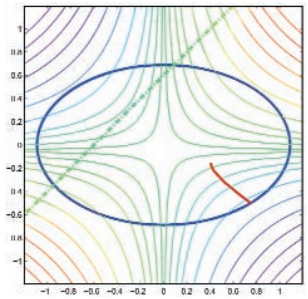
AMERICAN MATHEMATICAL SOCIETY

Current Events in Mathematics

An AMS Special Session Organized by AMS President

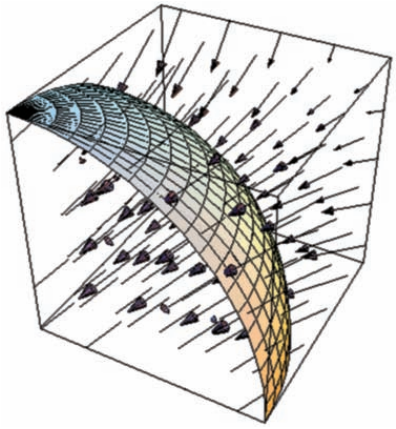
David Eisenbud

FRIDAY, JANUARY 9 - 1:00 TO 5:10 P.M.
PHOENIX CIVIC PLAZA, ROOM 41



The Interior-Point Revolution in Optimization: History, Recent Developments and Lasting Consequences

MARGARET H. WRIGHT



What Is Motivic Integration?

THOMAS C. HALES

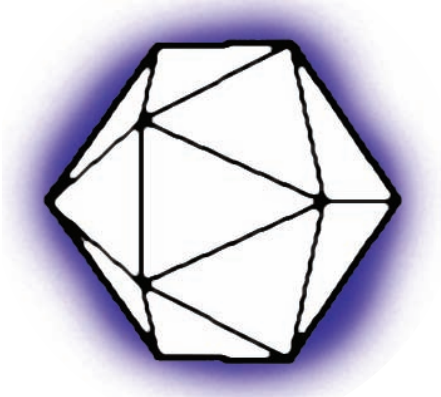
"SEVEN AND A HALF LOGS SHOULD DO IT!"

It Is Easy to Determine Whether or Not a Given Integer Is Prime

ANDREW GRANVILLE

Perelman's Recent Work on the Classification of 3-Manifolds

JOHN W. MORGAN



The Interior-Point Revolution in Optimization: History, Recent Developments, and Lasting Consequences

Margaret H. Wright
Computer Science Department
Courant Institute of Mathematical Sciences
New York University

AMS Special Session on Current Events
Joint Mathematics Meeting
January 9, 2004

Abstract

Interior methods are a pervasive feature of the optimization landscape today, but it was not always so. Although interior-point techniques, primarily in the form of barrier methods, were widely used during the 1960s for problems with nonlinear constraints, their use for the fundamental problem of linear programming was unthinkable because of the total dominance of the simplex method. During the 1970s, barrier methods were superseded, nearly to the point of oblivion, by newly emerging and seemingly more efficient alternatives such as augmented Lagrangian and sequential quadratic programming methods. By the early 1980s, barrier methods were almost universally regarded as a closed chapter in the history of optimization.

This picture changed dramatically in 1984, when Narendra Karmarkar announced a fast polynomial-time interior method for linear programming; in 1985, a formal connection was established between his method and classical barrier methods. Since then, interior methods have continued to transform both the theory and practice of constrained optimization. We present a condensed, unavoidably incomplete look at classical material and recent research about interior methods.

1 Overview

REVOLUTION:
(i) a sudden, radical, or complete change;
(ii) a fundamental change in political organization, especially the overthrow or renunciation of one government or ruler and the substitution of another.[1]

It can be asserted with a straight face that the field of continuous optimization has undergone a revolution since 1984 in the sense of the first definition, and that the second

[1] *Merriam–Webster Collegiate Dictionary*, Seventh Edition, 1965.

definition applies in a philosophical sense: Because the interior-point presence in optimization today is ubiquitous, it is easy to lose sight of the magnitude and depth of the shifts that have occurred during the past twenty years. Building on the implicit political metaphor of our title, successful revolutions eventually become the status quo.

The interior-point revolution, like many other revolutions, includes old ideas that are rediscovered or seen in a different light, along with genuinely new ideas. The stimulating interplay of old and new continues to lead to increased understanding as well as an ever-larger set of techniques for an ever-larger array of problems, familiar (Section 4.4) and heretofore unexplored (Section 5).

Because of the vast size of the interior-point literature, it would be impractical to cite even a moderate fraction of the relevant references, but more complete treatments are mentioned throughout. The author regrets the impossibility of citing all important work individually.

2 Linear and Nonlinear Programming: Separated from Birth

Prior to 1984, there was, to first order, *no connection* between linear and nonlinear programming. For historical reasons that seem puzzling in retrospect, these topics, one a strict subset of the other, evolved along two essentially disjoint paths. Even more remarkably, this separation was a fully accepted part of the culture of optimization—indeed, it was viewed by some as inherent and unavoidable. For example, in a widely used and highly respected textbook [24] published in 1973, the author comments in the preface that "Part II [unconstrained optimization] ... is independent of Part I [linear programming]" and that "except in a few isolated sections, this part [constrained optimization] is also independent of Part I". To provide an accurate reflection of this formerly prevailing viewpoint, we give separate background treatments for linear and nonlinear programming.

2.1 Linear programming

2.1.1 Problem statement and optimality conditions

The linear programming (LP) problem involves minimization of a linear (affine) function subject to linear constraints, and can be represented in various mathematically equivalent ways. The two forms of interest here are the all-inequality form,

$$\underset{x}{\text{minimize}}\ c^T x \quad \text{subject to} \quad Ax \geq b, \tag{1}$$

and *standard form*,

$$\underset{x}{\text{minimize}}\ c^T x \quad \text{subject to} \quad Ax = b, \quad x \geq 0, \tag{2}$$

where A is $m \times n$. In the standard-form problem (2), the only inequalities are the simple bound constraints $x \geq 0$, leading to the crucial (and sometimes overlooked) property that x plays two distinct roles—as the variables and the values of the constraints. It is customary in standard-form problems to assume that A has full rank.

A point is *feasible* if it satisfies the problem constraints. The feasible point x^* is a solution of the standard-form LP (2) if and only if, for some m-vector y^* and n-vector z^*,

$$c = A^T y^* + z^*, \quad z^* \geq 0, \quad \text{and} \quad z_i^* x_i^* = 0 \quad \text{for } i = 1, \ldots, n, \tag{3}$$

where z^* is the Lagrange multiplier for the bound constraints and y^* is the Lagrange multiplier for the equality constraints $Ax = b$.

2.1.2 The simplex method

A fundamental property of linear programs is that, if the optimal objective value is finite, a *vertex* minimizer must exist. (For details about linear programming and its terminology, see, e.g., [5], [31], and [19].) The simplex method, invented by George B. Dantzig in 1947, is an iterative procedure for solving LPs that completely depends on this property. The starting point for the simplex method must be a vertex. Thereafter, every iteration moves to an adjacent vertex, decreasing the objective as it goes, until an optimal vertex is found. The underlying motivation for the simplex method is easy to understand, but its simplicity is sometimes obscured by a focus on algebraic details.

Almost from the beginning, the simplex method (and, by association, linear programming) acquired an array of specialized terminology and notation, such as "basic feasible solution", "min ratio test", and the tableau. During the early years of the simplex method, simplex steps were carried out by performing unsafeguarded rank-one updates to the explicit inverse of the square basis matrix. As an aside, use of this risky technique shows that mainstream linear programming was widely separated not only from nonlinear programming, but also from numerical linear algebra; fortunately, during the 1960s, the simplex method became more closely connected with state-of-the-art linear algebraic techniques.

Although "non-simplex" strategies for LP were suggested and tried from time to time between 1950 and the early 1980s, such techniques never approached the simplex method in overall speed and reliability. Furthermore, the mindset induced by the dominance of the simplex method held sway to such an extent that even techniques labeled as non-simplex were at heart based on the same motivation as the simplex method: to identify the active inequality constraints by staying on a changing subset of exactly-satisfied constraints while reducing the objective function.

2.1.3 Concerns about complexity

In practice, because the simplex method routinely and efficiently solved very large linear programs, it retained unquestioned preeminence as the solution method of choice. However, the simplex method was viewed with nagging discontent by those interested in computational complexity, a field whose importance increased during the 1960s and 1970s. An underlying tenet of theoretical computer science is that any "fast" algorithm must be *polynomial-time*, meaning that the number of arithmetic operations required to solve the problem should be bounded above by a polynomial in the problem size.

Although the simplex method almost always converges on real-world problems in a number of iterations that is a small multiple of the problem dimension, it is known that the simplex method can visit *every vertex* of the feasible region—for example, on the famous

Klee-Minty "twisted cube" LP; see [31] and [17] for two formulations of this problem. Consequently the worst-case complexity of the simplex method is *exponential* in the problem dimension, which means that the simplex method must be a "bad" algorithm. The disconnect between observed speed and theoretical inefficiency was widely known, and there were several claims, subsequently shown to be false, that a provably polynomial-time method for LP had been discovered.

The first polynomial-time LP algorithm was devised in 1979 by Leonid Khachian of the then Soviet Union, in work that made newspaper headlines around the world. Khachian's *ellipsoid method* is based on specialization of general *nonlinear* approaches developed earlier by other Soviet mathematicians, notably Shor, Yudin and Nemirovskii. In particular, Khachian's method does not rely, as the simplex method does, on existence of a vertex solution or, more generally, the finiteness/combinatorial features of the LP problem. Polynomiality of the ellipsoid method arises from two bounds: an outer bound that guarantees existence of an initial (huge) ellipsoid enclosing the solution, and an inner bound that specifies how small the final ellipsoid must be to ensure sufficient closeness to the exact solution. See, for example, [31] for details about Khachian's method.

Despite its favorable complexity, the performance of the ellipsoid method in practice, i.e., its actual running time, was extremely slow—much slower than the simplex method. In fact, in complete contrast to the simplex method, the number of iterations of the ellipsoid method tended to be comparable to its enormous (albeit polynomial) upper bound. Thus the simplex method remained "the only game in town" for solving linear programs, leading to a puzzling and deeply unsatisfying anomaly in which an exponential-time algorithm was consistently and substantially faster than a polynomial-time algorithm. Even after Khachian's breakthrough, the quest continued for an LP algorithm that was not only polynomial, but also efficient in practice.[2] The linear programming story will continue in Section 3.

2.2 Nonlinear programming

2.2.1 Problem statement and optimality conditions

The generic nonlinear programming, or nonlinear optimization, problem involves minimization of a nonlinear function subject to nonlinear constraints. Special cases of nonlinear programming arise when, for example, the objective function is quadratic, the constraints are bounds, or the constraints are linear (equalities or inequalities). Here we consider only the all-inequality version of a nonlinear programming problem:

$$\underset{x\in\mathcal{R}^n}{\text{minimize}} \quad f(x) \quad \text{subject to} \quad c(x) \geq 0, \tag{4}$$

where $c(x)$ has m component functions, and f and $\{c_i\}$ are smooth. (Observe that (4) is analogous in form to the all-inequality linear program (1).) The n-vector $g(x)$ denotes the gradient of f; the matrix of second partial derivatives will be denoted by $H(x)$. The gradient and Hessian of $c_i(x)$ will be denoted by $a_i(x)$ and $H_i(x)$. The $m \times n$ Jacobian matrix of $c(x)$ is denoted by $A(x)$, whose ith row is $a_i(x)^T$. The *Lagrangian function* associated with (4) is

[2] A side comment: Recent work on "smoothed complexity" provides a fascinating explanation of why the simplex method is usually a polynomial-time algorithm; see [32].

$L(x, \lambda) = f(x) - \lambda^T c(x)$, where λ normally represents a vector of Lagrange multipliers. The Hessian of the Lagrangian with respect to x, denoted by W, is $W(x, \lambda) = H - \sum_{j=1}^{m} \lambda_j H_j(x)$.

The constraint $c_i(x) \geq 0$ is said to be *active* at $\bar{x}$ if $c_i(\bar{x}) = 0$ and inactive if $c_i(\bar{x}) > 0$. Let $\hat{A}(x)$ denote the Jacobian of the active constraints at x, and let $N(x)$ denote a matrix whose columns form a basis for the null space of $\hat{A}$.

Throughout the remainder of the paper, we assume the following conditions, which are sufficient to ensure that x^* is an isolated constrained minimizer of (4):

1. $c(x^*) \geq 0$ and $\hat{A}(x^*)$ has full rank;
2. $g(x^*) = A(x^*)^T \lambda^*$, with $\lambda_j^* \geq 0$ and

$$\lambda_j^* c_j(x^*) = 0, \quad j = 1, \ldots, m; \tag{5}$$

3. $\lambda_j^* > 0$ if $c_j(x^*) = 0$, $j = 1, \ldots, m$;
4. $N(x^*)^T W(x^*, \lambda^*) N(x^*)$, the *reduced Hessian of the Lagrangian*, is positive definite.

Relation (5), that each pairwise product of constraint and multiplier must be zero, is called *complementarity*. Condition 3, *strict* complementarity, requires that one of $c_j(x^*)$ and λ_j^* must be positive.

2.2.2 Newton's method

Newton's method occurs in multiple forms throughout optimization. When solving the nonlinear equations $\Phi(z) = 0$, let $J(z)$ denote the Jacobian matrix of Φ. If z_k is the current point and $J(z_k)$ is nonsingular, the *Newton step* δ_k is the solution of the linear system

$$J(z_k)\delta_k = -\Phi(z_k), \tag{6}$$

so that δ_k is the step from z_k to a zero of the local affine Taylor-series model of Φ.

For unconstrained minimization of $f(x)$ starting from x_k, the Newton step p_k is designed to minimize a local Taylor-series quadratic model of $f(x_k + p)$, namely $f(x_k) + g(x_k)^T p + \frac{1}{2} p^T H(x_k) p$. If the current Hessian $H(x_k)$ is positive definite, p_k solves the linear system

$$H(x_k)p = -g(x_k). \tag{7}$$

When minimizing $f(x)$ subject to m linear equality constraints $Ax = b$, the Newton step p_k should minimize the local Taylor-series quadratic model of f subject to also satisfying the constraints $A(x_k + p_k) = b$, so that p_k is a solution of the quadratic program

$$\underset{p \in \mathcal{R}^n}{\text{minimize}} \quad \tfrac{1}{2} p^T H_k p + g_k^T p \quad \text{subject to} \quad Ap = b - Ax_k, \tag{8}$$

where $H_k = H(x_k)$ and $g_k = g(x_k)$. Under appropriate conditions, p_k and a "new" multiplier y_{k+1} satisfy the following $n + m$ linear equations:

$$\begin{pmatrix} H_k & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} p_k \\ -y_{k+1} \end{pmatrix} = \begin{pmatrix} -g_k \\ b - Ax_k \end{pmatrix}, \tag{9}$$

where y_{k+1} is an estimate of the Lagrange multipliers for the equality constraints. The matrix in (9) is nonsingular if A has full rank and the reduced Hessian $N_A^T H_k N_A$ is positive definite, where N_A is a basis for the null space of A. If $Ax_k = b$, the second equation in (9) becomes $Ap_k = 0$, implying that p_k must lie in the null space of A.

A "pure" Newton method for zero-finding begins with an initial point z_0, and generates a sequence of Newton iterates $\{z_k\}$, where $z_{k+1} = z_k + \delta_k$, with δ_k defined by (6), and similarly for minimization, using (7) and (9). Under various conditions that can be quite restrictive, a pure Newton method converges quadratically to a solution.

One way to encourage convergence from a general starting point is to perform a *line search* in which the new iterate is defined by $z_{k+1} = z_k + \alpha_k \delta_k$, where the positive scalar α_k is chosen to decrease a *merit function* that measures progress. In unconstrained optimization, the merit function is typically the objective function. Standard line search acceptance criteria that ensure convergence are discussed in, for example, [29, 28]. A second strategy is based on defining a *trust region* around the current iterate within which the local model can be trusted. In optimization, the step in a trust-region method is typically chosen to minimize (approximately) the local Taylor-series quadratic model subject to remaining within a (normally ℓ_2) trust region.

2.2.3 Barrier methods for constrained optimization

The 1960s were the heyday of unconstrained optimization, and, as a result, it was common practice to convert constrained problems into unconstrained subproblems or sequences of unconstrained subproblems. Penalty and barrier methods were especially popular, both motivated by minimizing a composite function that reflects the original objective function as well as the influence of the constraints. Modern interior methods are closely related to "classical" (1960s) barrier methods, which we now describe.

The *logarithmic barrier function* associated with the problem (4) is

$$B(x, \mu) = f(x) - \mu \sum_{j=1}^{m} \ln c_j(x), \tag{10}$$

where μ is a positive scalar called the *barrier parameter*. The logarithmic terms are well defined at points x for which $c(x) > 0$, but become unbounded above as x approaches any point where a constraint is zero, and are undefined if $c_j(x) < 0$ for any j. (This behavior constitutes an obvious rationale for the descriptors "barrier" and "interior".) Numerous properties of $B(x, \mu)$ are known; see, for example, the classic reference [9] or [36, 12].

