

HIRSCH CONJECTURE AND THE SIMPLEX ALGORITHM

SUNIL ABRAHAM
ST. JOHN'S COLLEGE

3320 THORNHILL DR.
ADRIAN, MI 49221
USA

I declare that this essay is work done as part of the Part III Examination. I have read and understood the Statement on Plagiarism for Part III and Graduate Courses issued by the Faculty of Mathematics, and have abided by it. This essay is the result of my own work, and except where explicitly stated otherwise, only includes material undertaken since the publication of the list of essay titles, and includes nothing which was performed in collaboration. No part of this essay has been submitted, or is concurrently being submitted, for any degree, diploma or similar qualification at any university or similar institution.

Signed: May 5, 2011

Contents

1	Introduction	3
2	Disproving Hirsch's Conjecture	5
3	Bounds	8
3.1	Unrestricted Paths	8
3.2	Monotone Paths	11
4	Special Polytopes	13
4.1	(0,1)-Polytopes	13
4.2	Network Flow Polytopes	14
4.3	Leontief substitution systems	16
5	Concluding thoughts	18
5.1	Random Polytopes	18
5.2	Recent Work	19
5.2.1	Subexponential lower bounds for random pivot rules	19
5.2.2	Polymath3	19
5.3	Further Directions	20

1 Introduction

The final test of a theory is its capacity to solve the problems which originated it.
 — George Dantzig

Consider the general problem of minimizing a linear objective function, in n variables, subject to m linear equality constraints. We express this as the primal problem,

$$(P): \text{minimize } \{c^T x : Ax = b, x \geq 0\}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. We assume A is of full row rank, m , and that (P) has a feasible solution. We can then express the feasible region of (P) geometrically as the intersection of an $n - m$ dimensional affine subspace, as defined by $Ax \leq b$, with the positive orthant, as $x \geq 0$. This is the immediate geometric interpretation, but interestingly Dantzig's reformulation of the feasible set as the convex hull of points $A_j = (a_j, c_j)$, where a_j is the j th column of A and c_j is the j th component of c , led him to reconsider the *simplex algorithm* as an efficient solver of (P) [6, Chapter 7]. For now, let us satisfy ourselves with basics, first a definition of the *feasibility polyhedron*, P :

$$P := \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}.$$

It may be useful at times to denote P as

$$P = P(A, b).$$

Also for convenience, we say P is in the *class* (d, n) , if P is a d -dimension polytope with n facets. The triple (P, x, y) consists of a polytope P and vertices $x, y \in P$. (P, x, y) is in the *class* (d, n) if P is in the class (d, n) .

By our assumption of A being rank m , we have that P is a polyhedron with vertices. We abuse standard notation here a little, as a polyhedron has extreme points, but the polyhedron's graph has vertices, which naturally correspond to the polyhedron's extreme points. For the sake of simplicity, the notion of a polytope and its graph are treated as the same in this paper. As we are trying to minimize a linear function over a affine space, we can limit our search to the space's extremities, the polyhedron's vertices. This is effectively what the simplex algorithm does. Each vertex of P is a basic feasible solution, and the algorithm moves along adjacent vertices of P until it strikes an optimal solution. The simplex algorithm's pivot rule decides which adjacent vertex to choose. The pivot rule is locally determined, by the adjacent vertices, not globally determined, by all of the polyhedron's vertices, so, a priori, we do not know how many pivots the algorithm will require.

Most deterministic rules have been found to have a exponential worst-case complexity. In 1969, Klee and Minty were the first to give an example of such a case [20]. For Dantzig's classic rule of most reduced cost per pivot, they prove that $2^m - 1$ pivots may be required, where m is both the number of inequality constraints and nonnegative variables. (As an aside, most sources indicate that Klee and Minty demonstrated this in 1972, but in fact they presented this result in 1969 at the Third Symposium on Inequalities. The confusion likely lies in the proceedings of the conference was published later, in 1972.) Somewhat surprisingly then, non-deterministic pivot rules have been discovered with sub-exponential, but super-polynomial, worst-case complexity. Kalai [16] and, later, independently, Matoušek et al. [24] presented a randomized sub-exponential pivot rule and proved that the expected number of pivots and the expected number of arithmetic operations, both in the worst case, are bounded by $m^{O(\sqrt{n}/\log n)}$ and $\exp(K\sqrt{n} \log m)$, respectively, for (P) in n variables and m inequalities (not equalities).

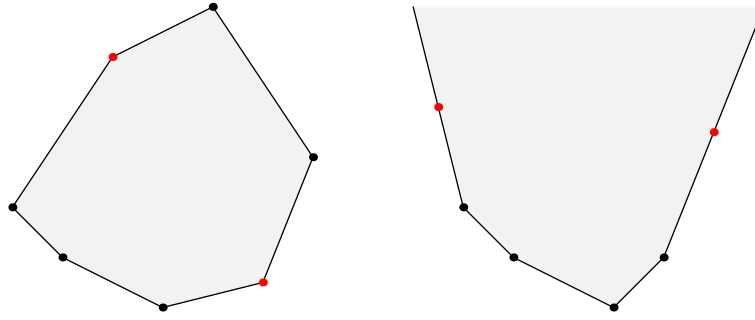


Figure 1: $\Delta(2, n) = \lfloor n/2 \rfloor$ and $\Delta_u(2, n) = n - 2$. The above is an example of $f = 7$.

Nonetheless, linear programming is a class P problem. In 1979, Khachian [17] demonstrated that the ellipsoid method of Shor [35] and Judin and Nemirovskii [13], [14] solves (P) polynomially in its worst case. However, in practice, the ellipsoid method is rarely used; through observation, it became clear that variants of the simplex, albeit exponential, were very efficient in practice, whereas the ellipsoid method, though polynomial, was very inefficient for the same problem [26].

The simplex algorithm typically takes between m to $3m$ pivots to terminate, for an m -equation problem with m different variables in the final basic set. This empirical observation was made as early as 1963 by Dantzig [6] and Wolfe and Culter [40].

Efforts to explain the success of the simplex algorithm have looked at the *combinatorial diameter* of a polytope, P , which we define as

$$\delta(P) := \text{the minimum number of edges in a path between any two vertices in } G$$

Further let us define

$$\Delta(d, n) := \text{the maximum diameter of a polytope in class } (d, n), \text{ and}$$

$$\Delta_u(d, n)_u := \text{the maximum diameter of an unbounded polytope in class } (d, n).$$

Then, if the simplex algorithm chose vertices optimally, but were initiated at the worst vertex, the algorithm would require Δ or Δ_u pivots. In the 1957, Hirsch conjectured to Dantzig through correspondence that $\Delta(d, n) \leq n - d$ and $\Delta_u(d, n) \leq n - d$ [6]. For this paper, we separate this into two cases:

Conjecture 1.1. (Bounded Hirsch's Conjecture) $\Delta(d, n) \leq n - d$.

Conjecture 1.2. (General Hirsch's Conjecture) $\Delta_u(d, n) \leq n - d$.

In relation to our interest in linear programming, Hirsch's conjecture is connected with the complexity of the simplex algorithm. In particular, it connects to the question of whether there can exist a pivot rule such that the simplex algorithm is a strongly polynomial algorithm for solving (P).

This essay provides an overview of Hirsch's Conjecture from its proposition over 50 years ago to current work presented earlier this year. In Section 2, we set review the developments that led to the disproof of Hirsch's Conjecture. Because we know the Hirsch Conjecture to be false, in Section 3, we highlight the best bounds on general polytope diameter that we do know. Particular polytopes such as the assignment polytope, network flow polytope, and shortest route polytopes are of particular interest to academics and practitioners, and because of these

polytopes are highly structured, we know very good results for their diameters. We will discuss these specific cases in Section 4. Finally, we overview recent work and the direction of future work, particularly of the polymath 3 project.

2 Disproving Hirsch's Conjecture

I will first examine the difference between imagination and pure understanding. When I imagine a triangle, for example, I do not merely understand that it is a figure bounded by three lines, but at the same time I also see the three lines with my mind's eye as if they were present before me; and this is what I call imagining. But if I want to think of a chiliagon, although I understand that it is a figure consisting of a thousand sides just as well as I understand the triangle to be a three-sided figure, I do not in the same way imagine the thousand sides or see them as if they were present before me...in doing this I notice quite clearly that imagination requires a peculiar effort of mind which is not required for understanding; this additional effort of mind clearly shows the difference between imagination and pure understanding.
— René Descartes [7]

We should admit that we now know that Hirsch's conjecture both for bounded and unbounded polytopes is false. However, how we came to know this is instructive. So we go through some of the major and interesting results that led to the disproof.

Claim 2.1. To find Δ , it is sufficient to consider simple polytopes.

