{k \cdot p_r! + p_i : 1 \leq i \leq r\}$ are all coprime.*

The proof is trivial and omitted.

Proposition 3 *For any d there exists N such that any $n = n_1 \geq n_2 \geq \ldots \geq n_d$ hypercuboid with $n < 2n_d$ and $n > N$ can be split into 2^d hypercuboids each having longest side at most three times its shortest side and each having a common factor of at least $n^{\frac{1}{2^d+1}}$ between its side lengths.*

PROOF. Let $p_1, p_2, \ldots p_{2^d}$ be the first 2^d primes. Further let k be such that $k \cdot p_{2^d}! > n^{\frac{1}{2^d+1}}$ and $k \cdot p_{2^d}! + p_{2^d} < 2n^{\frac{1}{2^d+1}}$, which we can ensure if N, and thus n, is large enough. For $1 \leq i \leq 2^d$ let $q_i = k \cdot p_{2^d}! + p_i$. Thus, the above condition implies that $n^{\frac{1}{2^d+1}} < q_i < 2n^{\frac{1}{2^d+1}}$ for each i.

For $1 \leq j \leq d$ let I_j be the subset of $\{0,1\}^d$ with jth coordinate 1. Note that all the sets $J_1 \cap J_2 \cap \cdots \cap J_d$, where each J_i is equal to I_i or I_i^c, have size one.

By Lemma 2 we know that the numbers q_i are coprime. Therefore, for each j, we can find integers α_j and β_j such that

$$\alpha_j \prod_{i \in I_j} q_i + \beta_j \prod_{i \notin I_j} q_i = n_j.$$

Since we also get a solution if we add $\prod_{i\notin I_j} q_i$ to α_j and subtract $\prod_{i\in I_j} q_i$ from β_j we may insist that $\alpha_j \prod_{i\in I_J} q_i$ satisfies

$$\frac{n_j}{2} - 2^{2^d} n^{2^d/2^d+1} \leq \frac{n_j}{2} - \prod_{i=1}^{2^d} q_i \leq \alpha_j \prod_{i\in I_j} q_i \leq \frac{n_j}{2} + \prod_{i=1}^{2^d} q_i \leq \frac{n_j}{2} + 2^{2^d} n^{2^{d-1}/2^d+1}$$

and, thus, that the same inequalities hold for $\beta_j \prod_{i\notin I_j} q_i$.

Divide each side of the hypercuboid into two pieces, the jth side being divided into pieces of length $\alpha_j \prod_{i\in I_j} q_i$ and $\beta_j \prod_{i\notin I_j} q_i$. This divides the hypercuboid into 2^d sub-hypercuboids. By the definition of the I_j each of these has side-lengths with a common factor of q_r for some $1 \leq r \leq 2^d$: in particular the sidelengths of each sub-hypercuboid have a common factor of at least $n^{\frac{1}{2^d+1}}$. Thus, to complete the proof all we need to do is bound the ratio of the longest side to the shortest side. Suppose that s is the length of a side. Then

$$\frac{n}{4} - 2^{2^d} n^{2^d/2^d+1} \leq \frac{n_j}{2} - 2^{2^d} n^{2^d/2^d+1} \leq s \leq \frac{n_j}{2} + 2^{2^d} n^{2^{d-1}/2^d+1} \leq \frac{n}{2} + 2^{2^d} n^{2^{d-1}/2^d+1}.$$

Since this is true for any side we see that, provided that N and thus n is large enough, all the hypercuboids have longest side at most three times their shortest side. □

Theorem 4 *Let $f(n)$ denote the maximum number of hypercubes required for an $n = n_1 \geq n_2 \geq \ldots \geq n_d$ hypercuboid with $n_d \geq n/2$. Then $f(n)$ is polylogarithmic.*

PROOF. Suppose that N is as in Proposition 3 and that $n > N$. We can split the hypercuboid into 2^d sub-hypercuboids each with longest side (after the common factor has been removed) at most $n^{1-\frac{1}{2^d+1}}$. However, the ratio of the longest side to the shortest side may be greater than two, but it is at most three. We can split all the sides which are longer than twice the shortest side into two as exactly as possible (this is done after the highest common factor has been removed). Thus, each of the sub-hypercuboids is split into at most 2^{d-1} pieces *and* each of these pieces does satisfy the condition that the longest side is at most twice the shortest.