For small μ, unconstrained minimizers of $B(x, \mu)$ are related in an intuitively appealing way to the solution x^* of (4). Given that x^* satisfies the sufficient optimality conditions given in Section 2.2.1, then, for a sequence of monotonically decreasing and sufficiently small values of μ, there is an associated sequence $\{x_\mu\}$ of isolated local unconstrained minimizers of the barrier function (10) such that

$$\lim_{\mu \to 0} x_\mu = x^* \quad \text{and} \quad \lim_{\mu \to 0} \frac{\mu}{c_j(x_\mu)} = \lambda_j^*. \tag{11}$$

Under suitable assumptions of smoothness, the points $\{x_\mu\}$ define a smooth curve, called either the *barrier trajectory* or the *central path*, that converges to x^* *non-tangentially*, from

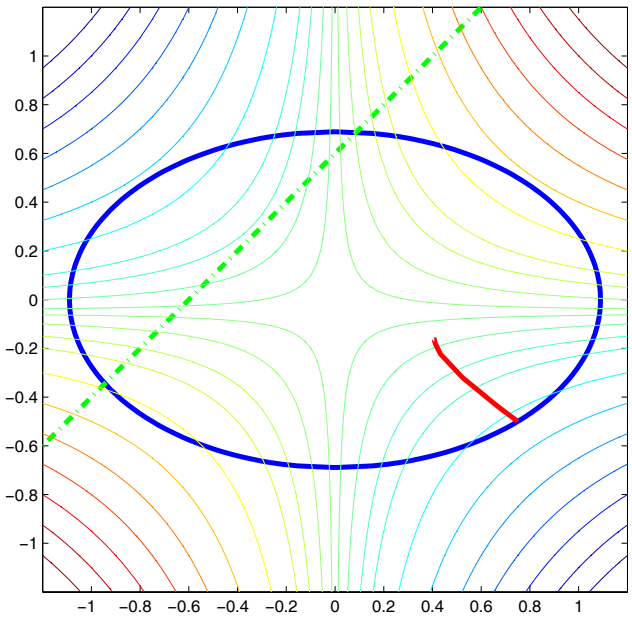


Figure 1: The contours of the nonconvex objective function (12) are shown, along with the boundaries of the ellipsoidal constraint and linear constraint of (13). A trajectory of local unconstrained minimizers of the logarithmic barrier function, shown in red, begins at the strictly feasible analytic center of the feasible region, corresponding to $\mu = \infty$, and converges to the boundary as $\mu \to 0$.

the strict interior of the feasible region—not along the boundary. For proofs and additional details, see, for example, [36, 12].

To illustrate the behavior of the log barrier function, we consider the two-variable inequality-constrained problem:

$$\begin{aligned} &\text{minimize} && \tfrac{10}{3}x_1x_2 + \tfrac{1}{6}x_1 && (12)\\ &\text{subject to} && \tfrac{19}{16} - x_1^2 - \tfrac{5}{2}x_2^2 \geq 0 \quad \text{and} \quad x_1 - x_2 + \tfrac{3}{5} \geq 0. && (13) \end{aligned}$$

The first (nonlinear) constraint is satisfied inside an ellipse centered at the origin; the second (linear) constraint cuts off part of the ellipse. Figure 1 shows the contours of f, which is unbounded below, and the boundaries of these two constraints; the feasible region lies inside the ellipse, to the right of the line.

The figure makes clear that there are two local minimizers of f in the feasible region. At the isolated constrained minimizer $x^* = (\frac{3}{4}, \frac{-1}{2})$, the first constraint is active. The path of barrier minimizers converging to x^* is shown in red. The strictly feasible starting point of the path of barrier minimizers corresponds to the minimizer of $-\sum \ln c_i(x)$—in effect, to an infinite value of μ, so that the objective function has no effect.

The gradient of the barrier function (10), denoted by g_B, is

$$g_B(x,\mu) = g(x) - \sum_{j=1}^{m} \frac{\mu}{c_j(x)}\, a_j(x) = g(x) - \mu A^T(x) C^{-1}(x)\mathbf{1}, \tag{14}$$

where $\mathbf{1} = (1, \ldots, 1)^T$. The final form in (14) uses the widely established convention in the interior-point literature that an uppercase version of a letter denoting a vector means the diagonal matrix whose diagonal elements are those of the vector. The barrier Hessian, denoted by H_B, has the form

$$H_B(x, \mu) = H(x) - \sum_{j=1}^{m} \frac{\mu}{c_j(x)} H_j(x) + \mu A^T(x) C^{-2}(x) A(x). \tag{15}$$

Since $g_B(x_\mu)$ vanishes, x_μ can be interpreted as a highly special point at which the objective gradient is a nonnegative linear combination of the constraint gradients. Further, the coefficients in the linear combination have a specific relationship with μ and the constraint values, i.e.,

$$g(x_\mu) = \sum_{j=1}^{m} \frac{\mu}{c_j(x_\mu)} a_j(x_\mu) = A^T(x_\mu) \lambda_\mu(x_\mu), \tag{16}$$

where the multiplier estimate λ_μ satisfies

$$\lambda_\mu(x) = \mu C(x)^{-1} \mathbf{1}. \tag{17}$$

A rearranged component-wise version of (17) is

$$\big(\lambda_\mu(x_\mu)\big)_i c_i(x_\mu) = \mu. \tag{18}$$

This relationship is similar to complementarity (5), which holds at x^* between λ^* and $c(x^*)$, and is sometimes called *perturbed complementarity* or the *centering* property of x_μ.

To move from the current point x to x_μ, a straightforward strategy is to apply Newton's method to a local quadratic model of the barrier function. Omitting arguments, the resulting $n \times n$ *primal Newton barrier equations* are

$$H_B p = -g + \mu A^T C^{-1} \mathbf{1}, \tag{19}$$

where "primal" refers to the original problem variables x.

Although barrier methods were widely used during the 1960s, they suffered a severe decline in popularity in the 1970s for various reasons, including perceived inefficiency compared to alternative strategies and worries about inherent ill-conditioning. With respect to the latter, it was observed in the late 1960's (see [23, 25]) that, if $1 \leq \hat{m} < n$, then $\operatorname{cond} H_B(x_\mu, \mu) = \Theta(1/\mu)$, so that the barrier Hessian becomes arbitrarily ill-conditioned at points lying on the barrier trajectory as $\mu \to 0$. Although it is impossible after a gap of more than 20 years to determine precisely why barrier methods became unpopular, concerns about ill-conditioning clearly played a role; see Section 4.3.

As penalty and barrier methods faded from the scene, the dominant approaches tended to be based directly on the optimality conditions for constrained optimization, in particular on properties of the Lagrangian function. Augmented Lagrangian methods and sequential quadratic programming (SQP) methods became especially popular, and remain so today. For further details about these methods, see, for example, [16, 10, 28].

3 The Revolution Begins

3.1 Karmarkar's method

In 1984, Narendra Karmarkar [21] announced a polynomial-time LP method for which he reported solution times that were consistently 50 times faster than the simplex method. This event, which received publicity around the world throughout the popular press and media, marks the beginning of the interior-point revolution.

Karmarkar's method had several unusual properties: a special, non-standard form was assumed for the linear program; nonlinear projective geometry was used in its description; and no information was available about the implementation. Amid the frenzy of interest in Karmarkar's method, it was shown in 1985 (and published the next year [15]) that there was a formal equivalence between Karmarkar's method and the classical logarithmic barrier method applied to the LP problem. Soon researchers began to view once-discarded barrier methods in a previously unthinkable context: as the source of polynomial-time algorithms for linear programming.

For several years the tie between Karmarkar's method and barrier methods was contentious and controversial. Researchers argued about whether the two approaches were fundamentally different, or very similar, or something in between. Now that the dust has settled, derivations of interior methods typically involve barrier functions or their properties, such as perturbed complementarity (18). Readers interested in Karmarkar's method should consult his original paper [21], or any of the many comprehensive treatments published since 1984 (e.g., [19, 30, 41, 35, 44]).

Beyond establishing the formal connection between Karmarkar's method and barrier methods, [15] reported computational results comparing a state-of-the-art (in 1985) simplex code, MINOS [26], and an implementation of the primal Newton barrier method on a widely available set of test problems. To the astonishment of many who believed that nothing could beat the simplex method, the barrier method was faster on several of the problems, and competitive on many others.

At this stage, the interior-point revolution gathered momentum and accelerated in several directions, to be described in Sections 4 and 5. First, however, we describe the derivation of the primal Newton barrier method for LP given in [15].

3.2 The primal Newton barrier method for LP

To make the connection between linear programming and a barrier method, consider a standard-form linear program—minimize $c^T x$ subject to $Ax = b$ and $x \geq 0$—with three properties: (a) the set of x satisfying $Ax = b$ and $x > 0$ is nonempty; (b) the set (y, z) satisfying $A^T y + z = c$ and $z > 0$ is nonempty; and (c) $\mathrm{rank}(A) = m$. Because the only inequality constraints are the bounds $x \geq 0$, the associated logarithmic barrier function (see (10)) is

$$B(x, \mu) = c^T x - \mu \sum_{j=1}^{n} \ln x_j, \tag{20}$$

and the barrier subproblem is to minimize (20) subject to satisfying the equalities $Ax = b$:

$$\text{minimize} \quad c^T x - \mu \sum_{j=1}^{n} \ln x_j \quad \text{subject to} \quad Ax = b. \tag{21}$$

The gradient and Hessian of the barrier function (20) have particularly simple forms:

$$g_B = c - \mu X^{-1} e \ \text{ and } \ H_B = \mu X^{-2}. \tag{22}$$

The barrier subproblem (21) has a unique minimizer if (b) is satisfied. At the optimal solution of (21), there exists y such that

$$g_B(x, \mu) = c - \mu X^{-1}\mathbf{1} = A^T y, \ \text{ so that } \ c = A^T y + \mu X^{-1}\mathbf{1}. \tag{23}$$

The *central path* (barrier trajectory) for a standard-form LP is defined by vectors x_μ and y_μ satisfying

$$\begin{aligned} Ax_\mu &= b, \quad x_\mu > 0; \\ A^T y_\mu + \mu X_\mu^{-1}\mathbf{1} &= c. \end{aligned} \tag{24}$$

The central path has numerous properties of interest; see, e.g., [20], [19],[35],[41], and [44].

Assume that we are given a point $x > 0$ for which $Ax = b$. Using (22), the Newton equations (9) for problem (21) are

$$\begin{pmatrix} \mu X^{-2} & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} p \\ -y \end{pmatrix} = \begin{pmatrix} -c + \mu X^{-1}\mathbf{1} \\ 0 \end{pmatrix}, \tag{25}$$

so that the Newton step p in x satisfies

$$\mu X^{-2} p + c - \mu X^{-1}\mathbf{1} = A^T y \tag{26}$$

for some Lagrange multiplier vector y. Multiplying (26) by AX^2 and using the relation $Ap = 0$ to eliminate p, we obtain

$$AX^2 A^T y = AX^2 c - \mu AXe = AX(Xc - \mu\mathbf{1}). \tag{27}$$

Since A has full rank and $x > 0$, the matrix AX^2A^T is positive definite, so that (27) has a unique solution y. Using (26), p is defined in terms of y as

$$p = x + \frac{1}{\mu} X^2 (A^T y - c).$$

Because $Ax = b$ and $Ap = 0$, the new point $x + \alpha p$ will continue to satisfy the equality constraints for any α. However, α may need to be less than one in order to retain strict feasibility with respect to the bound constraints.

4 The Revolution Advances

Following the announcements of Karmarkar's method and its connection with the logarithmic barrier method, researchers began to develop other interior LP methods with improved complexity bounds, and to derive properties of barrier methods applied to linear programs. Furthermore, since barrier methods (unlike the simplex method) were originally intended for nonlinear problems, it was evident that they could be applied not just to linear programming, but to other optimization problems, such as quadratic programming, linear and nonlinear complementarity, and nonlinear programming.

4.1 A change in perspective

The interior-point revolution has led to a fundamental shift in thinking about continuous optimization. Today, in complete contrast to the era before 1984, researchers view linear and nonlinear programming from a *unified* perspective; the magnitude of this change can be seen simply by noting that no one would seriously argue today that linear programming is independent of nonlinear programming.

Beyond a broadened perspective, one might wonder whether the revolution has made a substantive difference: is the net result simply that the log barrier method was rediscovered and applied to new problems? The answer to this is an emphatic "No". As we shall try to indicate in the remainder of the paper, there have been fundamental advances in complexity theory, algorithms, linear algebra, and solvable problems, all as a result of the interior revolution.

4.2 Complexity

A signature of interior methods is the existence of continuously parameterized families of approximate solutions that asymptotically converge to the exact solution; see, for example, [20]. As the parameter approaches its limit, these paths trace smooth trajectories with geometric properties (such as being "centered" in a precisely defined sense) that can be analyzed and exploited algorithmically. These paths also play a critical role in complexity analyses of interior algorithms.

The elements in a typical proof of polynomial complexity for an interior method are:

- Characterizing acceptable closeness to the solution through a stopping rule. Such a rule is needed because an interior method that generates strictly feasible iterates cannot produce, within a finite number of iterations, a solution that lies exactly on a constraint;
- Defining a computable measure of closeness to the parameterized path associated with the problem and the algorithm;
- Showing that a Newton step, or a suitably small number of Newton steps, taken from a point close to the path will stay sufficiently close to the path;
- Decreasing the controlling parameter at a rate that allows a polynomial upper bound on the number of iterations needed to become close enough to the solution.

Innumerable papers have been written about complexity issues in interior methods; the surveys [19, 30, 35, 44] (among others) provide details and further references.

Every discussion of the analysis of interior methods should pay tribute to the work of Nesterov and Nemirovskii, whose work in the late 1980s extended the scope of polynomial-time complexity results to a wide family of convex optimization problems; for details, see [27]. One of their major contributions was to define *self-concordant barrier functions*. A convex function ϕ from a convex region $\mathcal{F}^0 \in \mathcal{R}^n$ to $\mathcal{R}$ is κ-self-concordant in $\mathcal{F}^0$ if (i) ϕ is three times continuously differentiable in $\mathcal{F}^0$ and (ii) for all $y \in \mathcal{F}^0$ and all $h \in \mathcal{R}^n$, the following inequality holds:

$$| \nabla^3\phi(y)[h,h,h] | \leq 2\kappa\big(h^T\nabla^2\phi(y)h\big)^{3/2},$$

where $\nabla^3\phi(y)[h,h,h]$ denotes the third differential of ϕ at y and h. The logarithmic barrier functions associated with linear and convex quadratic programming are self-concordant with $\kappa = 1$. Existence of a self-concordant barrier function for a convex problem is closely related to existence of polynomial-time algorithms. Using the concept of self-concordance, new barrier functions have been devised for certain convex programming problems, such as semidefinite programming, that were previously considered computationally intractable; see Section 5.

Despite the polynomial bounds typically associated with interior methods, a mystery remains similar to that still holding for the simplex method: interior methods almost invariably require a number of iterations that is much smaller than the (very large) polynomial upper bound. The reasons for these disparities are not yet understood, but perhaps one day they will be.

4.3 Barrier methods revisited

The problem of ill-conditioning, as noted earlier, has haunted interior methods since the late 1960s, but there has been substantial recent progress in understanding this issue. A detailed analysis was given in [37] of the structure of the primal barrier Hessian (15) in an entire neighborhood of the solution. Several papers ([13], [11], [40], [42]) have analyzed the stability of specific factorizations for various interior methods.

Very recently, the (at first) surprising result was obtained ([39], [43]) that, under conditions normally holding in practice, ill-conditioning of certain key matrices in interior methods for nonlinear programming does not noticeably degrade the accuracy of the computed search directions. In particular, in modern primal-dual methods (see Section 4.4.1), if a backward-stable method is used to solve the condensed primal-dual system (the analogue of the barrier Hessian), the computed solution has essentially the same accuracy as that of the well-conditioned full primal-dual system (33). However, this result crucially depends on the special structure of the relevant ill-conditioned matrices, in particular their asymptotic relationship with $\hat{A}$ and the reduced Hessian of the Lagrangian. A similar kind of analysis applies to the search direction computed with the primal barrier Hessian. Consequently, ill-conditioning in interior methods undeniably exists, but will tend to be benign.

It turns out that ill-conditioning is not the only defect of primal barrier methods. Even if the Newton direction is calculated with perfect accuracy, primal barrier methods suffer from inherently poor scaling of the search direction during the early iterations following a

reduction of the barrier parameter; see [38, 6]. Thus, unless special precautions are taken, a full Newton step cannot be taken immediately after the barrier parameter is reduced. This fundamentally undesirable property implies that the classical primal barrier method will be unavoidably inefficient.

A fascinating but unresolvable question is whether the loss of popularity of barrier methods in the 1970s was unfairly blamed on ill-conditioning (which is often a genuine villain); the observed inefficiencies were probably attributable to the just-mentioned flaw of the primal barrier method rather than to ill-conditioning.

4.4 New algorithms for old problems

Leading candidates for the most popular algorithms to emerge from the interior revolution belong to the *primal-dual* family. Although there is no precise, universally accepted definition of a primal-dual method, these methods are almost always based on applying Newton's method to nonlinear equations stated in terms of the original ("primal") problem variables, along with "dual" variables representing the Lagrange multipliers.

4.4.1 Primal-dual methods for linear programming

The optimal solution x of the barrier subproblem (21) for a standard-form LP satisfies the condition $c = A^T y + \mu X^{-1}\mathbf{1}$ for some m-vector y (see (23)). Defining the n-vector z as $\mu X^{-1}\mathbf{1}$, we may replace this condition by the following two equations:

$$c = A^T y + z \quad \text{and} \quad Xz = \mu e. \tag{28}$$

The second relation in (28) has a clear resemblance to the perturbed complementarity condition (18) that holds along the barrier trajectory between the inequality constraints (here, the variables x) and Lagrange multiplier estimates.

