Jahresber Dtsch Math-Ver (2010) 112: 71–72
DOI 10.1365/s13291-010-0003-6

Vorwort Heft 2-10

Hans-Christoph Grunau

Online publiziert: 11. Juni 2010
© Vieweg+Teubner und Deutsche Mathematiker-Vereinigung 2010

In ihrem Übersichtsartikel „An update on the Hirsch conjecture“ lassen Edward D. Kim und Francisco Santos die aktuelle Entwicklung zu einer Vermutung aus der Polytop-Geometrie Revue passieren. Warren D. Hirsch hat 1957 vermutet, dass man je zwei Ecken eines d-dimensionalen Polytops mit n Facetten über maximal $n - d$ Kanten verbinden kann. Die Autoren erläutern zunächst den Hintergrund und die Bedeutung dieser Vermutung im Zusammenhang mit linearen Optimierungsproblemen und gehen dann auf bewiesene Spezialfälle, Zweifel an der Allgemeingültigkeit, Modifikationen und bekannte Schranken ein. Der vorliegende Beitrag ist ganz überblicksartig gehalten; Details kann man in einer ausführlicheren Version auf dem `arXiv`-Server finden. Während des letzten Korrekturgangs für dieses Heft erreicht mich die ganz aktuelle Nachricht von Francisco Santos, dass er ein Gegenbeispiel zur Hirsch-Vermutung gefunden hat. Genaueres finden Sie in einer „note added in proof“ am Ende des Übersichtsartikels.

In seinem historischen Beitrag „Some direct and remote relations of Gauss with Belgian mathematicians“ beleuchtet Jean Mawhin zunächst den Besuch Adolphe Quetelets – eines späteren Sekretärs der königlichen Akademie von Brüssel – bei Gauß. Ein sich anschließender Briefwechsel wirft ein interessantes Licht auf Gauß und den Wissenschaftsbetrieb des 19. Jahrhunderts. Im zweiten Teil rekapituliert der Autor den Beweis und dessen Vorgeschichte des von Gauß vermuteten Primzahlsatzes durch den belgischen Mathematiker Charles-Jean de La Vallée Poussin. Schließlich wird an Hand von de La Vallée Poussins Vortrag aus dem Jahre 1939 zum 100. Geburtstag von Gauß’ Werk zur Potentialtheorie dargelegt, wie intensiv dieses nachfolgende Wissenschaftlergenerationen inspiriert und geprägt hat.

H.-Ch. Grunau (✉)
Institut für Analysis und Numerik, Fakultät für Mathematik, Otto-von-Guericke-Universität,
Postfach 4120, 39016 Magdeburg, Deutschland
e-mail: hans-christoph.grunau@ovgu.de

Beim Durchblättern dieses Heftes wird Ihnen sofort ein neues Erscheinungsbild auffallen. Der Verlag Vieweg + Teubner und die DMV haben sich auf ein neues Produktionsverfahren verständigt, das für die Leserinnen und Leser des Jahresberichtes eine Reihe von Verbesserungen bringen wird. Der Jahresbericht wird Bestandteil von SpringerLink und damit unter den dort üblichen Bedingungen unmittelbar online verfügbar. Die DMV pflegt weiterhin ihre Webseiten zum Jahresbericht, auf denen für jedermann sofort die bibliographischen Informationen und längere Zusammenfassungen und nach Ablauf von zwei Jahren sogar die Originaldateien der Beiträge kostenfrei verfügbar sind. Der Verlag Vieweg + Teubner ist der DMV sehr entgegengekommen; gemeinsam hoffen wir, dass diese Verbesserungen Ihre Freude am und auf den Jahresbericht steigern werden.

Jahresber Dtsch Math-Ver (2010) 112: 73–98
DOI 10.1365/s13291-010-0001-8

ÜBERSICHTSARTIKEL

An Update on the Hirsch Conjecture

Edward D. Kim · Francisco Santos

Received: 3 July 2009 / Published online: 3 June 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The Hirsch conjecture was posed in 1957 in a question from Warren M. Hirsch to George Dantzig. It states that the graph of a d-dimensional polytope with n facets cannot have diameter greater than $n-d$. The number n of facets is the minimum number of closed half-spaces needed to form the polytope and the conjecture asserts that one can go from any vertex to any other vertex using at most $n-d$ edges.

Despite being one of the most fundamental, basic and old problems in polytope theory, what we know is quite scarce. Most notably, no polynomial upper bound is known for the diameters that are conjectured to be linear. In contrast, very few polytopes are known where the bound $n-d$ is attained. This paper collects known results and remarks both on the positive and on the negative side of the conjecture. Some proofs are included, but only those that we hope are accessible to a general mathematical audience without introducing too many technicalities.

Keywords Graph diameter · Hirsch conjecture · Linear programming · Polytopes

Mathematics Subject Classification (2000) 05C12 · 52B05 · 90C08

E.D. Kim was supported in part by the Centre de Recerca Matemàtica, NSF grant DMS-0608785 and NSF VIGRE grants DMS-0135345 and DMS-0636297. F. Santos was supported in part by the Spanish Ministry of Science through grant MTM2008-04699-C03-02.

E.D. Kim
Department of Mathematics, University of California, Davis, Davis, CA 95616, USA
e-mail: ekim@math.ucdavis.edu
url: http://www.math.ucdavis.edu/~ekim/

F. Santos (✉)
Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, 39005 Santander, Spain
e-mail: francisco.santos@unican.es
url: http://personales.unican.es/santosf/

1 Introduction

Convex polytopes generalize convex polygons (of dimension two). More precisely, a *convex polyhedron* is any intersection of finitely many affine closed half-spaces in $\mathbb{R}^d$. A *polytope* is a bounded polyhedron. The long-standing Hirsch conjecture is the following very basic statement about the structure of polytopes. Besides its implications in linear programming, which motivated the conjecture in the first place, it is one of the most fundamental open questions in polytope theory.

Conjecture 1.1 (Hirsch conjecture) *Let $n > d \geq 2$. Let P be a d-dimensional polytope with n facets. Then* $\operatorname{diam}(G(P)) \leq n - d$.

Facets are the faces of dimension $d - 1$ of P, so that the number n of them is the minimum number of half-spaces needed to produce P as their intersection (assuming P is full-dimensional). The number $\operatorname{diam}(G(P)) \in \mathbb{N}$ is the *diameter of the graph of P*. Put differently, the conjecture states that we can go from any vertex of P to any other vertex traversing at most $n - d$ edges.

Consider the following examples; all of them satisfy the inequality strictly, except for the cube where it is tight:

P	n	d	$\operatorname{diam}(G(P))$
Polygon	n	2	$\lfloor n/2 \rfloor$
Cube	6	3	3
Icosahedron	20	3	3
Soccer-ball	32	3	9

Polytopes and polyhedra are the central objects in the area of geometric combinatorics, but they also appear in diverse mathematical fields: From the applications' point of view, a polyhedron is the *feasibility region* of a *linear program* [14]. This is the context in which the Hirsch conjecture was originally posed (see below). In *toric geometry*, to every (rational) polytope one associates a certain projective variety (see, e.g., [50]). The underlying interaction between combinatorics and algebraic geometry has proved extremely fruitful for both areas, leading for example to a complete characterization of the possible numbers of faces (vertices, edges, facets, ...) that a *simplicial polytope* can have. The same question for arbitrary polytopes is open in dimension four and higher [60]. Polytopes with special symmetries, such as regular ones and variations of them arise naturally from Coxeter groups and other algebraic structures [6, 25]. Last but not least, counting integer points in polytopes with integer vertex coordinates has applications ranging from number theory and representation theory to cryptography, integer programming, and statistics [5, 16].

In this paper we review the current status of the Hirsch conjecture and related questions. Some proofs are included, and many more appear in an appendix which is available electronically [38]. Results whose proof can be found in [38] are marked with an asterisk. An earlier survey of this topic, addressed to a more specialized audience, was written by Klee and Kleinschmidt in 1987 [40].

1.1 A Bit of Polytope Theory

We now review several concepts that will appear throughout this paper. For further discussion, we refer the interested reader to [18, 59].

A *polyhedron* is the intersection of a finite number of closed half-spaces and a polytope is a bounded polyhedron. A *polytope* is, equivalently, the convex hull of a finite collection of points. Although the geometric objects are the same, from a computational point of view it makes a difference whether a certain polytope is represented as a convex hull or via linear inequalities: the size of one description cannot be bounded polynomially in the size of the other, if the dimension d is not fixed. The *dimension* of a polytope is the dimension of its affine hull $\mathrm{aff}(P)$. A d-dimensional polytope is called a *d-polytope*.

If H is a closed half-space containing P, then the intersection of P with the boundary of H is called a *face* of P. Every non-empty face is the intersection of P with a supporting hyperplane. Faces are themselves polyhedra of lower dimension. A face of dimension i is called an *i-face*. The 0-faces are the *vertices* of P, the 1-faces are *edges*, the $(d-2)$-faces are *ridges*, and the $(d-1)$-faces are called *facets*. In its irredundant description, a polytope is the convex hull of its vertices, and the intersection of its facet-defining half-spaces.

For a polytope P, we denote by $G(P)$ its *graph* or 1-*skeleton*, consisting of the vertices and edges of P: the vertices of the graph $G(P)$ are indexed by the vertices of the polytope P, and two vertices in the graph $G(P)$ are connected by an edge exactly when their corresponding vertices in P are contained in a 1-face. The *distance* between two vertices in a graph is the minimum number of edges needed to go from one to the other, and the *diameter* of a graph is the maximum distance between its vertices. (For an unbounded polyhedron, the graph contains only the *bounded* edges. The unbounded 1-faces are called *rays*.)

Example 1.2 Examples of polytopes one can build in every dimension are the following:

1. The d-simplex. The convex hull of $d+1$ points in $\mathbb{R}^d$ that do not lie on a common hyperplane is a d-dimensional simplex. It has $d+1$ vertices and $d+1$ facets. Its graph is complete, so its diameter is 1.
2. The d-cube. The vertices of the d-cube, the product of d segments, are the 2^d points with ± 1 coordinates. Its facets are given by the $2d$ inequalities $-1 \le x_i \le 1$. Its graph has diameter d: the number of steps needed to go from a vertex to another equals the number of coordinates in which the two vertices differ.
3. Cross polytope. This is the convex hull of the d standard basis vectors and their negatives, which generalizes the 3-dimensional octahedron. It has 2^d facets, one in each orthant of $\mathbb{R}^d$. Its graph is almost complete: the only edges missing from it are those between opposite vertices.

See Fig. 1.

Their numbers m of vertices, n of facets, dimension d and diameter are:

P	m	n	d	$\mathrm{diam}(G(P))$
d-Simplex	$d+1$	$d+1$	d	1
d-Cube	2^d	$2d$	d	d
d-Crosspolytope	$2d$	2^d	d	2

Of special importance are the simple and simplicial polytopes. A d-polytope is called *simple* if every vertex is the intersection of exactly d facets. Equivalently, a d-polytope is simple if every vertex in the graph $G(P)$ has degree exactly d. We note that the d-simplices and d-cubes are simple, but cross-polytopes are not simple starting in dimension three. Any polytope or polyhedron P, given by its facet-description, can be perturbed to a simple one P' by a generic and small change in the coefficients of its defining inequalities. This will make non-simple vertices "explode" and become clusters of new vertices, all of which will be simple. This process can not decrease the diameter of the graph, since we can recover the graph of P from that of P' by collapsing certain edges. Hence, to study the Hirsch conjecture, one only needs to consider the simple polytopes:

Lemma 1.3 *The diameter of any polytope P is bounded above by the diameter of some simple polytope P' with the same dimension and number of facets.*

Graphs of simple polytopes are better behaved than graphs of arbitrary polytopes. Their main property in the context of the Hirsch conjecture is that if u and v are vertices joined by an edge in a simple polytope then there is a single facet containing u and not v, and a single facet containing v and not u. That is, at each step along the graph of P we enter a single facet and leave another one.

Every polytope P (containing the origin in its interior, which can always be assumed by a suitable translation) has a polar polytope P^* whose vertices (respectively facets) correspond to the facets (respectively vertices) of P. More generally, every $(d-i)$-face of P^* corresponds to a face of P of dimension $i-1$, and the incidence relations are reversed.

The polars of simple polytopes are called *simplicial*, and their defining property is that every facet is a $(d-1)$-simplex. As an example, the d-dimensional cross polytope is the polar of the d-cube. Since cubes are simple polytopes, cross polytopes are simplicial. The polar of a simplex is a simplex, and simplices are the only polytopes

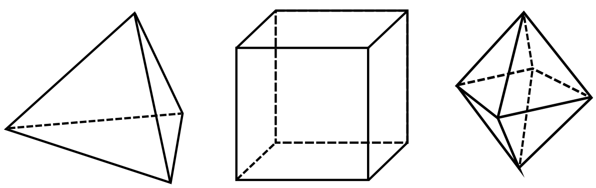


Fig. 1 Basic examples of polytopes

of dimension greater than two which are at the same time simple and simplicial. Since all faces of a simplex are themselves simplices, all faces of a simplicial polytope are simplices. From this viewpoint, one can forget the geometry of P^* and look only at the combinatorics of the simplicial complex formed by its faces, the boundary of P^*. Topologically, this simplicial complex is a sphere of dimension $d-1$.

For simplicial polytopes we can state the Hirsch conjecture as asking how many ridges do we need to cross in order to walk between two arbitrary facets, if we are only allowed to move from one facet to another via a ridge. This suggests defining the *dual graph* $G^{\Delta}(P)$ of a polytope: The undirected graph having as nodes the facets of P and in which two nodes are connected by an edge if and only if their corresponding facets intersect in a ridge of P. In summary, $G^{\Delta}(P) = G(P^*)$.

1.2 Relation to Linear Programming

The original motivation for the Hirsch conjecture comes from its relation to the simplex algorithm for linear programming. In linear programming, one is given a system of linear equalities and inequalities, and the goal is to maximize (or minimize) a certain linear functional. Every such problem can be put in the following standard form, where A is an $m \times n$ real matrix A, and $\mathbf{b} \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ are two real vectors:

$$\text{Maximize } c \cdot \mathbf{x}, \text{ subject to } A\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x} \geq 0.$$

Suppose the matrix A has full row rank $m \leq n$. Then, the equality $A\mathbf{x} = \mathbf{b}$ defines a d-dimensional affine subspace ($d = n - m$), whose intersection with the linear inequalities $\mathbf{x} \geq 0$ gives the *feasibility polyhedron* P:

$$P := \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x} \geq 0\}.$$

One typically desires not only the maximum value of $c \cdot \mathbf{x}$ but also (the actual coordinates of) a vector $\mathbf{x} \in P$ where the maximum is attained. It is easy to prove that such an $\mathbf{x}$, if it exists, can be found among the vertices of P. If P is unbounded and $c \cdot \mathbf{x}$ does not have an upper bound on it one considers the problem "solved" by describing a ray of P where the value $c \cdot \mathbf{x}$ goes to infinity.

In 1979, Hačijan [30] proved that linear programming problems can be solved in polynomial time via the so-called *ellipsoid method*. In 1984, Karmarkar [37] devised a different approach, the *interior point method*. Although the latter is more applicable (easier to implement, better complexity) than the former, still to this day the most commonly used method for linear programming is the *simplex method* devised by G. Dantzig in 1947. For a complete account of the complexity of linear programming, see the survey [48] by Megiddo.