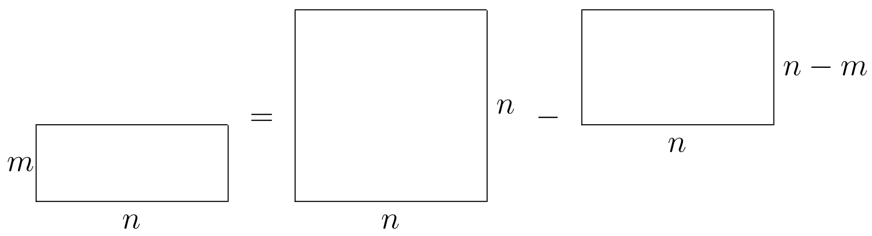
This splits the original hypercuboid into a total of at most 4^d hypercuboids with side lengths (after the common factors have been removed) at most $n^{1-\frac{1}{2^d+1}}$. Thus, we have $f(n) \leq 4^d F(n^{1-\frac{1}{2^d+1}})$ and Lemma 1 applies. □

The exponent in Theorem 4 can be explicitly calculated. Lemma 1 gives the exponent as $\log(1/k)/\log\alpha$. Since $k = 4^d$ we have $\log k = \theta(d)$. Since $\alpha = 1 - \frac{1}{2^d+1}$ for large d we get $\log\alpha = \theta(-2^{-d})$. We find that the bound is $f(n) = O((\log n)^{Cd2^d})$ for some constant C independent of d; i.e., the exponent grows exponentially with dimension. Pór pointed out ([5] personal communication) that the above method could be improved. Instead of splitting each side into two pieces split it into r pieces. Chooses $p_1, p_2, \ldots, p_r$ primes of size about $n^{1/r}$ and write each side as $\sum_{i=1}^{r} \alpha_i (\prod_{j=1}^{r} p_j)/p_i$. Since the terms $(\prod_{j=1}^{r} p_j)/p_i$ are jointly coprime this can be done. Each of the subcubes has a common factor of $n^{\frac{r-d}{r}}$ between its sides. This gives a bound of the form $f(n) < 2^d r^d F(n^{d/r})$. Working through this gives an exponent of size $O(d)$. This is quite close to the lower bound we prove on the exponent of $\Omega(d/2)$.

We get the result for the plus/minus problem almost immediately (recall that in this case we do not restrict the side lengths so this is not a total triviality).

Corollary 5 *Let $f(n)$ denote the maximum number of hypercubes required to write an $n = n_1 \geq n_2 \geq \ldots \geq n_d$ hypercuboid as a plus/minus sum. Then $f(n)$ is polylogarithmic.*

PROOF. We can write the hypercuboid as the plus/minus sum of 2^d hypercuboids with longest side lengths less than twice their shortest side lengths. To see that we can do this we use the following inductive method. Suppose we have a side of length $m < n/2$. Then look at the two-dimensional cross-section given by the side of length n and this side. We can write the $n \times m$ rectangle as $(n \times n) - (n \times (n - m))$ as shown:



Projecting these along the other sides of the hypercuboid we get two hypercuboids each with one less side too short. Repeating the argument on each of these hypercuboids we obtain the desired expression.

Writing each of the hypercuboids in this plus/minus sum as a sum of a polylogarithmic number of hypercubes gives the required expression. $\square$

If we combine the above (trivial) proof in the two-dimensional case with the result of Kenyon then we get a logarithmic upper bound for the plus/minus problem in two dimensions.

However in higher dimensions the best known bound before Corollary 5 was very weak. In fact for a $n \times 1 \times 1$ cuboid it was not even known if it could be done with fewer than n cubes.