Proof. For any polytope in class (d, n) , we construct a simple polytope in class (d, n) with diameter at least as large as the original's.

Consider a polytope $P = P(A, b)$ in class (d, n) . There exists b' , a perturbation of b , such that $P' = P(A, b')$ is simple. To make it explicit, let $b'_i = b_i + \epsilon^i$ and take ϵ to be small. By a natural bijection between P and P' , P' is then also a (d, n) class polytope; however, by perturbing the hyperplanes, increased the number of vertices and reduced the number of facets adjacent to some vertices.

These new vertices cannot have reduced the polytope's diameter though; they were only added between vertices in P that were adjacent to more than d facets and their neighbors. Hence, to find Δ , and prove Hirsch's conjecture, it is sufficient to consider simple polytopes.

Note: This claim is made in Klee and Walkup [21] and Ziegler [42] but no proof is given; this argument is the author's work, so any mistake is the author's alone. \square

Working on simple polytopes, we have the following conjecture by Klee [19]:

Conjecture 2.2. (Nonrevisiting path conjecture) Any two vertices of a simple polytope can be joined by a path that does not visit a facet more than once.

Clearly, if the nonrevisiting path conjecture were true, then Hirsch's conjecture would be too. To see this, note that the initial vertex lies on d facets, and each subsequent vertex in the nonrevisiting path is adjacent to at least one facet that has not been visited. Hence the path cannot be greater than $n - d$. Hence, we have that the nonrevisiting path conjecture implies Hirsch's conjecture.

A special case of Hirsch's conjecture is when $n = 2d$:

Conjecture 2.3. (Bounded d -step Conjecture) $\Delta(d, 2d) \leq d$.

Conjecture 2.4. (General d -step Conjecture) $\Delta_u(d, 2d) \leq d$.

Surprisingly, the the bounded d -step conjecture and the nonrevisiting path conjecture are both equivalent to the bounded Hirsch conjecture [21].

Theorem 2.5. *The following are equivalent.*

1. (Bounded Hirsch's Conjecture) $\Delta(d, n) \leq n - d$.
2. (Bounded d -step Conjecture) $\Delta(d, 2d) \leq d$.
3. (Nonrevisiting path conjecture) Any two vertices of a simple polytope can be joined by a path that does not visit a facet more than once.

To prove Theorem 2.5 we introduce some definitions in the convention of Klee and Walkup [21].

The *product* of d_1 dimensional polyhedron P_1 and d_2 dimensional polyhedron is $d_1 + d_2$ dimensional polyhedron $P_1 \times P_2$.

Now we define the *wedge* of a (d, n) polytope P over foot k -face K . Take C to be the product of L and P , where $L = [0, \infty)$. Take H to be a hyperplane in the space spanned by C such that $K = H \cap P$ and H is in the interior of C . H divides C into two $d + 1$ -dimensional polyhedra. By construction, only one of these polyhedra contains P . Take W be the polyhedron that contains K . Now, W is called the *wedge* of P over K . Observe that if P is bounded or simple, then W is too. Define the *lower base* and *upper base* of W to be P and $H \cap C$, respectively. Now for a vertex x of P not in K , define x' to be $x \times (L \cap H)$. Call x' to be the vertex *over* x in relation to W , as would follow from our intuition of the natural projection of the upper base onto the lower base. Note that the lower and upper bases are combinatorially equivalent. The wedge clearly is a $d + 1$ -dimensional polytope, but what of its facets? Each facet of P save K is extended, so we there we have $n - 1$ facets of W . We also have the lower and upper bases as facets, so in total, we have $n + 1$ facets. Hence, W is in the class $(d + 1, f + 1)$. See Figure 2 for an illustration for a polytope P in class $(2, 4)$.

Lemma 2.6. The following is useful in our proof of Theorem 2.5

1. A path, l , in a face F of polytope P is nonrevisiting if and only if l is nonrevisiting in P .
2. A path, l in the product of two polytopes P_1 and P_2 is nonrevisiting if and only the projections of l in P_1 and P_2 are nonrevisiting in their respective polytopes.
3. If a path, l , in a wedge W over P is nonrevisiting, then the projection of l onto P is nonrevisiting. [21]

We are now ready to prove Theorem 2.5.

Proof. From earlier we have that the nonrevisiting path conjecture implies Hirsch's conjecture. As the d -step conjecture is a special case of Hirsch's conjecture, the latter implies the former. All that is left is to prove that the d -step conjecture implies the nonrevisiting path conjecture, which is what we do now.

Suppose the d -step conjecture is true, that $\Delta(d, 2d) \leq d$. Then consider a bounded and simple triple (P, x, y) in the class $(m, 2m)$. Take $y_0 = y$. Consider the face of smallest dimension such

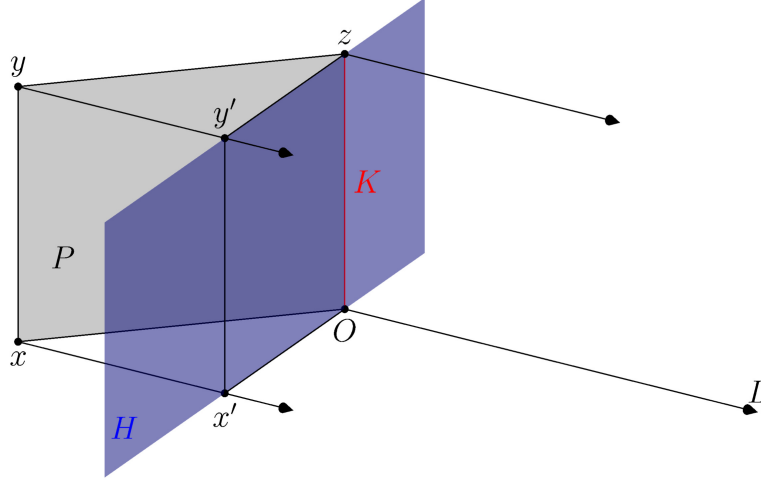


Figure 2: Wedge, W , of a 2 dimensional polytope, P , over its facet, K . P is a quadrilateral, of which K is an edge.

that x and y_0 are incident to it. Call this face P_0 . Then (P_0, x, y_0) is a bounded and simple triple in the class $(d_0, d_0 + m_0)$. Note that $m_0 = d_0 + k_0$, where k_0 is the number of facets of P_0 that are incident neither to x nor y_0 . If $k > 0$, take K_0 to be a facet that is neither incident to x nor y_0 . Now let P_1 be the wedge of P_0 over K_0 , and take y_1 to be the vertex over y_0 in relation to this wedge. We have that (P_1, x, y_1) is a bounded and simple triple in the class $(d_1, d_1 + m_0)$, where $d_1 = d_0 + 1$. However, what is key is that P_1 has $k_1 = k_0 - 1$ facets that are incident to neither x and y_1 , and none incident to both. Hence we iterate this process k_0 times until we get the bounded and simple triple (P_{k_0}, x, y_{k_0}) in the class $(d_0 + k_0, d_0 + k_0 + m_0) = (m_0, 2m_0)$.

By our construction, every facet of (P_{k_0}, x, y_{k_0}) is incident to either x or y_{k_0} and none incident to both. P_{k_0} is simple, so each vertex is adjacent to m_0 edges and facets, but there are $2m_0$ facets. A path from x to y_{k_0} then certainly visits every facet. Hence, such a path must have at least $2m_0 - m_0 = m_0$ edges. However, we have assumed $\Delta(d, 2d) \leq d$, so there is a path from x to y_{k_0} with at most m_0 edges. Hence, this path has exactly m_0 edges, and, because P_{k_0} is simple, this path must be nonrevisiting. We now project this path onto P_{k_0-1} to get a nonrevisiting path from x to y_{k_0-1} . Iterate until a nonrevisiting path from x to y in P_0 .

As there exists a nonrevisiting path from x to y in P_0 a face of P , there exists a nonrevisiting path from x to y in P . \square

Klee and Walkup [21] conclude by disproving the general d -step conjecture for $d \geq 4$. They construct an example to illustrate that $\Delta_u(4, 8) \geq 5$. Taking products of this counterexample, they demonstrate the following

Proposition 2.7. (Klee and Walkup [21]) $\Delta_u(d, n) \geq n - d + \min\{\lfloor d/4 \rfloor, \lfloor (n - d)/4 \rfloor\}$.

Recently, Santos [32] gives us an improved lower bound for large n .

Proposition 2.8. (Santos [32]) $\Delta_u(d, n) \geq (1 + \epsilon)(d + n)$, where d and ϵ are fixed and n is arbitrarily large.