In geometric terms, the simplex method first finds an arbitrary vertex in the feasibility polyhedron P. Then, it moves from vertex to adjacent vertex in such a way that the value $c \cdot \mathbf{x}$ of the linear functional increases at every step. These steps are called *pivots* and the rule used to choose one specific adjacent vertex is called the *pivot rule*. When no pivot step can increase the functional, convexity implies that we have arrived to the global maximum.

Clearly, a lower bound for the performance of the simplex method under *any* pivot rule is the diameter of the polyhedron P. The converse is not true, since knowing that

P has a small graph diameter does not in principle tell us how to go from one vertex to another in a small number of steps. In particular, many of the results on diameters of polyhedra do not help for the simplex method.

In fact, the complexity of the simplex method depends on the local rule (the *pivot rule*) chosen to move from vertex to vertex. The *a priori* best pivot rule, the one originally proposed by Dantzig, is "move along the edge with maximum gradient", but Klee and Minty [41] showed in 1972 that this can lead to paths of exponential length, even in polytopes combinatorially equivalent to cubes. The same worst-case exponential behavior has been proved for essentially every deterministic rule devised so far, although there are subexponential, but yet not polynomial, randomized pivot algorithms (see Theorem 2.8). However, the simplex algorithm is highly efficient *in practice* on most linear optimization problems.

There is another reason why investigating the complexity of the simplex method is important, even if we already know polynomial time algorithms. The algorithms of Khachiyan and Karmarkar are polynomial in the *bit length* of the input; but it is of practical importance to know whether a polynomial algorithm for linear programming in the *real number machine* model of Blum, Cucker, Shub, and Smale [7] exists. That is, is there an algorithm that uses a number of arithmetic operations that is polynomial on the number of coefficients of the linear program, rather than on their total bit-length; or, better yet, a *strongly polynomial algorithm*, i.e., one that is polynomial both in the arithmetic sense and the bit sense? These two related problems were included by Smale in his list of "mathematical problems for the next century" [54]. A polynomial pivot rule for the simplex method would solve them in the affirmative.

In this context, the following *polynomial version* of the conjecture is relevant, if the linear one turns out to be false. See, for example, [35]:

Conjecture 1.4 (Polynomial Hirsch conjecture) *Is there a polynomial function* $f(n,d)$ *such that for any polytope* (*or polyhedron*) P *of dimension* d *with* n *facets*, $\operatorname{diam}(G(P)) \le f(n,d)$*?*

1.3 Overview of This Paper

Our initial purpose with this paper was two-fold: on the one hand, we thought it is about time to have in a single source an overview of the state of the art concerning the Hirsch conjecture and related issues, putting up to date the 20 year old survey by Klee and Kleinschmidt [40]. On the other hand, since there seems to be agreement on the fact that the Hirsch conjecture is probably false (with opinions about the polynomial version of it being divided) we wanted to give a fresh look at the past attempts to *disprove* the conjecture.

These two goals turned out to be in conflict, or at least too ambitious, so the first version of the paper was too long and too technical for the intended readership. After wise comments from our editor Jörg Rambau and an anonymous referee, we decided to take most of the proofs out of the main paper, and compiled them in the companion paper [38]. Results whose proof can be found in [38] are marked with an asterisk.

Our two-fold intentions are still reflected in the two quite distinct parts that the paper has, Sects. 2 and 3. The first one is devoted to positive results, comprising general upper bounds for polytope diameters, special cases where the Hirsch conjecture is known, etc., and the second one contains mainly constructions and results aimed at disproving the Hirsch conjecture.

The two sections differ in several respects: Sect. 2 is written in an informative style. No proofs are included (although some appear in [38]) since this section covers quite different topics and the techniques and ideas used are too technical and, more importantly, too diverse. In Sect. 3, on the contrary, we provide proofs for essentially all the results (some here and some in [38]); on the one hand the tools needed are more homogeneous and elementary; on the other hand, in this section we feel that having a new look at old results is useful. We have tried to identify the basic ingredients in each construction, obtaining in some cases much simpler (to our taste) proofs and expositions than the original ones. In particular, the main novelty in this survey, if any, is probably in our descriptions of the non-Hirsch polyhedron and Hirsch-sharp polytope found by Klee and Walkup in 1967 (Sect. 3.3) and of the non-Hirsch sphere found by Mani and Walkup in 1980 (Sect. 3.6).

Another difference between the two sections is that Sect. 2 contains several very recent developments, while all of Sect. 3, with the single exception of Theorem 3.11, refers to results that are at least 25 years old.

Let us now give a brief roadmap for the paper.

Section 2.1 lists the pairs of parameters (n, d) such that the Hirsch Conjecture is known to hold for all d-polytopes with n facets. That is, denoting $H(n, d)$ the maximum diameter of d-polytopes with n facets, we list all pairs for which the Hirsch inequality $H(n, d) \leq n - d$ is known to hold. This comprises the cases $d \leq 3$ (Klee [39]), $n - d \leq 6$ (Bremner and Schewe [11]), and $(n, d) \in \{(11, 4), (12, 4)\}$ (Bremner et al. [10]).

The next section lists general upper bounds on $H(n, d)$: a linear one in fixed dimension (Barnette and Larman) [4, 44] and a quasi-polynomial one of $n^{\log_2(d)+1}$ (Kalai-Kleitman [36]). These bounds hold not only for diameters of polytopes but also for much more abstract and general objects. A very recent development by Eisenbrand, Hähnle, Razborov and Rothvoß [24] is the identification of one such class for which the proofs of these two bounds work but which admit objects with quadratic diameter. This may be considered evidence against the Hirsch conjecture.

In Sect. 2.3 we concentrate on algorithmic aspects. For example, we state two algorithmic analogues of the two bounds mentioned above: the proof by Meggido [47] that linear programming can be done in linear time if the dimension is fixed, and randomized pivot rules for the simplex method in arbitrary dimension that finish in $O(\exp(K\sqrt{d \log n}))$ (Kalai [34], Matoušek, Sharir and Welzl [46]).

We then turn our attention to special polytopes for which good bounds are known. Polytopes with 0-1 coordinates and linear programming duals of transportation polytopes are known to satisfy the Hirsch conjecture (Naddef [49], Balinski [3]). Diameters of network-flow polytopes, which include transportation polytopes, have quadratic bounds [13, 28, 51].

Section 2 finishes with an account of recent work of Deza, Terlaky and Zinchenko [20–22] on a continuous analogue of the Hirsch conjecture that arises in the context of interior point methods for linear programming.

Almost all of the results in Sect. 3 revolve around two basic ingredients. The first one is the *wedge* operation, which we describe in Sect. 3.1. Wedging is a very simple operation that increases both the dimension and number of facets of a polytope by one maintaining (or increasing) its diameter. Using this, it is easy to prove the following fundamental result:

Theorem 1.5 (Klee and Walkup [42]) $H(d+k,d) \le H(2k,k)$, *with equality if* (*but not only if*) $k < d$.

In particular, to prove (or disprove) the Hirsch conjecture one can concentrate on the case where the number of facets equals twice the dimension. This case is sometimes referred to as the *d-step Conjecture*, since the Hirsch conjecture is saying that we can go from any vertex to any other vertex in d-steps. Via wedging, the Hirsch conjecture is also equivalent to the following *non-revisiting Conjecture*: If u and v are two arbitrary vertices of a simple polytope P, then there is a path from u to v which at every step enters a facet of P that was not visited before. We prove the equivalence of the three conjectures (Hirsch, d-step and non-revisiting) in Sect. 3.2.

The second ingredient is the construction by Klee and Walkup [42] of a 4-dimensional polytope with 9 facets that meets the Hirsch bound with equality (that is, whose diameter equals $9-4=5$). Polytopes with this property are called *Hirsch-sharp*. They are easy to construct with a number of facets not exceeding twice their dimension (e.g., cubes). Klee and Walkup's Hirsch-sharp polytope is the smallest "non-trivial" Hirsch-sharp polytope, with more facets than twice its dimension. In fact, it is also the starting block to the construction of every other Hirsch-sharp polytope with $n > 2d$ known to date. In Sect. 3.3 we give our own description and coordinatization (much smaller than the original one) of the Klee-Walkup polytope.

In Sect. 3.4 we recount the state of the art on the existence of Hirsch-sharp polytopes, following work of Fritzsche, Holt and Klee [26, 31, 32]. Their results, combined with what is known for small dimension or number of facets, are summarized in Table 1, which gives a "plot" of the function $H(n,d)-(n-d)$. The horizontal coordinate is $n-2d$, so that the column labeled "0" corresponds to the polytopes relevant to the d-step conjecture. The cases where we know $H(n,d)$ exactly are marked "=" or "<" depending on whether Hirsch-sharp polytopes exist or not. The cases where Hirsch-sharp polytopes are known to exist but for which the Hirsch conjecture is not proved are marked "≥". Cases where we neither know the Hirsch conjecture nor the existence of Hirsch-sharp polytopes are marked "?" and appear only in dimensions 4, 5 and 6. The diagonal dots in the left column reflect the equality case of Theorem 1.5.

The Klee-Walkup polytope is also instrumental in the construction by Klee, Walkup and Todd [42, 56] of counter-examples to two generalizations of the Hirsch conjecture that are quite natural in the context of linear programming: the case of perhaps-*unbounded* polyhedra (which was the original conjecture by Hirsch) and a *monotone* version in which we look at the maximum number of monotone steps with respect to a given linear function that are needed to go from any vertex of a polytope P to an optimal vertex. We show these constructions in Sect. 3.5, and show in Sect. 3.6 a counter-example, by Mani and Walkup [45] to a third, topological, version of the conjecture.

Table 1 $H(n, d)$ versus $n - d$, the state of the art

d \ $n-2d$	$\cdots$	0	1	2	3	4	5	6	7	$\cdots$
2		=	<	<	<	<	<	<	<	$\cdots$
3	$\iddots$	=	<	<	<	<	<	<	<	$\cdots$
4	$\iddots$	=	=	<	<	<	?	?	?	$\cdots$
5	$\iddots$	=	=	?	?	?	?	?	?	$\cdots$
6	$\iddots$	=	?	?	?	?	?	?	?	$\cdots$
7	$\iddots$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\cdots$
8	$\iddots$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\cdots$
$\vdots$	$\iddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

2 Bounds and Algorithms

In this section, we present special cases for which the Hirsch conjecture holds, upper bounds for diameters of polytopes and subexponential complexity results for the simplex method. We also summarize recent work on analogues of the conjecture for hyperplane arrangements and for paths of interior point methods.

2.1 Small Dimension or Few Facets

The following statements exhaust all pairs (n, d) for which the maximum diameter $H(n, d)$ of d-polytopes with n facets is known. We omit the cases $n < 2d$, because $H(d + k, d) = H(2k, k)$ for all $k < d$ (see Theorem 1.5), and the trivial case $d \leq 2$. Remember that an asterisk in front of a statement denotes the proof can be found in [38].

*__Theorem 2.1__ (Klee [39]) $H(n, 3) = \lfloor \frac{2n}{3} \rfloor - 1$.

Theorem 2.2

- $H(8, 4) = 4$ (Klee [39]).
- $H(9, 4) = H(10, 5) = 5$ (Klee-Walkup [42]).
- $H(10, 4) = 5$, $H(11, 5) = 6$ (Goodey [29]).
- $H(11, 4) = H(12, 6) = 6$ (Bremner-Schewe [11]).
- $H(12, 4) = 7$ (Bremner et al. [10]).

Since $\max_d H(d + k, d) = H(2k, k)$ (see Theorem 1.5 again), the results for $H(8, 4)$, $H(10, 5)$ and $H(12, 6)$ imply:

Corollary 2.3 *The Hirsch conjecture holds for polytopes with at most six facets more than their dimension.*

It is easy to generalize one direction of Theorem 2.1, giving the following lower bound for $H(n, d)$. Observe that the formula gives the exact value of $H(n, d)$ for $d \in \{1, 2, 3\}$.

*__Proposition 2.4__

$$H(n,d) \geq \left\lfloor \frac{d-1}{d} n \right\rfloor - (d-2).$$

2.2 General Upper Bounds on Diameters

Diameters of polytopes admit a *linear* upper bound when the dimension d is fixed. This was first noticed by Barnette [4] and then improved by Larman [44]:

*__Theorem 2.5__ (Larman [44]) *For every* $n > d \geq 3$, $H(n,d) \leq n2^{d-3}$.

But when the number of facets is not much bigger than d, a much better upper bound was given by Kalai and Kleitman [36], with a surprisingly simple and elegant proof (the paper is just two pages!).

*__Theorem 2.6__ (Kalai-Kleitman [36]) *For every* $n > d$, $H(n,d) \leq n^{\log_2(d)+1}$.

The proofs of Theorems 2.6 and 2.5 use very limited properties of graphs of polytopes. For example, Klee and Kleinschmidt (see Sect. 7.7 in [40]) show that Theorem 2.5 holds for the ridge-graphs of all pure simplicial complexes, and even more general objects. In the same vein, Eisenbrand, Hähnle, Razborov and Rothvoß [24] have recently shown the following generalization of Theorems 2.6 and 2.5:

Theorem 2.7 (Eisenbrand et al. [24]) *Let G be a graph whose vertices are certain subsets of size d of an n-element set. Assume that between every pair of vertices u and v in G there is a path using only vertices that contain $u \cap v$.*

Then, $\mathrm{diam}(G) \leq \min\{n^{1+\log d}, n2^{d-1}\}$.

The novelty in [24] is that the authors show that there are graphs with the hypotheses of Theorem 2.7 and with $\mathrm{diam}(G) \geq cn^2/\log n$, for arbitrarily large n and a certain constant c. It is not clear whether this is support against the Hirsch conjecture or it simply indicates that the arguments in the proofs of Theorems 2.6 and 2.5 do not take advantage of properties that graphs of polytopes have and which prevent their diameters from growing. For example, observe that *any* connected graph is valid for the case $d = 1$ of Theorem 2.7.

2.3 Subexponential Simplex Algorithms

Since the Hirsch conjecture is strongly motivated by the simplex algorithm of linear programming, it is natural to ask about the number of iterations needed under particular pivot rules. Most of the proofs for the upper bounds in the previous sections do not give a clue on how to find a short path towards the vertex maximizing a given functional, or even an explicit path between any pair of given vertices.

Kalai [34] and, independently, Matoušek, Sharir and Welzl [46] proved the existence of randomized pivot rules for the simplex method with subexponential running time for arbitrary linear programs.

Theorem 2.8 (Kalai [34], Matoušek, Sharir and Welzl [46]) *There exist randomized simplex algorithms where the expected number of arithmetic operations needed in the worst case is at most* $\exp(K\sqrt{d\log n})$, *where* K *is a fixed constant.*

If we consider Theorem 2.8 as an algorithmic analogue of Theorem 2.6, then the following result of Megiddo is the analogue of Theorem 2.5. It says that linear programming can be performed in linear time in fixed dimension:

Theorem 2.9 (Meggido [47]) *There are pivot rules for the simplex algorithm that run in* $O(2^{2^d} n)$ *time.*

It is also known that random polytopes have polynomial diameter, which explains why the simplex method seems to work well in practice. The first results in this direction were proved by Borgwardt [8] and, independently, Smale [53], who analyzed the "average case" complexity of the simplex method. Average case means that we are looking at a linear program

$$\text{Maximize } c \cdot \mathbf{x}, \text{ subject to } A\mathbf{x} = b \text{ and } \mathbf{x} \geq 0,$$

but the entries of A, b and c are considered random variables with respect to certain spherically symmetric probability distributions. In Borgwardt's model the simplex method runs in expected polynomial time in the size of the input. Smale shows that in his model, if one of the parameters d and $n-d$ is fixed and the other is allowed to grow, the expected running time is only polylogarithmic. The latter was improved to constant by Megiddo, see [48] for details.