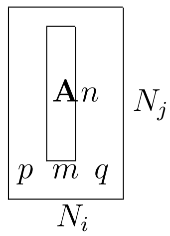
In some circumstances a slightly stronger result is useful. The above result gives no information about where the hypercubes are: it only states that they are in $\mathbb{R}^d$. (Note that the proof does give some information but in some circumstances this is not sufficient.) This lack of information can be inconvenient. One example is when applying the result to discrepancy theory. Suppose that all hypercubes "near" the hypercuboid have small discrepancy. Does that imply that the hypercuboid has small discrepancy? If we could prove that we could write the hypercuboid as the plus/minus sum of a polylogarithmic number of hypercubes which are near to the hypercuboid, then the result would be immediate.

Suppose that our hypercuboid A with side lengths $n_1 \geq n_2 \geq \cdots n_d$ is contained in a hypercuboid B with side lengths $N_1, N_2, \ldots, N_d$. We aim to prove that we can write our hypercuboid as the plus/minus sum of at most a polylogarithmic number of hypercubes all contained in B. Obviously the longest side of B has length at least n_1. If the shortest side of B is less than n_1/C then we will need at least C cubes. Thus, we will insist that all sides of B have length at least n_1/C for some C.

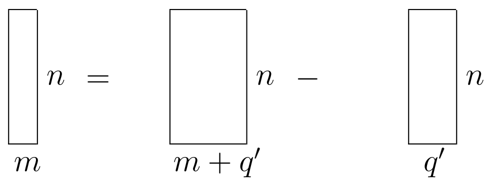
Corollary 6 *Suppose that A is a hypercuboid with sidelengths $n = n_1 \geq n_2 \geq \ldots \geq n_d$ and that A is contained in a hypercuboid B with side lengths $N_1, N_2, \ldots, N_d$. Further suppose that $N_i \geq n/C$ for each i and some fixed constant C. Then we can write A as a plus/minus sum of a polylogarithmic number of hypercubes contained in B.*

PROOF. We show that we can write the hypercuboid as the plus/minus sum of 2^d hypercuboids with longest side lengths less than $3C$ times their shortest side lengths which are contained in B (we will call a side "too short" if it is less than $n/3C$).

To see that we can do this, we use the following inductive method. Suppose we have a side of length $m < n/3C$. Then look at the two-dimensional cross-section given by the side of length n and this side. Suppose that the side lengths of B in this cross section are N_i and N_j. Thus, we have the following picture.



One of p and q (the distances of A from the sides of B) must be at least $n/3C$ (since $p+m+q = N_i \geq n/C$); we may assume $q \geq n/3C$. Thus, we can write the $n \times m$ rectanglular cross-section as



where $q' = q$ unless $q > n-m$ (so $m+q$ would be greater than n) in which case $q' = n-m$. Since we assumed that $q \geq n/3C$, these rectangles do not have too short a side. Projecting these along the other sides of the hypercuboid, we get two hypercuboids each with one less side too short. Since each of the rectangles above is in the cross-section of B, these hypercuboids are in B. Repeating on each of these hypercuboids we see that we can write A as the plus/minus sum of at most 2^d hypercuboids contained in B none of which has too short a side. Writing each of these hypercuboids as a sum of a polylogarithmic number of hypercubes gives the required expression (since all the hypercubes in a *sum* have to be in the hypercuboid, they will all be in B). □

An application of this result is given in [7]

2.2 The Projection Problem.

In this section we investigate a simpler problem also asked by Laczkovich. Suppose that R is a d-dimensional hypercuboid. Pick a side s. The union $\mathcal{H}$ of a collection of hypercubes with disjoint interior will be called a *tiling in the projection sense* if, for every hyperplane P perpendicular to s, the $(d-1)$-dimensional volume of the intersection of P with R is equal to the $(d-1)$-dimensional volume of the intersection of P with $\mathcal{H}$. Obviously a true tiling of R is a tiling in the projection sense since in this case $\mathcal{H} = R$. However, a tiling in the projection sense is much weaker: we are not concerned with how the hypercubes fit together, only with their size.

The motivation for this problem was the fact that very little was known about the original problem in more than two dimensions and that this problem might provide a simpler approach. In particular a lower bound for this problem would be a lower bound for the original problem. Moreover, it was thought that tilings of a hypercube with few hypercubes might not exist for number-theoretic reasons and that the projection problem separated these difficulties from those of how hypercubes fit together.