The bounded Hirsch conjecture was disproven last year by Santos in the same paper [32]. Santos constructs his counterexample by looking at a special class of polytope he calls *spindles*. A d -*spindle* is a d -polytope P having two vertices x and y such that every facet of P is incident

to either x or y but not both. The *length* of the spindle is the distance between x and y . Note that in our proof of Theorem 2.5, (F_{k_0}, x, y_{k_0}) is an m_0 -spindle. Santos proves the following d -step theorem for spindles:

Theorem 2.9. Santos [32]) *If P is a spindle in the class (d, n) and has length l , there exists another spindle, P' , in the class $(d - n, 2d - 2n)$ with length at least $l + n - 2d$.*

Then disproving the bounded Hirsch conjecture amounts to finding a spindle whose length exceeds d . Santos does just this by finding a spindle in class $(5, 48)$ of length 6. Hence, there is a spindle in class $(43, 86)$ with diameter at least 44, and the bounded Hirsch conjecture was proved false [32]. Since the original counterexample, Santos has since found a counterexample of a smaller dimension: a polytope in class $(43, 86)$ with 44 vertices [33]. Interestingly, Santos's counterexample is not a simple polytope; as we showed earlier, we can reduce the problem of maximum diameter of a polytope to maximum diameter of a simple polytope, but clearly here there was great value in thinking about the problem more generally.

Even more amazing, building on Santos's work, Weibel [39] announced 23 and 20 dimension polytope counterexamples. Weibel's goal was to find a low dimension counterexample to Hirsch conjecture and to minimize the number of vertices of the polytope. He does this by considering the *Minkowski sum* of particular polytopes. For polytopes P_1 and P_2 , their Minkowski sum is $P_1 + P_2 := \{x_1 + x_2 : x_1 \in P_1, x_2 \in P_2\}$. More to the point, Weibel utilizes that when P_1 and P_2 have no facets parallel with the other, then any facet of P_1 or P_2 has a corresponding parallel facet of the same size and shape in $P_1 + P_2$. He calls these facets *pure facets*. Weibel claims that counterexamples can be built from any Minkowski sum in which the pure facets from different polytopes do not share any vertices in the sum polytope. He further claims that this can only happen when each of the summand polytopes has dimension at least 4.

3 Bounds

*You are a lover; borrow Cupid's wings,
And soar with them above a common bound.*
— Mercutio [34]

3.1 Unrestricted Paths

The result of the Hirsch conjecture being false is fascinating, but has little practical significance. We know that $n - d$ is not an upper bound for Δ or Δ_u , but for all we know, the upper bound could still be linear in n and d . The best known upper bounds are lacking.

Lemma 3.1. (Larman [23]) For any two vertices x and y in a d -dimensional polyhedron, there is a path that visits each face at most 2^{d-3} times .

We do have Δ_u bounded linearly in n and exponentially for d .

Theorem 3.2. (Larman [23]) $\Delta_u(d, n) \leq n2^{d-3}$.

Proof. Let x and y be vertices on a d -dimensional polyhedron, and let K be a path that begins with x and ends with y and visits each face of P at most 2^{d-3} times. Such a path exists by Lemma 3.1. For any two consecutive vertices in K , there is a facet that contains one but not the other, and conversely. Then, with each subsequent edge in K , there is a facet that

has just been visited, and one that is newly visited. If the length of K were greater than $n2^{d-3}$, there would then be at least one facet that was visited more than 2^{d-3} times, which is a contradiction. Hence, the path is less than $n2^{d-3}$, and because x and y were arbitrary, given P was d -dimensional, we have $\Delta_u(d, n) \leq n2^{d-3}$. \square

For classes of polynomials in which n and d are close together, Kalai and Kleitman [15] derived a better bound of $n^{\log_2(d)+2}$. However, we can sharpen their result to n^{\log_d+1} .

Lemma 3.3. (Kalai and Kleitman [15]) $\Delta_u(d, n) \leq \Delta_u(d-1, n-1) + 2\Delta_u(d, \lfloor n/2 \rfloor) + 2$.

Proof. Take x and y to be vertices on a polyhedron P in class (d, n) . Let k_x be the maximal number such that the union of the vertices in paths of length k_x starting from x is incident to no more than $n/2$ facets. Define k_y similarly. There is a facet F such that paths to F from x and y are lengths $k_x + 1$ and $k_y + 1$.

We prove that $k_x \leq \Delta_u(d, \lfloor n/2 \rfloor)$. Create a polyhedron Q from P by deleting the inequality constraints that correspond to facets that cannot be reached from a path of length k_x starting from x . So Q is in class $(d, \lfloor n/2 \rfloor)$. Consider a vertex z in P such that the distance from x to z in P is k_x . Then the distance from x to z in Q is less than k_x . We show this by supposing that the distance in Q is less than k_x . Consider a path in which this is the case. Then there must be an edge that is not in P . Consider the first occurrence of such an edge, E , in the path. As E is in Q but not P , E must intersect one of the hyperplanes, H , corresponding to a constraint we ignored to extract Q from P . Then, in P , the distance from x to H is less than k_x . This is a contradiction, by our construction, the distance from x to an ignored facet must be more than k_x . Hence, $k_x \leq \Delta_u(d, \lfloor n/2 \rfloor)$.

Then, the distance from x to F is bounded by $\Delta_u(d, \lfloor n/2 \rfloor) + 1$. The distance between any two vertices in F is bounded by $\Delta_u(d-1, n-1)$. And, the distance from F to y is bounded by $\Delta_u(d, \lfloor n/2 \rfloor) + 1$. Hence we have the inequality, $\Delta_u(d, n) \leq \Delta_u(d-1, n-1) + 2\Delta_u(d, \lfloor n/2 \rfloor) + 2$, which gives us our result. \square

Now we are prepared to prove a better bound. The following is entirely original work. At first, the author thought no one had derived this result, as the bound of $n^{\log_2(d)+2}$ is what is cited in the literature. However, since our initial derivation of the sharpened bound, $n^{\log_2(d)+1}$, we have found a paper from 2009 Kim and Santos [18] that used similar techniques to get the same result.

Theorem 3.4. $\Delta_u(d, n) < n^{\log_2(d)+1}$.

Proof. We have

$$\Delta_u(d, n) \leq \Delta_u(d-1, n-1) + 2\Delta_u(d, \lfloor n/2 \rfloor) + 2,$$

but since Δ is monotonic increasing in d and f , we work with a cruder bound of

$$\Delta_u(d, n) \leq \Delta_u(d-1, n) + 2\Delta_u(d, \lfloor n/2 \rfloor) + 2,$$

Let's define

$$h(n, i) := \lfloor \lfloor n/2 \rfloor / \dots \rfloor / 2 \rfloor,$$

where we have i divisions by 2. We bound $\Delta_u(d, n)$ by a standard telescoping argument.

$$\begin{aligned}
\Delta_u(d, n) - 2\Delta_u(d, \lfloor n/2 \rfloor) &\leq \Delta_u(d-1, n) + 2 \\
2\Delta_u(d, \lfloor n/2 \rfloor) - 4\Delta_u(d, h(n, 2)) &\leq 2\Delta_u(d-1, \lfloor n/2 \rfloor) + 4 \\
&\vdots \leq \vdots \\
2^{\lfloor \log_2 \frac{n}{d} \rfloor} \left(\Delta_u(d, h(n, \lfloor \log_2 \frac{n}{d} \rfloor)) - 2\Delta_u(d, h(n, \lfloor \log_2 \frac{n}{d} \rfloor + 1)) \right) &\leq 2^{\lfloor \log_2 \frac{n}{d} \rfloor} (\Delta_u(d-1, h(n, \lfloor \log_2 \frac{n}{d} \rfloor)) + 2)
\end{aligned}$$

Hence we get

$$\Delta_u(d, n) \leq \sum_{i=0}^{\log_2(n/d)} 2^i (\Delta_u(d-1, h(n, i)) + 2).$$

Before we continue, let us explain where some of the terms come from. We note that we have boundary conditions $\Delta_u(2, n) = n-2$ and $\Delta_u(d, n) = 0$, the latter for $n \leq d$. Then in the above telescoping inequalities, we have at most $\log_2(n/d)$ inequalities, as $h(n, \lfloor \log_2(n/d) \rfloor + 1) \leq d$, so $\Delta_u(d, h(n, \lfloor \log_2(n/d) \rfloor + 1)) = 0$.

Let

$$g(d, n) := (2d)^{\log_2 f} = f^{1+\log_2 d}.$$

Note that the last equality comes from $x^{\log y} = y^{\log x}$. This holds because $\log(y^{\log(x)}) = \log(x) \log(y) = \log(x^{\log(y)})$.