The primal Newton barrier algorithm described in Section 3.2 is formulated in terms of only primal variables x; the Lagrange multiplier estimate y of (25) arises as a byproduct of the equality-constrained Newton subproblem. One could alternatively seek primal variables x and dual variables y (for the equalities) and z (for the inequalities) satisfying the central-path conditions (24) rewritten to include z:

$$Ax = b, \;\; x > 0, \;\; A^T y + z = c, \;\; z > 0, \;\; \text{and} \;\; Xz = \mu\mathbf{1}. \tag{29}$$

Note that only the third equation in (29) is nonlinear.

Applying Newton's method (6) to these $2n+m$ equations, we obtain the following linear system for Newton steps in x, y, and z:

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} b - Ax \\ c - A^T y - z \\ \mu e - XZe \end{pmatrix}. \tag{30}$$

Eliminating p_z and p_x gives the linear system

$$AZ^{-1}XA^T p_y = AZ^{-1}X(c - \mu X^{-1}e - A^T y) + b - Ax, \tag{31}$$

where $AZ^{-1}XA^T$ is symmetric and positive definite, with the form AD^2A^T for a nonsingular diagonal matrix D. Once p_y is known, p_z and p_x may be calculated directly from the second and third block rows of (30) without solving any equations.

Primal-dual methods for linear programming have been enormously successful in practice. For a detailed discussion of many aspects of primal-dual methods, see [41].

A striking effect of the interior revolution has been the magnitude and extent of performance improvements in the simplex method, which was (wrongly) thought in 1984 to have already reached its speed limit. In LP today, interior methods are faster than simplex for some very large problems, the reverse is true for some problems, and the two approaches are more or less comparable on others; see [2]. Consequently, commercial LP codes routinely offer both options. Further analysis is still needed of the problem characteristics that determine which approach is more appropriate. Unless a drastic change occurs, both approaches are likely to remain viable into the foreseeable future.

4.4.2 Primal-dual methods for nonlinear programming

Because of inherent flaws in primal barrier methods (see Section 4.3), primal-dual methods based on properties of x_μ are increasingly popular for solving general nonlinear programming problems; see, for example, the recent papers [8, 4, 11, 7, 14]. As in primal-dual methods for LP, the original (primal) variables x and the dual variables λ (representing the Lagrange multipliers) are treated as independent.

The usual motivation for primal-dual methods is to find (x, λ) satisfying the equations that hold at x_μ. In the spirit of (16) and (17), $(x_\mu, \lambda_\mu(x_\mu))$ satisfy the following $n + m$ nonlinear equations:

$$g = A^T\lambda \quad \text{and} \quad c_i\lambda_i = \mu, \quad i = 1, \ldots, m. \tag{32}$$

Applying Newton's method, we obtain the (full) $n + m$ *primal-dual equations* for Newton steps in x and λ:

$$\begin{pmatrix} W & -A^T \\ \Lambda A & C \end{pmatrix} \begin{pmatrix} p_x \\ p_\lambda \end{pmatrix} = \begin{pmatrix} -g + A^T\lambda \\ \mu\mathbf{1} - C\lambda \end{pmatrix} \tag{33}$$

where W is the Hessian of the Lagrangian evaluated at (x, λ).

All primal-dual methods are based on more or less the idea just described, which is sometimes presented in terms of the logarithmic barrier function (hence leading to properties of x_μ), or else in terms of perturbed complementarity (18) as a desired property in itself. Naturally, the equations (33) do not begin to constitute a complete algorithm for nonlinear programming. Primal-dual methods are the object of active research today, and span a wide range of approaches to algorithmic details, including

1. formulation of the constraints;
2. solution of the linear system that defines the Newton steps;
3. treatment of indefiniteness;
4. strategies for encouraging progress toward the solution from an arbitrary starting point;

5. treatment of equality constraints (an option needed for a general-purpose nonlinear programming method).

4.5 Linear algebra

Interior methods would not be fast or reliable without efficient, numerically stable linear algebraic techniques for solving the associated distinctive, specially structured linear systems. Great advances have taken place since 1984 in sparse Cholesky-based techniques for factorizing matrices of the form $A^T D^2 A$, where D is diagonal and is becoming ill-conditioned in a specified manner—either some elements of D are becoming infinite while the others are $\Theta(1)$, or else some are approaching zero while the remainder are $\Theta(1)$. In addition, techniques for sparse symmetric indefinite factorizations of matrices of the form

$$\begin{pmatrix} W & A^T \\ A & D^2 \end{pmatrix}, \tag{34}$$

where D is diagonal and ill-conditioned as just described, are important. See, for example, [11, 13, 40, 42].

5 New Problems

The flowering of interior methods, and in particular the realization that efficient algorithms exist for a wide class of convex optimization problems, has led to the application of interior methods to a broad range of problems that were previously considered to be computationally intractable.

Certain problems involving *eigenvalue optimization* have been particularly amenable to solution by interior methods; for details, see the excellent survey [22]. In the next section we summarize a few key ideas in *semidefinite programming* (SDP), an area of intense research during the past few years.

5.1 The semidefinite programming problem

Semidefinite programming may be viewed as a generalization of linear programming, where the variables are $n \times n$ symmetric matrices, denoted by X, rather than n-vectors. In SDP, we wish to minimize an affine function of a symmetric matrix X subject to linear constraints and semidefiniteness constraints, the latter requiring (in words) that "X must be positive semidefinite". This relation is typically written as $X \succeq 0$, a form that strongly resembles inequality constraints in ordinary continuous optimization. (When X is a symmetric matrix, the condition $X \succ 0$ means "X is positive definite".)

Let $\mathcal{S}^n$ denote the set of real $n \times n$ symmetric matrices, let C and $\{A_i\}$ be real symmetric $n \times n$ matrices, and let b be a real m-vector. The semidefinite programming problem is the following:

$$\begin{aligned} \underset{X \in \mathcal{S}^n}{\text{minimize}} \quad & \operatorname{trace}(CX) && (35) \\ \text{subject to} \quad & \operatorname{trace}(A_i X) = b_i, \quad i = 1, \ldots, m && (36) \\ & X \succeq 0. && (37) \end{aligned}$$

When the SDP problem is written in this form, its similarity to a standard-form LP (2) is hard to miss, but, not surprisingly, many extra complications arise in SDP. For example, the feasible region defined by (36) and (37) is not polyhedral, so there is no analogue of the simplex method. Furthermore, several major regularity assumptions are needed to obtain duality results analogous to those in LP. These assumptions will not be stated here; see [22] for details.

Nesterov and Nemirovskii [27] show that the function $\log\det(X)$ is a self-concordant barrier function for the semidefinite programming problem, which means that the SDP (35)–(37) can be solved in polynomial time via a sequence of barrier subproblems parameterized by μ:

$$\underset{X\in\mathcal{S}^n}{\text{minimize}} \qquad \text{trace}(CX) - \mu \log\det X \tag{38}$$

$$\text{subject to} \qquad \text{trace}(A_i X) = b_i, \quad i = 1, \ldots, m. \tag{39}$$

Under suitable regularity assumptions, there is a unique sequence $\{X_\mu, y_\mu\}$, where X_μ is a symmetric positive definite matrix satisfying the constraints (36) and (37), and y_μ is an m-vector, such that X_μ and y_μ together satisfy the following "perturbed complementarity" condition

$$X(C - \sum_{i=1}^{m} y_i A_i) = \mu I, \tag{40}$$

with $C - \sum_{i=1}^m y_i A_i \succeq 0$. Newton's method cannot be applied directly to solve (36) and (40) because the matrix on the left-hand side of (40) is not symmetric. A primal approach, first suggested in [1], is to replace (40) by the relation

$$X(C - \sum y_i A_i) + (C - \sum y_i A_i)X = 2\mu I.$$

An analogous primal-dual method, called the "$XZ + ZX$ method" for obvious reasons, is defined by finding (X_μ, y_μ, Z_μ), where $X_\mu \succ 0$ and $Z_\mu \succ 0$, such that

$$\text{trace}(A_i X) = b_i, \quad Z = C - \sum y_i A_i, \quad \text{and} \quad XZ + ZX = 2\mu I. \tag{41}$$

Note the strong parallel between the two final equations in (41) and the primal-dual equations (28) in linear programming.

Semidefinite programming is an extremely lively research area today, producing new theory, algorithms, and implementations; see the surveys [33] and [34].

5.2 New applications of interior methods

Interior methods are playing major roles in at least two areas: approximation techniques for NP-hard combinatorial problems, and system and control theory.

In the former, it has recently been shown that certain semidefinite programs and NP-hard problems are closely related in the following way: solution of the semidefinite program leads to an approximation whose objective value is provably within a known factor of the optimal objective value for the associated NP-hard problem. For example, a semidefinite program formulation leads to an approximate solution of the max-cut problem whose objective value is within a factor of 1.14 of the optimal value; see [18]. This kind of relationship

guarantees that good approximate solutions to NP-hard problems can be computed in polynomial time.

Interior methods are important in system and control theory because of their connection with *linear matrix inequalities*, which have the forms

$$F_0 + \sum_{i=1}^{p} x_i F_i \succ 0 \quad \text{or} \quad F_0 + \sum_{i=1}^{p} x_i F_i \succeq 0, \tag{42}$$

where x is a p-vector and $\{F_i\}$ are real symmetric matrices. Many constraints in system and control theory, including convex quadratic inequalities, matrix norm inequalities, and Lyapunov matrix inequalities, can be expressed as linear matrix inequalities. It is straightforward to see that the forms (42) allow the variables to be symmetric matrices.

Numerous problems in system and control theory involve optimization of convex functions of matrix arguments subject to linear matrix inequalities. Because these are convex programming problems, it is possible to apply polynomial-time interior methods. For details, the reader should consult [3].

6 Summary

The interior point revolution has had many highly positive results, including

- a deeper and more unified understanding of constrained optimization problems;
- continuing improvements to theory and methods;
- more algorithmic options for familiar problems, even for linear programming;
- the ability to solve new problems.

One could argue, however, not entirely whimsically, that the interior-point revolution has had some negative consequences. For example, both teaching linear programming and solving linear programs are much more complicated than they used to be. With respect to the former, instructors in linear programming courses face increased pedagogical challenges. Before 1984, it was perfectly acceptable simply to describe the simplex method; today, any credible treatment of linear programming needs to include interior methods. Similarly, someone with an LP to solve can no longer be content with mindless application of the simplex method.

On balance, the interior revolution has energized and expanded the field of constrained optimization. Although the revolutionary pace has (inevitably) slowed down since its first heady days, ample opportunities remain for many further years of lively and innovative research.

References

[1] F. Alizadeh, J.-P. Haeberly, and M. L. Overton (1998). Primal-dual interior-point methods for semidefinite programming: convergence rates, stability, and numerical results, *SIAM J. Opt.* 8, 746–768.

[2] R. E. Bixby (2002). Solving real-world linear programs: a decade or more of progress, *Operations Research* 50, 3–15.

[3] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan (1994). *Linear Matrix Inequalities in System and Control Theory*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

[4] R. H. Byrd, J. C. Gilbert, and J. Nocedal (2000). A trust region method based on interior point techniques for nonlinear programming, *Math. Prog.* 89, 149–185.

[5] V. Chvátal (1983). *Linear Programming*, W. H. Freeman, New York.

[6] A. R. Conn, N. I. M. Gould, and P. L. Toint (1994). A note on using alternative second-order models for the subproblems arising in barrier function methods for minimization, *Num. Math.* 68, 17–33.

[7] A. R. Conn, N. I. M. Gould, and P. L. Toint (2000). A primal-dual trust-region algorithm for minimizing a non-convex function subject to bound and linear equality constraints, *Math. Prog.* 87, 215–249.

[8] A. S. El-Bakry, R. A. Tapia, T. Tsuchiya, and Y. Zhang (1996). On the formulation and theory of the Newton interior-point method for nonlinear programming, *J. Opt. Theory Appl.* 89, 507–541.

[9] A. V. Fiacco and G. P. McCormick (1968). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley and Sons (New York). Republished by Society for Industrial and Applied Mathematics, Philadelphia, 1990.

[10] R. Fletcher (1987). *Practical Methods of Optimization* (second edition), John Wiley and Sons, Chichester.

[11] A. Forsgren and P. E. Gill (1998). Primal-dual interior methods for nonconvex nonlinear programming, *SIAM J. Opt.* 8, 1132–1152.

[12] A. Forsgren, P. E. Gill, and M. H. Wright (2002). Interior methods for nonlinear optimization, *SIAM Review* 44, 525–597.

[13] A. Forsgren, P. E. Gill, and J. R. Shinnerl (1996). Stability of symmetric ill-conditioned systems arising in interior methods for constrained optimization, *SIAM J. Matrix Anal. Appl.* 17, 187–211.

[14] D. M. Gay, M. L. Overton, and M. H. Wright (1998). A primal-dual interior method for nonconvex nonlinear optimization, in *Advances in Nonlinear Programming*, (Y. Yuan, ed.), Kluwer Academic, Dordrecht, 31–56.

[15] P. E. Gill, W. Murray, M. A. Saunders, J. A. Tomlin and M. H. Wright (1986). On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method, *Math. Prog,* 36, 183–209.

[16] P. E. Gill, W. Murray and M. H. Wright (1981). *Practical Optimization*, Academic Press, London and New York.

[17] P. E. Gill, W. Murray and M. H. Wright (1991). *Numerical Linear Algebra and Optimization, Volume 1*, Addison-Wesley, Redwood City.

[18] M. X. Goemans and D. P. Williamson (1995). Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *J. ACM* 42, 1115–1145.

[19] D. Goldfarb and M. J. Todd (1989). Linear programming, in *Optimization* (G. L. Nemhauser, A. H. G. Rinnooy Kan and M. J. Todd, eds.), North Holland, Amsterdam and New York, 73–170.

[20] C. C. Gonzaga (1992). Path following methods for linear programming, *SIAM Review* 34, 167–224.

[21] N. K. Karmarkar (1984). A new polynomial time algorithm for linear programming, *Combinatorica* 4, 373–395.

[22] A. S. Lewis and M. L. Overton (1996). Eigenvalue optimization, *Acta Numerica 1996*, 149–190.

[23] F. A. Lootsma (1969). Hessian matrices of penalty functions for solving constrained optimization problems, *Philips Res. Repts.* 24, 322–331.

[24] D. G. Luenberger (1973). *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Menlo Park.

[25] W. Murray (1971). Analytical expressions for the eigenvalues and eigenvectors of the Hessian matrices of barrier and penalty functions, *J. Opt. Theory Appl.* 7, 189–196.

[26] B. A. Murtagh and M. A. Saunders (1987). MINOS 5.1 User's Guide, Report SOL 83-20R, Department of Operations Research, Stanford University, Stanford, California.

[27] Y. Nesterov and A. Nemirovskii (1994). *Interior-Point Polynomial Algorithms in Convex Programming*, Society for Industrial and Applied Mathematics, Philadelphia.

[28] J. Nocedal and S. J. Wright (1999). *Numerical Optimization*, Springer-Verlag, New York.

[29] J. M. Ortega and W. C. Rheinboldt (1970). *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, London and New York.

[30] C. Roos, T. Terlaky, and J.-Ph. Vial (1997). *Theory and Algorithms for Linear Optimization: An Interior Point Approach*, John Wiley & Sons, New York.

[31] A. Schrijver (1987). *Theory of Linear and Integer Programming*, John Wiley and Sons, New York.

[32] D. A. Spielman and S.-H. Tang (2001). Why the simplex method usually takes polynomial time, in *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, 296–305.

[33] M. J. Todd (2001). Semidefinite optimization, *Acta Numerica 2001*, 515–560.

[34] L. Vandenberghe and S. Boyd (1996). Semidefinite programming, *SIAM Review* 38, 49–95.

[35] R. J. Vanderbei (1997). *Linear Programming: Foundations and Extensions*, Kluwer Academic Publishers, Boston.

[36] M. H. Wright (1992). Interior methods for constrained optimization, *Acta Numerica 1992*, 341–407.

[37] M. H. Wright (1994). Some properties of the Hessian of the logarithmic barrier function, *Math. Prog.* 67, 265–295.

[38] M. H. Wright (1995). Why a pure primal Newton barrier step may be infeasible, *SIAM J. Opt.* 5, 1–12.

[39] M. H. Wright (1998). Ill-conditioning and computational error in interior methods for nonlinear programming, *SIAM J. Opt.* 9, 81-111.

[40] S. J. Wright (1995). Stability of linear equation solvers in interior-point methods, *SIAM J. Matrix Anal. Appl.* 16, 1287–1307.

[41] S. J. Wright (1997). *Primal-Dual Interior-Point Methods*, Society for Industrial and Applied Mathematics, Philadelphia.

[42] S. J. Wright (1999). Modified Cholesky factorizations in interior-point algorithms for linear programming, *SIAM J. Opt.* 9, 1159–1191.

[43] S. J. Wright (2001). Effects of finite-precision arithmetic on interior-point methods for nonlinear programming, *SIAM J. Opt.* 12, 36–78.

[44] Y. Ye (1997). *Interior Point Algorithms, Theory and Analysis*, John Wiley & Sons, New York.

WHAT IS MOTIVIC MEASURE?

THOMAS C. HALES

Abstract. This article gives an exposition of the theory of arithmetic motivic measure, as developed by J. Denef and F. Loeser.

1. Preliminary Concepts

There is much that is odd about motivic measure if it is judged by measure theory in the sense of twentieth century analysis. It does not fit neatly with the tradition of measure in the style of Hausdorff, Haar, and Lebesgue. It is best to view motivic measure as something new and different, and to recognize that when it comes to motivic measure, the term 'measure' is used loosely.