We obtain a polylogarithmic upper bound in all dimensions but this time with a fixed exponent (i.e., the exponent does not change with dimension). This should be contrasted with the lower bound given in the next section; it indicates that to prove that lower bound (or any improvement on it) we require some idea about how the cubes fit together, not just their sizes. The proof of this upper bound is very similar to that of Theorem 4. First we describe a different problem and show that an upper bound for this implies an upper bound for the projection problem.

Suppose that R is an $m \times N$ rectangle (we use this notation to emphasise that N is going to be much bigger than m). We will call a rectangle with side length k in the m direction and k^{d-1} in the N direction a d-hypersquare. (The motivation for this definition is that a hypercube with side length k will contribute k^{d-1} when projected onto the m side.) A *d-hypertiling* is a dissection of R into d-hypersquares.

Suppose that H is a hypercuboid with side lengths $n_1 \geq n_2 \geq \ldots \geq n_d = m$. Let R be an $n_1 n_2 \cdots n_{d-1} \times m$ rectangle. Then it is easy to see that any d-hypertiling of R with k hypersquares allows us to find a tiling of H in the projection sense using k hypercubes. Indeed, for each hypersquare in the hypertiling, place a hypercube somewhere in the corresponding cross-section through H so that it does not intersect any of the other hypercubes. Since we have unlimited space, this is easy to achieve.

We prove an upper bound for the number of hypersquares necessary to hypertile a rectangle.

Theorem 7 *Let $f(m)$ be the maximum number of hypersquares required to d-hypertile an $m \times N$ rectangle with $m^{d-1} \leq N \leq (2m)^{d-1}$. Then $f(m)$ is polylogarithmic with exponent independent of d.*

PROOF. We are trying to cover an $m \times N$ rectangle with $k \times k^{d-1}$ rectangles. An $m'h \times N'h^{d-1}$ rectangle can be hypertiled using the same number of hypersquares as an $m' \times N'$ rectangle. (For each hypersquare we multiply its short side by h and its long side by h^{d-1}. Thus, since it has scaled correctly it remains a hypersquare.)

As in Theorem 4 we show that there exists M such that any $m \times N$ rectangle with $m > M$ can be split into four smaller rectangles each of which has sides with a large common factor (where common factor means a number p such that p divides the short side and p^{d-1} divides the long side).

By Lemma 2 pick p, q, r, s coprime with $m^{\frac{1}{5}} < p, q, r, s < 2m^{\frac{1}{5}}$, which we can do provided that M and thus m is large enough. Then we can find integers α, β such that $\alpha pq + \beta rs = m$. Since we get another solution if we add rs to α and subtract pq from β we may insist that

$$\frac{m}{2} - pqrs \leq \alpha pq \leq \frac{m}{2} + pqrs$$

and thus that the same inequalities hold for βrs.

Similarly, we can find integers γ, δ such that

$$\gamma p^{d-1} r^{d-1} + \delta q^{d-1} s^{d-1} = N.$$

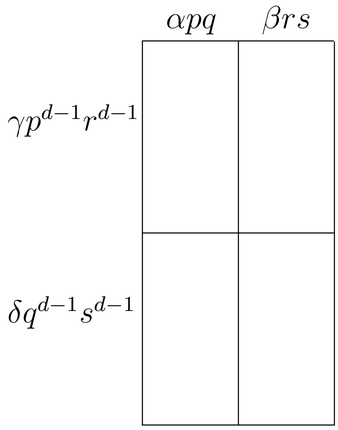
If we add $q^{d-1}s^{d-1}$ to γ and subtract $p^{d-1}r^{d-1}$ from δ we still have a solution, so we may insist that

$$\frac{N}{2} - (pqrs)^{d-1} \leq \gamma p^{d-1} r^{d-1} \leq \frac{N}{2} + (pqrs)^{d-1}$$

and thus that $\delta q^{d-1} s^{d-1}$ satisfies the same inequalities.