We claim that $\Delta_u(d, n) \leq g(d, n)$. We proceed by strong induction. First we have that $\Delta_u(2, f) = f-2 \leq f^2 = g(2, f)$, and $\Delta_u(n, f) = 0 \leq f \cdot f^{\log_2 d}$, for $f \leq d$. We induct on d ,

$$\begin{aligned}
\Delta_u(d, n) &\leq \sum_{i=0}^{\log_2(n/d)} 2^i (\Delta_u(d-1, h(n, i)) + 2) \\
&\leq \sum_{i=0}^{\log_2(n/d)} 2^i (g(d-1, h(n, i)) + 2) \\
&\leq \sum_{i=0}^{\log_2(n/d)} 2^i (g(d-1, n/2^i) + 2) \\
&= \sum_{i=0}^{\log_2(n/d)} 2^i (2(d-1))^{\log_2 n/2^i} + 2 \sum_{i=0}^{\log_2(n/d)} 2^i \\
&= f(d-1)^{\log_2 f} \sum_{i=0}^{\log_2(n/d)} \left(\frac{1}{d-1} \right)^i + 2(2 \cdot \frac{n}{d} - 1) \\
&< f(d-1)^{\log_2 f} \cdot \frac{d-1}{d-2} + \frac{4f-2d}{d} \\
&= f(d-1)^{\log_2 f} \cdot \frac{d-1}{d-2} + \frac{4f-2d}{d}.
\end{aligned}$$

We claim that this final expression is less than $g(d, n)$. To see this, note that

$$\frac{(d-1)^{1+\log_2 n}}{d^{1+\log_2 n} + 2d/n - 4} < \frac{d-2}{d}, \text{ for } d \geq 3, n > d.$$

This implies that

$$nd(d-1)^{1+\log_2 n} < (d-2) (nd^{1+\log_2 n} + 2d - 4n).$$

Finally, we have

$$n(d-1)^{\log_2 n} \cdot \frac{d-1}{d-2} + \frac{4n-2d}{d} < nd^{\log_2 n} = n^{1+\log_2 d} = g(d, n).$$

This completes the inductive step, and we thus have, $\Delta_u(d, n) < n^{\log_2(d)+1}$. \square

3.2 Monotone Paths

We consider now a slightly different question, one that more directly tied to linear programming.

Conjecture 3.5. (Monotone Hirsch conjecture) If ϕ is a linear form on a polyhedron P , then each vertex of P can be joined by an monotone increasing nonrevisiting path to an optimal vertex.

The conjecture is a more restrictive form of Hirsch's conjecture. The conjecture says that for a vertex x in P there is a nonrevisiting path to x_{\max} in P where x_{\max} maximizes cx over P . As we showed earlier, a nonrevisiting path is equivalent to the diameter being at most $n - d$. The monotone Hirsch conjecture was shown to be true for 3-dimensional polyhedra by Klee [19], but was eventually disproved by Todd [37]. Todd constructs a 4-dimensional polytope with 8 facets in which every monotone path is at least length 5. This idea of monotone paths is certainly interesting, as it seems closer to the realization of a deterministic pivot rule.

Take $H(d, n)$ and $H_u(d, n)$ to be smallest number of edges needed for a monotone increasing path from any vertex x to x_{\max} for bounded and unbounded polyhedra, respectively. Todd [37] demonstrates that $n - d + \min\{\lfloor d/4 \rfloor, \lfloor (n - d)/4 \rfloor\} \leq H(d, n) \leq H_u(d, n)$.

We even have a bound on H_u , thus giving us a pseudopolynomial monotone path to the optimal vertex.

Theorem 3.6. (Kalai [16], Ziegler [42]) $H_u(d, n) \leq 2f^{1+\log_2(d)} = 2(2d)^{\log_2 n}$.

Proof. The argument hinges on the idea of an *active facet*, which we define here. For a vertex v of a polyhedron P , and a linear function cx , a facet of P is *active* if it contains a point that is higher than v ; that is, either the facet is unbounded with respect to cx or it has a top vertex w such that $cv < cw$. To be clear about which vertex and polyhedron we are considering a facet to be *active*, we write F is an *active facet* of $P(v)$.

Define $\bar{H}(d, n)$ to be the length of a path required from a vertex v that has at most f active facets and an arbitrary number of nonactive facets. Then $H(d, n) \leq \bar{H}(d, n)$.

We can easily derive exact values for \bar{H} on boundaries. We have

$$\bar{H}(2, f) = f,$$

as all edges in the path to the optimal vertex are active, and we may have to traverse all of them. Effectively, we are adding vertices to unbounded edges here, starting along one unbounded edge, and finishing along the other. We continue to consider vertices on the unbounded edges throughout the rest of this proof. Next, we can see that

$$\bar{H}(d, 0) = \dots = \bar{H}(d, d - 2) = 0,$$

as if a vertex is not optimal, there is an edge that is increasing; this edge lies on $d - 1$ facets, all of which are necessarily active.

Now we want to derive a recursion for $\bar{H}(d, n)$. We do this using four key claims.

Given any set \mathcal{F} of k active facets of $P(v)$, there is a monotone path from v to the optimal vertex of P or some vertex in some facet of \mathcal{F} that has length at most $\bar{H}(d, n - k)$.

If v lies on a facet of in \mathcal{F} , we are immediately finished. Consider then when v does not lie on a facet in \mathcal{F} . Let $P = P(A, b)$ be a minimal system in that we do not have any redundant constraints; then each constraint defines a facet of P . Now take P' to be the polyhedron defined by deleting constraints that correspond to facets in \mathcal{F} . Then v is a vertex of P' and then has at most $n - k$ active facets in $P'(v)$. Now consider the shortest monotone path in P' from v to the optimal vertex in P' . Such a path has length at most $\bar{H}(d, n - k)$. Take this path back in P , to get that a monotone path of length at most $\bar{H}(d, n - k)$ from v must either reach or not reach a facet in \mathcal{F} . If it does reach, which happens when the path in P' enters an edge that was not in P , our claim is complete. If it does not reach, then the optimal vertex in P' is in fact the optimal vertex in P ; this path in P' is the same as the path in P , as else we hit a facet in \mathcal{F} , hence our claim is complete.

The collection \mathcal{G} of all active facets for which there is a path from v with length at most $\bar{H}(d, n - k)$ has at least $n - k + 1$ active facets.

Note that there are $\binom{n}{k}$ total sets of k active facets. Since there is at least one facet, say F_1 within $\bar{H}(d, n - k)$, we can add F_1 to \mathcal{G} and note there are now $\binom{n-1}{k}$ sets of k active facets that do not include F_1 . Continuing in this manner until we have at least $\{F_1, \dots, F_k\} \subset \mathcal{G}$. There is $\binom{k}{k} = 1$ subset which does not contain a facet in $\{F_1, \dots, F_k\}$. Again, this subset contains an active facet within $\bar{H}(d, n - k)$, so we have $\{F_1, \dots, F_{k+1}\}$, as desired.

There exists a monotone path from v to the optimal vertex w_0 in \mathcal{G} with length at most $\bar{H}(d, n - k) + \bar{H}(d - 1, n - 1)$.

This follows immediately from above. It takes at most $\bar{H}(d, n - k)$ edges to reach the active facet w_0 is on; then, because the facet is $d - 1$ -dimensional and has at most $n - 1$ facets, it takes at most $\bar{H}(d - 1, n - 1)$ edges to get to w_0 .

There is a monotone path from w_0 to the optimal vertex with length at most $\bar{H}(d, k - 1)$.

Since we are at the optimal vertex in \mathcal{G} , none of the facets in \mathcal{G} is active for $P(w_0)$. Then, w_0 has at most $n - (n - k + 1) = k - 1$ active facets. Hence, the monotone path from w_0 to the optimal vertex has length at most $\bar{H}(d, k - 1)$, as desired.

Collecting the inequalities, we have

$$\bar{H}(d, n) \leq \bar{H}(d, n - k) + \bar{H}(d - 1, n - 1) + \bar{H}(d, k - 1)$$

for any k .