Motivic measure will be easier to understand, once two of its peculiarities are explained. The first peculiarity is that the measure is not real-valued. Rather, it takes values in a scissor group. An introductory section on scissor groups for polygons will recall the basic facts about these groups. The second peculiarity is that rather than a boolean algebra of measurable sets, we work directly with the underlying boolean formulas that define the sets. The reasons for working directly with boolean formulas will be described in a second introductory section.

After these two introductory remarks, we will describe 'motivic counting' in Section 2. Motivic counting is to ordinary counting what motivic measure is to ordinary measure. Motivic counting will lead into motivic measure.

1.1. **Scissor Groups for polygons.** Motivic volume is defined by a process that is similar to the scissor-group construction of the area of polygons in the plane. To draw out the similarities, let us recall the construction. It determines the area of a polygons without taking limits.

Any polygon in the plane can be cut into finitely many triangles that can be reassembled into a rectangle of unit width. Figure 1 illustrates three steps (2, 3, and 4) of the general algorithm. The algorithm consists of 5 elementary transformations. (1) Triangulate the polygon. (2) Transform triangles into rectangles. (3) Fold long rectangles in half. (4) Rescale each rectangle to give it an edge of unit width. (5) Stack all the unit width rectangles end to end. The length of the unit width rectangle is the area.

An abelian group encodes these cut and paste operations. Let F be the free abelian group on the set of polygons in the plane.

We impose two families of relations:

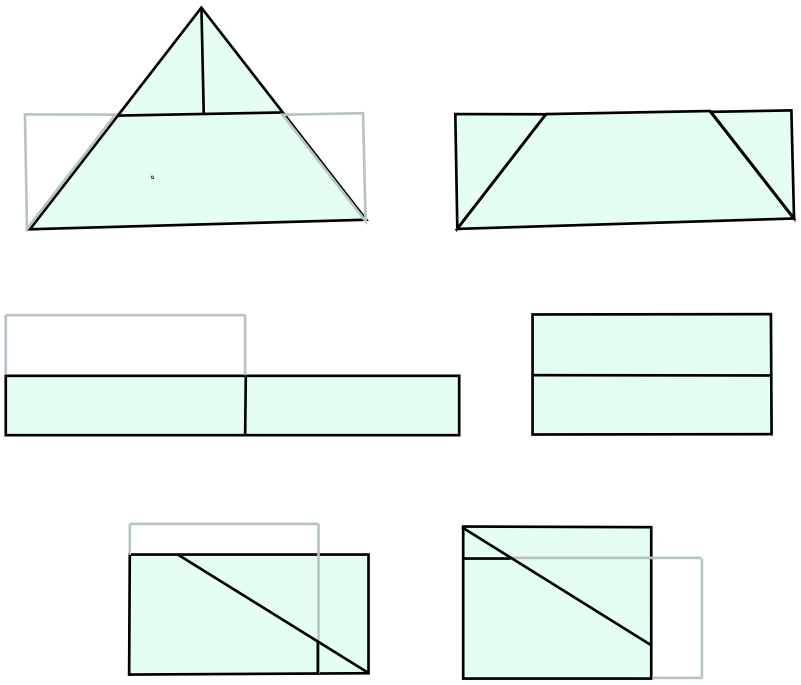


FIGURE 1. Triangles transform into unit width rectangles by scissor and congruence relations. Later, we will transform ring formulas into algebraic varieties by scissor and congruence relations.

Scissor relations. If P is a polygon that can be cut into polygons P_1 and P_2, then

$$[P] = [P_1] + [P_2]$$

Congruence relations. If P and P' are congruent polygons then

$$[P] = [P'].$$

The scissor group $\mathbb{S}_{\text{poly}}$ of polygons is defined as the free abelian group subject to these two families of relations. In some sense, this entire article is an exploration of scissor and congruence relations in diverse contexts. By and by, we will construct several closely related scissor groups $\mathbb{S}_{\text{poly}}$, $\mathbb{S}_{\text{count}}$, $\mathbb{S}_{\text{ring}}$, $\mathbb{S}_{\text{cover}}$, and $\mathbb{S}_{\text{mot}}$, each constructed as a free abelian group modulo scissor and congruence relations.

Theorem 1.1. *The polygon scissor group $\mathbb{S}_{\text{poly}}$ of polygons is isomorphic to the additive group of real numbers $\mathbb{R}$. Under this isomorphism, the real number attached to the class $[P]$ of a polygon is its area.*

Proof. A group homomorphism from $\mathbb{S}_{\text{poly}}$ to $\mathbb{R}$ sends each class $[P]$ to its area. It is onto, because there are polygons of every positive real area, and negations of polygons of every negative real area. By scissor and congruence

relations, every element of the scissor group is represented by the difference of two unit width rectangles. To be in the kernel, the two rectangles must have the same area; but then they are congruent, and their difference is zero element of $\mathbb{S}_{\text{poly}}$. Thus, the homomorphism is also one-to-one. □

The area function on the set $\{P\}$ of polygons thus factors through $\mathbb{S}_{\text{poly}}$.

$$\begin{array}{ccccc} \{P\} & \to & \mathbb{S}_{\text{poly}} & \to & \mathbb{R} \\ P & \mapsto & [P] & \mapsto & \text{area}(P) \end{array} \tag{1}$$

We might ponder which of these two maps ($P \mapsto [P]$ or $[P] \mapsto \text{area}(P)$) captures the greater part of the area-taking process. Motivic measure commits to a position on this issue: the first stage ($P \mapsto [P]$) is identified as the area-taking process and the second stage $[P] \mapsto \text{area}(P)$ is a *specialization* of the area. In this case, specialization is an isomorphism. Our approach to measure in this article is decidedly unsophisticated: taking the measure of something consists in mapping that thing into its scissor group, $P \mapsto [P]$.

1.2. **The measure of a formula.** Traditionally, we take the measure of a set $X = \{X \mid \phi(x)\}$ (say a subset of a locally compact space), but we do not take the measure of the formula ϕ defining a set. With motivic measure, we take the measure of the formula directly. Concretely, the formula

$$\text{`}x^2 + y^2 = 1\text{'} \tag{2}$$

defines the circle

$$\{(x, y) \mid x^2 + y^2 = 1\}. \tag{3}$$

With motivic measure, we take the measure of the equation of the circle (Equation 2) rather than the measure of the circle itself (Equation 3). Attention shifts from sets to formulas.

What purpose does it serve to measure formulas rather than the underlying set? As algebraic geometers are eager to remark, each formula defines an infinite collection of sets. For instance, for each finite field $\mathbb{F}_q$, we can take the set of $\mathbb{F}_q$ points on the circle:

$$\{(x, y) \in \mathbb{F}_q^2 \mid x^2 + y^2 = 1\}. \tag{4}$$

We will see that the motivic measure of the formula is a universal measure in the sense that the value it attaches to the formula does not commit us to any particular field. And yet if we are supplied with a particular field, it will be possible to recover the traditional measure of a set from the motivic measure of its defining formula. In this sense, motivic measure is to traditional measures what an algebraic variety is to its set of solutions.

2. Counting measures and Finite Fields

Counting is the fountainhead of all measure. The measure of a finite set is its cardinality. At the risk of belaboring the point, in preparation for what

is to come, let us recast ordinary counting. The scissor relation for disjoint finite sets is

$$[X \cup Y] = [X] + [Y].$$

More generally, if we allow the sets to intersect, it is

$$[X \cup Y] = [X] + [Y] - [X \cap Y]. \tag{5}$$

The congruence relation asserts that

$$[X] = [X'].$$

whenever there is a bijection between X and X'. The scissor group $\mathbb{S}_{\text{count}}$ is the quotient of the free abelian group on finite sets satisfying the scissor and congruence relations. It is is isomorphic to $\mathbb{Z}$. The cardinality $\#X$ of a finite set X factors through the scissor group

$$X \mapsto [X] \mapsto \#[X] \in \mathbb{Z}.$$

Of course, if our only purpose were to count elements in finite sets, this construction is overkill. The first motivic measure that we present is an analogue of this approach to counting. We call it the *motivic counting measure*. The scissor relation will be similar to Equation 5.

2.1. **Ring formulas.** Traditional measure calls for a full discussion of the class of measurable sets. Since we work with formulas rather than sets, our approach calls for a full discussion of the class of formulas to be measured.

We allow all syntactically correct formulas built from a countable collection of variables x_i, parentheses, and the symbols

$$\forall, \quad \exists, \quad \vee, \quad \wedge, \quad \neg, \quad 0, \quad 1, \quad (+), \quad (-), \quad (*), \quad (=) \tag{6}$$

More precisely, we allow all formulas in the first-order language of rings. A formula that has been constructed from these symbols will be called a *ring formula*. We avail ourselves of the usual mathematical abbreviations and renamings of variables. We write 3 for $1 + (1 + 1)$, x^n for $x * x * x \cdots * x$ (n times), xy for $x * y$, $a + b + c$ for $a + (b + c)$, and so forth.

With usual abbreviations,

$$`\forall x\ y\ z.\ x^3 + y^3 = z^3,$$

is a ring formula, because its syntax is correct. But

$$`))\forall + \forall = 2\forall(('$$

and

$$`\ \wedge \vee \wedge \vee \wedge\ '$$

are not ring formulas.

2.2. **The scissor group of ring formulas.** We imitate the construction of the scissor groups $\mathbb{S}_{\text{poly}}$ and $\mathbb{S}_{\text{count}}$ to build the scissor group of ring formulas.

Take the free abelian group on the set of ring formulas.

We impose two families of relations. The scissor relation takes the form established in Equation 5 for unions.

Scissor relations. If $\phi_1 \vee \phi_2$ is a disjunction of two formulas, then

$$[\phi_1 \vee \phi_2] = [\phi_1] + [\phi_2] - [\phi_1 \wedge \phi_2]. \tag{7}$$

To describe the congruence relation, we must decide what it should mean for two ring formulas to be congruent. By way of analogy, in the case of polygons, two are congruent if there is a bijection between the two sets that is induced by an isometry. Our first guess at the congruence relation for ring formulas is that two ring formulas are congruent if there is a bijection between the sets of solutions for each finite field $\mathbb{F}_q$. (We limit ourselves to *finite* fields because we are attempting to imitate the counting measure of *finite* sets.) However, there are two modifications that we must make to this first guess to arrive at a workable relation.

The first modification is to use pseudo-finite fields rather than finite fields. A *pseudo-finite field* is an infinite perfect field such that every absolutely irreducible variety over the field has a rational point and such that there is a unique field extension of each finite degree (inside a fixed algebraic closure of the field). The defining properties of a pseudo-finite field are properties possessed by finite fields (except the part about being infinite). Moreover, logicians have found that the behavior of pseudo-finite fields is essentially no different from the generic behavior of finite fields, but they avoid the hassles that appear in positive characteristic. For those seeing pseudo-finite fields for the first time, it would not be a severe distortion of the facts to ignore the 'pseudo' and to work instead with finite fields.

The second modification is to require that the bijection between the solutions come from a ring formula that is independent of the underlying field. We are now ready to state the congruence relations.

Congruence relations.

$$[\phi] = [\phi']$$

if there exists a ring formula ψ such that for every pseudo-finite field K of characteristic zero, the interpretation of ψ gives a bijection between the tuples in K satisfying ϕ and the tuples in K satisfying ϕ'.

Example 2.1. The congruence relation gives

$$[\text{`}\exists x. \quad x^2 + bx + c = 0\text{'}] = [\text{`}\exists X. \quad X^2 = B^2 - 4C\text{'}]$$

The formula ψ realizing the congruence and the bijection at the level of points is

$$\text{`}(b = B) \wedge (c = C)\text{'}.$$

That is, in every pseudo-finite field of characteristic zero, a monic quadratic polynomial has a root if and only if its discriminant is a square.

Definition 2.2. The scissor group $\mathbb{S}_{\text{ring}}$ of ring formulas is defined as the free abelian group subject to the scissor and congruence relations.

2.2.1. *Counting measure.*

Definition 2.3. The *counting measure* of a ring formula ϕ is its class $[\phi]$ in the scissor group of ring formulas.

2.2.2. *Fubini and Products.* There is a trivial sort of Fubini theorem for finite sets: the cardinality of a Cartesian product of two sets is the product of the cardinalities of the two sets. To make sense of a Fubini theorem for ring formulas, it is necessary to introduce products to the scissor groups; that is, we need a *scissor ring*. This is easy to arrange. If $\phi_1(x_1, \ldots, x_n)$ is a formula with free variables $x_1, \ldots, x_n$ and $\phi_2(y_1, \ldots, y_m)$ is a formula with free variables $y_1, \ldots, y_m$, and if the free variables of ϕ_1 are distinct from the free variables of ϕ_2, then we declare the product to be

$$\phi_1(x_1, \ldots, x_n) \wedge \phi_2(y_1, \ldots, y_m).$$

This induces a well-defined product[1] on the scissor group

$$[\phi_1(x)][\phi_2(y)] = [\phi_1(x) \wedge \phi_2(y)]. \tag{8}$$

Under this product, the scissor group becomes a ring. Equation 8 asserts that counting measure satisfies a rather trivial Fubini theorem for ring formulas – at least for ring formulas without any shared free variables.

2.2.3. *The universal nature of the counting measure.* The counting measure $[\phi]$ of a ring formula ϕ is designed to be the *universal counting measure* for ring formulas. For every finite field $\mathbb{F}_q$, there is a counting measure on ring formulas:

$$\phi \mapsto \#_q(\phi) = \#\{(x_1, \ldots, x_n) \in \mathbb{F}_q^n \mid \phi^{\mathbb{F}_q}(x_1, \ldots, x_n)\}. \tag{9}$$

It gives the number of solutions to the ring formula over a particular finite field. In contrast, the counting measure of a ring formula takes values in a scissor ring whose construction bundles all pseudo-finite fields together.

We can be precise about the way in which the counting measure bundles the counting measures $\#_q(\phi)$. Each formula ϕ gives a function $q \mapsto \#_q(\phi)$, an integer-valued function on the set of prime powers. Let F be the ring of all integer-valued functions on the set $\{p^r\}$ of prime powers. Declare two functions equivalent, if they take the same value at p^r for all r and for all but finitely many p. Write $F/\sim$ for the quotient of F under this equivalence relation.

[1]We have a moving lemma: the congruence relation on the scissor group can be used to relabel the free variables of a formula, so that free variables of the two factors are always distinct.

Theorem 2.4. *There exists a ring homomorphism N from the scissor ring $\mathbb{S}_{\text{ring}}$ to $F/\sim$ that respects counting: $\#_*(\phi) = N([\phi])$.*

In other words, with only a finite amount of ambiguity, the counting measure specializes to counting solutions to ring formulas over finite fields. To say that N is a ring homomorphism is to say that it is compatible with products and Fubini. Unlike the earlier isomorphisms for polygons $\mathbb{S}_{\text{poly}} \cong \mathbb{R}$ and finite sets $\mathbb{S}_{\text{count}} \cong \mathbb{Z}$, here we make no claim of isomorphism between the scissor group $\mathbb{S}_{\text{ring}}$ and the target ring $F/\sim$.

The proof of the theorem relies on ultraproducts, a standard tool in logic.

2.3. **Improving the scissor ring.** The shortcoming of the scissor ring $\mathbb{S}_{\text{ring}}$ is that is too much about it has been left inexplicit. In our discussion of the area of planar polygons, we found a handy set of generators (unit width rectangles). Our current aim is to find a handy set of generators of a somewhat modified scissor ring $\mathbb{S}_{\text{mot}}$. The idea is to take a ring formula, and through a process of "quantifier elimination" arrive at an equivalent ring formula that does not involve any quantifiers (that is, the symbols $\forall$, $\exists$ will be eliminated). A formula without quantifiers belongs less to the realm of logic than to the realm of algebraic geometry. A suggestive example of a quantifier-free formula is

$$(f_1 = 0) \wedge (f_2 = 0) \wedge \cdots \wedge (f_n = 0).$$

That is, the zero set of an affine variety. In fact, we will find that the improved scissor ring is defined as a quotient of the free abelian group on the set of varieties over $\mathbb{Q}$. The details of this construction will reveal what is so *motivic* about motivic measure.

2.4. **A scissor ring for coverings.** Each ring homomorphism $f : \mathbb{S}_{\text{ring}} \to R$ defines a specialization of the counting measure

$$\phi \mapsto [\phi] \to f[\phi] \in R.$$

The ring $F/\sim$ is one of many possible specializations R.

Another specialization of $\mathbb{S}_{\text{ring}}$ comes from n-sheeted covers:

Definition 2.5. We say that one formula $\phi(x)$ is an *n-sheeted cover* of another formula $\phi'(x')$ if there exists a ring formula $\psi(x, x')$ such that for every pseudo-finite field of characteristic zero, ψ gives an n to 1 correspondence between the solutions x of $\phi(x)$ and the solutions x' of $\phi(x')$.

Example 2.6. Let $\phi(x)$ be the formula '$x \neq 0$' and let $\phi'(y)$ be the formula

$$\text{'}\exists\ z.\ (z^2 = y) \wedge (y \neq 0)\text{'}.$$

The formula $\psi(x, y)$ given by

$$\text{'}x^2 = y\text{'},$$

presents ϕ as a 2-sheeted cover of ϕ'.

The congruence condition for $\mathbb{S}_{\text{ring}}$ asserts that if ϕ is a 1-sheeted cover of ϕ', then they give the same class in $\mathbb{S}_{\text{ring}}$. A broader congruence condition can be given as follows.