We split the rectangle into four pieces with side lengths $\alpha pq, \beta rs, \gamma p^{d-1} r^{d-1}$

and $\delta q^{d-1}s^{d-1}$ as shown.



Substituting the bounds on p, q, r, s and N we see that

$$\frac{m}{2} - 2^4 m^{\frac{4}{5}} \leq \alpha pq \leq \frac{m}{2} + 2^4 m^{\frac{4}{5}}$$

and

$$\frac{m^{d-1}}{2} + 2^{4d} m^{\frac{4(d-1)}{5}} \leq \gamma p^{d-1} r^{d-1} \leq 2^{d-2} m^{d-1} + 2^{4d} m^{\frac{4(d-1)}{5}}.$$

Combining these inequalities we have

$$(\alpha pq)^{d-1} \leq \gamma p^{d-1} r^{d-1} \leq 2^{2d} (\alpha pq)^{d-1}.$$

Provided that M, and thus m, is large enough, we can remove at most 2^{2d} rectangles of size $\alpha pq \times (\alpha pq)^{d-1}$ to get back to a rectangle with the correct ratio between its side lengths. Cancelling the common factor we reduce to a rectangle with side lengths m' and N' where $m' \leq m^{\frac{4}{5}}$ and $m'^{d-1} \leq N' \leq (2m')^{d-1}$. This can be hypertiled using at most $F(m^{\frac{4}{5}})$ rectangles.

Since the same is true for each of the other rectangles we have

$$f(m) \leq 4F(m^{\frac{4}{5}}) + 2^{2d+2}$$

for $m > M$. This is of the form required by Lemma 1, and by noting that the exponent it gives is independent of the constant (which does depend on dimension), we get the required result. □

Theorem 8 *Let $f(n)$ denote the maximum number of hypercubes needed to tile, in the projection sense, a hypercuboid with sides of length $n = n_1 \geq n_2 \geq \ldots \geq n_d > n/2$. Then f is polylogarithmic with exponent independent of dimension.*

PROOF. Let $m = n_d$. Initially we assume we are projecting onto this (the shortest) side. Let $N = n_1 n_2 \dots n_{d-1}$. Thus, we have

$$(m)^{d-1} \leq N \leq (2m)^{d-1}. \tag{$*$}$$

Then by Theorem 7 and the remarks proceeding it we see that we can tile the hypercuboid, in the projection sense, with a polylogarithmic number of hypercubes where the exponent is independent of dimension.

To extend to an arbitrary side (i.e., not necessarily the shortest side) we just split the new side into a (bounded but dependent on dimension) number of bits such that each of the new rectangles formed does satisfy ($*$). This affects the constant multiplying the logarithm but not the exponent. □

3 Lower Bounds.

Using very standard methods we prove a logarithmic lower bound in two dimensions and then use this to get polylogarithmic lower bounds in higher dimensions.

Suppose we have a tiling of an $n \times m$ rectangle with k squares. The main idea is that a tiling of a rectangle by squares can be regarded as a resistor network. This idea was first used to prove that a squared rectangle has side lengths with rational ratio and to construct examples of squared squares. (For more on these results and a general overview of resistor networks see Chapter 2 of *Graph Theory.* [1]). Suppose that we have a tiling of a rectangle with squares. We view the rectangle as being made of a resistive lamina. The resistance from one side of a rectangle to the opposite side is then just the ratio of the side lengths. In particular the resistance of any square is one. We do not affect the resistance of the lamina if we cut down the sides of the squares which are in the direction of the current flow (no current is flowing across the cut) or if we put perfect conductors along the sides of the square at constant potential (no current flows along these sides). Since each square has resistance one, this gives us a resistor network with all resistances one and total resistance the ratio of the side lengths of the rectangle.