To reduce this inequality, we make use of the fact that \bar{H} is increasing in f and take $k = \lceil n/2 \rceil$. Hence,

$$\bar{H}(d, n) \leq \bar{H}(d - 1, n - 1) + 2\bar{H}(d, \lfloor n/2 \rfloor).$$

Take

$$g(d, t) := 2^{-t} \bar{H}(d, 2^t), \text{ for } t \geq 0 \text{ and } d \geq 2.$$

Then we have

$$\begin{aligned}
g(d, t) &= 2^{-t} \bar{H}(d, 2^t) \\
&\leq 2^{-t} (H(d-1, 2^t-1) + 2\bar{H}(d, \lfloor 2^{t-1} \rfloor)) \\
&= 2^{-t} H(d-1, 2^t-1) + 2^{-(t-1)} \bar{H}(d, 2^{t-1}) \\
&\leq g(d-1, t) + g(d, t-1).
\end{aligned}$$

Transforming the boundary conditions, we have

$$\begin{aligned}
g(2, t) &= 1 \leq \binom{t}{1} \text{ for } t \geq 1 \\
g(d, 0) &= 0 = \binom{d-2}{d-1} \text{ for } d \geq 3,
\end{aligned}$$

from which we get, by inducting on $t \geq 0$ and $d \geq 2$,

$$g(d, t) \leq \binom{d+t-2}{d-1}.$$

Finally, we get

$$\begin{aligned}
\bar{H}(d, n) &\leq \bar{H}(d, 2^{1+\lfloor \log_2 n \rfloor}) \\
&= 2^{1+\lfloor \log_2 n \rfloor} g(d, 1 + \lfloor \log_2 n \rfloor) \\
&\leq 2n \binom{d + \lfloor \log_2 n \rfloor - 1}{d-1} \\
&\leq 2nd^{\log_2 n} \\
&= 2n^{1+\log_2 d}
\end{aligned}$$

for $n, d \geq 2$. The last inequality follows by $\binom{a+b}{b} \leq (a+1)^b$, by inducting over $a, b \geq 0$.

Thus, we have $H_u(d, n) \leq \bar{H}(d, n) \leq 2(2d)^{\log_2 n}$. □

4 Special Polytopes

Nous sommes tous des cas exceptionnels. Nous voulons tous faire appel de quelque chose!

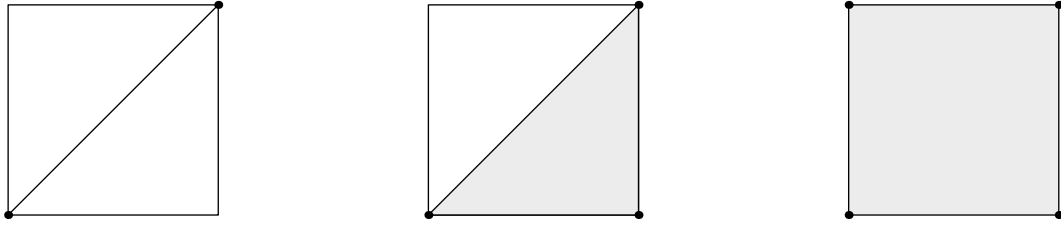
— Albert Camus, La Chute

Now we restrict our consideration to special classes of polytopes. We look at the maximal diameter of these polytopes, and in doing so, whether they satisfy the Hirsch conjecture, and in some cases, discuss how the simplex algorithm performs in these special circumstances.

4.1 (0,1)-Polytopes

Let P be a (0,1)-polytope, such that P is the convex hull of $V \subseteq \{0, 1\}^d$. Naddef [28] proved the Hirsch conjecture for (0,1)-polytopes. We have simplified the proof here a little by requiring less machinery (in fact, no heavy machinery at all) in the following lemma.

Lemma 4.1. $\delta(P) \leq \dim P$. We have equality when P is isomorphic to the cube.

Figure 3: The $(0,1)$ -polytopes in dimension 2, up to isomorphism

Proof. We proceed by induction on d . This is certainly true for $d = 2$. Now consider two vertices u and v of P such that $\delta(u, v) \geq d$. By symmetry of the cube, we can assume that $u = 0$ and $v \in \{0, 1\}^d$. Further, we can assume that $\dim P = d$, as otherwise we can project P into a lower dimension space, to be easier, into $\{0, 1\}^{d-1}$, and by induction we have our claim. So $\dim P = d$.

Suppose now that $v_i = 0$ for some i . Then $u = 0$ and v share a facet, specifically, the facet defined by $x_i \geq 0$. Thus, we get $\delta(u, v) = \delta(0, v) \leq d - 1$, by the inductive hypothesis. Hence, we now consider the case when $v_i \neq 0$, for any i . That is, $v = 1$. Now, by induction we can show that any vertex adjacent to $v = 1$, say w , that has k 0-components have $\delta(0, w) \leq k$. From this we get, $(u, v) = (0, 1) \leq (0, w) + (w, 1) \leq d - k + 1$.

So if w has $k > 1$ 0-components, we have $\delta(u, v) < d$. Hence, if $\delta(0, 1) \geq d$, then all vertices adjacent to 1 have only one 0-component. But 1 has at least d adjacent vertices; so set of adjacent vertex is in fact those with exactly one 0-component. By induction on d , these vertices are distance $d - 1$ from 0 in the facet they share with 0. This facet though is the convex hull of points such that the respective component is 0. That is, the facet is a $d - 1$ cube. Assembling these $d - 1$ -cubes, we can see that the convex hull of the points in P form the d -cube, and that $\delta(0, 1) = d$, as desired. \square

Theorem 4.2. *The Hirsch conjecture holds for $(0,1)$ -polytopes.*

Proof. Again, we induct on the dimension of P . We can see the hypothesis holds for $\dim P = 2$. Assume the hypothesis holds for $\dim P = d$. Now take P such that $\dim P = d + 1$. Take u and v such that their distance is the diameter of P . Then there are at least $d + 1$ facets incident to u and $d + 1$ facets incident to v . Suppose there is a facet incident to both. Then by the inductive hypothesis, we are finished. Then consider if there is not at least one facet incident to both u and v . Then there are at least $2(d + 1)$ facets and P is $d + 1$ dimensional, so $n - d = d + 1 \geq \delta(u, v)$, by the above lemma. \square

Despite $(0, 1)$ -polytopes satisfaction of the Hirsch conjecture, a polynomial deterministic simplex algorithm has not yet been found.

Kleinschmidt and Onn [22] generalize a bound for the diameter of $\{0, \dots, k\}^d$ polytopes.

Theorem 4.3. *If $V \subseteq \{0, \dots, k\}^d$, then the diameter of the convex hull of V is at most k times the dimension of V .*

4.2 Network Flow Polytopes

We specifically consider the minimum cost flow problems of transportation and assignment. The transportation polytope is defined as

$$T_{m,n}(a, b) := \{x \in \mathbb{R}^{m \times n} : \sum_i^m x_{i,j} = b_j, \sum_j^n x_{i,j} = a_i, x_{i,j} \geq 0\}.$$

Brightwell et al. [5] show that the diameter of any $m \times n$ transportation polytope is at most $8(m + n - 2)$. Kim and Santos [18] cites the following bound received by correspondence. The author of this paper has been unable to find a paper or pre-print to verify this.

Theorem 4.4. (Hurkens [12]) The diameter of any $m \times n$ transportation polytope is at most $3(m + n - 1)$

Balinski [1] considers the dual of the transportation problem and demonstrates the following:

Theorem 4.5. (Balinski [1]) *The diameter of the dual $m \times n$ transportation polytope is at most $(m - 1)(n - 1)$. This bound is optimal.*

As the dual transportation polytope is in class $(m + n - 1, mn)$, this proves the Hirsch conjecture for dual transportation polytopes. In [1], Balinski establishes a combinatorial characterization of the polytopes vertices; he calls this characterization a *signature*. We mention this because through this characterization, he makes progress with the simplex algorithm on the assignment problem in Balinski [2].

Recall the assignment problem: We are given a set of n people and a set of n tasks and a cost $c_{i,j}$ for the i th person to conduct the j th task. Let $f_{i,j}$ be an indicator function, taking 1 if the i th person does task j , and 0 otherwise. Then we wish to minimize $\sum_i^n \sum_j^n c_{i,j} f_{i,j}$, subject to $\sum_i^n f_{i,j} = 1$ and $\sum_j^n f_{i,j} = 1$, that is, subject to a person taking on one and only one task. The $n \times n$ polytope, P_n , is derived from the assignment problem. $P_n := T_{n,n}(1, 1)$. As we mentioned earlier, Balinski [2] derived a result for the simplex algorithm on the assignment problem.

Theorem 4.6. (Balinski [2]) *There is a simplex implementation on the assignment problem whose worst-case behavior terminates in $\binom{n-1}{2}$ pivots.*

This is interesting as this shows that the assignment problem is one of the few times in which the simplex algorithm is competitive with non-simplex procedures in worst case behavior. We now go back to considering the geometry of the assignment polytope.

Balinski and Russakoff [3] first demonstrated that the assignment polytope satisfies the Hirsch conjecture by demonstrating that it has diameter 2. We give a simpler proof of this fact using permutations.