Even more surprising is the fact that *every* linear program can be slightly perturbed to one that can be solved in polynomial time. Let us formalize this. Let P be the feasibility polyhedron

$$P = \{\mathbf{x} \in \mathbb{R}^d \mid \langle a_i, \mathbf{x}\rangle \leq b, (i = 1, \ldots, n)\}$$

of a certain linear program. If we replace the vectors $a_i \in \mathbb{R}^d$ and $b \in \mathbb{R}^n$ with independent Gaussian random vectors with means $\mu_i = a_i$ and $\mu = b$ (respectively), and standard deviations $\sigma \max_i \|(\mu_i, \mu)\|$ we say that we have *perturbed* P *randomly within a parameter* σ. In [55], Spielman and Teng proved that the expected diameter of a linear program that is perturbed within a parameter σ is polynomial in d, n, and σ^{-1}. In [57], Vershynin improved the bound to be polylogarithmic in n and polynomial only in d and σ^{-1}.

Theorem 2.10 (Vershynin [57]) *If a linear program is perturbed randomly within a parameter* σ, *then its expected diameter of its feasibility polyhedron is in* $O(log^7 n(d^9 + d^3\sigma^{-4}))$.

As mentioned above, this result is not only structural. The simplex method can *find* a path of that expected length in the perturbed polyhedron.

2.4 Some Polytopes from Combinatorial Optimization

There are some classes of polytopes of special interest and for the diameters of which we know polynomial upper bounds.

2.4.1 Small Integer Coordinates

Of special importance in combinatorial optimization are the 0-1 polytopes, in which every vertex has coordinates 0 or 1.They satisfy the Hirsch conjecture.

****Theorem 2.11** (Naddef [49]) *If P is a 0-1 polytope then*

$$\mathrm{diam}(P) \le \#facets(P) - \dim(P).$$

As a generalization, Kleinschmidt and Onn [43] prove the following bound on the diameter of lattice polytopes in $[0,k]^d$. A polytope is called a *lattice polytope* if every coordinate of every vertex is integral.

Theorem 2.12 (Kleinschmidt and Onn [43]) *The diameter of a lattice polytope contained in $[0,k]^d$ cannot exceed kd.*

However, existence of a polynomial pivot rule for the simplex method in 0-1 polytopes is open. The proof of Theorem 2.11 constructs a short path from u to v only assuming that we know the coordinates of both.

2.4.2 Network-Flow Polytopes

A network flow polytope is defined by an arbitrary directed graph $G = (V, E)$ with weights given to its vertices. Negative weights represent demands and positive weights represent supplies. A flow is an assignment of non-negative numbers to the edges so as to cancel all the demands and supplies. See [13], [28], and [51] for details.

For any network G with e edges and v vertices, every sufficiently generic set of vertex weights produces a simple $(e - v + 1)$-dimensional polytope with at most $2e$ facets. Its diameter has the following almost quadratic upper bound. The proof yields a polynomial time pivot rule for the simplex method on these polytopes.

Theorem 2.13 [13, 28, 51] *The diameter of the network flow polytope on a directed graph $G = (V, E)$ is in $O(ev \log v)$. This, in turn, is in $O(n^2 \log n)$, where n is the number of facets of the polytope.*

The matrices defining network flow polytopes are examples of *totally unimodular* matrices, meaning that all its subdeterminants are 0, 1, or -1. Polytopes defined by these matrices still have polynomially bounded diameters, although the degree in the bound is much worse than the one for network flow polytopes:

Theorem 2.14 (Dyer and Frieze [23]) *For any totally unimodular $n \times d$ matrix A the diameter of the polyhedron $\{\mathbf{x} \in \mathbb{R}^d : A\mathbf{x} \le c\}$ is in $O(d^{16} n^3 (\log(dn))^3)$.*

2.4.3 Transportation and Dual Transportation Polytopes

Given vectors $a \in \mathbb{R}^p$ and $b \in \mathbb{R}^q$, the $p \times q$ transportation polytope defined by a and b is the set of all $p \times q$ non-negative matrices with row sums given by a and column sums given by b:

$$T_{p,q}(a,b) = \left\{ (x_{ij}) \in \mathbb{R}^{p\times q} \;\Big|\; \sum_j x_{ij} = a_i, \sum_i x_{ij} = b_j, x_{ij} \geq 0 \right\}.$$

As an example, the Birkhoff polytope, whose vertices are the permutation matrices, is the transportation polytope obtained with $p = q$ and $a = b = (1, \ldots, 1)$.

It is easy to show that (generically) $T_{p,q}(a,b)$ is a $(p-1)(q-1)$-dimensional polytope with at most pq facets. Thus, the Hirsch conjecture translates to its diameter being at most $p + q - 1$.

Transportation polytopes are a special case of network flow polytopes; they arise when the network is a complete bipartite graph on p and q nodes with all edges directed in the same direction. In particular, Theorem 2.13 gives an almost quadratic bound for their diameters. But Brightwell et al. [12] have recently proved a *linear* bound, with a multiplicative factor of eight. This has now been improved to:

Theorem 2.15 (Hurkens [33]) *The diameter of any $p \times q$ transportation polytope is at most* $3(p + q - 1)$.

In the context of linear programming, for every d-polyhedron with n facets there is a dual $(n - d)$-polyhedron with the same number of facets. Every linear program can be solved in its "primal" or "dual" polyhedron. The optimum achieved is the same in both, but the complexity of the algorithm may not.

The linear programming duals of $p \times q$ transportation polytopes are $(p + q - 1)$-polyhedra with pq facets. Balinski [3] proved the Hirsch conjecture for them.

Theorem 2.16 (Balinski [3]) *Let C be a $p \times q$ matrix. The diameter of the dual transportation polytope $D_{p,q}(C)$ is at most $(p-1)(q-1)$. This bound is the best possible and it yields a polynomial time dual simplex algorithm.*

2.4.4 3-Way Transportation Polytopes

Another seemingly harmless generalization of transportation polytopes comes from considering 3-way tables $(x_{ijk})_{ijk} \in \mathbb{R}^{p\times q\times r}$ instead of matrices. In fact, there are two different such generalizations. An *axial* 3*-way transportation polytope* consists of all non-negative tables with fixed sums in 2-d slices. A *planar* 3*-way transportation polytope* consists of all non-negative tables with fixed sums in 1-d slices. (The names seem wrong, but they reflect the fact that an axial transportation polytope is defined by three *vectors* of sizes p, q and r, and a planar one by three *matrices* of sizes $p \times q$, $p \times r$, and $q \times r$.)

Despite their definition being so close to that of transportation polytopes, 3-way transportation polytopes are *universal* in the following sense:

Theorem 2.17 (De Loera and Onn [17]) *Let P be a rational convex polytope.*

- *There is a 3-way planar transportation polytope Q isomorphic to P.*
- *There is a 3-way axial transportation polytope Q which has a face F isomorphic to P.*

In both cases there is a polynomial time algorithm to construct Q (and F).

Isomorphic here means affinely (and rationally) equivalent. In particular, that the polytope Q or its face F have the same edge-graph as P. Thus, it was interesting to try to apply to the 3-way case the methods that gave polynomial upper bounds for the graphs of transportation polytopes. This was attempted in [15], where a quadratic upper bound was obtained but only for axial transportation polytopes. A generalization of this result to *faces* of them or to *planar* transportation polytopes would prove the polynomial Hirsch conjecture.

Theorem 2.18 (De Loera, Kim, Onn, Santos [15]) *The diameter of every 3-way axial $p \times q \times r$ transportation polytope is at most $2(p+q+r-3)^2$.*

2.5 A Continuous Hirsch Conjecture

Here we summarize some recent work of Deza, Terlaky and Zinchenko [20–22] in which they propose continuous analogues of the Hirsch and d-step conjectures related to the central path method—a variant of interior point methods—of linear programming. (For a complete description of the method we refer the reader to [9, 52].) The analogy comes from analyzing the total curvature $\lambda_c(P)$ of the central path with respect to a certain cost function c for the polyhedron P. By analogy with $H(n,d)$, let $\Lambda(n,d)$ denote the largest total curvature of the central path over all polytopes P of dimension d defined by n inequalities and over all linear objective functions c.

It had been conjectured that $\lambda_c(P)$ is bounded by a constant for each dimension d, and that it grows at most linearly with varying d. Deza et al. have disproved both statements: in [20], they construct polytopes for which $\lambda_c(P)$ grows exponentially with d. More strongly, in [22] they construct a family of polytopes that show that λ_c cannot be bounded only in terms of d:

Theorem 2.19 [22] *For every fixed dimension $d \geq 2$, $\liminf_{n\to\infty} \frac{\Lambda(n,d)}{n} \geq \pi$.*

Deza et al. consider this result a continuous analogue of the existence of Hirsch-sharp polytopes. Motivated by this they pose the following conjecture:

Conjecture 2.20 (Continuous Hirsch conjecture) $\Lambda(n,d) \in O(n)$. *That is, there is a constant K such that $\Lambda(n,d) \leq Kn$ for all n and d.*

Theorem 2.19 says that if the continuous Hirsch conjecture is true, then it is tight, modulo a constant factor. Deza et al. also conjecture a continuous variant of the d-step conjecture, and show it to be equivalent to the continuous Hirsch conjecture, thus providing an analogue of Theorem 3.2:

Conjecture 2.21 (Continuous d-step conjecture) *The function* $\Lambda(2d,d)$ *grows linearly in its input. That is to say,* $\Lambda(2d,d)$ *is in* $O(d)$.

Theorem 2.22 [21] *The continuous Hirsch conjecture is equivalent to the continuous* d*-step conjecture. That is, if* $\Lambda(2d,d)\in O(d)$ *for all* d*, then* $\Lambda(n,d)\in O(n)$ *for all* d *and* n.

The best upper bound known for $\Lambda(n,d)$ is a bound in $O(n^d)$, derived from Theorem 2.23 below. This theorem refers to the central path curvature for hyperplane arrangements, as studied by Dedieu, Malajovich and Shub [19].

An arrangement $\mathcal{A}$ of n hyperplanes in dimension d is called *simple* if every n hyperplanes intersect at a unique point. It is easy to show that any simple arrangement of n hyperplanes in $\mathbb{R}^d$ has exactly $s=\binom{n-1}{d}$ bounded full-dimensional cells. For a simple arrangement $\mathcal{A}$ with bounded cells $P_1,\dots,P_s$ and a given objective function c, Dedieu et al. consider the quantity: $\lambda_c(\mathcal{A})=\frac{1}{s}\sum_{i=1}^{s}\lambda_c(P_i)$. That is, the average total curvature of central paths of all bounded cells in the arrangement. They prove:

Theorem 2.23 [19] $\lambda_c(\mathcal{A})\le 2\pi d$, *for every simple arrangement.*

Put differently, even if individual cells can give total curvature linear in n by Theorem 2.19, the average over all cells of a given arrangement is bounded by a function of d alone.

Turning the analogy back to polytope graphs, Deza et al. [22] consider the average diameter of the graphs of all bounded cells in a simple arrangement $\mathcal{A}$. Denote it $\mathrm{diam}(\mathcal{A})$ and let $\mathcal{H}(n,d)$ be the maximum of $\mathrm{diam}(\mathcal{A})$ over all simple arrangements defined by n hyperplanes in dimension d. They relate $\mathcal{H}(n,d)$ to the Hirsch conjecture, as follows:

Proposition 2.24 [22] *The Hirsch conjecture implies* $\mathcal{H}(n,d)\le d+\frac{2d}{n-1}$.

3 Constructions

We now move to interesting constructions of polytopes motivated by or related to the Hirsch conjecture. All the proofs that are not included in this section, plus additional comments, can be found in [38].

3.1 The Wedge Operation

Wedging is a very basic, yet extremely fruitful, operation that one can do to a polytope. Its simplicial counter-part is the *one-point suspension* (see [18, 38]).

Roughly speaking, the wedge of P at a facet F of it is the polytope, of one dimension more, obtained gluing two copies of P along F. See Fig. 2 for an example. More formally, let $f(x)\le b$ be an inequality defining the facet F. The wedge of P over F is the polytope

$$\mathrm{W}_F(P):=P\times[0,\infty)\cap\{(x,t)\in\mathbb{R}^d\times\mathbb{R}: f(x)+t\le b\}.$$

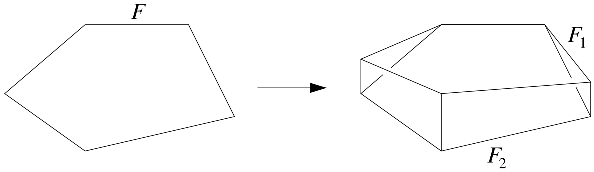


Fig. 2 A 5-gon and a wedge on its top facet

Put differently, $\mathrm{W}_F(P)$ is formed by intersecting the half-cylinder $C := P \times [0, \infty)$ with a closed half-space J in $\mathbb{R}^{d+1}$ such that:

- the intersection $J \cap C$ is bounded and has nonempty interior, and
- $\partial J \cap C = F$.

Lemma 3.1 *Let P be a d-polytope with n facets. Let $\mathrm{W}_F(P)$ be its wedge on a certain facet F. Then, $\mathrm{W}_F(P)$ has dimension $d+1$, $n+1$ facets, and*

$$\mathrm{diam}(\mathrm{W}_F(P)) \geq \mathrm{diam}(P).$$

Proof The wedge increases both the dimension and the number of facets by one. Indeed, $\mathrm{W}_F(P)$ has a vertical facet projecting to each facet of P other than F, plus the two facets that cut the cylinder $P \times \mathbb{R}$, and whose intersection projects to F. The diameter of $\mathrm{W}_F(P)$ is at least that of P, since every edge of $\mathrm{W}_F(P)$ projects either to an edge of P or to a vertex of P. □

In particular, if P is Hirsch-sharp then $\mathrm{W}_F(P)$ is either Hirsch-sharp or a counterexample to the Hirsch conjecture. The properties that P would need for the latter to be the case will be made explicit in Remark 3.5.