Suppose we have a resistor network with k resistors (of resistance one) and total resistance m/n. Now by standard results on resistor networks the current in a resistor is given by

$$\frac{N(s,t,a,b) - N(s,t,b,a)}{N},$$

where a current of size one enters the network at s and leaves at t, $N(s,t,a,b)$ denotes the number of spanning trees of the network with ab an edge occuring in that order on the (unique) st path, and N denotes the total number of spanning trees. Now since each resistor has unit resistance, the potential difference across it is equal to the current through it which is an integer times $1/N$. Thus, the total potential drop across the network is an integer times $1/N$. Since the total current is one, the resistance of the network is an integer times $1/N$. Recall that the resistance of the network is m/n. Thus, if m and n are coprime we require $N \geq n$.

However, the total number of subgraphs is 2^k (k the number of resistors as above) since each edge is either present or absent. The number of spanning trees is less than the number of subgraphs so $N < 2^k$. Thus, if n and m are coprime, $n \leq 2^k$; i.e., $k \geq \log_2 n$. We have proved the following (first proved by Kenyon [3] using the same method):

Theorem 9 *If n and m are coprime then at least $\log_2 n$ squares are required to tile an $n \times m$ rectangle.*

Recall that Kenyon also proved a logarithmic upper bound and so this bound is tight.

Notice that if we know how the squares "fit together" (i.e., which other squares each one touches) then we know the resistor network and thus the size of each of the squares and the ratio of the rectangle. We can construct such a resistor network for the two-dimensional projection problem giving a logarithmic lower bound for this simpler problem.

We now extend this to higher dimensions. We observe that not very many of the hypercubes can be big. This allows us to find a cross-section through the hypercuboid with "few" hypercubes in it. We then apply Theorem 9 to this cross-section. We may assume that the hypercuboid has no common factor between its side lengths.

Lemma 10 *Suppose that $n = n_1 \geq n_2 \geq \ldots \geq n_d$ are jointly coprime integers. Then there exists j such that the highest common factor of n and n_j is at most $n^{1-\frac{1}{d-1}}$.*

PROOF. Since every prime factor of n must be "missing" from some n_j, we have

$$\prod_{j=2}^{d} \text{hcf } (n, n_j) \le \frac{n^{d-1}}{n} = n^{d-2}.$$

Therefore we can choose j such that hcf $(n, n_j) \le n^{\frac{d-2}{d-1}} = n^{1-\frac{1}{d-1}}$ as required. □

Theorem 11 *Given an $n = n_1 \ge n_2 \ge \ldots \ge n_d$ hypercuboid with $n_d \ge n/2$. Suppose the n_i, $1 \le i \le d$, are (jointly) coprime. Then $\Omega((\log n)^{d/2})$ hypercubes are required to tile the hypercuboid.*

PROOF. By Lemma 10 we can choose j such that the highest common factor of n and n_j is at most $n^{1-\frac{1}{d-1}}$. Now define a *slice* to be a two dimensional cross-section through the hypercuboid parallel to the $1j$ edge. There are at least $(n/2)^{d-2}$ of them. The number of slices that a hypercube of side length m appears in is m^{d-2}. Therefore the average number of cubes in a slice is at most

$$\sum_{i=1}^{k} \left(\frac{2m_i}{n}\right)^{d-2},$$

(where k is the number of hypercubes and $m_1, m_2, \ldots, m_k$ are their side lengths). However, by equating volumes, we have $\sum_{i=1}^{k} m_i^d \le n^d$, or equivalently $\sum_{i=1}^{k} \left(\frac{m_i}{n}\right)^d \le 1$. By convexity this constraint shows that

$$\sum_{i=1}^{k} \left(\frac{m_i}{n}\right)^{d-2} \le k^{2/d}.$$

Thus, the average number of cubes in a slice is at most $2^{d-2}k^{2/d}$. Fix a slice with at most this many hypercubes in it. It is a rectangle with side lengths n and n_j. After we cancel the common factor we get an $n' \times m'$ rectangle with $n' \ge n^{\frac{1}{d-1}}$. Applying Theorem 9 to this slice (and observing that hypercubes have square cross-sections) we have $2^{d-2}k^{2/d} \ge \log(n^{\frac{1}{d-1}})$. Rearranging gives the required bound. □

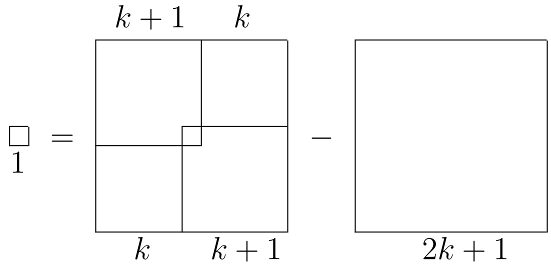
As mentioned earlier we see that this is bigger than the upper bound for the projection problem. The choosing of a slice is the part that is not possible in that case. Indeed, since there does not seem to be any natural reduction of the projection problem in high dimensions to the problem in two dimensions, we do not even have a logarithmic lower bound.