Congruence (Covers). If ϕ is an n-sheeted cover of ϕ' for some n, then

$$[\phi] = n[\phi'].$$

We may form a new scissor ring $\mathbb{S}_{\text{cover}}$ with this broader congruence condition and the old scissor relation. We have a canonical surjection $\mathbb{S}_{\text{ring}} \to \mathbb{S}_{\text{cover}}$.

2.5. **The scissor group of motives. Generators**. Let $\text{Var}_{\mathbb{Q}}$ be the category of varieties over the field of rational numbers $\mathbb{Q}$. We take the free abelian group generated by the objects of $\text{Var}_{\mathbb{Q}}$.

An example of a element of the free abelian group is $[\mathbb{A}^1]$, the generator attached to the affine line. This particular generator will be of special importance in the constructions that follow. We write $\mathbb{L} = [\mathbb{A}^1]$ for this element and for its image in various scissor groups. (The 'L' is for *Lefschetz*, as in Lefschetz motive.)

There are two types of relations: scissor relations and congruence relations. Our scissor relation will be rather crude, but justifiably so, since the Zariski topology is a coarse topology that limits the possibilities for a scissor relation. The only cutting that will be permitted is that of partitioning a variety into a closed subvariety and its complement.

Scissor Relation. If Z is a closed subvariety of X, then

$$[X] = [Z] + [X \setminus Z].$$

The congruence relation is more involved than the scissor relation. If we make a direct translation of the congruence relation for the scissor group of ring formulas, we might guess that the congruence condition between two varieties X and Y should be the existence of a correspondence Ψ between X and Y that induces a bijection between $X(K)$ and $Y(K)$ for every pseudo-finite field of characteristic zero. This first guess is suggestive: the congruence relation should involve an algebraic correspondence. This suggestion lands us deep in the territory of motives. Here is the precise definition of the congruence relation.

Congruence Relation.

$$[X] = [Y]$$

whenever X and Y are nonsingular projective varieties that give the *same virtual Chow motive*. We will uncoil this definition a bit below. All that is 'motivic' about motivic measure stems from this particular congruence relation.

Definition 2.7. The quotient of the free abelian group by the scissor and congruence relations is the motivic scissor ring $\mathbb{K}$. (The letter 'K' is the standard notation for a Grothendieck group, which for our purposes is just another name for a scissor group.) The localized version $\mathbb{K}[\mathbb{L}^{-1}] \otimes \mathbb{Q}$ will be called the *localized motivic scissor ring* and denoted $\mathbb{S}_{\text{mot}}$. (It will become clear in Section 3.6.3 why it is useful to invert $\mathbb{L}$.)

It is time to uncoil the definition of this congruence relation. There is a category of Chow motives. To describe this category, we assume familiarity with the Chow groups $A^i(X)$ of a variety X. They are groups of cycles of a given codimension i modulo the subgroup of cycles that are rationally equivalent to 0. A detailed treatment of cycles, rational equivalence, and Chow groups can be found in [8]. Other good treatments of Chow motives can be found in [16] and [9].

An object in the category of Chow motives is a triple (X, p, m) where X is a smooth projective variety of dimension d, p is an element in the Chow ring $A^d(X \times X)$ that is a projector ($p^2 = p$), and m in an integer. The set of morphisms from (X, p, m) to (X, p', m') is defined to be the set

$$p' A^{d+n-m}(X \times Y) p.$$

Varieties that are not isomorphic as varieties can very well become isomorphic when viewed as Chow motives. For example, isogenous elliptic curves are isomorphic as Chow motives.

There is a canonical morphism from the Grothendieck ring of the category $\text{Var}_{\mathbb{Q}}$ to the Grothendieck ring of the category of Chow motives. We let $\mathbb{K}$ be the image of this morphism. To say that two varieties are equal as virtual Chow motives is to say that they have the same class in $\mathbb{K}$.

2.6. **The motivic counting measure.** The following theorem follows from a deep investigation of Chow motives, and the theory of quantifier elimination for pseudo-finite fields.

Theorem 2.8. *There exists a unique ring homomorphism* $\mathbb{S}_{\text{cover}} \to \mathbb{S}_{\text{mot}}$ *that satisfies the following property (Zero Sets).*

Zero Sets. If ϕ is a ring formula that is given by the conjunction of polynomial equations, then $[\phi]$ is sent to the affine variety defined by those polynomial equations.

There are ring homomorphisms $\mathbb{S}_{\text{count}} \to \mathbb{S}_{\text{cover}} \to \mathbb{S}_{\text{mot}}$. We use the notation $\phi \mapsto [\phi]$ for the class of ϕ in any of these rings, depending on the context.

Definition 2.9. The composite map $\phi \mapsto [\phi] \in \mathbb{S}_{\text{mot}}$ will be called the *motivic counting measure* of the formula ϕ.

The motivic counting measure of a ring formula is thus represented by a rational linear combination of varieties over $\mathbb{Q}$. I like to think of the motivic

counting measure as counting the number of solutions of the ring formula over finite fields in a way that does not depend on the finite field. Instead of giving the answer as a particular number, it gives the answer in terms of a formal combination of varieties having the same number of solutions over a finite field. Here is the precise statement.

Theorem 2.10. *Let ϕ be a ring formula, and let $\sum a_i[X_i]$ be a representative of the motivic counting measure $[\phi]$ as a formal linear combination of varieties. Choose a model of each X_i over $\mathbb{Z}$. For all r and for all but finitely many primes p, the number of solutions of ϕ in $\mathbb{F}_{p^r}$ is equal to*

$$\sum a_i \# X_i(\mathbb{F}_{p^r}).$$

Example 2.11. As an example, let us calculate the motivic counting measure of the 'set' of nonzero cubes. The formula is given by

$$\phi(x): \quad \text{`}\exists y.\ (y^3 = x)\ \wedge\ (x \neq 0)\text{'}.$$

The scissor relation can be used to break ϕ into two disjoint pieces $\phi = \phi_1 \vee \phi_2$: the part ϕ_1 on which -3 is a square and the part ϕ_2 on which it is not. Let $\mathbb{M}$ be the class in $\mathbb{S}_{\mathrm{mot}}$ corresponding to the zero-dimensional variety $x^2 + 3 = 0$. The class $\mathbb{M}$ has two solutions or no solutions according as -3 is a square or not. When -3 is a square, the cube roots of unity lie in the field, so that the nonzero points on the affine line give a 3-fold cover of ϕ_1 (under $y \mapsto y^3$). Thus, ϕ_1 has measure

$$\left(\frac{\mathbb{L}-1}{3}\right)\frac{\mathbb{M}}{2}.$$

On the other hand, if -3 is not a square, each non-zero element of a pseudo-finite field of characteristic zero is a cube, so that ϕ_2 has measure

$$(\mathbb{L}-1)\left(1-\frac{\mathbb{M}}{2}\right).$$

The sum of these two terms is the measure of ϕ in $\mathbb{S}_{\mathrm{mot}}$.

3. Locally Compact Fields and Haar Measures

This section makes the transition from finite fields to locally compact fields and from counting measures to additive Haar measures.

In Section 2, we developed a universal counting measure for ring formula. It may be viewed as counting solutions to the ring formula over a finite field in a way that does not depend on the finite field.

Counting measures are a rather simple and uninteresting type of measure. In this section, we construct a universal (motivic) measure with ties to locally compact fields. This new measure may be viewed as the volume expressed in a way that does not depend on the locally compact field. To carry out the construction, we must work with a different collection of formulas (called DVR formulas) that are better adapted to locally compact fields. 'DVR' is an acronym for *discrete valuation ring*.

3.1. **Examples of rings.** To make the transition from finite fields to locally compact fields, we wish to replace ring formulas with formulas in a language that has a rich assortment of locally compact structures.

Example 3.1. Let $\mathbb{C}[[t]]$ be the ring of formal power series with complex coefficients. A typical element of this ring has the form

$$x = \sum_{i=k}^{\infty} a_i t^i$$

(with no constraints on the convergence of the series). Pick the initial index k so that $a_k \neq 0$ (if $x \neq 0$).

The *valuation* of x is defined to be the integer k:

$$\mathrm{val}(x) = k.$$

The *angular component* of x is defined to be the complex number a_k.

$$\mathrm{ac}(x) = a_k \in \mathbb{C}^\times.$$

(In the special case $x = 0$, we set $\mathrm{val}(0) = \infty$ and $\mathrm{ac}(0) = 0$.)

The name *angular component* is not meant to suggest any precise connection to angles. The name is based on a loose analogy with the polar coordinate representation of a complex number: just as the angular component θ of a nonzero complex number $re^{i\theta}$ distinguishes among complex numbers of the same magnitude (or valuation) r, so the angular component of a formal power series helps to distinguish among formal power series of a given valuation k.

There are many other rings with similar functions, ac and val. For example, we can change the coefficient ring of the formal power series from $\mathbb{C}$ to any other field k to obtain $k[[t]]$. Or we can take the field of fractions of $k[[t]]$, which is the field of formal Laurent series with coefficients in k:

$$k((t)) = \{\sum_{-N}^{\infty} a_i t^i \mid a_i \in K\}.$$

For each prime p, there are valuation and angular component functions defined on the field of rational numbers. If x is a nonzero rational number, pick integers a, b, c, N so that

$$x = ap^N + \frac{bp^{N+1}}{c},$$

where c is not divisible by p, and $a \in \{1, \ldots, p-1\}$. The integers a and N are uniquely determined by this condition. Define the valuation of x to be $\mathrm{val}_p(x) = N \in \mathbb{Z}$ and the angular component of x to be image of a modulo p in $\mathbb{F}_p$.

Example 3.2. If $p = 2$ and $x = 17/8$, then

$$17/8 = 1.2^{-3} + 2, \quad \mathrm{val}_2(17/8) = -3, \quad \mathrm{ac}(17/8) = 1 \in \mathbb{F}_2.$$

Other examples, can be obtained from this one by completion. For each p,

$$d(x,y) = (1/2)^{\mathrm{val}_p(x-y)}$$

is a metric on the set of rational numbers. The completion is a locally compact field, called the field of p-adic numbers $\mathbb{Q}_p$. The valuation val_p and angular component function ac functions extend to the completion.

3.2. **The DVR language.** We have seen by example that there are many rings with functions val and ac. In each case, there are three separate rings that come into play: the domain of the functions val and ac, the range of the function val (which we augment with a special symbol $\{\infty\}$ for the valuation of 0), and the range of the function ac. We call these rings the valued ring, the value group, and the residue field, respectively.

We formalize this relationship as a language in first-order logic with function symbols val and ac. We allow ourselves to build syntactically well-formed expressions with variables, parentheses, quantifiers, the function symbols val and ac, the usual ring operations $(0, 1, (+), (-), (*), (=))$ on the valued ring and residue field, and the usual group operations and inequalities on the value group $(0, (+), (\leq))$. These formulas will use variables of three different types x_i for the value ring, m_i for the value group, and ξ_i for the residue field. Quantifiers $\forall$, $\exists$ can be used to bind all three sorts of variables.

The construction of first-order languages is commonplace in logic, but even without any background in logic, it is not hard to guess whether a formula is syntactically correct. We allow standard mathematical abbreviations similar to those introduced above for ring formulas.

$$\text{`}\forall y.\ (\exists x.\ x^2 = y) \implies (\exists m.\ 2m = \mathrm{val}(y))\text{.'}$$

is syntactically correct. But

$$\text{`}\forall f.\ \forall x.\ \forall y. f(y, \mathrm{ac}(y))\text{'}$$

is not well-formed, because quantifiers are not allowed over higher-order relations f in a first-order language. Also,

$$\text{`}\forall x\ \xi.\ (0 \leq x) \vee (\mathrm{ac}(x) = \xi)\text{'}$$

is not well-formed, because of a type error; the variable symbol x appears once as an integer $0 \leq x$ and again as variable in the valued field $\mathrm{ac}(x)$.

A syntactically correct formula is called a DVR formula. *The aim of motivic measure is to compute the "volume" of a DVR formula in a universal way; that is, in a way that does not depend on the underlying locally compact field.*

3.3. **Assumptions on the ring.** The various examples that we have mentioned are all structures for the DVR language: rings of formal power series $k[[t]]$, fields of formal Laurent series $k((t))$. For each prime p, $(\mathbb{Q}, \mathrm{ac}, \mathrm{val}_p)$ is a structure for the language, as well as its completion $(\mathbb{Q}_p, \mathrm{ac}, \mathrm{val}_p)$.

We will temporarily restrict the set of examples to structures $(K, k, \mathrm{ac}, \mathrm{val})$ that satisfy the following conditions.

- K is a valued field of characteristic zero, with valuation function $\mathrm{val} : K \to \mathbb{Z} \cup \{\infty\}$ and angular component functions $\mathrm{ac} : K \to k$.
- The residue field k has characteristic zero.
- K is henselian. (We review the definition below.)

Examples that satisfy these conditions include the fields $k((t))$, where k has characteristic zero. The analogy that will guides us is that these fields stand in the same relation to locally compact DVR fields, as pseudo-finite fields do to finite fields.

3.4. **Henselian field.** There is only one plausible definition for a henselian field: A field is *henselian* if the field satisfies Hensel's lemma.

Hensel's lemma gives checkable conditions on a polynomial that insure that it has a root in a given neighborhood. Hensel's lemma occupies same ground in the realm of DVR rings that the intermediate value theorem occupies in the realm of real numbers. (The intermediate value theorem also gives checkable conditions on a polynomial that insure that it has a real root in a given neighborhood.)

Our experience with motivic counting measures has alerted us to the importance of quantifier elimination, that is, the process of replacing a formula with quantifiers $\forall, \exists$ with an equivalent formula that does not contain quantifiers. The simplest case of quantifier elimination is the determination of when there exists a root of a polynomial. Without a criterion for the existence of roots to polynomials, quantifier elimination would be impossible. For the pseudo-finite fields, this is handled through the defining property of pseudo-finite fields that "every absolutely irreducible variety has a root." For real fields, quantifier elimination is based on the intermediate value theorem. For henselian fields, quantifier elimination is based on Hensel's lemma.

Lemma 3.3. *(Hensel's lemma) For every monic polynomial $f \in K[x]$, whose coefficients have non-negative valuation, and for every x such that*

$$\mathrm{val}(f(x)) > 0$$

and

$$\mathrm{val}(f'(x)) = 0,$$

there exists $y \in K$ such that $f(y) = 0$ and $\mathrm{val}(y - x) > 0$.

This is stated as a lemma, but we view it as a condition on the field K and its valuation. It can be proved that the fields $k((t))$ and $\mathbb{Q}_p$ are henselian by showing that under the hypotheses of Hensel's lemma, Newton's approximations to the roots

$$\begin{aligned} x_0 &= x \\ x_{n+1} &= x_n - f(x_n)/f'(x_n) \end{aligned}$$

converge to a root.

3.5. **Quantifier elimination.**

Theorem 3.4. *(Pas* [14]*) Let K be a field satisfying the other conditions enumerated in 3.3 with residue field k . Let ϕ be a DVR formula. Then there is another formula ϕ' without quantifiers of the valued field sort such that*

$$\forall (x, \xi, m) \in K^n \times k^m \times (\mathbb{Z} \cup \{\infty\})^r. \quad \phi^K(x, \xi, m) = \phi'^K(x, \xi, m).$$

Moreover, the formula ϕ' can be chosen to be independent of the structure K.

3.6. **Outer measure of a DVR formula.** As a first step toward constructing the measure of a DVR formula, we will define an outer measure of a formula. To motivate this construction, it might be helpful first to describe an analogous construction in Euclidean space.

3.6.1. *An outer measure in Euclidean space.* Fix a positive integer m. Tile Euclidean space with cubes of width $1/2^m$ whose vertices are centered at points a with coordinates $a_i \in \mathbb{Z}/2^m$.

According to the Calculus 101 approach to volume, we can approximate the volume of a set by counting the number of cubes that it meets. Let A be a bounded set in $\mathbb{R}^n$. Let $C_m(A)$ be the set of cubes in this tiling that meet A. In our naive approach to measure, let us define the outer measure of A at level m to be

$$\frac{\#C_m(A)}{2^{mn}}, \tag{10}$$

that is the number of cubes divided by the scaling factor 2^{mn}. (If doing so did not involve logical circularity, we would identify $1/2^{mn}$ with the volume of cube and the entire expression as the volume of the set $C_m(A)$ of cubes.)

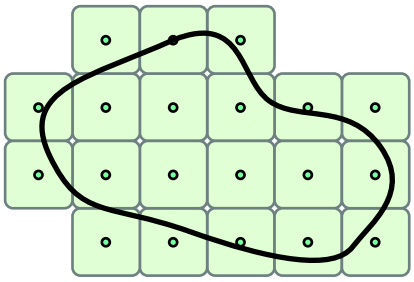


FIGURE 2. Volumes of DVR formulas can be approximated in Calculus 101 fashion by counting centers of cubes that meet a given formula, scaled according to the size of the cubes.

The outer motivic measure of a DVR formula will be formed in an entirely analogous way. Of course, we will need to decide what to use for cubes, how to count the number of cubes that "meet" a given formula, and what scaling

factor to use. Once we make these decisions, the formula for outer measure will take precisely the same form as Equation 10.

In the planar case, we gave a construction of area of polygons as taking values in a scissor group $\mathbb{S}_{\text{poly}}$. The outer approximation of any bounded planar set A by squares gives a value in the scissor group of polygons. Here too, if our outer approximation to a DVR formula is with a ring formula, then the value of the outer measure of the DVR formula will be in a scissor ring $\mathbb{S}_{\text{mot}}$.