As a corollary of Lemma 3.1 we get that in order to prove (or disprove) the Hirsch conjecture it is sufficient to restrict attention to the case when the number of facets equals twice the dimension:

Theorem 3.2 (Klee-Walkup [42]) *$H(k+d, d) \leq H(2k, k)$, with equality if (but not necessarily only if) $k < d$.*

Proof By Lemma 3.1,

$$H(n, d) \leq H(n+1, d+1), \quad \forall n, d \tag{1}$$

so we only need to show that

$$H(n, d) \leq H(n-1, d-1) \quad \forall n < 2d. \tag{2}$$

Let P be a polytope with $n < 2d$ and let u and v be vertices of it. Since each vertex is incident to at least d facets, u and v lie in a common facet. This facet F has dimension $d-1$, and each facet of it is the intersection of F with another facet of P. Hence, F has at most $n-1$ facets itself. Since every path on F is also a path on P, we get (2). □

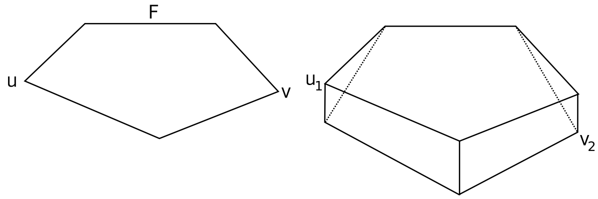


Fig. 3 The pentagon P and the wedge $P' = \mathrm{W}_F(P)$ over its facet F: the upper pentagonal facet of P' is F_1 and the lower pentagonal facet is F_2

3.2 The d-Step and Non-revisiting Conjectures

The intuition behind the Hirsch conjecture is that to go from vertex u to vertex v of a polytope P, one does not expect to have to enter and leave the same facet several times. This suggests the following conjecture:

Conjecture 3.3 (The non-revisiting conjecture) *Let P be a simple polytope. Let u and v be two arbitrary vertices of P. Then, there is a path from u to v which at every step enters a facet of P that was not visited before.*

Paths with the conjectured property, that they do not revisit any facet, are called *non-revisiting paths*. (In the literature, they are also called W_v paths and Conjecture 3.3 is also known as the W_v conjecture.) Non-revisiting paths are never longer than $n - d$: at each step, we must enter a different facet, and the d facets that the initial vertex lies in cannot be among them. Thus, the non-revisiting conjecture implies the Hirsch conjecture. It turns out both are equivalent. A first step in the proof is the following analogue of Theorem 3.2 for the non-revisiting conjecture:

Theorem 3.4 *If all k-polytopes with $2k$ facets has the non-revisiting property, then the same holds for all d-polytopes with $d + k$ facets, for all d.*

Proof Let P be a polytope with $n \neq 2d$ and suppose it does not have the non-revisiting property. That is, there are vertices u and v such that every path from u to v revisits some facet that it previously abandons. We will construct another polytope P' without the non-revisiting property and with:

- One less facet and dimension than P if $n < 2d$, and
- One more facet and dimension than P if $n > 2d$.

In the first case, u and v lie in a common facet F and we simply let $P' = F$. In the second case, let F be a facet not containing u nor v and let $P' = \mathrm{W}_F(P)$ be the wedge over F. Let F_1 and F_2 be the two facets of P' whose intersection projects to F. Let u_1 and v_2 be the vertices of P' that project to u and v and lie, respectively, on F_1 and F_2 (see Fig. 3). Now, consider a path from u_1 to v_2 on P' and project it to a path from u to v on P:

- If the path on P revisits a facet (call it G) other than F, then the path on P' revisits the facet that projects to G.
- If the path on P revisits F, then the path on P' revisits either F_1 or F_2.

□

Remark 3.5 In the proof of Lemma 3.1 we noted that, applied to a Hirsch-sharp polytope P the wedge operator produced either another Hirsch-sharp polytope or a counterexample to the Hirsch conjecture. The last proof shows that the latter can happen only if P does not have the non-revisiting property.

Theorems 3.2 and 3.4 say that both in the Hirsch and the non-revisiting conjectures the crucial case is that of $n = 2d$. It is not surprising then that they are equivalent, since in this case they both almost restrict to the following:

Conjecture 3.6 (The d-step conjecture) *Let P be a simple d-polytope with $2d$ facets and let u and v be two complementary vertices* (*i.e.*, *vertices not lying in a common facet*). *Then*, *there is a path of length d from u to v.*

There is still something to be proved, though. If u and v are not complementary vertices in a d-polytope with $2d$ facets then the d-step conjecture does not directly imply the other two. But in this case u and v lie in a common facet, so the proof of equivalence is not hard to finish via induction:

***Theorem 3.7** (Klee-Walkup [42]) *The Hirsch*, *non-revisiting*, *and d-step Conjectures 1.1*, *3.3*, *and 3.6 are equivalent.*

3.3 The Klee-Walkup Polytope Q_4

In their seminal 1967 paper [42] on the Hirsch conjecture and related issues, Klee and Walkup describe a 4-polytope Q_4 with nine facets and diameter five. Innocent as this might look, this first "non-trivial" Hirsch-sharp polytope is at the basis of the construction of every remaining Hirsch-sharp polytope known to date (see Sect. 3.4.2). It is also instrumental in disproving the unbounded and monotone variants of the Hirsch conjecture, which we will discuss in Sect. 3.5. Moreover, its existence is something of an accident: Altshuler, Bokowski and Steinberg [2] list all combinatorial types of simplicial spheres with nine vertices (there are 1296, 1142 of them polytopal); among them, the polar of Q_4 is the only one that is Hirsch-sharp.

Here we describe Q_4 in the polar view. That is, we will describe a simplicial 4-polytope Q_4^* with nine vertices and show that its ridge-diameter is five. The vertices of Q_4^* are:

$$\begin{aligned} a &:= (-3, 3, 1, 2), & e &:= (3, 3, -1, 2),\\ b &:= (3, -3, 1, 2), & f &:= (-3, -3, -1, 2),\\ c &:= (2, -1, 1, 3), & g &:= (-1, -2, -1, 3),\\ d &:= (-2, 1, 1, 3), & h &:= (1, 2, -1, 3),\end{aligned}$$
$$w := (0, 0, 0, -2).$$

[The simple polytope Q_4 is obtained converting each vertex v of Q_4^* into an inequality $v \cdot \mathbf{x} \le 1$. For example, the inequality corresponding to vertex a above is $-3x_1 + 3x_2 + x_3 + 2x_4 \le 1$.]

The key property of this polytope is that:

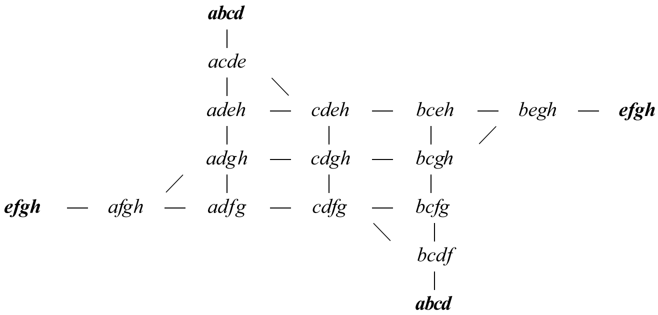


Fig. 4 The dual graph of the subcomplex K

Theorem 3.8 (Klee-Walkup [42]) *Any path in Q_4^* from the tetrahedron abcd to the tetrahedron $efgh$ needs at least five steps.*

To prove this, you may simply input these coordinates into any software able to compute the (dual) graph of a polytope. Our suggestion for this is `polymake` [27]. But we believe that fully understanding this polytope can be the key to the construction of counter-examples to the Hirsch conjecture, so it is worth presenting a hybrid computer-human proof. It is worth mentioning that the coordinates we use for Q_4^*, much smaller than the original ones in [42], were obtained as a by-product of the description of Q_4^* contained in this proof.

Proof Paths through some intermediate tetrahedron containing the vertex w necessarily have at least five steps: apart of the step that introduces w, four more are needed to introduce, one by one, the four vertices e, f, g and h.

This means we can concentrate on the subcomplex K of ∂Q_4^* consisting of tetrahedra that do not use w. This subcomplex is called the *anti-star* of w in ∂Q_4^*. We claim (without proof, here is where you need your computer) that this subcomplex consists of the 15 tetrahedra in Fig. 4.

Figure 4 shows adjacencies among tetrahedra; that is, it shows the dual graph of K. The two tetrahedra $abcd$ and $efgh$ that we want to join are in boldface and appear repeated in the figure, to better reflect symmetry. The proof finishes by noticing that there is no intermediate tetrahedron that can be reached in two steps from both $abcd$ and $efgh$. Hence, five steps are needed to go from one to the other. $\square$

3.4 Many Hirsch-Sharp Polytopes?

Recall that we call a d-polytope (or polyhedron) with n facets *Hirsch-sharp* if its diameter is exactly $n - d$, as happens with the Klee-Walkup polytope of the previous section. Here we describe several ways to construct them.

3.4.1 Trivial Hirsch-Sharp Polytopes

Constructing Hirsch-sharp d-polytopes with any number of facets not exceeding $2d$ is easy. For this reason we call such Hirsch-sharp polytopes *trivial*:

1. *Product.* If P and Q are Hirsch-sharp, then so is their Cartesian product $P \times Q$. Indeed, the dimension, number of facets, and diameters of $P \times Q$ are the sum of

those of P and Q. For the diameter, if we want to go from vertex (u_1, v_1) to vertex (u_2, v_2) we can do so by going from (u_1, v_1) to (u_2, v_1) along $P \times \{v_1\}$ and then to (u_2, v_2) along $\{u_2\} \times Q$; there is no better way.

In particular, any product of *simplices* of any dimension is Hirsch-sharp. The dimension of $\Delta_{i_1} \times \cdots \times \Delta_{i_k}$, where Δ_i denotes the i-simplex, is $\sum_{j=1}^{k} i_j$, its number of facets is $\sum_{j=1}^{k}(i_j + 1)$, and its diameter is k.

Corollary 3.9 *For every $d < n \leq 2d$ there are simple d-polytopes with n facets and diameter $n - d$.*

Proof Let $k = n - d \leq d$ and let $i_1, \ldots, i_k$ be any partition of d into k positive integers (that is, $i_1 + \cdots + i_k = d$). Let $P = \Delta_{i_1} \times \cdots \times \Delta_{i_k}$. □

2. *Intersection of two affine orthants.* Let $k = n - d \leq d$ and let u be the origin in $\mathbb{R}^d$. Let $v = (1, \ldots, 1, 0, \ldots, 0)$ be the point whose first k coordinates are 1 and whose remaining $d - k$ coordinates are 0.

   Consider the polytope P defined by the following $d + k$ inequalities:

$$x_i \geq 0, \quad \forall i; \qquad \psi_j(\mathbf{x}) \geq 0, \quad j = 1, \ldots, k,$$

   where the ψ_i are affine linear functionals that vanish at v and are positive at u. No matter what choice we make for the ψ_j's, as long as they are sufficiently generic to make P simple, P will have diameter (at least) k; to go from v to u we need to enter the k facets $x_j = 0$, $j = 1, \ldots, k$, and each step gets you into at most one of them. In principle, P may be an unbounded polyhedron; but if one of the ψ_j's is, say, $k - \sum x_i$, then it will be bounded.

Hirsch-sharp unbounded polyhedra with any number of facets are also easy to obtain: setcounterpropositionn9

***Proposition 3.10** *For every $n \geq d$ there are simple d-polyhedra with n facets and diameter $n - d$.*

3.4.2 Non-trivial Hirsch-Sharp Polytopes

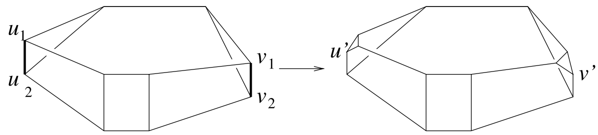
***Theorem 3.11** (Fritzsche-Holt-Klee [26, 31, 32]) *Hirsch-sharp d-polytopes with n facets exist in at least the following cases*: (1) $n \leq 3d - 3$; *and* (2) $d \geq 7$.

Both parts are proved using the Klee-Walkup polytope Q_4 as a starting block, from which more complicated polytopes are obtained. In a sense, *Q_4 is the only non-trivial Hirsch-sharp polytope we know.*

The case $n \leq 3d - 3$ was first proved in 1998 [32], and follows from the iterated application of the next lemma to the Klee-Walkup polytope Q_4. Part 2 was proved in [26] for $d \geq 8$ and was improved to $d \geq 7$ in [31]. We sketch its proof in [38].

Lemma 3.12 (Holt-Klee [32]) *If there are Hirsch-sharp d-polytopes with $n > 2d$ facets, then there are also Hirsch-sharp $(d + 1)$-polytopes with $n + 1$, $n + 2$, and $n + 3$ facets.*

Fig. 5 After wedging in a Hirsch-sharp polytope, we can truncate twice



Proof Let u and v be vertices at distance $n-d$ in a simple d-polytope with n-facets. Let F be a facet not containing any of them, which exists since $n > 2d$. When we wedge on F we get two edges u_1u_2 and v_1v_2 with the properties that the distance from any u_i to any v_i is again (at least) d. We can then truncate one or both of u_1 and v_1 to obtain one or two more facets in a polytope that is still Hirsch-sharp. See Fig. 5. □

Hirsch-sharp polytopes of dimensions two and three exist only when $n \leq 2d$ (see Sect. 2.1). Existence of Hirsch-sharp polytopes with many facets in dimensions four to six remains open. We do know that they do not exist in dimension four with 10, 11 or 12 facets (see Theorem 2.2), which may well indicate that Q_4 is the only Hirsch-sharp 4-polytope.

3.5 The Unbounded and Monotone Hirsch Conjectures

In the Hirsch conjecture as we have stated it, we only consider bounded polytopes. However, in the context of linear programming the feasible region may well not be bounded, so the conjecture is equally relevant for *unbounded* polyhedra. In fact, that is how W. Hirsch originally posed the question.

Moreover, for the simplex method in linear programming one follows *monotone paths*: starting at an initial vertex u one does pivot steps (that is, one crosses edges) always increasing the value of the linear functional ϕ to be maximized, until one arrives at a vertex v where no pivot step gives a greater value to ϕ. Convexity then implies that v is the global maximum for ϕ in the feasible region. This raises the question whether a *monotone* variant of the Hirsch conjecture holds: given two vertices u and v of a polyhedron P and a linear functional that attains its maximum on P at v, is there a ϕ-monotone path of edges from u to v whose length is at most $n-d$?

Both the unbounded and monotone variants of the Hirsch conjecture fail, and both proofs use the Klee-Walkup Hirsch-sharp polytope Q_4 described in Sect. 3.3. In fact, knowing the mere existence of such a polytope is enough. The proofs do not use any property of Q_4 other than the fact that it is Hirsch-sharp, simple, and has $n > 2d$. Simplicity is not a real restriction since it can always be obtained without decreasing the diameter (Lemma 1.3). The inequality $n > 2d$, however, is essential.

Theorem 3.14 (Klee-Walkup [42]) *There is a simple unbounded polyhedron $\tilde{Q}_4$ with eight facets and dimension four and whose graph has diameter* 5.

Proof Let Q_4 be the simple Klee-Walkup polytope with nine facets, and let u and v be vertices of Q_4 at distance five from one another. By simplicity, u and v lie in (at most) eight facets in total and there is (at least) one facet F not containing u nor v. Let $\tilde{Q}_4$ be the unbounded polyhedron obtained by a projective transformation

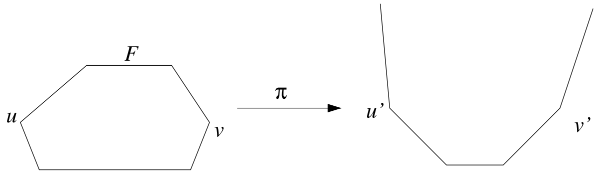


Fig. 6 Disproving the unbounded Hirsch conjecture

that sends this ninth facet to infinity. The graph of $\tilde{Q}_4$ contains both u and v, and is a subgraph of that of $\tilde{Q}_4$, hence its diameter is still at least five. See Fig. 6 for a schematic rendition of this idea. □

Remark 3.15 The "converse" of the above proof also works: from any non-Hirsch unbounded 4-polyhedron $\tilde{Q}$ with eight facets, one can build a bounded 4-polytope with nine facets and diameter five, as follows:

Let u and v be vertices of $\tilde{Q}$ at distance five from one another. Construct the polytope Q by cutting $\tilde{Q}$ with a hyperplane that leaves all the vertices of $\tilde{Q}$ on the same side. This adds a new facet and changes the graph, by adding new vertices and edges on that facet. But u and v will still be at distance five: to go from u to v either we do not use the new facet F that we created (that is, we stay in the graph of $\tilde{Q}_4$) or we use a pivot to enter the facet F and at least another four to enter the four facets containing v.