Next we turn to the plus/minus problem. Theorem 9 goes through as before; we view a minus square as a square of resistance -1. It can be checked that all the work for resistor networks is still valid. We replace N by a weighted function N^* defined to be the sum over all spanning trees of the product of the conductances of the tree's edges (similarly for $N(s,t,a,b)$). Since all resistances are ± 1 we still have the bound as before. However, there are some small differences. Firstly N^* may be zero. This corresponds to the case of a zero resistance (e.g. a $+$ and $-$ resistor in parallel). Since this corresponds to a tiling of a line we ignore this case. Secondly we do not have a unique current flow. For example we can have a current flowing around a cycle. (This is expected; we can place a square of arbitrary size and then cancel it with a minus square of the same size and position. The arbitrary size of this square corresponds to an arbitrary current.) However, this does not affect the result: although we can deform the arrangement (i.e., change the size of some of the squares) we cannot change the overall ratio of the rectangle.

Theorem 12 *If n and m are coprime then at least $\log n$ squares are required to write an $n \times m$ rectangle as a plus/minus sum of them.*

Combining this result with Kenyon's result (in the form observed after Corollary 5) we get an exact logarithmic bound on the number of squares that are required.

However, we cannot modify the proof of Theorem 11 to work in this case. There can be many hypercubes in each slice. (In particular, in the plus case "many" hypercubes have to be small; in the plus/minus case they can all be large, even if the sum is still efficient. This can be achieved in two dimensions by writing each square as the plus/minus sum of five arbitrarily large squares as shown.



This only increases the number of squares by a factor of five.) The only result we have in this case is the following weak theorem.

Theorem 13 *Let H be a hypercuboid with sidelengths $n = n_1 \geq n_2 \geq \ldots \geq n_d$. Suppose that the n_i, $1 \leq i \leq d$, are (jointly) coprime. Then $\Omega(\log n)$ hypercubes are required to write H as a plus/minus sum of them.*

PROOF. By Lemma 10 we can find j such that the highest common factor of n and n_j is at most $n^{1-\frac{1}{d-1}}$. Each cross-sectional slice parallel to the $1j$ edge has side lengths n and n_j. After we cancel the common factor we get an $n' \times m'$ rectangle with $n' \geq n^{\frac{1}{d-1}}$. Applying Theorem 12 to any such slice we find that the number of hypercubes used is $\Omega(\log(n^{\frac{1}{d-1}})) = \Omega(\log n)$ as required. □

This is one place where the upper and lower bounds differ significantly. The best lower bound for the plus/minus problem in any number of dimensions is logarithmic. The upper bound is $O((\log n)^d)$ (using Pór's improvement). It would definitely be interesting to know whether the exponent of the logarithm tends to infinity. We do not even know that it is not always one.

References

[1] B. Bollobás, Graph Theory, Springer Verlag. 1979

[2] M. Dehn, Über die Zerlegung von Rechtecken in Rechtecke, Math. Ann. 57. (1903) 314-322.

[3] R. Kenyon, Tiling a rectangle with the fewest squares, J. Combin. Theory. Ser. A. 76. (1996) 272-291.

[4] M. Laczkovich, Personal Communication.

[5] A. Pór, Personal Communication.

[6] I. Z. Ruzsa, The discrepancy of rectangles and squares, Grazer Math. Ber. 318. (1992) 135-140.

[7] M. J. Walters, PhD Thesis "Ramsey Theory, Discrepancy Theory and Related Areas". University of Cambridge (1999)