Given all our preliminaries, it almost goes without saying at this point that rather that the counting of cubes that takes place in the numerator of Equation 10 will be replaced with the motivic counting measure of a ring formula.

3.6.2. *Cubes.* What is a cube? Well, it is a product of equal width intervals. In DVR formulas, a cube centered at a of "width" m is again a product of intervals:

$$\{(x_1, \ldots, x_n) \in K^n \mid \text{val}(x_i - a_i) \geq m, \quad \text{for } i = 1, \ldots, n\}.$$

If $K = k[[t]]$, then the interval around a formal power series a is the set of all formal power series with the same leading terms. Shaking (wagging) the tails of the power series fills out the interval. In other words, we can make precise the idea of covering a DVR formula with cubes by replacing each solution to the DVR formula with a bigger set where the tails of the solutions are allowed to vary.

Let us make this precise. We have truncation map

$$k[[t]] \to k[[t]]/(t^m) \simeq k^m$$

$$\textstyle\sum_0^\infty a_i t^i \mapsto \sum_0^{m-1} a_i t^i \mapsto (a_0, \ldots, a_{m-1}).$$

In the opposite direction, given $b \in K^m$, there is a polynomial with those coefficients

$$p(b, t) = \sum_0^{m-1} b_i t^i \in k[[t]]$$

Definition 3.5. Let ϕ be a DVR formula with free variables $(x_1, \ldots, x_n)$ and no free variables of other sorts. An *outer ring formula* ϕ_m approximation to ϕ *(at level m)* is a ring formula in nm free variables u_{ij} such that over every field k:

$$\{u \in k^{nm} \mid \phi_m(u)\} = \\ \{u \in k^{nm} \mid \exists a_1, \ldots, a_n.\ \phi(a_1, \ldots, a_n) \ \wedge \ \text{val}(a_i - p(u_{ij}, t)) \geq m\}.$$

This set is the set of centers of cubes that contain a solution to ϕ.

Theorem 3.6. *Outer ring formula approximations exist for every DVR formula ϕ at every level m.*

The proof of this theorem uses quantifier elimination results to eliminate the quantifiers that bind variables ranging over the valued field. It uses results of Presburger on quantifier elimination to eliminate the quantifiers that range over the additive group of integers. The quantifiers that bind variables in the residue field remain as quantifiers in the ring formula ϕ_m.

3.6.3. *Scaling Factors.* How is the scaling factor chosen in Equation 10 for Euclidean outer measures? The scaling factor $1/2^{nm}$ is the unique constant that has the property that if the set A is itself a union of properly aligned cubes (of width m'), then the outer measure of A is independent of m for all $m \geq m'$.

To find the scaling factor for DVR formulas, we work a simple example in which the DVR formula is itself a union of cubes of width m' (that is, its set of solutions is stable under perturbation of the power series tails).

Example 3.7. Let $\phi(x_1, \ldots, x_n) = \mathbb{T}$, a formula that is true for all values of the free variables x_i. In this case the outer ring formula approximation is exact. Substitute polynomials $p(u_{i\cdot}, t)$ for each x_i and expand in terms of mn distinct free variables u_{ij} to get

$$\phi_m(u_{ij}) = \mathbb{T}$$

for all input values u_{ij}. The number of solutions of ϕ_m over a finite field $\mathbb{F}_q$ is q^{nm}. If we take the motivic counting measure of ϕ_m, we find that the variety that counts the points of ϕ_m over any finite field is the affine space of dimension nm:

$$\#\mathbb{A}^{nm}(\mathbb{F}_q) = q^{nm}.$$

The class of ϕ_m in $\mathbb{K}[\mathbb{L}^{-1}] \otimes \mathbb{Q}$ is

$$[\mathbb{A}^{nm}] = [\mathbb{A}^1]^{nm} = \mathbb{L}^{nm}.$$

From this one example, we see that the scaling factor for DVR formulas must be $1/\mathbb{L}^{nm}$.

Definition 3.8. Let ϕ be a DVR formula. Let the *outer measure* of ϕ at level m be given by

$$\frac{[\phi_m]}{\mathbb{L}^{nm}} \in \mathbb{K}[\mathbb{L}^{-1}] \otimes \mathbb{Q} = \mathbb{S}_{\mathrm{mot}}.$$

This formula is analogous to Formula 10 for the Euclidean outer measure at level m. The numerator counts the number of centers of cubes that contain a solution to the DVR formula.

Definition 3.9. Let the *motivic measure (or motivic volume)* of ϕ be given by

$$\lim_{m\to\infty} [\phi_m]\mathbb{L}^{-nm},$$

whenever that limit exists. (The limit must be taken in a completion of $\mathbb{S}_{\mathrm{mot}}$.)

3.7. **The universal nature of motivic measure.** Just as the motivic counting measure counts solutions to ring formulas over finite fields in a field independent way, so the motivic measure takes the volume of a DVR formula over locally compact fields in a field independent way.[2]

There is a good theory of measure on locally compact fields. This is the Haar measure, which is translation invariant. Given a DVR formula ϕ and a locally compact structure K with ring of integers O_K, we can take the volume of the set of solutions to the DVR formula

$$\mathrm{vol}(\{x \in O_K^n \mid \phi^K(x)\}, dx). \tag{11}$$

The measure dx can be given a canonical normalization by requiring that it assigns volume 1 to the full set O_K^n.

We are now ready to state the main result on motivic measure. Like all the other principal results in this article, the result is due to J. Denef and F. Loeser.

Theorem 3.10. *The motivic volume of ϕ is universal in the following sense. Let $\sum a_i[X_i]\mathbb{L}^{-N_i}$ be any representative of the motivic volume of ϕ as a convergent formal sum of varieties over $\mathbb{Q}$. Pick models for the varieties over $\mathbb{Z}$. After discarding finitely many primes, for any locally compact structure of the DVR language, the K-volume of the formula is given by a convergent sum (in $\mathbb{R}$)*

$$\sum a_i \#X(\mathbb{F}_q) q^{-N_i},$$

where $\mathbb{F}_q$ is the residue field of K.

This wonderful result states that the Haar measures on all locally compact fields have an deep underlying unity. The volumes of sets can be expressed geometrically in a way that is independent of the underlying field.

Moreover, there are effective procedures to calculate the varieties X_i and the coefficients a_i, N_i that represent the outer motivic volume at level m. If the outer ring formula approximations ϕ_m converge at some finite level m to the DVR formula ϕ, then we obtain effective procedures to calculate the motivic volume of the formula.

4. Applications and Conclusions

What good is motivic measure? Here are a few examples.

4.1. **Invariants of ring formulas.** The group $\mathbb{S}_{\mathrm{mot}}$ is generated by varieties $\mathrm{Var}_{\mathbb{Q}}$. Many geometrical invariants of varieties (such as Euler characteristics and Hodge polynomials) can be reformulated as invariants of the ring $\mathbb{S}_{\mathrm{mot}}$. This gives a novel way to attach invariants to every ring formula ϕ: take a geometric invariant of $[\phi] \in \mathbb{S}_{\mathrm{mot}}$. In particular, ring formulas have

[2]It is impossible for the structure K both to be locally compact and to have a residue field k of characteristic zero, as required by Condition 3.3. In these final paragraphs, we allow the residue field to have positive characteristic. The residue field of a locally compact field is always finite.

Euler characteristics and Hodge polynomials! For example, the formula for the nonzero squares in a field

$$\text{`}\exists y.\ (y^2 = x)\ \wedge\ (x \neq 0)\text{'}$$

has Euler characteristic zero.

4.2. **Geometry of varieties.** There is a motivic change-of-variables formula that is similar to the standard change of variables formula in calculus. Using this formula, it is sometimes possible to show that two birationally equivalent varieties have the same motivic volume. This has deep implications for the geometry of the two varieties. In particular, the motivic volume determines the Hodge polynomial of the varieties.

This approach was followed by Kontsevich, who used a change-of-variables calculation to show that birationally equivalent projective Calabi-Yau manifolds have the same Hodge numbers [10]. Applications to orbifolds appear in [13].

4.3. **Computation of p-adic integrals.** Many integrals over p-adic fields are notoriously difficult to calculate. Motivic measure exposes the underlying similarities between volumes on different p-adic fields. It gives a decision procedure to calculate p-adic integrals (at least when the data defining the integral can be expressed as DVR formulas that can be reproduced at some finite level m). In particular, this means that a computer can be programmed to compute a large class of p-adic integrals.

4.4. **Generating Functions.** Motivic counting gives a way of counting that is independent of the finite field. Let

$$Z_p(t) = \sum_{i=0}^{\infty} a_i^{(p)} t^i$$

be a generating function, where the constants $a_i^{(p)}$ are obtained by counting solutions to a formula in some p-dependent way. (Each generating function depends on a single prime p.) Motivic measure can often give a way of forming a p-independent series

$$Z_{\mathrm{mot}}(t) = \sum_{i=0}^{\infty} [a_i] t^i$$

taking values in $\mathbb{S}_*[[t]]$ and specializing for almost all p to the p-dependent series $Z_p(t)$. The motivic series collects the behavior of the various series $Z_p(t)$ into a single series.

Denef and Loeser have studied motivic versions of Hasse-Weil series, Igusa series, and Serre series. They have used the general motivic series to prove that various properties of these series are independent of the prime p. See [5].

4.5. **Concluding Remarks.** This article is an exposition of a particular version of motivic integration, called *arithmetic motivic integration*. Proofs of results stated in this article can be found in [6] and [4]. Motivic integration has been developing at a break-neck pace, ever since Kontsevich gave the first lecture on the topic in 1995. The version of motivic integration developed in the late nineties goes by the name of *geometric motivic integration*. Geometric motivic integration is a coarser theory, but is sufficient for many applications. Good introduction are [1] and [12]. Some articles on geometric motivic integration include [3] and [7]. Another version of motivic integration has been developed by J. Sebag for formal schemes [15]. See also [11]. Cluckers and Loeser are in the final stages of preparation of an ultimate version of motivic integration that subsumes both geometric and arithmetic motivic integration [2].

We began this article by stating that motivic measure does not fit neatly into the tradition of Hausdorff, Haar, and Lebesgue. However, a major result states that the motivic measure specializes to the additive Haar measure on locally compact fields (Theorem 3.10). Thus, the motivic measure is perhaps not so peculiar after all. In fact, in many respects it is strikingly similar to the additive Haar measure on locally compact fields. It has been my experience when I calculate motivic volumes to lose track – mid-calculation – of which measure is being used.

References

[1] A. Craw, An introduction to motivic integration, 1999.

[2] R. Cluckers and F. Loeser, *Fonctions constructibles et intégration motivique I, II*, in preparation.

[3] J. Denef and F. Loeser, *Germs of arcs on singular algebraic varieties and motivic integration*, Inventiones Mathematicae 135, 201-232 (1999).

[4] J. Denef and F. Loeser, Definable sets, motives and p-adic integrals, JAMS, 14, No. 2, 429–469, 2000.

[5] J. Denef and F. Loeser, On some rational generating series occuring in arithmetic geometry, to appear.

[6] J. Denef and F. Loeser, Motivic integration and the Grothendieck group of pseudo-finite fields, Proc. ICM, Vol II (Beijing, 2002), 13-23, Higher Education Press 2002.

[7] J. Denef and F. Loeser, Motivic Igusa functions, Journal of Algebraic Geometry 7, 505-537 (1998).

[8] W. Fulton, Intersection Theory, second edition, Springer, 1998.

[9] G. van der Geer and B. Moonen, Abelian Varieties, preliminary version of Chapter XI. The Fourier transform and the Chow ring, July 2003, `http://turing.wins.uva.nl/~bmoonen/boek/BookAV.html`.

[10] M. Kontsevich, lecture at Orsay, Dec 1995.

[11] F. Loeser and J. Sebag, Motivic integration on smooth rigid varieties and invariants of degenerations. Duke Mathematical Journal, 119, 315-344 (2003).

[12] E. Looijenga, Motivic measures. Séminaire Bourbaki, Vol. 1999/2000. Astérisque No. 276 (2002), 267–297.

[13] E. Lupercio and M. Poddar, The Global McKay-Ruan correspondence via motivic integration, preprint, 2002.

[14] J. Pas, Uniform p-adic cell decomposition and local zeta functions, J. reine angew. Math. 399 (1989), 137–172.
[15] J. Sebag, Intégration motivique sur les schémas formels, preprint.
[16] Scholl, Classical Motives, in Motives, U. Jannsen, S. Kleiman, J.-P. Serre Ed., Proc. Symp. Pure Math., Vol 55 Part 1 (1994), 163-187.

Version: Dec 8, 2003.
Work supported by the NSF

IT IS EASY TO DETERMINE WHETHER A GIVEN INTEGER IS PRIME

Andrew Granville

Dedicated to the memory of W. 'Red' Alford, friend and colleague.

Abstract. "The problem of distinguishing prime numbers from composite numbers, and of resolving the latter into their prime factors is known to be one of the most important and useful in arithmetic. It has engaged the industry and wisdom of ancient and modern geometers to such an extent that it would be superfluous to discuss the problem at length. Nevertheless we must confess that all methods that have been proposed thus far are either restricted to very special cases or are so laborious and difficult that even for numbers that do not exceed the limits of tables constructed by estimable men, they try the patience of even the practiced calculator. And these methods do not apply at all to larger numbers ... It frequently happens that the trained calculator will be sufficiently rewarded by reducing large numbers to their factors so that it will compensate for the time spent. Further, *the dignity of the science itself seems to require that every possible means be explored for the solution of a problem so elegant and so celebrated* ... It is in the nature of the problem that *any* method will become more complicated as the numbers get larger. Nevertheless, in the following methods the difficulties increase rather slowly ... The techniques that were previously known would require intolerable labor even for the most indefatigable calculator."

from article 329 of *Disquisitiones Arithmeticae* (1801) by C. F. GAUSS.

In August 2002, three Indian computer scientists, Manindra Agrawal, Neeraj Kayal and Nitin Saxena, constructed a "polynomial time primality test", a much sought-after but elusive goal of researchers in the algorithmic number theory world. Most shocking was the simplicity and originality of their test ... whereas the "experts" had made complicated modifications on existing tests to gain improvements, these authors rethought the direction in which to push the usual ideas with stunning success. Their algorithm is based on the following elegant characterization of prime numbers.

Agrawal, Kayal and Saxena (2004). *For given integer $n \geq 2$, let r be a positive integer for which n has order $> (\log n)^2$ modulo r. Then n is prime if and only if*

- *n is not a perfect power,*
- *n does not have any prime factor $\leq r$,*
- *$(x+a)^n \equiv x^n + a \pmod{(n, x^r - 1)}$ for each integer $a, 1 \leq a \leq \sqrt{r} \log n$.*

In this talk we will explain their test, with complete proofs, and put the result and ideas in appropriate historical context. Details will be elaborated on in a forthcoming article.

1.1. Our objective is to find a "quick" foolproof algorithm to determine whether a given integer is prime. Everyone knows *trial division*, when we try to divide n by every integer m in the range $2 \le m \le \sqrt{n}$. The number of steps in this algorithm will be at least the number of integers m we consider, which is something like $\sqrt{n}$, in the worst case (when n is prime). Note that $\sqrt{n}$ is roughly $2^{d/2}$ where d is the number of digits of n when written in binary (and d is roughly $(\log n)/(\log 2)$).

The *objective* in this area has been to come up with an algorithm which works in no more than cd^A steps in the worst case, where c and A are some fixed positive constants; that is, an algorithm which works in *Polynomial Time* (which is often abbreviated as `P`). With such an algorithm one expects that one can rapidly determine whether any "reasonably sized" integer is prime.

Before the work of Agrawal, Kayal and Saxena the fastest algorithm worked in about $d^{\log\log d}$ steps. Their algorithm works in about $d^{7.5}$ steps (and thus "*Seven-and-a-half logs suffice*"); and a modification by Lenstra and Pomerance in about d^6 steps.

1.2. Recognizing primes. Are there ways to recognize primes other than by trial division? One way that comes to mind is by using

Wilson's Theorem (1770). *Integer $n \ge 2$ is prime if and only if n divides $(n-1)!+1$.*

The problem here though is that there is no obvious way to compute $(n-1)!$ rapidly (or even $(n-1)! \pmod n$). Another idea is to use

Matijasevič's polynomial (1970). *There exists a polynomial $f(x_1, x_2, \dots, x_{26}) \in \mathbb{Z}[x_1, x_2, \dots, x_{26}]$ of degree 25, with the property that the set of positive values $f(m_1, m_2, \dots, m_{26})$ where $m_1, \dots, m_{26}$ are all taken to be positive integers, is the same as the set of primes.*

We might hope to somehow quickly identify whether a given integer is a value of f, but no one has yet figured out how.

There are many places that primes come up in the mathematical literature, and many of these might suggest a way to identify primes — some of us who are interested in primality testing always look at anything new that we learn with one eye open to this application. However, for the remainder of this first half of my talk I want to focus on one classical approach.

1.3. Prime numbers have many interesting properties. One of the most amazing is known as

Fermat's Little Theorem (1637). *If n is a prime then n divides $a^n - a$ for all integers a.*

Conversely, if integer n *does not divide* $a^n - a$ for some integer a, then n is composite.

For example[1], taking $a = 2$ we calculate that

$$2^{1001} \equiv 123 \pmod{1001},$$

so we know that 1001 is composite.

We might ask whether this always works. In other words,

Is it true that *if n is composite then n does not divide $2^n - 2$*?

For, if so, we have a very nice way to distinguish primes from composites. Unfortunately the answer is "no" since, for example,

$$2^{341} \equiv 2 \pmod{341},$$

but $341 = 11 \times 31$. Note though that by taking $a = 3$ above we get

$$3^{341} \equiv 168 \pmod{341},$$

so we can use these ideas to prove that 341 is composite.