Theorem 4.7. *The diameter of P_n is 2.*

Proof. First note that there are $n!$ ways to arrange the tasks amongst the people, so P_n has $n!$ vertices. This inspires us to consider a correspondence between the vertices and the set of permutations on n distinct letters. Recall from introductory algebra that any permutation can be decomposed into a product of disjoint cycles. Say we have a permutation

$$(j_1, \dots, j_n) = (j_1, \dots, j_{n_1})(j_{n_1+1}, \dots, j_{n_2}) \cdots (j_{n_k} + 1, \dots, j_n),$$

then note that we can express the permutation as the product of two cycles

$$(j_1, \dots, j_n) = (j_1, \dots, j_n)(j_n, j_{n_k}, \dots, j_{n_1}).$$

Let $Q'(n)$ be the set of permutations on n letters that can be expressed in cycle notation by a single cycle. Let us further take p to map the natural correspondence between vertices of P to

permutations on m letters. That is, p maps permutation matrices to their natural permutation. It can be shown that vertices x and y of P_n are neighbors if and only if there is a $z \in Q'(m)$ such that $p(x) = p(y)p(z)$, [3] and [41]. Hence the diameter of P_n is greater than 1, for $n \geq 4$.

Consider any two vertices, x and y , in P . We have from above that there exists $w, v \in Q(n)$ such that $p(x)^{-1}p(y) = p(u)p(v)$. Supposing u is not the identity, y then is a neighbor of the vertex corresponding to $p(x)p(u)$, say w . We then immediately have that w is a neighbor of x , as $p(w) = p(x)p(u)$. So x and y are distance at most 2 apart. Hence the diameter of P_n is 2, for $n \geq 4$. \square

4.3 Leontief substitution systems

A Leontief matrix, $A \in \mathbb{R}^{m \times n}$, has rank m and exactly one positive element in each column. A pre-Leontief matrix, $M \in \mathbb{R}^{m \times n}$, has rank m and has at most one positive element in each column. $P = \{x : Ax = b, x \geq 0\}$, where b is a constant positive m -vector, is a Leontief substitution system, and $Q = \{x : Mx = b, x \geq 0\}$, where b is a constant non-negative m -vector, is a pre-Leontief substitution system.

Not only is the Hirsch conjecture true for pre-Leontief systems (and hence Leontief systems), but the monotone Hirsch conjecture is true for pre-Leontief systems. In fact, the bound we can achieve is slightly lower than that of the monotone Hirsch conjecture.

Theorem 4.8. (Grinold [10]) *Take x and y to be vertices of the pre-Leontief substitution system, Q . If both x_i and y_i are positive for k columns in B , then the distance between x and y is at most $m - k$. Further, if y is the optimal vertex, a path can be chosen from x to y such that the objective function is nonincreasing.*

Note that the distance between any two vertices is at most $m - k$. As $k \geq 1$, we immediately have Hirsch's conjecture satisfied. Also note that the shortest route polytope is a pre-Leontief system, so while it was previously shown by Saigal [31] that the shortest route problem satisfies the Hirsch conjecture, we now get it for free.

Proof. First we show the claim holds for Leontief systems. Let \mathcal{B}' be the set of all $m \times m$ Leontief submatrices of A . Consider $B \in \mathcal{B}$. Take $B-$ to be the submatrix of A with the columns of B deleted. Denote x_B to be the m -vector made of the components of x in which the columns of A appear in B . By Veinott [38], we say that x , for which $Ax = b$, is *determined* by a matrix B if B is a basis of A and $(x_B, x_{B-}) = (B^{-1}b, 0)$. This is a little confusing, so we give an example. Take

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ and } B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

We thus get

$$x_B = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ and the determined } x = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

By Veinott [38], $x \in P$ is a vertex if and only if x is determined by a $B \in \mathcal{B}$.

We assume that y is the optimal vertex for the linear program $\min\{cx : Ax = b, x \geq 0\}$. As y is a vertex, we have that there is some $D \in \mathcal{B}$ that determines y , and since D is nonnegative,

so is its inverse, which implies D maintains its optimality as a basis no matter what values b take, so long as b is nonnegative.

Now take $B_0 \in \mathcal{B}$ such that B_0 determines x . B_0 shares at least $l \geq k$ columns with D . We now try to construct an $m - l + 1$ sequence of submatrices of A , that is, bases, B_0, \dots, B_{m-l} . We hope to construct the sequence in such a way that $D = B_{m-l}$, adjacent bases in the sequence share $m - 1$ columns in common, and the objective function is nonincreasing along the sequence of vertices determined by the bases. If we can do this, we have our result.

We have three cases: first, the nondegenerate case in which b is strictly positive and y is the unique optimal vertex, next the degenerate case in which b has some 0 elements, and finally, the degenerate case in which there are multiple optimal vertices.

First we take care of the nondegenerate case, so assume that b is strictly positive and y is the unique optimal vertex. That b is strictly positive implies there is a bijection between vertices and bases and that each basis is in \mathcal{B} . Consider a basis $B_j \in \mathcal{B}$. We then have that the dual variables corresponding with this basis are $c_{B_j}(B_j)^{-1}$, and the value of the objective function with this basis is $c_{B_j}(B_j)^{-1}b$. Now say there is a column of D , d_i , with cost c_i , that is eligible to enter the current basis, B_j . That is, $c_i < c_{B_j}(B_j)^{-1}b$. If we swap d_i with a column in B_j , we must ensure the new basis corresponds to a vertex. The only way to do that is, as we mentioned earlier, force the basis to be in \mathcal{B} . The only way to do that is to force the swap to be with the column of B_j that has its positive element in the same row as d_i 's positive element. We conduct the swap and obtain the new basis B_{j+1} . Note that B_{j+1} shares one more column in common with D and its dual variables cannot be greater than those of B_j , hence the objective function can only decrease from the change in basis. If now column of D was eligible to enter the basis, the current basis is optimal, as we have assumed D is the optimal solution. Hence, we have built the desired sequence of bases, and the claim holds for the nondegenerate case.

Consider the degenerate case in which b is not strictly positive. Consider the same problem but replace b with $b' > 0$. The inverse of every $B \in \mathcal{B}$ is nonnegative, so the sequence as we constructed above is feasible for any nonnegative b . However, we lose that the objective function is strictly decreasing along the path; instead we satisfy ourselves that it is nonincreasing along the path.

Finally, consider the degenerate case in which y is not the optimal solution. This is the easiest to deal with, as we use the original argument of the nondegenerate case. In the event we enter a basis that is also optimal before we enter the basis that determines y , we terminate the sequence, hence obtaining a sequence that is strictly less than that proposed. Thus, our sequence of bases has still at most $m - l + 1$ bases.

We have exhausted the result for Leontief substitution systems. Let us move to pre-Leontief substitution systems. We can reduce the pre-Leontief case to the Leontief case by making heavy use of results found in Veinott [38].

Suppose we have a pre-Leontief system, then we can partition M and b such that

$$M = \begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix}, b = \begin{pmatrix} b^1 \\ b^2 \end{pmatrix}$$

where A_1 is Leontief and A_2 is sub-Leontief (the definition of which is not important for this discussion, but if interested, do see [38]), and where b^i is permuted to correspond with the form of presented for A . Partitioning x to $x = (x^1, x^2)$, we see that $Q^1 = \{x^1 : A_1 x^1 = b^1, x^1 \geq 0\}$ is a Leontief substitution system. Our original system, $Q = \{x : Mx = b, x \geq 0\}$ is non empty if and only if $b^2 = 0$. Further, x is a vertex of Q if and only if x^1 is a vertex of Q^1 and $x^2 = 0$. Hence, for any vertices in Q , we map to vertices in the corresponding pre-Leontief

system, which we have already shown holds for our claim. Hence we have shown the result for pre-Leontief substitution systems, so our claim is complete. \square

A *Leontief flow problem* is the linear programming problem

$$\min\{cx : Ax = b, x \geq 0\}$$

where A is an integral Leontief matrix with all positive entries equal to 1 and the cost vector c is integral as well.

Ng and Wagner [29] give a result for the simplex algorithm's performance on Leontief flow problems.

Theorem 4.9. ([29]) *The simplex algorithm, using Dantzig's pivot rule, solves the Leontief flow problem in $\mathcal{O}(n^2 U \log(npU))$ pivots, where p is the entry of largest magnitude in the Leontief matrix $A \in \mathbb{R}^{m \times n}$, and U is an upper bound on any variable in x in any solution on a vertex.*

In general this bound is not polynomial, as U , for example, U can easily be exponential. However, there are some cases, such as the shortest-path problem, in which $U = m$ and $p = 1$. For the shortest-path problem, however, Orlin [30] gives a better bound of $\mathcal{O}(m^2 n \log m)$ pivots.

5 Concluding thoughts

Not with a bang but a whimper.
— T.S. Eliot, The Hollow Men

We finish off with a consideration of what “average case” behavior may mean for the simplex algorithm and discuss some recent results and developments.