We now turn to the monotone variant of the Hirsch conjecture:

***Theorem 3.16** (Todd [56]) *There is a simple* 4*-polytope P with eight facets, two vertices u and v of it, and a linear functional ϕ such that*:

1. *v is the only maximal vertex for ϕ.*
2. *Any edge-path from u to v and monotone with respect to ϕ has length at least five.*

In both the constructions of Theorems 3.14 and 3.16 one can glue several copies of the initial block Q_4 to one another, increasing the number of facets by four and the diameter by five, per Q_4 glued:

Theorem 3.17 (Klee-Walkup, Todd)

1. *There are unbounded* 4*-polyhedra with $4+4k$ facets and diameter $5k$, for every $k \geq 1$.*
2. *There are bounded* 4*-polyhedra with $4+4k$ facets and vertices u and v of them with the property that any monotone path from u to v with respect to a certain linear functional ϕ maximized at v has length at least $5k$.*

This leaves the following open questions:

- Improve these constructions so as to get the ratio "diameter versus facets" bigger than 5/4. A ratio bigger than two for the unbounded case would probably yield counter-examples to the bounded Hirsch conjecture.

- Ziegler [59, p. 87] poses the following *strict* monotone Hirsch conjecture: "for every linear functional ϕ on a d-polytope with n facets there is a ϕ-monotone path of length at most $n - d$ from the minimal to the maximal vertex". Put differently, in the monotone Hirsch conjecture we add the requirement that not only v but also u has a supporting hyperplane where ϕ is constant.

3.6 The Topological Hirsch Conjecture is False

Another natural variant of the Hirsch conjecture is topological. Since (the boundary of) every simplicial d-polytope is a topological triangulation of the $(d-1)$-dimensional sphere, we can ask whether the simplicial version of the Hirsch conjecture, the one where we walk from simplex to simplex rather than from vertex to vertex, holds for arbitrary triangulations of spheres. The first counterexample to this statement was found by Walkup in 1979 (see [58]), and a simpler one was soon constructed by Walkup and Mani [45].

Both constructions are based on the equivalence of the Hirsch conjecture to the non-revisiting conjecture (Theorem 3.7). The proof of the equivalence is purely combinatorial, so it holds true for topological spheres. Walkup's initial example is a 4-sphere without the non-revisiting property, and Mani and Walkup's is a 3-sphere:

****Theorem 3.18** (Mani-Walkup [45]) *There is a triangulated* 3*-sphere with* 16 *vertices and without the non-revisiting property. Wedging on it eight times yields an* 11*-sphere with* 24 *vertices and with ridge-graph diameter at least* 13.

This triangulated 3-sphere would give a counterexample to the Hirsch conjecture if it were *polytopal*. That is, if it were combinatorially isomorphic to the boundary complex of a four-dimensional simplicial polytope. Altshuler [1] has shown that (for the explicit completion of the subcomplex K given in [45]) this is not the case. As far as we know, it remains an open question to show that *no completion of* $K \cup \{abcd, mnop\}$ to the 3-sphere is polytopal, but we believe that to be the case. Even more strongly, we believe that $K \cup \{abcd, mnop\}$ cannot be embedded in $\mathbb{R}^3$ with linear tetrahedra, a necessary condition for polytopality by the well-known Schlegel construction [59].

As in the monotone and bounded cases, several copies of the construction can be glued to one another. Doing so provides triangulations of the 11-sphere with $12 + 12k$ vertices and diameter at least $13k$, for any k.

Note Added in Proof

On May 10, 2010, while this paper was already in press, the second author has announced a counter-example to the Hirsch Conjecture: there is a 43-dimensional polytope with 86 facets and whose graph has diameter at least 44. See http://gilkalai.wordpress.com/2010/05/10/ for the announcement. By the time this paper is published a preprint will most probably be available at http://www.arxiv.org.

Acknowledgements This paper grew out of several conversations during the second author's sabbatical leave at UC Davis in 2008 and the authors' participation in the I-Math DocCourse on Discrete and Computational Geometry at the *Centre de Recerca Matemàtica* in 2009. We thank both institutions for hosting us and the financial support from the National Science Foundation and the Spanish Ministry of Science.

We also thank David Bremner, Jesús De Loera, Antoine Deza, Jörg Rambau, Günter M. Ziegler and the anonymous referee for their valuable comments on the first version of the paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Altshuler, A.: The Mani-Walkup spherical counterexamples to the W_v-path conjecture are not polytopal. Math. Oper. Res. **10**(1), 158–159 (1985)
2. Altshuler, A., Bokowski, J., Steinberg, L.: The classification of simplicial 3-spheres with nine vertices into polytopes and non-polytopes. Discrete Math. **31**, 115–124 (1980)
3. Balinski, M.L.: The Hirsch conjecture for dual transportation polyhedra. Math. Oper. Res. **9**(4), 629–633 (1984)
4. Barnette, D.: W_v paths on 3-polytopes. J. Comb. Theory **7**, 62–70 (1969)
5. Beck, M., Robins, S.: Computing the Continuous Discretely: Integer-point Enumeration in Polyhedra. Undergraduate Texts in Mathematics. Springer, Berlin (2007)
6. Björner, A., Brenti, F.: Combinatorics of Coxeter Groups. Graduate Texts in Mathematics, vol. 231. Springer, Berlin (2005)
7. Blum, L., Cucker, F., Shub, M., Smale, S.: Complexity and Real Computation. Springer, Berlin (1997)
8. Borgwardt, K.H.: The average number of steps required by the simplex method is polynomial. Z. Oper. Res. **26**, 157–177 (1982)
9. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
10. Bremner, D., Deza, A., Hua, W., Schewe, L.: More bounds on the diameters of convex polytopes. Preprint, 8 p. (2009). arXiv:0911.4982v1
11. Bremner, D., Schewe, L.: Edge-graph diameter bounds for convex polytopes with few facets. Preprint, 9 p. (2008). arXiv:0809.0915v3
12. Brightwell, G., van den Heuvel, J., Stougie, L.: A linear bound on the diameter of the transportation polytope. Combinatorica **26**(2), 133–139 (2006)
13. Cunningham, W.H.: Theoretical properties of the network simplex method. Math. Oper. Res. **4**, 196–208 (1979)
14. Dantzig, G.B.: Linear Programming and Extensions. Princeton University Press, Princeton (1963)
15. De Loera, J.A., Kim, E.D., Onn, S., Santos, F.: Graphs of transportation polytopes. J. Comb. Theory Ser. A **116**(8), 1306–1325 (2009)
16. De Loera, J.A.: The many aspects of counting lattice points in polytopes. Math. Semesterber. **52**(2), 175–195 (2005)
17. De Loera, J.A., Onn, S.: All rational polytopes are transportation polytopes and all polytopal integer sets are contingency tables. In: Proc. 10th Ann. Math. Prog. Soc. Symp. Integ. Prog. Combin. Optim. (Columbia University, New York, NY, June 2004). Lec. Not. Comp. Sci., vol. 3064, pp. 338–351. Springer, New York (2004)
18. De Loera, J.A., Rambau, J., Santos, F.: Triangulations: Structures for Algorithms and Applications. Algorithms Comput. Math. vol. 25 (to appear), Springer-Verlag
19. Dedieu, J.-P., Malajovich, G., Shub, M.: On the curvature of the central path of linear programming theory. Found. Comput. Math. **5**, 145–171 (2005)
20. Deza, A., Terlaky, T., Zinchenko, Y.: Central path curvature and iteration-complexity for redundant Klee-Minty cubes. Adv. Mech. Math. **17**, 223–256 (2009)
21. Deza, A., Terlaky, T., Zinchenko, Y.: A continuous d-step conjecture for polytopes. Discrete Comput. Geom. **41**, 318–327 (2009)
22. Deza, A., Terlaky, T., Zinchenko, Y.: Polytopes and arrangements: diameter and curvature. Oper. Res. Lett. **36**(2), 215–222 (2008)

23. Dyer, M., Frieze, A.: Random walks, totally unimodular matrices, and a randomised dual simplex algorithm. Math. Program. **64**, 1–16 (1994)
24. Eisenbrand, F., Hähnle, N., Razborov, A., Rothvoß, T.: Diameter of polyhedra: limits of abstraction. Preprint, available at http://people.cs.uchicago.edu/~razborov/research.html. A preliminary version appeared in: Proceedings of the 25th Annual ACM Symposium on Computational Geometry (SoCG'09), ACM, New York (2009)
25. Fomin, S., Zelevinsky, A.: Y-systems and generalized associahedra. Ann. Math. **158**(2), 977–1018 (2003)
26. Fritzsche, K., Holt, F.B.: More polytopes meeting the conjectured Hirsch bound. Discrete Math. **205**, 77–84 (1999)
27. Gawrilow, E., Joswig, M.: Polymake: a framework for analyzing convex polytopes. In: Polytopes—combinatorics and computation (Oberwolfach, 1997). DMV Sem., vol. 29, pp. 43–73. Birkhäuser, Basel (2000)
28. Goldfarb, D., Hao, J.: Polynomial simplex algorithms for the minimum cost network flow problem. Algorithmica **8**, 145–160 (1992)
29. Goodey, P.R.: Some upper bounds for the diameters of convex polytopes. Isr. J. Math. **11**, 380–385 (1972)
30. Hačijan, L.G.: A polynomial algorithm in linear programming. Dokl. Akad. Nauk SSSR **244**(5), 1093–1096 (1979) (in Russian)
31. Holt, F.B.: Blending simple polytopes at faces. Discrete Math. **285**, 141–150 (2004)
32. Holt, F., Klee, V.: Many polytopes meeting the conjectured Hirsch bound. Discrete Comput. Geom. **20**, 1–17 (1998)
33. Hurkens, C.: Personal communication (2007)
34. Kalai, G.: A subexponential randomized simplex algorithm. In: Proceedings of the 24th Annual ACM Symposium on the Theory of Computing, pp. 475–482. ACM, New York (1992)
35. Kalai, G.: Online blog http://gilkalai.wordpress.com. See, for example, http://gilkalai.wordpress.com/2008/12/01/a-diameter-problem-7/, December 2008
36. Kalai, G., Kleitman, D.J.: A quasi-polynomial bound for the diameter of graphs of polyhedra. Bull. Am. Math. Soc. **26**, 315–316 (1992)
37. Karmarkar, N.: A new polynomial time algorithm for linear programming. Combinatorica **4**(4), 373–395 (1984)
38. Kim, E.D., Santos, F.: Companion to "An update on the Hirsch conjecture". Preprint, 22 p. (2009). arXiv:0912.4235v1
39. Klee, V.: Paths on polyhedra II. Pac. J. Math. **17**(2), 249–262 (1966)
40. Klee, V., Kleinschmidt, P.: The d-step conjecture and its relatives. Math. Oper. Res. **12**(4), 718–755 (1987)
41. Klee, V., Minty, G.J.: How good is the simplex algorithm? In: Inequalities, III: Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin, pp. 159–175. Academic Press, New York (1972)
42. Klee, V., Walkup, D.W.: The d-step conjecture for polyhedra of dimension $d < 6$. Acta Math. **133**, 53–78 (1967)
43. Kleinschmidt, P., Onn, S.: On the diameter of convex polytopes. Discrete Math. **102**(1), 75–77 (1992)
44. Larman, D.G.: Paths of polytopes. Proc. Lond. Math. Soc. **20**(3), 161–178 (1970)
45. Mani, P., Walkup, D.W.: A 3-sphere counterexample to the W_v-path conjecture. Math. Oper. Res. **5**(4), 595–598 (1980)
46. Matoušek, J., Sharir, M., Welzl, E.: A subexponential bound for linear programming. In: Proceedings of the 8th Annual Symposium on Computational Geometry, pp. 1–8 (1992)
47. Megiddo, N.: Linear programming in linear time when the dimension is fixed. J. Assoc. Comput. Mach. **31**(1), 114–127 (1984)
48. Megiddo, N.: On the complexity of linear programming. In: Bewley, T. (ed.) Advances in Economic Theory: Fifth World Congress, pp. 225–268. Cambridge University Press, Cambridge (1987)
49. Naddef, D.: The Hirsch conjecture is true for (0, 1)-polytopes. Math. Program. **45**, 109–110 (1989)
50. Oda, T.: Convex Bodies and Algebraic Geometry. Springer, Berlin (1988)
51. Orlin, J.B.: A polynomial time primal network simplex algorithm for minimum cost flows. Math. Program. **78**, 109–129 (1997)
52. Renegar, J.: A Mathematical View of Interior-Point Methods in Convex Optimization. SIAM, Philadelphia (2001)
53. Smale, S.: On the average number of steps of the simplex method of linear programming. Math. Program. **27**, 241–262 (1983)

54. Smale, S.: Mathematical problems for the next century. In: Mathematics: Frontiers and Perspectives, pp. 271–294. American Mathematics Society, Providence (2000)
55. Spielman, D.A., Teng, S.: Smoothed analysis of algorithms: why the simplex algorithm usually takes polynomial time. J. ACM **51**(3), 385–463 (2004)
56. Todd, M.J.: The monotonic bounded Hirsch conjecture is false for dimension at least 4. Math. Oper. Res. **5**(4), 599–601 (1980)
57. Vershynin, R.: Beyond Hirsch conjecture: walks on random polytopes and smoothed complexity of the simplex method. In: IEEE Symposium on Foundations of Computer Science. NN, vol. 47, pp. 133–142. IEEE, New York (2006)
58. Walkup, D.W.: The Hirsch conjecture fails for triangulated 27-spheres. Math. Oper. Res. **3**, 224–230 (1978)
59. Ziegler, G.M.: Lectures on Polytopes. Graduate Texts in Mathematics, vol. 152. Springer, Berlin (1995)
60. Ziegler, G.M.: Face numbers of 4-polytopes and 3-spheres. In: Proceedings of the International Congress of Mathematicians, vol. III, Beijing, 2002, pp. 625–634. Higher Ed. Press, Beijing (2002)

Edward D. Kim is finishing his Ph.D. studies at the University of California, Davis under the supervision of Jesús A. De Loera. His research is in combinatorial optimization, convex polytopes, Lagrangian matroids, and multi-objective optimization. He is about to start his postdoctoral career with positions at Technische Univ. Delft in the Netherlands and at POSTECH (Pohang Univ. of Science and Technology) in Korea.

Francisco Santos is full professor of mathematics at the Universidad de Cantabria in Spain. He received the Young Researcher award from the Universidad Complutense de Madrid in 2003 and was an invited speaker in the Combinatorics Section of the International Congress of Mathematicians in 2006. He is a specialist in Geometric Combinatorics, especially in the theory of polytopes and triangulations. He is an editor of the journal “Discrete and Computational Geometry”.