But then we might ask whether this always works, whether there is always *some* value of a that helps us prove a composite n is indeed composite.

In other words,

Is it true that *if n is composite then there is some integer a for which n does not divide $a^n - a$?*

Again the answer is "no" since 561 divides $a^{561} - a$ for all integers a, yet $561 = 3 \times 11 \times 17$. Composite integers n which divide $a^n - a$ for all integers a are called *Carmichael numbers*, 561, 1105 and 1729 being the smallest three examples. Carmichael numbers are a nuisance, masquerading as primes like this, though computationally they only appear rarely. Unfortunately it was recently proved that there are infinitely many of them, and that when we go out far enough they are not so rare as it first appears.

1.4. Square Roots. In a field, a non-zero polynomial of degree d has at most d roots. For the particular example $x^2 - 1$ this implies that 1 has just two squareroots mod p, a prime > 2, namely 1 and -1.

If we consider odd composite n then we quickly find $1^2 \equiv 4^2 \equiv 11^2 \equiv 14^2 \pmod{15}$, that is, there are four squareroots of 1 (mod 15). In general if odd n is divisible by two distinct primes then we have *at least* four distinct squareroots of 1 (mod n). Thus we might try to prove n is composite by finding a squareroot of 1 (mod n), which is neither 1 nor -1.

Now, by Fermat's Little Theorem, if p is prime then $(a^{\frac{p-1}{2}})^2 = a^{p-1} \equiv 1 \pmod{p}$, so $a^{\frac{p-1}{2}} \pmod{p}$ is a squareroot of 1 mod p and must be 1 or -1. Therefore if $a^{\frac{n-1}{2}} \pmod{n}$ is neither 1 nor -1 then n is composite. Let's try an example: We have $64^{948} \equiv 1 \pmod{949}$, and the squareroot $64^{474} \equiv 1 \pmod{949}$. Hmmmm, we failed to prove 949 is

[1]A few definitions for the uninitiated: We say that $a \equiv b \pmod{m}$ if and only if m divides $b - a$; the main advantage of this notation is that we can do most regular arithmetic operations (mod m). The *order* of $n \pmod{m}$ is the least positive integer k for which $n^k \equiv 1 \pmod{m}$.

composite like this but, wait a moment, since 474 is even so we can take the squareroot again, and a calculation reveals that $64^{237} \equiv 220 \pmod{949}$, so that 949 is composite. In general, integer a is a *witness* to n being composite if the finite sequence

$$a^{n-1} \pmod n,\ a^{(n-1)/2} \pmod n, \dots,\ a^{(n-1)/2^u} \pmod n$$

(where $n-1 = 2^u v$ with v odd) is not equal to either $1, 1, \dots, 1$ or $1, 1, \dots, 1, -1, *, \dots, *$.

It is known that, for all odd composite n, at least three-quarters of the integers a, $1 \le a \le n$ are witnesses for n. So can we find a witness "quickly" if n is composite?

• One idea is to try $a = 2, 3, 4, \dots$ consecutively until we find a witness. We believe that there is a witness $\le 2(\log n)^2$, though we cannot prove this except under the assumption of a big tool, the Generalized Riemann Hypothesis.

• Pick integers a in $\{1, 2, 3, \dots, n\}$ at random until we find a witness. By what we wrote above, if n is composite then the probability that none of the first k integers chosen are witnesses is $< 1/4^k$. Thus with a hundred or so such tests we get a probability that is so small that it is inconceivable that it could occur in practice; so we believe that any integer n for which none of a hundred randomly chosen a's is a witness, is prime. We call such n "*industrial strength primes*".

The big problem with the above method is that although we strongly believe that an industrial strength prime is indeed a prime, we have no proof, and mathematicians like proof. Indeed if you claim such integers are prime, without proof, then a cynic might not believe that your randomly chosen a are so random, or that you are unlucky, or ... No, what we need is a proof that a number is prime when we think that it is.

1.5. Proofs and the complexity class NP.

At the 1903 meeting of the American Mathematical Society, F.N. Cole came to the blackboard and without a word wrote down

$$2^{67} - 1 = 193707721 \times 761838257287,$$

long-multiplying the numbers out on the right side of the equation, and determining the decimal expansion of $2^{67} - 1$ to prove that he was indeed correct. Afterwards he said that figuring this out had taken him "three years of Sundays". The moral of this tale is that although it took Cole a great deal of work and perseverance to find these factors, it did not take him long to justify his result to a room full of mathematicians (and, indeed, to give a proof that he was correct). Thus we see that one can provide a short proof, even if finding that proof takes a long time.

In general one can exhibit factors of a given integer n to give a short proof that n is composite (such proofs are called *certificates*). By "short" we mean that the proof can be verified in polynomial time, and we say that such problems are in class NP ("*non-deterministic polynomial time*[2]"). We are not suggesting that the proof can be found in polynomial time, only that the proof can be checked in polynomial time; indeed we have

[2]Note that NP is **not** "non-polynomial time", a common source of confusion.

no idea whether it is possible to factor numbers in polynomial time, and this is now the outstanding problem of this area.

What about primality testing? If someone gives you an integer and asserts that it is prime, can you check that this is so is in polynomial time? Can they give you better evidence than their say-so that it is a prime number? Can they provide some sort of "certificate" that gives you all the information you need to verify that the number is indeed a prime? It is not, as far as I can see, obvious how to do so; certainly not as obvious as with the factoring problem. It turns out that some old remarks of Lucas from the 1870's can be modified for this purpose:

First note that n is prime if there are precisely $n-1$ integers a in the range $1 \le a \le n-1$ which are coprime to n. Therefore if we can show the existence of $n-1$ such integers then we have a proof that n is prime. In fact if n is prime then these values form a cyclic group, and so have a generator g; that is, there exists an integer g for which $1, g, g^2, \dots, g^{n-2}$ are all coprime to n and distinct mod n. Thus to show that n is prime we need simply exhibit g and prove that these numbers are distinct mod n. In fact g is a generator if and only if g has order $n-1 \pmod n$. It can be shown that any such order must divide $n-1$, and so one can show that if g is not a generator then $g^{(n-1)/q} \equiv 1 \pmod p$ for some prime q dividing $n-1$. Thus a "certificate" to show that n is prime would consist of g and $\{q$ prime : q divides $n-1\}$, and the checker would need to verify that $g^{n-1} \equiv 1 \pmod n$ whereas $g^{(n-1)/q} \not\equiv 1 \pmod p$ for all primes q dividing $n-1$, something that can be accomplished in polynomial time.

There is a problem though: One needs certification that each such q is prime. The solution is to iterate the above algorithm; and one can show that no more than $(\log n)/(\log 2)$ odd primes need to be certified after one has iterated all the way down. Thus we have a polynomial time certificate (short proof) that n is prime, and so primality testing is in the class NP.

But isn't this the algorithm we seek? Doesn't this give a polynomial time algorithm for determining whether a given integer n is prime? The answer is "no" because along the way we would have to factor $n-1$ quickly, something no-one knows how to do.

1.6. Random polynomial time algorithms.

In section 1.4 we introduced the notion of "industrial strength primes". In fact if our given integer is composite then there is a probability of at least $1/2$ that each application of that "witness" test succeeds in providing a short certificate verifying that the number is composite (the certificate provides a witness a). This is a *random polynomial time* test for compositeness (complexity class RP). As we noted it is almost certain to provide such a proof in 100 runs of the test if n is indeed composite, so if it fails then it is very likely that n is prime. Our main objection was that this doesn't provide a proof that n is prime.

One objective, just short of finding a polynomial time test for primality, is to find a random polynomial time test for primality. This was achieved by Adleman and Huang in 1992 using a method of counting points on elliptic and hyperelliptic curves over finite fields (based on ideas of Goldwasser and Kilian). Although beautiful in structure, their test is very complicated and almost certainly impractical, as well as being rather difficult to justify theoretically in all its details. It does however provide a short certificate verifying

that a given prime is prime, and proves that primality testing is in complexity class RP.

1.7. An old beginning.

The new work of Agrawal, Kayal and Saxena is much simpler than many of the more recent developments in this subject. Their starting point is the following result, which is good exercise for an elementary number theory course.

Theorem 1. *Integer n is prime if and only if $(x+1)^n \equiv x^n + 1 \pmod n$ in $\mathbb{Z}[x]$.*

Proof. Since $(x+1)^n - (x^n+1) = \sum_{1\le j\le n-1} \binom{n}{j} x^j$, we have that $x^n + 1 \equiv (x+1)^n \pmod n$ if and only if n divides $\binom{n}{j}$ for all j in the range $1 \le j \le n-1$. If $n = p$ is prime then p appears in the numerator of $\binom{p}{j}$ but is larger than, and so does not divide, any term in the denominator.

If n is composite let p be a prime dividing n. In the expansion $\binom{n}{p} = n(n-1)(n-2)\dots(n-(p-1))/p!$ we see that the only terms p divides are the n in the numerator and the p in the denominator, and so if p^k is the largest power of p dividing n then p^{k-1} is the largest power of p dividing $\binom{n}{p}$; and therefore n does not divide $\binom{n}{p}$. □

This simple theorem is the basis of the new primality test: Why don't we compute $(x+1)^n - (x^n+1) \pmod n$ and determine whether or not n divides each coefficient? This is a valid primality test, but computing $(x+1)^n \pmod n$ is obviously slow since it will involve storing n coefficients!

Since our difficulty is that the answer here involves many coefficients (as the degree is so high), one idea is to compute mod some small degree polynomial as well as mod n, so that neither the coefficients nor the degree get large. The simplest polynomial of degree r is perhaps $x^r - 1$. So we could verify whether

$$(x+1)^n \equiv x^n + 1 \pmod{(n, x^r-1)}.$$

This can be computed rapidly, and it is true for any prime n (as a consequence of the theorem above), but it is unclear whether this fails for all composite n and thus provides a primality test. The main theorem (at the start of the talk) provides a modification of this congruence, which can be shown to succeed for primes and fail for composites, thus providing a polynomial time primality test. In the second part of our talk we shall investigate this in detail.

1.Appendix. Fast Exponentiation.

An astute reader might ask how we can raise something to the nth power "quickly" (where by "quickly" we mean that the number of steps is bounded by a power of $\log n$). This problem was beautifully solved by computer scientists long ago:

We wish to compute $(x+a)^n \pmod{(n, x^r-1)}$ quickly. Define $f_0(x) = (x+a)$ and then $f_{j+1}(x) \equiv f_j(x)^2 \pmod{(n, x^r-1)}$ for $j \ge 0$ (at each step we determine $f_j(x)^2$ and then reduce mod $x^r - 1$ so the degree of the resulting polynomial is $< r$, and then reduce mod n to obtain f_{j+1}). Note that $f_j(x) \equiv (x+a)^{2^j} \pmod{(n, x^r-1)}$.

Writing n in binary, say as $n = 2^{a_1} + 2^{a_2} + \cdots + 2^{a_\ell}$ with $a_1 > a_2 > \cdots > a_\ell \geq 0$, let $g_1(x) = f_{a_1}(x)$ and then $g_j(x) \equiv g_{j-1}(x) f_{a_j}(x) \pmod{(n, x^r - 1)}$ for $j = 1, 2, \ldots, \ell$. Therefore

$$g_\ell(x) \equiv (x+a)^{2^{a_1}+2^{a_2}+\cdots+2^{a_\ell}} = (x+a)^n \pmod{(n, x^r-1)}.$$

Thus we have computed $(x+a)^n \pmod{(n, x^r-1)}$ in $a_1 + \ell \leq 3 \log n$ such steps, where a step involves multiplying two polynomials of degree $< r$ with coefficients in $\{0, 1, \ldots, n-1\}$, and reducing $\pmod{(n, x^r-1)}$.

2. Proof of the Theorem. The AKS algorithm

In the second half of the talk we will prove the theorem of Agrawal, Kayal and Saxena. We will assume that we are given an odd integer n which we know is not a perfect power, and has no prime factor $\leq r$. For simplicity we will assume that n has order $> 9(\log n)^2$ modulo r, and that[3]

$$(x+a)^n \equiv x^n + a \pmod{(n, x^r-1)} \tag{1}$$

for each integer $a, 1 \leq a \leq A$ where we take $A = 3\sqrt{r}\log n$. (One can replace each of the constants "9" and "3" by "1" with extra work). By Theorem 1 we know that these hypotheses hold if n is prime, so we must show that they cannot hold if n is composite.

Let p be a prime dividing n. We can factor $x^r - 1$ into irreducibles in $\mathbb{Z}[x]$, as $\prod_{d|r} \Phi_r(x)$, where $\Phi_r(x)$ is the rth cyclotomic polynomial, whose roots are the primitive rth roots of unity. Let $h(x)$ be an irreducible factor of $\Phi_r(x) \pmod{p}$. Then (1) implies that

$$(x+a)^n \equiv x^n + a \pmod{(p, h(x))} \tag{2}$$

for each integer $a, 1 \leq a \leq A$, since $(p, h(x))$ divides $(n, x^r - 1)$.

The congruence classes $\pmod{(p, h(x))}$ are really the elements of the ring $\mathbb{Z}[x]/(p, h(x))$, which is isomorphic to the field of p^m elements (where m is the degree of h). In particular the non-zero elements form a cyclic group of order $p^m - 1$. (Below it will be occasionally convenient to suppress the "mod" notation.)

Let G be the (cyclic) subgroup generated by $x+1, x+2, \ldots, x+A$. Notice that if $g(x) = \prod_{1 \leq a \leq A} (x+a)^{e_a} \in G$ then

$$g(x)^n = \prod_a ((x+a)^n)^{e_a} \equiv \prod_a (x^n + a)^{e_a} = g(x^n) \pmod{(p, h(x))}.$$

We define S to be the set of integers k for which $g(x^k) = g(x)^k$ for all $g \in G$. Note that $p, n \in S$.

Our plan is to give upper and lower bounds on the size of G to establish a contradiction.

2.1. Upper bounds on $|G|$.

[3]We write $f(x) \equiv g(x) \pmod{(m, h(x))}$ where m is an integer and $f(x), g(x), h(x) \in \mathbb{Z}[x]$ if there exists $u(x), v(x) \in \mathbb{Z}[x]$ for which $f(x) - g(x) = mu(x) + h(x)v(x)$

Lemma 1. *If $a, b \in S$ then $ab \in S$.*

Proof. (Here we work in the ring $\mathbb{Z}[x]/(p, h(x))$). If $g(x) \in G$ then $g(x^a) = g(x)^a \in G$ since G is a group. Therefore $g((x^a)^b) = g(x^a)^b$ as $b \in S$ and so

$$g(x)^{ab} = (g(x)^a)^b = g(x^a)^b = g((x^a)^b) = g(x^{ab}).$$

Lemma 2. *If $a, b \in S$ and $a \equiv b \mod r$ then $a \equiv b \mod |G|$.*

Proof. For any $g(x) \in \mathbb{Z}[x]$ we have that $u - v$ divides $g(u) - g(v)$. Therefore $x^r - 1$ divides $x^{a-b} - 1$, which divides $x^a - x^b$, which divides $g(x^a) - g(x^b)$. Reducing mod p, so that $h(x)$ divides $x^r - 1$, we deduce that $g(x)^a = g(x^a) = g(x^b) = g(x)^b$ in $\mathbb{Z}[x]/(p, h(x))$ for all $g \in G$, and so $g(x)^{a-b} = 1$ for all $g \in G$. Now G is a cyclic group, so taking g to be a generator of G we deduce that $|G|$ divides $a - b$.

Let R be the subgroup of $(\mathbb{Z}/r\mathbb{Z})^*$ generated by n and p. Since n is not a power of p, the integers $n^i p^j$ with $i, j \geq 0$ are distinct. There are $> |R|$ such integers with $0 \leq i, j \leq \sqrt{|R|}$ and so two must be congruent $\pmod r$, say

$$n^i p^j \equiv n^I p^J \pmod r.$$

By Lemma 1 these integers are both in S. By Lemma 2 their difference is divisible by $|G|$, and therefore

$$|G| \leq |n^i p^j - n^I p^J| \leq (np)^{\sqrt{|R|}} \leq n^{2\sqrt{|R|}}. \tag{3}$$

2.2. Lower bounds on $|G|$.

We wish to show that there are many distinct elements of G. If $f(x), g(x) \in \mathbb{Z}[x]$ with $f(x) \equiv g(x) \pmod{(p, h(x))}$ then we can write $f(x) - g(x) \equiv h(x)k(x) \mod p$ for some polynomial $k(x) \in \mathbb{Z}[x]$. Thus if f and g both have smaller degree than h then $k(x) \equiv 0 \pmod p$ and so $f(x) \equiv g(x) \pmod p$. Thus all polynomials of the form $\prod_{1 \leq a \leq A} (x - a)^{e_a}$ of degree $< m$ (the degree of $h(x)$) are distinct elements of G, and this gives a lower bound for G. One can show that m is the order of $p \pmod r$ and so if one can show that this value is large then we can get good lower bounds on G.

This was what Agrawal, Kayal and Saxena did in their first preprint, and to prove such r exist they needed to use deep tools of analytic number theory. In their second preprint, inspired by a remark of Lenstra, they were able to replace m by $|R|$ in this result, which allows them to give an entirely elementary proof of their theorem, and to get a stronger result when they do invoke the deeper estimates.