5.1 Random Polytopes

Let's break from special polytopes and think about why the simplex algorithm may work well generally. A natural idea to explore this is to set our constraints to be random variables under nice or natural distributions. This is exactly what we do.

For the linear program we originally introduced, (P): minimize $\{c^T x : Ax = b, x \geq 0\}$, we now take entries of A , b , and c to be entries of random variables. We hope that we can bound the number of expected pivots for a given distribution and implementation of the simplex.

Here, we state results of the average behavior of the simplex algorithm over random polytopes and discuss some of the ideas used in the analysis.

We first consider the work done by Smale [36]. In the paper, Smale proposes a spherically symmetric probabilistic model for analysis of the seln-dual simplex. This model is equivalent to taking the coefficient random variables to be standard normal distributed Megiddo [25]. Let $\rho(m, n)$ be the expected number of pivots required of the seln-dual simplex when A is an $m \times n$ matrix. Smale demonstrates that $\rho(m, n) \leq c(m) \log n^{m(m+1)}$, where $c(m)$ is a constant dependent on m .

What is key is that Smale shows that we can express ρ as the sum of volumes of four different types of matrices, each of which correspond to different types of bases. The volume of a matrix is its determinant, but a notion the author had not encountered that is made use of is the

Gaussian volume of a matrix, $GV(A)$ for $A \in \mathbb{R}^{m \times n}$. $GV(M)$ is the probability that a random vector distributed standard n -normal lines in the convex cone spanned by M 's columns. This is relevant because Smale expresses ρ in terms of probabilities of random rays being in specific random cones.

Other work in this area was done independently by Borgwardt [4]. Borgwardt worked under a different probabilistic model than did Smale, and under this model Borgwardt tests a special variant of the simplex of his design. Under this model, Borgwardt proves a bound for expected pivots $\phi(m, n) \leq cmn^2(n+1)^2$, where c is a constant equal to $(e\pi/4)(\pi/2 + 1/e)$.

5.2 Recent Work

The recent IPAM conference “Efficiency of the Simplex Method: Quo vadis Hirsch conjecture?” and Gil Kali’s proposal of the polynomial Hirsch conjecture as the Polymath3 project have resulted in a slew of results and ideas in the past few months. Here we cover a few.

5.2.1 Subexponential lower bounds for random pivot rules

At the IPAM conference, subexponential lower bounds of the form 2^{n^α} were proved for two basic randomized pivot rules for the simplex Freedman et al. [9]. The two pivot rules they consider are RANDOM-EDGE and RANDOM-FACE. The former is the most natural random pivot rule, while the latter is theoretically the fastest pivot rule currently known. Before this result, no non-polynomial lower bounds existed for the rules’ performance, and finding such bounds was an open problem for decades.

Freedman et al. [9] construct concrete linear programs for which they obtain expected subexponential number of pivots. For RANDOM-EDGE, they construct *Parity games*, a family of deterministic 2-player games, for which RANDOM-EDGE takes an expected subexponential number of pivots. They transform the parity games into a Markov Decision Processes, a family of stochastic 1-player games. Finally, they transform the Markov Decision Process into linear programs. For RANDOM-FACET, the process is the same, but the construction differs. The key idea lies in the construction themselves, and thinking about the problem in terms of parity games. As the authors note, had they considered Markov Decision Process constructions directly, they likely would not have succeeded.

5.2.2 Polymath3

The polynomial Hirsch conjecture, for a polytope in class (d, n) , Δ is bounded above by a polynomial in d and f , was chosen as the Polymath3 project last September. The approach to tackle a combinatorial formulation drew great interest. We present the combinatorial abstraction given here:

Conjecture 5.1. (Combinatorial polynomial Hirsch conjecture) Take t non-empty disjoint families, F_1, \dots, F_t , of subsets of the first n naturals, $\{1, \dots, n\}$. We place the further restriction on the families, that for all $i < j < k$, for any $S \subset F_i$ and any $T \subset F_k$ we have some $R \subset F_j$ such that $S \cap T \subset R$. Take $f(n)$ to be the largest value of t for the input n . The conjecture is that $f(n)$ is polynomial in n .

The combinatorial polynomial Hirsch conjecture implies the polynomial Hirsch conjecture.

Theorem 5.2. *A simple polytope with n facets has diameter bounded by $f(n)$.*

The result is given by Eisenbrand et al. [8], but the proof we give is by Kalai. We notice the proof is similar in flavor to the proof of Kalai and Kleitman’s bound.

Proof. Let P be a simple polytope with n facets. Let x be a vertex of P . Define S_x to be the set of facets that contain x . Choose a vertex w such that there exists a vertex v whose distance from w is the diameter of P . Now define F_i to be the family of sets that correspond to vertices that are distance $i + 1$ from w . That is, F_i contains S_x ’s, not vertices. Then the number of families is exactly the diameter of P . Hence, why we seek to maximize the number of families in the formulation.

Now we show why we have the further restriction. Take vertices v and u such that v is distance i from w and u is distance $k > i$ from w . Now take the shortest path from v to u in the smallest face containing both u and v . For every vertex z in this path we have that $S_u \cap S_v \subset S_z$. Now note in the path, the distance of adjacent vertices to w can change at most by 1, so there is some vertex in the path that is distance j , for $i < j < k$, away from w . Hence, we have showed the claim. \square

Other abstractions and continued discussions for the interested reader can be found on Gil Kalai’s blog¹ [27]. Also, a talk was given at IPAM on abstractions of the polynomial Hirsch conjecture. See Hahnle [11] for the presentation.

5.3 Further Directions

When Gil Kalai posted on his blog that Francisco Santos had disproved the Hirsch conjecture, there were a great many enthusiastic responses. However, amongst the happy congratulators there were a few dismayed doctoral students,

“That’s my whole PhD work going to trash!”
— Tarantino from Stanford

“I was working on the same problem. I guess I need to find a new research topic.”
— A PhD student from CMU

drawing responses from Santos and Kalai,

“To Tarantino: I hope that is not true. As Anand says, I think the overall question ‘how much can the diameter of a polytope grow’ is as interesting now as it was before.”
— Francisco Santos

“To the Stanford and CMU PhD students, [sic]
While this is certainly a major development, there are quite a few remaining interesting questions in this research area, and probably reading Santos’s paper will raise additional problems and ideas.”
— Gil Kalai

¹As of the middle of April, Kalai has been trying a topological abstraction, inspired by Santos’s work in the counterexample to the Hirsch conjecture. Since our review of the subject is concerned with connections to the Simplex algorithm, and is entirely combinatorial, we do not include this very recent line of pursuit in this overview. While interest in the topologic formulation is not as great as that for the combinatorial formulation—the latter had active participation and discussion from renowned combinatorialists—that it uses insights from Santos’s counterexample and is very elegant demonstrates promise.

The main reason of highlighting the exchange is to elicit amusement from the reader, but it also addresses where to go next. Despite the Hirsch conjecture being disproven, it is still possible that the upper bound on a polyhedron's diameter is linear. Fruitful results are more likely to be found in exploring the polynomial Hirsch conjecture, for which a combinatorial reformulation has already produced many ideas. Reformulation is a viable attack on bounds for expected iterations of simplex algorithm, as well. Simply, the bounds we have for polyhedron diameter and expected iterations of the simplex are gigantic. Any refinement or reformulation of either problem, or insight in how these functions grow is still relevant and of theoretical and practical importance.