Jahresber Dtsch Math-Ver (2010) 112: 99–116
DOI 10.1365/s13291-010-0002-7

Some Direct and Remote Relations of Gauss with Belgian Mathematicians

Jean Mawhin

Received: 30 December 2009 / Published online: 27 May 2010
© Vieweg+Teubner und Deutsche Mathematiker-Vereinigung 2010

Abstract We describe the personal and epistolar relations between Carl-Friedrich Gauss and the Belgian astronomer and statistician Adolphe Quetelet, and Gauss' influence upon the work of the Belgian mathematician Charles-Jean de La Vallée Poussin on the prime number theorem and in potential theory.

Keywords Gauss · Quetelet · De La Vallée Poussin · Terrestrial magnetism · Gauss curve · Prime number theorem · Potential theory

Mathematics Subject Classification (2000) 01A55 · 01A60 · 01A74

1 Introduction

The reputation of Carl-Friedrichs Gauss (1777–1855) as mathematician, astronomer and physicist was already established during his lifetime, and his legacy is immense. Although he was somewhat reluctant to travel, he remained in contact with many scientists in Germany and in Europe through his correspondence, and through the many visits he received in Göttingen.

However, with the exception of a recent interesting paper of Gert Schubring [46] devoted to the correspondence between Gauss and Adolphe Quetelet (1796–1874), very little has been written about either the direct relations of Gauss with Belgian mathematicians in the first part of the XIXth century, or about its direct influence upon their work or the one of their followers. For example, Adolphe Quetelet, who, as we will see, visited Gauss and shared several of his scientific interests, is only

An expanded version of the historical lecture preceding the DMV-Gauss-Vorlesung in Otto-von-Guericke-Universität, Magdeburg, April 20th, 2009.

J. Mawhin (✉)
Département de Mathématique, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium
e-mail: jean.mawhin@uclouvain.be

mentioned three time in Dunnington's substantial biography of Gauss [9]. The first citation, on p. 20, is part of a general statement on the influence of Gauss's theory of least squares:

> The writings of [. . .] Liagre and Quetelet in Belgium, [. . .] have brought the method [of Gauss] to a high degree of perfection in all its branches and have caused it to be universally adopted by scientific men as the only proper method for the discussion of observations.

The next citation, on p. 154, quotes Gauss' letter from October 12, 1829 to Olbers about Quetelet's visit in Göttingen:

> The acquaintance of M. Quetelet has been very pleasant to me; in my yard we put on with his splendid apparatus various series of experiments on the intensity of magnetic force which granted an agreement scarcely expected by me.

The last citation, on p. 234, mentions that Gauss found in Quetelet's *Annuaire* the ratio of widows to existing weddings as one to four. In the more recent biographies of Gauss by Hall [23], Reich [43] and Bühler [3], the name of Quetelet is not even mentioned. The first part of this paper is devoted to a more careful study of the relations between Gauss and the most famous representative of science in Belgium during the XIXth century.

A remote but important connection between Gauss and another Belgian mathematician is more widely mentioned: the proof of Gauss' conjecture on the asymptotic distribution of prime numbers by Charles-Jean de La Vallée Poussin (1866–1962). As we will see, the name of Gauss surprisingly does not occur in de La Vallée Poussin's memoir of 1896, but only three years later in a work giving a formula for the error between the number of prime numbers smaller than a given integer and its asymptotic value. Maybe less known is the fact that Gauss' contribution to potential theory was very influential in some subsequent work of de La Vallée Poussin on the sweeping out method. In a lecture given in 1939 in front of the Belgian Mathematical Society, to celebrate the centenary of Gauss' famous memoir on potential theory, de La Vallée Poussin gave a deep and moving analysis of this paper, only published after his death in 1962 [6]. We devote the second part of this paper to an analysis of Gauss' inspiration on the work of this exceptional figure of Belgian mathematics.

2 An Astronomer in Quest of Observatories

The Belgian scientist Adolphe Quetelet, born in 1796 in Gent, had been in 1819 the first PhD in mathematics and physics of the University of Gent, an institution created in 1817 by King Wilhelm I, who reigned on the reunified Low Countries since Vienna's congress. After some work in geometry and optics, which opened him the doors of the Royal Academy of Brussels (renamed in 1845 as Royal Academy of Science, Letters and Arts of Belgium), Quetelet turned his interests to astronomy, meteorology, geophysics and to the application of statistics to physical and human sciences. He organized and chaired the first International Conference on Statistics in 1853 in Brussels, and is well remembered for his concept of 'average man' [26].

Fig. 1 Adolphe Quetelet (Académie Royale de Belgique)

The young Academician, maybe still better as an entrepreneur than as a scientist, decided to convince the authorities to create an astronomical observatory in Brussels. Sent to Paris in 1823 to visit the observatory and talk with French astronomers, he described his meeting with Laplace, Poisson and Fourier as a decisive step for the scientific orientation of his career. The political decision made in 1826 to create the Royal Observatory of Brussels was followed by another trip of Quetelet to Great Britain in 1826, to order instruments. Appointed astronomer at the (still non existing) Observatory in 1828, Quetelet took some distance from the vicissitudes of the construction of the observatory by visiting in 1829 its existing *alter ego* in Germany. A wise decision, if one remembers that the Brussels observatory was only ready in 1834!

In July 1829, Quetelet visited Altona, meeting Schumacher and Repsold, who joined him on a trip to Bremen to visit Olbers, seemingly more impressed by Mrs Quetelet than by the scientific triumvirate. In Berlin, Quetelet met Encke, Poggendorf, Crelle, Mitscherlich, whilst Mrs Quetelet, herself a good musician, became a friend of Fanny Henselt, Mendelssohn sister's. After visiting Dresden and Leipzig, Quetelet arrived in Weimar at the end of August, not to see an observatory but to meet a star: Goethe was celebrating his eightieth birthday. After Weimar, Quetelet visited Hansen in Gotha and arrived in Göttingen [36] (Fig. 1).

3 'Dr. Gauss, I Presume ?'

On some afternoon of fall 1829, Adolphe Quetelet was walking in the direction of Göttingen's Observatory, whose new building, finished in 1816, had been promised to Carl-Friedrich Gauss at the time of his appointment as Director, in 1807, to replace the old uncomfortable one.

Of course Quetelet was well aware of Gauss' scientific reputation. He just had enjoyed Schumacher's gossips about the great mathematician, and had admired his portrait in Altona. He therefore had no merit in recognizing Gauss in the person leaving the observatory, and politely asked him, in French, something like: 'Dr. Gauss

I presume'. The Herr Director answered in German that he did not understand the question, forcing Quetelet to pursue in his best German. Informed of Quetelet's visit through a letter of Schumacher, Gauss realized his mistake. He explained, in an excellent French, that he always proceeded in this uncivil way with people not recommended to him by friends. He then returned with Quetelet to the observatory and guided him through all rooms and facilities [37]. The story told by Quetelet in his own terms goes as follows [41]:

> My mind fully absorbed with the recent conversations with his friend, M. Schumacher, as well as with the peculiarities told about him, as well as with his portrait full of strength and energy that I had observed during my stay in Altona, I was walking, pensive, toward the observatory. Then, looking up a person passing near me, I thought seeing brought to life the portrait I had observed with so much attention. The impression was such that I could not refrain stopping the passer and asking him if I had the honor to speak to Mr. Gauss. This sudden address stopped the movement of the summoned person, who answered me in German that he did not understand me. Following my thoughts, I hasted to answer, in the same language, which did not allow me to express myself clearly, in which way my request was inconvenient. After I had expressed myself in the requested German language, the illustrious geometer understood immediately the incident: he answered me this time, in an excellent French, that M. Schumacher had kindly informed him of my visit and that, if he had spoken German, it was to follow a form he did not use with the persons who were affectionately recommended by his best friends. Despite of my protests, Gauss accepted to go back and to escort me to his home.

Since 1828, Quetelet was interested in measuring the terrestrial magnetism. The corresponding magnetic field is characterized, at one point, by its angle with the meridian at this point (magnetic declination), its angle with the horizontal at this point (magnetic inclination) and its intensity. Special instruments using a magnetized iron bar were used to measure the two angles, and the intensity could be deduced from the period of oscillations of the iron bar around the equilibrium. Quetelet had carried with him some small apparatus to measure the terrestrial magnetic field. So the two men met again the next day in the garden of the Observatory, to measure the intensity of the magnetic field, a new activity for Gauss at this time. In Quetelet's terms [41]:

> The illustrious astronomer accepted to show me the interest he took to those researches [on terrestrial magnetism] and expressed the desire to make, the next day, some observations with me. Those observations on the intensity of the magnetic field of the Earth, he said, were new for him, and he wanted to know how they were made and the precision one could hope to reach.

Both men made independent measurements, Quetelet preferring to measure the period of oscillation from the position with zero velocity, and Gauss from the position with maximal velocity. Gauss was quite surprised by the agreement and precision of their respective measures. Quetelet was candid enough to think that this garden party, so to speak, was determinant in Gauss' subsequent interest and work in geomagnetism [41]:

> Maybe was I wrong to conjecture since, that this precision could have influenced the ideas of this able observator, but it is three or four years later that his beautiful memoirs and joint work on geomagnetism have begun to be published, works which, together with those of Hansteen, have put so much light on this part of physics and have so much accelerated the progress of this science.

Although those joint observations with Quetelet may have played a role in exciting Gauss' interest for terrestrial magnetism, as confirmed by their subsequent correspondence, one must however remember that, as early as 1803, Gauss had consulted Olbers about the connections between magnetism and navigation, and in 1806, had asked Harding to provide him declination and inclination data of the magnetic field of the Earth. Weber's arrival in Göttingen in 1831 was surely more stimulating to revive Gauss' interest than Quetelet's visit. Gauss and Weber constructed a magnetic observatory in Göttingen in 1833, which became the coordination center for European data, and created a Magnetic Society. In 1837, the two scientists invented the two wires magnetometer to measure the intensity of the terrestrial magnetic field. It was Gauss also who, in 1840, initiated the study of its periodic variations by prescribing, in several German cities, Leyden and Brussels, the following scheme of observations: to measure each month, during twenty-four hours and every five minutes, the magnetic declination and inclination. Those observations showed that the same perturbations occurred simultaneously in the different cities. One year later, Humboldt increased the network to other cities of Europe, United States and British empire. This practical interest of Gauss for magnetism, doubled by a theoretical one to which we come back later, was put in concrete form through the publication of several treatises and of a specialized journal, the *Resultate aus den Beobachtungen des magnetischen Vereins*.

During the visit, Gauss described also to Quetelet his current work on capillarity and elliptic functions, not surpassed in his opinion by the recent work of Abel and Jacobi [36]:

> Mr. Gauss also possesses an unpublished work on analysis and on elliptic functions in particular, which contains as special cases the beautiful researches of MM. Abel and Jacobi. I had heard that Mr. Gauss has been preceded by those two young scientists in the publication of the discoveries we just spoke about; but it seems that the geometer of Göttingen still has some material to console himself from this contrariety.

Gauss offered to Quetelet a copy of his thesis on the fundamental theorem of algebra [14] and of his *Disquisitiones generales circa superficies curvas* [17] (Fig. 2).

4 An Asymmetric Exchange of Letters

Quetelet's encounter with Gauss was followed, between 1830 and 1842, by a regular exchange of letters, recently published by Gert Schubring [46]. Speaking of exchange is somewhat exaggerated, with thirteen letters from Quetelet against three letters only from Gauss. The occasion of the first two letters (April 3, 1830; June 26, 1832) was

Fig. 2 Carl-Friedrich Gauss (Braunschweig)

the sending of some publications of Quetelet, in particular the one including their joint measurements in Gauss' garden.

Appointed permanent Secretary of the Royal Academy of Brussels in 1834 (he justified the appellation by keeping the office for 40 years), Quetelet asked, in the third letter (February 22, 1837), for an exchange of publications between his Academy and the Royal Society of Göttingen. He also requested some memoirs of Gauss on magnetism, difficult to find in Brussels, and some information about Gauss-Weber's two wires magnetometer, in order to have one built in Berlin. Gauss answered this letter on April 12, 1837, by directing Quetelet to the permanent secretary of Göttingen's Royal Society Blumenbach,

> unfortunately subject to repeated losses of memory…

for the exchanges, to some library in Aachen for getting the memoirs, and to a paper in press for the description of the magnetometer [46]:

> I am surprised that you have some difficulty in finding German books in Brussels, residence of a king from German origin. If you find there no bookseller in relation with the ones from Germany, I would advise you to contact a neighboring German bookseller. Surely the one's of Aachen will be able to provide you very rapidly what you want.

He advised Quetelet to have the magnetometer constructed in Göttingen [46]:

> I doubt that Mr. von Humboldt will take care of constructing for you a magnetometer in Berlin. There exists in Berlin only a very small apparatus, that Mr. Encke had ordered in 1834, after having seen mine. [...] It is Mr. Meirstrein in Göttingen, who has constructed the greatest part of the existing magnetometers up to now.

Quetelet immediately acknowledged this somewhat unhelpful answer on May 7, 1837, and, on January 8, 1838, mentioned the reception of the magnetometer constructed in Göttingen. Having heard about the troubles at the University of Göttingen due to the restoration of the old conservative constitution, and in particular about the sacking of Weber and of Gauss' son in law Ewald from the University, Quetelet did not hesitate to offer to Gauss the hospitality of his Observatory in Brussels [46]:

> I have read with an infinite sorrow in the newspapers that you do not have, in this moment, all the needed peace for your important achievements. I have seen that several professors of Göttingen were forced to quit their home; if such a misfortune should hit one of the greatest illustration of Germany and of modern time, I dare to beseech you, Sir, not to forget that the Observatory of Brussels is entirely at your disposal. Here you will find a family which will be very happy and very honored to welcome you, and take care in providing you the respect due to your good character and your eminent talents.

Gauss did not react, neither to this generous offer, nor to five subsequent letters of Quetelet, from May 4, 1839, May 9, 1840, September 18, 1840, December 25, 1840, and July 17, 1841, joining the sending of some memoirs of Plateau and of various geomagnetic measurements, and providing some information about Quetelet's magnetic laboratory in Brussels.

But, on September 9, 1841, Gauss' second letter provided explanations and critics to some magnetic irregularities described in a letter of Quetelet to Hausmann, together with some priority claim [46]:

> In a letter that you have sent to our Secretary Mr. Hausmann, that he was kind enough to show to me, you speak about some anomalies in the amplitudes of the oscillations, mentioned earlier by myself: but you do not seem to have seen directly the corresponding excerpt. If I dare to make a conjecture, they correspond to the phenomena that I described widely in the volume 2 of the *Resultate*, p. 70, and recommended to the attention of the observers; however, until this day, I do not know any other observations than those of Göttingen.

In his immediate answer of September 27, 1841, Quetelet promised to take benefit from Gauss' remarks and suggestions. His next letter, from December 27, 1841, announced to Gauss his election as a corresponding member of the Royal Academy of Brussels. On February 4, 1842, Gauss thanked Quetelet and the Royal Academy of Brussels for this distinction. The two letters of Gauss are reproduced *in extenso* in Gauss' obituary at the Belgian Academy [41]. On March 9, 1842, Quetelet apologized for the delay in answering Gauss' letter and provided some information about the perturbation in the magnetic field of the Earth due to a polar aurora. This last letter of Quetelet to Gauss marked the end of their epistolary relations. It seems that

Gauss did not even acknowledge, in 1845, his promotion to associate member at the Royal Academy of Brussels.