References

- [1] M. L. Balinski. The hirsch conjecture for dual transportation polyhedra. *Mathematics of Operations Research*, 9(4):pp. 629–633, 1984. ISSN 0364765X. URL <http://www.jstor.org/stable/3689468>.
- [2] M. L. Balinski. Signature methods for the assignment problem. *Operations Research*, 33(3):pp. 527–536, 1985. ISSN 0030364X. URL <http://www.jstor.org/stable/170555>.
- [3] M. L. Balinski and Andrew Russakoff. On the assignment polytope. *SIAM Review*, 16(4):pp. 516–525, 1974. ISSN 00361445. URL <http://www.jstor.org/stable/2028692>.
- [4] K. H. Borgwardt. The average number of pivot steps required by the simplex-method is polynomial. *Mathematical Methods of Operations Research*, 26:157–177, 1982. ISSN 1432-2994. URL <http://dx.doi.org/10.1007/BF01917108>. 10.1007/BF01917108.
- [5] Graham Brightwell, Jan van den Heuvel, and Leen Stougie. A linear bound on the diameter of the transportation polytope*. *Combinatorica*, 26:133–139, 2006. ISSN 0209-9683. URL <http://dx.doi.org/10.1007/s00493-006-0010-5>. 10.1007/s00493-006-0010-5.
- [6] George Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, NJ, 1963.
- [7] René Descartes. *Mediations on first philosophy*. Cambridge University Press, 1996.
- [8] Friedrich Eisenbrand, Nicolai Hahnle, Alexander Razborov, and Thomas Rothvob. Diameter of polyhedra: Limits of abstraction. Preprint, 2010.
- [9] Oliver Freedman, Thomas Dueholm Hansen, and Uri Zwick. Subexponential lower bounds for randomized pivoting rules for solving linear programs. Preprint, 2011.
- [10] Richard C. Grinold. The hirsch conjecture in leontief substitution systems. *SIAM Journal on Applied Mathematics*, 21(3):pp. 483–485, 1971. ISSN 00361399. URL <http://www.jstor.org/stable/2099595>.
- [11] Nicolai Hahnle. An abstract view on the polynomial hirsch conjecture. Preprint, 2010.
- [12] C. Hurkens. noted in Kim, E., Santos F (2010) and De Loera, J, Kim, E., et al (2007) as personal correspondence, 2007.
- [13] D.B. Judin and A. S. Nemirovskii. Evaluation of the informational complexity of mathematical programming problem. *Matekon*, 13:3–24, 1976.
- [14] D.B. Judin and A. S. Nemirovskii. Informational complexity and efficient methods for the solution of convex extremal problems. *Matekon*, 13:25–45, 1976.

- [15] G. Kalai and D. J. Kleitman. A quasi-polynomial bound for the diameter of graphs of polyhedra. *ArXiv Mathematics e-prints*, March 1992.
- [16] Gil Kalai. A subexponential randomized simplex algorithm (extended abstract). In *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*, STOC '92, pages 475–482, New York, NY, USA, 1992. ACM. ISBN 0-89791-511-9. doi: <http://doi.acm.org/10.1145/129712.129759>. URL <http://doi.acm.org/10.1145/129712.129759>.
- [17] L. G. Khachian. A polynomial algorithm in linear programming. *Dokl. Akad. Nauk SSSR*, 244(1):pp. 1093–1096, 1979. English translation in *Soviet Math. Dokl.* 20, 191-194, 1979.
- [18] E. D. Kim and F. Santos. Companion to "An update on the Hirsch conjecture". *ArXiv e-prints*, December 2009.
- [19] Victor Klee. Paths on polyhedra. i. *Journal of the Society for Industrial and Applied Mathematics*, 13(4):pp. 946–956, 1965. ISSN 03684245. URL <http://www.jstor.org/stable/2946414>.
- [20] Victor Klee and George J. Minty. How good is the simplex algorithm? In Oved Shisha, editor, *Inequalities - III*, 1969.
- [21] Victor Klee and David Walkup. The d-step conjecture for polyhedra of dimension ≤ 6 . *Acta Mathematica*, 117:53–78, 1967. ISSN 0001-5962. URL <http://dx.doi.org/10.1007/BF02395040>. 10.1007/BF02395040.
- [22] Peter Kleinschmidt and Shmuel Onn. On the diameter of convex polytopes. *Discrete Math.*, 102:75–77, May 1992. ISSN 0012-365X. doi: [http://dx.doi.org/10.1016/0012-365X\(92\)90349-K](http://dx.doi.org/10.1016/0012-365X(92)90349-K). URL [http://dx.doi.org/10.1016/0012-365X\(92\)90349-K](http://dx.doi.org/10.1016/0012-365X(92)90349-K).
- [23] D. G. Larman. Paths on polytopes. *Proceedings of the London Mathematical Society*, s3-20(1):161–178, 1970. doi: 10.1112/plms/s3-20.1.161. URL <http://plms.oxfordjournals.org/content/s3-20/1/161.short>.
- [24] J. Matoušek, M. Sharir, and E. Welzl. A subexponential bound for linear programming. *Algorithmica*, 16:498–516, 1996. ISSN 0178-4617. URL <http://dx.doi.org/10.1007/BF01940877>. 10.1007/BF01940877.
- [25] Nimrod Megiddo. Improved asymptotic analysis of the average number of steps performed by the self-dual simplex algorithm. *Math. Program.*, 35:140–172, June 1986. ISSN 0025-5610. doi: 10.1007/BF01580645. URL <http://portal.acm.org/citation.cfm?id=9514.9517>.
- [26] Nimrod Megiddo. On the complexity of linear programming. In *Advances in Economic Theory*. Cambridge University Press, 1989.
- [27] Gil Kalai (moderator). Multiple/many contributors. Polymath3. Preprint, 2010.
- [28] Denis Naddef. The hirsch conjecture is true for $(0, 1)$ -polytopes. *Mathematical Programming*, 45:109–110, 1989. ISSN 0025-5610. URL <http://dx.doi.org/10.1007/BF01589099>. 10.1007/BF01589099.
- [29] P. H. Ng and D. K. Wagner. A simplex algorithm for a class of leontief flow problems. *Mathematical and Computer Modelling*, 24(7):69 – 75, 1996. ISSN 0895-7177. doi: DOI:10.1016/0895-7177(96)00128-8. URL <http://www.sciencedirect.com/science/article/B6V0V-426HSFG-7/2/78c4dfa8f9bb7934973837c18800d985>.

- [30] James B. Orlin. On the simplex algorithm for networks and generalized networks. *Mathematical Programming Studies*, 24(1):pp. 166–178, 1985.
- [31] Romesh Saigal. A proof of the hirsch conjecture on the polyhedron of the shortest route problem. *SIAM Journal on Applied Mathematics*, 17(6):pp. 1232–1238, 1969. ISSN 00361399. URL <http://www.jstor.org/stable/2099203>.
- [32] F. Santos. A counterexample to the Hirsch conjecture. *ArXiv e-prints*, June 2010.
- [33] Francisco Santos. A counter-example to the hirsch conjecture. poster, 2010. URL http://dcgprogram.epfl.ch/webdav/site/dcgprogram/users/184244/public/slides_francisco_santos.pdf.
- [34] William Shakespeare. *Romeo and Juliet*. Oxford University Press, 2000.
- [35] N. Z. Shor. Cut-off method with space extension in convex programming problems. *Cybernetics and Systems Analysis*, 13:94–96, 1977. ISSN 1060-0396. URL <http://dx.doi.org/10.1007/BF01071394>. 10.1007/BF01071394.
- [36] Steve Smale. On the average number of steps of the simplex method of linear programming. *Mathematical Programming*, 27:241–262, 1983. ISSN 0025-5610. URL <http://dx.doi.org/10.1007/BF02591902>. 10.1007/BF02591902.
- [37] Michael J. Todd. The monotonic bounded hirsch conjecture is false for dimension at least 4. *Mathematics of Operations Research*, 5(4):pp. 599–601, 1980. ISSN 0364765X. URL <http://www.jstor.org/stable/3689331>.
- [38] Arthur F. Jr. Veinott. Extreme points of leontief systems. *Linear Algebra and Its Applications*, 1(1):pp. 181–194, 1968.
- [39] Christophe Weibel. Low dimensions counter-examples to the hirsch conjecture. poster, 2011. URL <http://www.cs.dartmouth.edu/~weibel/posterHirsch.pdf>.
- [40] Philip Wolfe and Leola Culter. Experiments in linear programming. In *Recent Advances in Mathematical Programming*. Princeton University Press, 1963.
- [41] H. P. Young. On permutations and permutation polytopes. In R. W. Cottle, L. C. W. Dixon, B. Korte, T. L. Magnanti, M. J. Todd, E. L. Allgower, R. Bartels, V. Chvatal, J. E. Dennis, B. C. Eaves, R. Fletcher, J.-B. Hiriart-Urruty, M. Iri, R. G. Jeroslow, D. S. Johnson, C. Lemarechal, L. Lovasz, L. McLinden, M. W. Padberg, M. J. D. Powell, W. R. Pulleyblank, K. Ritter, R. W. H. Sargent, D. F. Shanno, L. E. Trotter, H. Tuy, R. J. B. Wets, C. Witzgall, E. M. L. Beale, G. B. Dantzig, L. V. Kantorovich, T. C. Koopmans, A. W. Tucker, P. Wolfe, M. L. Balinski, and A. J. Hoffman, editors, *Polyhedral Combinatorics*, volume 8 of *Mathematical Programming Studies*, pages 128–140. Springer Berlin Heidelberg, 1978. ISBN 978-3-642-00790-3. URL <http://dx.doi.org/10.1007/BFb0121198>. 10.1007/BFb0121198.
- [42] Gunter M. Ziegler. *Lectures on Polytopes*. Springer-Verlag, New York, NY, 1994. ISBN 0387943293.