A few months after Gauss' death, in July 1855, Quetelet suffered a stroke leaving him somewhat active, but his memory was substantially diminished. He remained however permanent Secretary of the Royal Academy of Belgium until his death in 1874, devoting his last years to the publication of new editions of his ancient work. It is somewhat surprising that, in the largest biography of Quetelet, due to Mailly [26], the name of Gauss is mentioned only once, on p. 184:

> After Weimar, Quetelet visited successively Gotha and Göttingen, whose Observatories has respectively for directors Hansen and Gauss.

5 The Bell-Shaped Curve: From Heavens to Earth

Besides geomagnetism, another strong common scientific interest of Gauss and Quetelet was of course statistics and the theory of errors. There is no evidence that they ever discussed those questions when they met (Quetelet surely would have mentioned it in [36]), and there is no trace of it in their correspondence.

In 1809, in his famous *Theoria Motus Corporum Coelestium in sectionibus Conicis Solem Ambientium* [15], Gauss described the method of least squares, already used by him for some fifteen years, and observed that the curve of normal distribution of errors

$$\phi(x) = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2}$$

was the only one to make this method fully coherent. Indeed, after introducing statistical considerations like maximum likehood and the postulate of the arithmetic mean, Gauss derived both the normal distribution and the principle of least squares. In 1823 [16], he put forward an alternative statistical assumption, namely the condition of least variance, to arrive again at the method of least squares. The validity of Gauss' assumptions in linking for the first time the method of least squares with the normal distribution has been widely discussed by historians of statistics [20, 21, 34, 45, 47, 50, 54]. The priority question for the method of least squares between Legendre and Gauss is fully analyzed in [10, 33, 49], and it is established that Gauss had the priority in discovery and Legendre in publication [24].

Quetelet's contacts in Paris with Laplace, Poisson and Fourier had raised a strong interest for statistics, and in particular for its application to natural phenomena and human sciences, a topic he named 'social physics'. There is not much mathematics in the first edition of his book on social physics [39], where Quetelet introduces his well-known concept of *average man*, which has been equally admired and criticized [2]. Gauss is nowhere mentioned in this book first published in 1835 and it does not appear that Quetelet benefited much from Gauss' method of least squares in the theory of errors of observation. In his first book on probability [35] of 1828, when dealing with n observations $x_1, x_2, \ldots, x_n$ having mean $\overline{x} := \frac{x_1+\cdots+x_n}{n}$, Quetelet exactly followed Fourier [11] by asserting, without proof, that the degree of approximation of the mean result was represented by the formula

$$g = \frac{1}{\sqrt{n}}\sqrt{2\left(\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n} - \overline{x}^2\right)}.$$

Quetelet expressed this formula in words and borrowed an unconvincing example from Fourier. Quetelet added mistakenly that this g was calculated according the rule of least squares. In [16], Gauss had introduced the variance as a measure of precision and respectively called quantities inversely proportional to the square root of the variance and to the variance itself the *precision* and the *weight* of the observation.

In 1846, after collecting intensively during twenty years meteorological, geophysical, astronomical and sociological data, Quetelet published his popular *Lettres à S.A.R. le grand-duc régnant de Saxe-Cobourg et Gotha* on probability theory applied to moral and political sciences [40]. The bell shaped curve, or some approximation, appears at several instances in the second part 'On mean and limits' of this book, and first on p. 103 when discussing the result of sequence of extractions from a bag containing a large equal number of white and black balls. Using a table published by the British army around 1840 in volume 13 of the *Edinburgh medical journal*, giving the size of the chest of about 4000 Scottish soldiers, Quetelet showed that its distribution around its mean value followed the same shape. He explained this fact by an accumulation of a large number of particular circumstances (heredity, nutrition, ...) providing deviations around the average size. From this time, the reign of the bell-shaped curve was extended from heaven to Earth. The equation of the normal law is given in the Notes of [40], collecting the mathematical aspects of the *Lettres* and a picture can be found on p. 396 in the context of the theory of errors.

Quetelet was the first one to name this function as the 'law of accidental causes'. Nowadays, although it appeared earlier in some overlooked work of De Moivre (see e.g. [31]), it is usually referred as Gaussian or Gauss' curve (after Bertrand [2] in 1889), or normal law (after Pearson [30] popularized in 1894 some earlier independent suggestions of Peirce, Lexis and Galton). Being the first one to introduce it in fields other than the theory of errors or theoretical probability, Quetelet contributed in an essential way to popularize Gauss' curve, whose use and abuse has shaped the applications of statistics to natural and human sciences [48].

In demography, the data given by Quetelet in the German edition of his famous monograph *Sur l'homme et le développement de ses facultés* [37] of 1835, were used by Gauss to check his empirical formula

$$x = 10{,}000 - A\sqrt[3]{n}$$

giving the number of newly born x who will live to be n months old, where 10,000 is the initial number of children born, and $\log A = 3.98273$. Gauss gave this formula in a letter to Schumacher [32], reproduced on p. 302 of the first volume of [42]:

> I took the liberty, in my letter to the private adviser Collin, to express a few wishes, and in particular to divide children's mortality in the early ages in shorter periods. I was led to express this wish by the remark, made a long time ago, that the table given by Ad. Quetelet (in his *Annuaire* of 1844, page 193,

> and of 1846, page 185) can be represented, for the first six months, through some formula with some almost marvelous precision. I have added in my letter another proposition, that I could somewhat modify, because I do not know exactly on which facts depend the data of the author. After having finished and sealed this letter, I found in the book of Quetelet *Sur l'homme*, page 144 of the German translation by Rieke, data relative to West Flanders, which seem to have been the basis of the data in the *Annuaire*. However, I did not want to reopen and change my letter. Maybe will you be interested in this formula, if I join it here. The last member is represented, for the six first months, by the expression $10000 - A\sqrt[3]{n}$ (where $\log A = 3.98273$ and n represents the number of months); it is given with a degree of exactness not usually found in ordinary mortality tables.

XIXth century was apparently a good time for empirical formulas, and the following one was proposed by Quetelet in 1832 in his *Recherches sur le penchement au crime* [37],

$$z = \frac{1 - \sin x}{1 + 2^{18-x}}$$

to express the inclination to crime z in term of the age x! Notice that in the mentioned monograph *Sur l'homme et le développement de ses facultés* [39], $18 - x$ was replaced by $x - 18$, and that the formula had disappeared in the revised version *Physique sociale* [42] of 1869.

Let us mention finally Quetelet's *obesity index*, defined as the ratio of the weight, in kilos, divided by the square of the height, in meters. In Quetelet's own words [38]:

> If we compare now entirely developed and regularly constructed individuals, to know the relations which can exist between the weight and the height, we will find that *the weights among developed individuals of different heights are about like the squares of the heights.*

An identical statement was repeated in [39].

6 A Belgian Proof for Gauss' Conjecture on Prime Numbers

The son of a professor of geology at the Catholic University of Louvain, Charles-Jean de La Vallée Poussin spent all his life, with the exception of the First World War, near or in this institution. Born in Louvain in 1866, he graduated in mathematics in 1891 and became associate professor the same year. He teached analysis and mechanics there during sixty years, until his voluntary retirement in 1951 at 85. He still enjoyed eleven years as emeritus professor, and died in Brussels in 1962. He is the author of very original work in analytic number theory, integration and measure, Fourier series, interpolation, approximation theory, quasi-analytic functions, conformal mappings and potential theory. His *Cours d'analyse infinitésimale*, with more than ten editions during some sixty years, has been highly influential, like his monographs on set functions and approximation. More biographical details can be found in [7]. De

La Vallée Poussin's first acquaintance with Gauss' work is of course his proof of the prime number conjecture.

Around 1792, the young Gauss found happiness in reading the mathematical tables offered to him by his protector, the Duke of Brunswick. Gauss wrote later, without little chance to be contradicted:

> You have no idea of the poetry contained in a table of logarithms.

Scrutinizing as well his table of prime numbers, he observed the highly nontrivial fact, that around some integer m, the proportion of prime numbers among integers is close to $\frac{1}{\log m}$, when m is sufficiently large. This observation implies that the number $\pi(x)$ of prime numbers smaller or equal to x can be, for x large, approximated by the integral logarithm function Li defined by

$$Li(x) := \lim_{\epsilon \to 0} \left(\int_0^{1-\epsilon} \frac{ds}{\log s} + \int_{1+\epsilon}^{x} \frac{ds}{\log s} \right),$$

in the sense that

$$\lim_{x \to +\infty} \frac{\pi(x)}{Li(x)} = 1. \tag{1}$$

Notice that since $\lim_{x \to +\infty} \frac{Li(x)}{\frac{x}{\log x}} = 1$, Gauss' assertion is equivalent to

$$\lim_{x \to +\infty} \frac{\pi(x)}{\frac{x}{\log x}} = 1.$$

The following Table 1 comparing $\pi(x)$, $Li(x)$ and $\frac{x}{\log x}$ makes Gauss' conjecture quite plausible. As it has been often the case, Gauss did not care publishing his observation, and just wrote it in one of the consulted tables. He mentioned it, some fifty years later, in a letter of 1849 to one of his former students, the astronomer Encke. Meanwhile, Legendre [25], condemned by the Gods of the mathematical Olympia to compete with Gauss, had stated a similar conjecture in 1808.

Despite of the efforts of talented mathematicians like Dirichlet, Chebychev, Riemann or Stieltjes, Gauss' conjecture remained open for about one century. In 1896, two proofs, somewhat similar and based upon Riemann's technique of localizing the non trivial zeros of the zeta function, were given independently by de La Vallée Poussin [4] and Hadamard [22]. It is curious to notice that the name of Gauss is absent from both papers, but the conjecture is stated and proved there in a third equivalent form, namely

$$\lim_{x \to \infty} \frac{\sum_{p<x} \log p}{x} = 1, \tag{2}$$

where the sum is extended to prime numbers p. The name of Gauss appeared for the first time in de La Vallée Poussin's subsequent memoir of 1899 [5], starting as follows:

> I have proved, for the first time [...], that the function $\zeta(s)$ has no roots of the form $1 + \beta i$. Mr. Hadamard, before knowing my research, had also found the

Table 1 Exact and approximate values of $\pi(x)$

x	$\pi(x)$	$Li(x)$	$x/\log x$
10	4	6	4
10^2	25	30	22
10^3	168	178	145
10^4	1.229	1.246	1.086
10^5	9.592	9.630	8.686
10^6	78.498	78.628	72.382
10^7	664.579	664.918	620.421
10^8	5.761.455	5.762.209	5.428.711
10^9	50.847.534	50.849.237	48.254.942
10^{10}	455.052.511	455.055.615	434.594.481
10^{11}	4.118.054.813	4.118.066.401	3.928.131.653
10^{12}	37.607.912.018	37.607.950.281	36.191.205.825

same theorem in a simpler way. The importance of this theorem is considerable, by the number of its asymptotic consequences. The one which maybe is the most interesting, judging by the large number of works that it has suggested, can be expressed by the following theorem: [...] $\pi(x)$ *has for asymptotical expression, when x is large,* $Li(x)$, *with an error which becomes infinitely small with respect to* $Li(x)$ *when x goes to infinity.* The proof of this theorem had been published for the first time in a paper of Mr. von Mangoldt [1898]. One also finds in this article historical informations that it is interesting to reproduce.

The mentioned paper of von Mangoldt [53] of 1898 had provided a new proof of the prime number theorem in the form (1) instead of (2). This paper contained a short note of de La Vallée Poussin proving directly the equivalence of (1) and (2). See [28] for details.

The full report of the Belgian mathematician Paul Mansion [27], recommending to the Academy the publication of the memoir of 1899, gives more historical details than La Vallée Poussin's account:

Legendre [25] seems to be the first one to publish a formula giving approximatively the number $\pi(x)$ of prime numbers smaller to a given limit x. This formula is the following one

$$\pi(x) = \frac{x}{\log x - 1.08366}.$$

He had obtained it in an empirical way. Gauss, in a letter of December 24, 1849 [19] to Encke, who also had imagined a formula for the same topics, makes the history of his trials in this question; those trials, empirical like those of Legendre, started in 1792 or 1793, had led him to the approximate formula

$$\pi(x) = Li(x).$$

He compares the three formulas, Legendre's one, Encke's one, and his one, to the data of the tables and suspects that, for x very large, his formula is the most

exact one. A formula equivalent to Gauss' one is contained in a handwritten note added by Dirichlet to one of his memoirs sent to the great geometer of Göttingen [8], and it follows from the context that he had reached this result analytically, and not by induction (1838). Unfortunately, he never published his researches on this topic and only announced them. [...]
In 1848 and 1850, Tchebychef presented to the Academy of St-Petersburg two memoirs about prime numbers [51, 52], whose main results [...] became rapidly classical. They contain many theorems making plausible the exactness of the relation $\pi(x) = Li(x)$ as an asymptotic law. [...]
A short but substantial note of Riemann, presented to the Berlin Academy [44] opened a new way of research in this area. Riemann [...] found that the sum

$$F(x) = \pi(x) + \frac{1}{2}\pi(x^{\frac{1}{2}}) + \frac{1}{3}\pi(x^{\frac{1}{3}}) + \cdots$$

is made of a first term remaining finite when x increases indefinitely, followed by $Li(x)$ and then by a series of periodic definite integrals containing the roots of some transcendental function $\xi(t)$, [which] can be expressed in terms of another one $\zeta(s)$, previously considered by Euler in the case of a real variable, and defined by the relation

$$\prod \frac{1}{1 - \frac{1}{p^s}} = \sum \frac{1}{n^s},$$

where the product is made over all prime numbers $p \geq 2$ and the sum over all positive integers n. [...] All important works on the number of prime numbers use the properties of this transcendental function $\zeta(s)$ since.

De La Vallée Poussin refined, in this 1899 memoir, Gauss conjecture in the form

$$\pi(x) = Li(x) + 0\left(\frac{x}{\log x}\sqrt{p\log x}\, e^{-\sqrt{p\log x}}\right),$$

where $p = 0,03282$, providing a relative error of the form $\sqrt{p\log x}\, e^{-\sqrt{p\log x}}$ for large x. Notice finally that a positive answer to Riemann's conjecture would lead to a relative error of the form a constant times $x^{-1/2}(\log x)^2$ (Fig. 3).

7 The Successful Legacy of Immature Fruits

De La Vallée Poussin's contributions to number theory essentially ended with the XIXth century, and the Belgian mathematician concentrated on many other aspects of analysis during the first half of XXth century. In November 1939, to celebrate the centenary of Gauss' famous memoir on potential theory [18], de La Vallée Poussin delivered at the Belgian Mathematical Society a lecture entitled 'Gauss et la théorie du potentiel'. Its preserved text was published posthumously in 1962, the year of de La Vallée Poussin's death [6].

The lecture started with some moving remarks on Gauss' solitude, possibly inspired by de La Vallée Poussin's own situation in Belgium:

Fig. 3 Charles-Jean de La Vallée Poussin (Université Catholique de Louvain)

> Now, secluded in his domain, the lonely man of Göttingen has retrenched himself in his own thoughts and his memoir on potential bears the sign of his strange selfish attitude. [...] In the work of Gauss, not a single name is quoted. Maybe the old master believes that he owes nothing to anybody, and maybe is he only half wrong. For the first time, potential theory is treated as a branch of pure mathematics, self-sufficient and without seeking any support on experimental facts. [...] Those existence theorems – he is the first one to see their necessity – for which he attempts defective proofs, make him, despite of their gaps and their failures, [...] closer to us than all his predecessors and immediate followers.

Gauss' interest for potential theory came from his work of 1813 on the attraction of an ellipsoid to a point according to Newton's law (to take account of the shape of the Earth in celestial mechanics), his researches on conformal mappings (1834–1839), and his mentioned interest in terrestrial magnetism. This last reason led him not to restrict himself to positive distributions or charges.

The first part of Gauss' memoir gives a strict mathematical form to the well known part of Newton's potential theory:

> Gauss considers potential theory *ab ovo*, as if nobody had ever written about it. The presentation starts exactly like the best contemporary treatise of Analysis.

> [...] We can ask now if, after one century, we have added anything to it. Many efforts have been spent, I fear, for a rather deceiving result. We have generalized a little bit the validity conditions of the formulas, but at the expense of loosing the beautiful simplicity of the proofs, and what the theory has gained in generality, has been lost in elegance.

The second part of Gauss' memoir is completely original and de La Vallée Poussin insisted upon the large number of fruitful techniques introduced there, like the reciprocity principle for mass distributions, the mean value property of harmonic functions (not explicitly stated but used in subsequent proofs), a form of maximum principle and two important theorems. The first one states that, given a positive mass (or charge) M spread out on a surface S bounding a bounded domain $D \subset \mathbb{R}^3$, there exists a distribution (or measure) $d\mu$ of it whose Newtonian potential $V(x) = \int_S \frac{d\mu(y)}{|x-y|}$ is constant on S. Such a distribution was called an *equilibrium distribution* and the corresponding constant potential an *equilibrium potential*. Gauss observed that it is given by the distribution $d\mu$ of M minimizing the energy integral $\int_S V\, d\mu = \int_S \int_S \frac{1}{|x-y|}\, d\mu(x)\, d\mu(y)$ over all measures $d\mu$ such that $\int_S d\mu = M$. Gauss restricted himself to the case of distributions $d\mu(x) = \sigma(x) dS_x$ having a density $\sigma(x)$ on S. Like Dirichlet, Lord Kelvin and Riemann later in other contexts, Gauss admitted as evident the existence of the minimizer. See more details in [1, 29].

The second theorem, as stated by Gauss, was incorrect but contained, as a special case, the famous sweeping method, so important in potential theory, consisting in replacing the masses included in a bounded domain by a layer spread out on its boundary, without changing the potential.

The rigorous proof of those two results had to wait 95 years and Frostman's thesis [12], based upon all resources of the modern theory of measure and set functions. Gauss' influence is clearly described in Frostman's *Introduction*:

> In Chapter II, we study the classical problem, already considered by Gauss, to distribute a given positive mass in a domain of space, in such a way that the potential becomes constant in this domain (equilibrium problem). [...] It will follow from our researches that such a distribution exists under very general conditions. Moreover our proof is essentially based upon a variational principle due to Gauss, which depends very little on the harmonic properties of the ordinary potential; one recognizes indeed that the maximum principle plays the fundamental role. It follows that the equilibrium theorem holds for a wide class of so-called generalized potentials. [...] In Chapter V we first present a sweeping out method for generalized potentials, and use it to establish the maximum principle in its whole generality. This sweeping out reduces in the case of Newtonian potential to the ordinary one, named after Poincaré, but already considered by Gauss.

What was hidden in Gauss' incomplete attempt is the fact that a mass distribution is not always determined by its density, in the same way as a continuous function needs not to be determined by its derivative, if it is not absolutely continuous.

De La Vallée Poussin ended his paper in a somewhat provocative manner, by rejoicing that Gauss' memoir on potential theory was a counter-example to his own motto *Pauca sed matura*:

> If Gauss has thought to remain always faithful to his *moto*, he happily was wrong. What we mathematicians request from our predecessors, is not so much to leave us achieved material than to help us in our work. It happened to Gauss to realize his desire and to produce works very close to perfection. One century after him, they are dead works, like those *Disquisitiones* in which we only see today a remarkable monument of the past. But, without realizing, he has also thrown seeds falling on a ground not mature to receive them. Later, other workers came and have turned up this unfruitful soil, have bettered it, brought in the needed feeding essences for its fecundity, and, some day, after a long sleep, the seed still alive has germinated. The plant that came out is young and living and we finally see in its fruits the deepness of the remote thought where it comes from: we are mostly grateful to this type of thought.

Gauss' legacy has been and will remain an invaluable treasure for generations of mathematicians.

References

1. Anger, G.: Die Entwicklung der Potentialtheorie im Hinblick auf ihre grundlegenden Existenzsätze. Jahresber. Dtsch. Math.-Ver. **64**, 51–78 (1961); 100–134
2. Bertrand, J.: Calcul des Probabilités. Gauthier-Villars, Paris (1889)
3. Bühler, W.K.: Gauss. A Biographical Story. Springer, Berlin (1981)
4. de La Vallée Poussin, Ch.J.: Recherches analytiques sur la théorie des nombres premiers. Première partie. La fonction $\zeta(s)$ de Riemann et les nombres premiers en général. Ann. Soc. Sci. Brux. 2 **20**, 183–256 (1896)
5. de La Vallée Poussin, Ch.J.: Sur la fonction $\zeta(s)$ de Riemann et le nombre des nombres premiers inférieurs à une limite donnée. Mém. Acad. R. Sci. Lett. Beaux-Arts Belg. **49**, 74 (1899)
6. de La Vallée Poussin, Ch.J.: Gauss et la théorie du potentiel. Rev. Quest. Sci. **133**, 314–330 (1962)
7. de La Vallée Poussin, Ch.J.: Collected works. In: Butzer, P., Mawhin, J., Vetro, P. (ed.) Oeuvres Scientifiques, vol. I. Académie Royale de Belgique/Circolo Matematico di Palermo, Bruxelles/Palermo (2000)
8. Kronecker, L. (ed.): G. Lejeune Dirichlet's Werke. Reimer, Berlin (1889), vol. 1, note 2 of note 2
9. Dunnington, G.W.: Carl Friedrich Gauss, Titan of Science. Hafner, New York (1955). Reprinted with additional material by J. Gray and F.-E. Dohse, Mathematical Association of America, Washington (2004)
10. Dutka, J.: On Gauss' priority in the discovery of the method of least squares. Arch. Hist. Exact Sci. **49**, 355–370 (1996)
11. Fourier, J.B.: Mémoire sur les résultats moyens déduits d'un grand nombre d'observations, in Recherches statistiques sur la ville de Paris et le département de la Seine, 4 vol., Paris, 1821–1820. Oeuvres de Fourier, G. Darboux ed. Gauthier-Villars, Paris (1890), vol. 2, 525–545
12. Frostman, O.: Potentiel d'équilibre et capacité des ensembles, avec quelques applications à la théorie des fonctions, Meddelanden fran Lunds Universitets Matematiska Seminarium 3. Gleerup, Lund (1935)
13. Carl Friedrich Gauss' Werke, 12 vol. Perthes/Teubner/Springer, Gotha/Leipzig/Berlin (1863–1933)
14. Gauss, C.F.: Demonstratio nova theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores primi vel secundi gradus resolvi posse, Doctoral dissertation. C.G. Fleckeisen, Helmstedt (1799). Werke III, pp. 1–30
15. Gauss, C.F.: Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium. Perthes und Besser, Hamburg (1809). Werke VII, pp. 1–280
16. Gauss, C.F.: Theoria Combinationis Observationum Erroribus Minimis Obnoxiae. Comment. (Göttingen) 5, pp. 33–62, 63–90 (1819–1822). Werke IV, pp. 95–108
17. Gauss, C.F.: Disquisitiones Generales Circa Superficies Curvas. Comment. (Göttingen) 6, pp. 99–146 (1823–1827). Werke IV, pp. 217–258

18. Gauss, C.F.: Allgemeine Lehrsätze in Beziehung auf die im verkehrten Verhältnisse des Quadrats der Entfernung wirkenden Anziehungs -und Abstossungs-Kräfte. In: Resultate aus den Beobachtingen des magnetischen Vereins im Jahre 1839, pp. 1–51. Werke V, pp. 195–242 (1840)
19. Gauss, C.F.: Werke II, pp. 444–447
20. Gnedenko, B.W.: Über die Arbeiten von C.F. Gauss zur Wahrscheinlichkeitsrechnung. In: Reichardt, H. (ed.) C.F. Gauss 1777–1855. Gedenkband anlässlich des 100. Todestages am 23. Februar 1955, pp. 194–204. Teubner, Leipzig (1957)
21. Goddard, R.: History of the least squares method and its link with the probability theory. In: Knobloch, E., Mawhin, J., Demidov, S.S. (eds.) Studies in History of Mathematics dedicated to A.P. Youschkevitch, pp. 237–250. Brepols, Turnhout (2002)
22. Hadamard, J.: Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques. Bull. Soc. Math. France **14**, 199–220 (1896)
23. Hall, T.: Carl Friedrich Gauss: A Biography. MIT Press, Cambridge (1970)
24. Legendre, A.M.: Nouvelles méthodes pour la détermination des orbites des comètes; Appendice: Sur la méthodes des moindres quarrés. Didot, Paris (1805)
25. Legendre, A.M.: Théorie des Nombres, 2^e éd. Didot, Paris (1808)
26. Mailly, E.: Essai sur la vie et les ouvrages de L.A.J. Quetelet. Ann. Acad. R. Belg. **41**, 109–297 (1875)
27. Mansion, P.: Sur la fonction $\zeta(s)$ de Riemann et le nombre des nombres premiers inférieurs à une limite donnée par Ch.-J. de la Vallée Poussin. Rapport de M. Paul Mansion, premier commissaire. Bull. l'Acad. R. Belg. (3) **36**, 10–21 (1898)
28. Mawhin, J.: A note of de La Vallée Poussin on the prime number theorem in a publication of von Mangoldt. In: Butzer, P., Mawhin, J., Vetro, P. (eds.) Ch. J. de La Vallée Poussin Collected Works, vol. 1, pp. 117–122. Académie Royale de Belgique/Circolo Matematico di Palermo, Bruxelles/Palermo (2000)
29. Monna, A.F.: Dirichlet's Principle. A Mathematical Comedy of Errors and its Influence on the Development of Analysis. Oosthoek, Scheltema & Holkema, Utrecht (1975)
30. Pearson, K.: Contributions to the mathematical theory of evolution. Philos. Trans. R. Soc. Lond. A **185**, 72 (1894)
31. Pearson, K.: Historical note on the original of the normal curve of errors. Biometrika **16**, 402–404 (1924)
32. Peters, C.A.F.: Briefwechsel zwischen C.F. Gauss und H.C. Schumacher, vol. 5, p. 325. G. Esch, Altona (1865)
33. Plackett, R.L.: Studies in the history of probability and statistics. XXIX. The discovery of the method of least squares. Biometrika **59**, 239–251 (1972)
34. Porter, T.M.: The Rise of Statistical Thinking. Princeton University Press, Princeton (1986)
35. Quetelet, Ad.: Instruction populaires sur le calcul des probabilités. Tarlier et Hayez, Bruxelles (1828)
36. Quetelet, Ad.: Notes extraites d'un voyage scientifique, fait en Allemagne pendant l'été de 1829. In: Correspondance Mathématique et Physique vol. 6, pp. 126–148, 161–178, 225–239 (1830)
37. Quetelet, Ad.: Recherches sur le penchant au crime aux différens âges. Nouv. Mém. Acad. R. Sci. B.-Lett. Brux. **7**, 87 (1832)
38. Quetelet, Ad.: Recherches sur le poids de l'homme aux différents âges. Nouv. Mém. Acad. R. Sci. B.-Lett. Brux. **7**, 44 (1833)
39. Quetelet, Ad.: Sur l'homme et le développement de ses facultés, ou Essai de physique sociale, vol. 2. Bachelier, Paris (1835). German translation Stuttgart: Schweizenbart (1838)
40. Quetelet, Ad.: Lettres à S.A.R. le duc régnant de Saxe-Cobourg et Gotha sur la théorie des probabilités appliquée aux sciences morales et politiques. Hayez, Bruxelles (1846)
41. Quetelet, Ad.: Sciences mathématiques et physiques chez les Belges au commencement du XIX^e siècle. Muquardt, Bruxelles (1867)
42. Quetelet, Ad.: Physique sociale ou Essai sur le développement des facultés de l'homme, vol. 2. Muquardt/Baillire/Jacques Issakogg, Bruxelles/Paris/Saint-Pétersbourg (1869)
43. Reich, K.: Carl Friedrich Gauss. Inter Nationes, Bonn-Bad Godesberg (1977)
44. Riemann, B.: Berliner Monatsberichte der Berliner Akademie, November 1859, 671–680. In: Bernhard Riemann's Gesammelte Mathematische Werke, H. Weber ed., 2nd edn. (1892), pp. 145–156
45. Schneider, I.: Die Arbeiten von Gauss im Rahmen der Warscheinlichkeitsrechnung: Methode der kleinsten Quadrate und Versiecherung, Carl Friedrich Gauss (1777–1855): Sammelband von Beiträgen zum 200. Geburtstag von C.F. Gauss, pp. 143–172. Minerva, München (1991)
46. Schubring, G.: Der Briefwechsel Quetelet-Gauss: Magnetismus und Sternschnuppen. Acta Hist. Leopold. **54**, 789–807 (2008)

47. Sheynin, O.B.: C.F. Gauss and the theory of errors. Arch. Hist. Exact Sci. **20**, 21–72 (1979)
48. Sheynin, O.B.: A. Quetelet as a statistician. Arch. Hist. Exact Sci. **36**, 281–325 (1986)
49. Stigler, S.M.: Gauss and the invention of least squares. Ann. Stat. **9**, 465–474 (1981)
50. Stigler, S.M.: The History of Statistics: The Measurement of Uncertainty before 1900. Belknap Press of Harvard University, Harvard (1986)
51. Tchebychef, P.L.: Sur la fonction qui détermine la totalité des nombres premiers inférieurs à une limite donnée. Mém. Acad. Saint-Pétersbourg (savants étrangers) **6**, 141–157 (1851); J. Math. Pures Appl. **17**, 341–365 (1852)
52. Tchebychef, P.L.: Mémoire sur les nombres premiers. Mém. Acad. Saint-Pétersbourg (savants étrangers) **7**, 14–23 (1854); J. Math. Pures Appl. **17**, 366–380 (1852)
53. von Mangoldt, H.: Über eine Anwendung der Riemann'sche Formel für die Anzahl der Primzahlen unter einer gegeben Grenze. J. Reine Angew. Math. **119**, 65–71 (1898)
54. Waterhouse, W.C.: Gauss' first argument for least squares. Arch. Hist. Exact Sci. **42**, 137–171 (1991)

Jean Mawhin is emeritus professor at the *Université Catholique de Louvain*. Fellow and former president of the *Académie royale de Belgique*, he has received three honorary degrees, the Humboldt Prize (1990–1991) and Bolzano medal (2002). His fields of interest are nonlinear functional analysis and differential equations, integration theory and history of mathematics.