Lemma 3. *Suppose that $f(x), g(x) \in \mathbb{Z}[x]$ with $f(x) \equiv g(x) \pmod{(p, h(x))}$, and $f, g \in G$. If f and g both have degree $< |R|$ then $f(x) \equiv g(x) \pmod p$.*

Proof. Let $\Delta(y) := f(y) - g(y)$. If $k \in S$ then

$$\Delta(x^k) = f(x^k) - g(x^k) \equiv f(x)^k - g(x)^k \equiv 0 \pmod{(p, h(x))}.$$

It can be shown that x has order r mod $(p, h(x))$ so that $\{x^k : k \in R\}$ are all distinct roots of $\Delta(y)$ mod $(p, h(x))$. Now, $\Delta(y)$ has degree $< |R|$, but $\geq |R|$ distinct roots mod $(p, h(x))$, and so $\Delta(y) \equiv 0$ mod $(p, h(x))$, which implies that $\Delta(y) \equiv 0 \pmod p$ since its coefficients are independent of x.

By definition R contains all the elements generated by $n \pmod r$, and so R is at least as large as the order of $n \pmod r$, which is $> 9(\log n)^2$ by assumption. Therefore A, $|R| > B$, where $B := [3\sqrt{|R|}\log n]$. Lemma 3 implies that the products $\prod_{a\in T}(x + a)$ for every $T \subset \{1, 2, \ldots, B\}$ give distinct elements of G, and so

$$|G| \geq 2^B > n^{2\sqrt{|R|}}$$

since $2^3 > e^2$, which contradicts (3). This completes the proof of the theorem of Agrawal, Kayal and Saxena.

2.3. Running time.

One can write an algorithm (using standard techniques) to test the steps of the theorem of Agrawal, Kayal and Saxena, which runs in roughly $r^{3/2}(\log n)^3$ steps (bit operations).

We must have $r > (\log n)^2$ since n must have order $> (\log n)^2$ mod r; and thus we would not expect the AKS algorithm to run in much fewer than $(\log n)^6$ steps. It is expected that there exists a prime r in $[1 + (\log n)^2, 2(\log n)^2]$ for which n is a primitive root $\pmod r$, and thus has order[4] $r - 1$ mod r. Therefore we expect (and it is borne out in practice) that we have a running time of around $(\log n)^6$. However we cannot **prove** that such r exist.

In the next section we will give an elementary proof that such an r exists with r around $(\log n)^5$, which thus leads to a running time of around $(\log n)^{10\frac{1}{2}}$ (since $10\frac{1}{2} = \frac{3}{2} \times 5 + 3$).

In the accompanying article I will show how basic tools of analytic number theory can be used to show that such an r exists with r around $(\log n)^{24/7}$ (using an old argument of Goldfeld), which leads to a running time of around $(\log n)^{8\frac{1}{7}}$

Using much deeper tools, a result of Fouvry[5] implies that such an r exists with r around $(\log n)^3$, which leads to a running time of around $\ll (\log n)^{7\frac{1}{2}}$. This can be improved using a recent result of Baker and Harman to $(\log n)^{7.49}$.

2.4. Large orders mod r.

The prime number theorem can be paraphrased as: *The product of the primes up to x is roughly e^x.* A weak explicit version states that the product of the primes between N and $2N$ is $\geq 2^N$ for all $N \geq 1$.

Lemma 4. *If $n \geq 6$ then there is a prime $r \in [(\log n)^5, 2(\log n)^5]$ for which the order of n mod r is $> (\log n)^2$.*

---

[4] In fact, it should suffice to restrict attention to primes r for which $(r - 1)/2$ is also prime.

[5] Fouvry's 1984 result was at the time immediately applied to prove a result about Fermat's Last Theorem (then an open problem). In the accompanying article we will see how other tools developed to attack Fermat's Last Theorem can be used on the problem here.

Proof. If not, then the order of n mod r is $\le I := (\log n)^2$ for every prime $r \in [N, 2N]$ with $N := (\log n)^5$, so that their product divides $\prod_{i \le I}(n^i - 1)$. But then

$$2^N \le \prod_{N \le r \le 2N} r \le \prod_{i \le I}(n^i - 1) < n^{\sum_{i \le I} i} < 2^{(\log n)^5},$$

for $n \ge 6$, giving a contradiction.

The bound on r here holds for all $n \ge 6$, and thus using this bound our running time analysis of AKS is *effective*; that is, one can explicitly bound the running time of the algorithm for all $n \ge 6$. In the better bounds on r discussed in the previous section, the proofs are not effective, in that they do not imply how large n must be in order for the given upper bound for r to hold.

3. Even more recent developments

Can we achieve the feasible $(\log n)^6$ running time? One approach is to achieve better lower bounds on the size of G (than were obtained in section 2.2): Although this has not yet been done in a way to achieve our goal, Voloch had the beautiful idea that one can bound how often different high degree products of $(x + a)$ can be equal $\pmod{(p, h(x))}$ by using the *abc*-theorem for polynomials. However this goal has now been achieved in a different manner: Lenstra and Pomerance have extended the idea in AKS (with extra complications) to obtain the desired running time of around $(\log n)^6$ steps, as we will discuss in the next section.

Following up on ideas of Berrizbeitia, Bernstein has modified AKS to obtain an algorithm that runs in around $(\log n)^4$ steps, but which only succeeds half the time in providing a certificate of primality; in other words this is an `RP` algorithm for primality testing which is faster, easier and more elegant than that of Adleman and Huang. In practice this makes the original AKS algorithm irrelevant. For if we run the "witness" test, which is an `RP` algorithm for compositeness, by day, and run the AKS-Berrizibeita-Bernstein `RP` algorithm for primality by night, then a number n is, in practice, certain to yield its secrets faster (in around $(\log n)^4$ steps) than by the original AKS algorithm!

3a. Stop the press: Lenstra and Pomerance achieve the desired running time.

Lenstra and Pomerance replace the polynomial $\Phi_r(x)$ in AKS by a polynomial $f(x)$ with certain properties: If $f(x)$ is a monic polynomial of degree m with integer coefficients and n is a positive integer for which

- $f(x^n) \equiv 0 \bmod (n, f(x))$,
- $x^{n^m} - x \equiv 0 \bmod (n, f(x))$,
- $x^{n^{m/q}} - x$ is a unit in $\mathbb{Z}[x]/(n, f(x))$ for all primes q dividing m,

then we say that $\mathbb{Z}[x]/(n, f(x))$ is a *pseudofield*. When n is prime and $f(x)$ is irreducible mod n then these criterion are all true, and $\mathbb{Z}[x]/(n, f(x))$ is a field.

Lenstra and Pomerance. *For given integer $n \geq 2$ let m be a positive integer $> 4(\log n)^2$ for which there exists a monic polynomial $f(x)$ of degree m with integer coefficients, such that $\mathbb{Z}[x]/(n, f(x))$ is a pseudofield. Then n is prime if and only if*

- *n is not a perfect power,*
- *n does not have any prime factor $\leq A := 2\sqrt{m}\log n$,*
- *$(x+a)^n \equiv x^n + a \ mod\ (n, f(x))$ for each integer $a, 1 \leq a \leq A$.*

Evidently for a given f one can quickly determine whether one gets a pseudofield, and if so check the criteria of the theorem. Thus if we can quickly find an f which gives a pseudofield this approach will lend itself to a quick primality test. Lenstra and Pomerance's construction of f comes back full circle to Gauss's *Disquisitiones*, and his construction of regular n-gons, in particular what are now known as "Gaussian periods". For $M := 4(\log n)^2$ their polynomial has degree $q_1 q_2 \dots q_k \in (M, 4M]$ where the q_i are coprime positive integers for which there exists a prime $r_i \leq M$ such that $n^{(r_i-1)/q_i}$ has order q_i mod r_i for each i. They show how to determine these numbers, and thus f, in less than $(\log n)^3$ bit operations, once n is bigger than some n_0, which can be effectively determined.

References

1. Leonard M. Adleman and Ming-Deh A. Huang, *Primality testing and abelian varieties over finite fields*, Lecture Notes in Mathematics, vol. 1512, Springer-Verlag, Berlin.
2. Leonard M. Adleman, Carl Pomerance and Robert S. Rumely, *On distinguishing prime numbers from composite numbers*, Annals of Mathematics **117** (1983), 173–206.
3. Manindra Agrawal, Neeraj Kayal and Nitin Saxena, *PRIMES is in P* (to appear).
4. W. R. Alford, Andrew Granville and Carl Pomerance, *There are infinitely many Carmichael numbers*, Annals of Mathematics **139** (1994), 703–722.
5. Roger C. Baker and Glynn Harman, *The Brun-Titchmarsh Theorem on average*, Progr. Math. **138,** (1996), 39-103.
6. D. J. Bernstein, *Proving primality in essentially quartic random time* (to appear).
7. Pedro Berrizbeitia, *Sharpening "PRIMES is in P" for a large family of numbers* (to appear).
8. Richard Crandall and Carl Pomerance, *Prime numbers. A computational perspective*, Springer-Verlag, New York, 2001.
9. Etienne Fouvry, *Theoreme de Brun-Titchmarsh; application au theoreme de Fermat*, Invent. Math **79** (1985), 383–407.
10. Dorian M. Goldfield, *On the number of primes p for which p+a has a large prime factor*, Mathematika **16** (1969), 23-27.
11. Shafi Goldwasser and Joe Kilian, *Almost all primes can be quickly certified*, Proceedings of the 18th annual ACM symposium on theory of computing, Association for Computing Machinery, New York, 1986.
12. Donald E. Knuth, *The art of computer programming, volume 2: Seminumerical algorithms*, Addison-Wesley, Reading, 1969.
13. H.W. Lenstra, Jr. and Carl Pomerance, *Primality testing with Gaussian periods* (to appear).
14. Yu. V. Matijasevich, *Hilbert's Tenth Problem.*, MIT Press, Cambridge, MA, 1993.
15. Paulo Ribenboim, *The new book of prime number records*, Springer-Verlag, New York, 1995.
16. José Felipe Voloch, *On some subgroups of the multiplicative group of finite rings* (to appear).

Département de Mathématiques et statistique, Université de Montréal, CP 6128 succ. Centre-Ville, Montréal QC H3C 3J7, Canada

E-mail address: andrew@dms.umontreal.ca

Perelman's work on the classification of 3-manifolds

John W. Morgan

December 7, 2003

1 Introduction

Motivated by what was, by then, well-known for surfaces, Poincaré formulated in 1905, a conjecture stating that a closed, simply connected three-manifold is diffeomorphic to S^3. Developing tools to attack this problem formed the basis for much of the work in 3-dimensional topology. In the 1980's Thurston studied 3-manifolds with riemannian metrics of constant negative curvature -1, so-called hyperbolic manifolds. He formulated general conjectures about when a 3-manifold admits such a metric and proved various important special cases of his conjecture. This study led him to formulate a conjecture about the existence of homogeneous metrics for all manifolds, the so-called **Geometrization Conjecture** for 3-manifolds. This conjecture includes the Poincaré Conjecture as a very special case. Thurston's conjecture has two advantages over the Poincaré Conjecture:

- It applies to all 3-manifolds.
- It posits a close relationship between topology and geometry in dimension three.

The conjectural existence of an especially nice metric on the three-manifold leads to a more analytic approach to the problem of classifying 3-manifolds. Hamilton [H1] formalized one analytic approach by introducing the Ricci flow on the space of riemannian metrics (on manifolds of any dimension). He then conjectured that starting with any metric on a compact three-manifold, the Ricci flow should produce a one-parameter family of metrics converging to the nice metric as postulated by Thurston's Geometrization Conjecture. There are many technical issues with this program – for example, one knows that in general the Ricci flow will develop singularities in finite time. Thus, a method for analyzing these singularities and continuing the flow past them must be found. Furthermore, even if the flow goes on for all time, there are many complicated issues about the nature of the limit at time $t = +\infty$.

Hamilton [H1, H2, H3, H4, H5] made much progress on this program and established many crucial analytic estimates for the evolving metric and curvatures. He also showed that in special cases the Ricci flow could indeed be used to establish the Geometrization Conjecture. More recently, building on this work, G. Perelman, in a series of three preprints [P2, P3, P4] has claimed to surmount all of the various technical and analytic difficulties to

complete the program and establish the Geometrization Conjecture and hence the Poincaré conjecture. Perelman's arguments are quite intricate and involve many strikingly original ideas. The mathematical community is still trying to digest the argument and ascertain whether it is indeed a complete and correct argument. So far it is holding up under this scrutiny. At this point one can say with assurance that Perelman has made tremendous advances in understanding the nature of Ricci flow.

2 The Background

Let us set the stage for Hamilton's and Perelman's work.

2.1 Homogeneous Geometries and Geometric Structures

A homogeneous riemannian manifold (M, g) is one whose group of isometries acts transitively on the manifold. Thus, the manifold looks the same metrically at any point. Examples of homogeneous manifolds are the round sphere S^n, Euclidean space $\mathbb{R}^n$, and hyperbolic space, $\mathbf{H}^n$. A locally homogeneous manifold is one whose universal covering is homogeneous. Said another way, a locally homogeneous manifold is a complete riemannian manifold such that given any two points x, y in M there are neighborhoods U_x of x and U_y of y and an isometry from U_x to U_y carrying x to y. We say that a locally homogeneous manifold is modelled on a homogeneous manifold if every point of the locally homogeneous manifold has a neighborhood isometric to an open set in the homogeneous model. We are primarily concerned with locally homogeneous manifolds of finite volume.

In dimension 2 there are four models for homogeneous geometries: S^2, $\mathbb{R}^2$, $\mathbb{H}^2$, and G where G is the group $\mathbb{R} \rtimes \mathbb{R}^*$ with the natural action of $\mathbb{R}^*$ on $\mathbb{R}$. It turns out that there are no finite volume locally homogeneous manifolds modelled on the fourth example, and hence we are concerned with only the first three models. The manifolds of the first type are the sphere and the projective plane, of the second type are the torus and the Klein bottle, and of the third type are complete hyperbolic surfaces of finite volume which can be either compact or non-compact.

This gives us the following well-known and classical theorem.

Theorem 2.1.1. *(Uniformization in dimension 2) Let X be a compact surface. Then X admits a locally homogeneous metric modelled on one of the constant curvature models above. The model will be positively curved if $\chi(X) > 0$, flat if $\chi(X) = 0$, and negatively curved or hyperbolic if $\chi(X) < 0$.*

2.2 Dimension 3

In dimension three, up to isomorphism, there are 8 homogeneous geometries property for which there are finite volume locally homogeneous examples. First, we have the constant (sectional) curvature examples:

- S^3 of constant curvature $+1$.

- $\mathbb{R}^3$, which is flat.
- $\mathbb{H}^3$ of constant curvature -1.

Any manifold with a geometric structure modelled on the first is a riemannian manifold of constant positive sectional curvature. These are of the form S^3/Γ where Γ is a finite subgroup of $SO(4)$ acting freely on S^3 and include $S^3, \mathbb{R}P^3$, lens spaces, as well as the quotients by the symmetry groups of the exceptional regular solids.

Any geometric structure of finite volume modelled on $\mathbb{R}^3$ has a flat metric, and hence is finitely covered by a flat T^3. In particular, these manifolds are all compact.

A geometric structure modelled on $\mathbb{H}^3$ is a complete, finite volume hyperbolic manifold, i.e., a riemannian manifold with all sectional curvatures equal to -1 that is complete and of finite volume. There are infinitely many such manifolds up to diffeomorphism, both compact and non-compact. By Mostow rigidity, a three-manifold admits at most one (up to isomorphism) complete, finite volume hyperbolic metric. If the manifold in question is not compact then each component of a neighborhood of infinity is diffeomorphic to $T^2 \times (1, \infty)$. Geometrically the torus sections become small exponentially fast as t increases. In particular, any finite volume hyperbolic three-manifold is diffeomorphic to the interior of a compact three-manifold with the property that every boundary component is a two-torus.

Next we have the reducible examples:

- $S^2 \times \mathbb{R}$.
- $\mathbb{H}^2 \times \mathbb{R}$.

Finite volume geometric structures based on $S^2 \times \mathbb{R}$ are automatically compact and are either isometric to $S^2 \times S^1$ or have this riemannian manifold as the double covering. In the later case, the manifold is diffeomorphic to $\mathbb{R}P^3 \# \mathbb{R}P^3$. Finite volume geometric structures based on $\mathbb{H}^2 \times \mathbb{R}$ are of the form $\Sigma \times S^1$ where Σ is a finite area hyperbolic surface or are finitely covered by such manifolds. In the later case, the manifold is Seifert fibered over a hyperbolic two-dimensional orbifold of finite area. The examples Seifert fibered over a hyperbolic base can be non-compact, but if they are then they are diffeomorphic to interiors of compact manifolds with every boundary component being a torus.

Lastly, we have the examples where M is a connected Lie group with a left-invariant metric. Then the group itself is embedded in the group of isometries of M which can be of larger dimension. Three-dimensional examples of this type admitting locally homogeneous examples of finite volume are:

- The unipotent group of strictly upper triangular three-by-three matrices N^3. Locally homogeneous manifolds modelled on this group are called nil-manifolds.
- The solvable group which is written as a semi-direct product $\mathbb{R}^2 \rtimes \mathbb{R}^*$ where the action of $t \in \mathbb{R}^*$ on $\mathbb{R}^2$ is diagonal and given by the matrix

$$\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}.$$

Locally homogeneous examples modelled on this group are called solv-manifolds.

- $G = \widetilde{PSL_2(\mathbb{R})}$, the universal covering group of $PSL_2(\mathbb{R})$. This manifold can also be viewed as the universal metric covering of the unit tangent bundle to $\mathbb{H}^2$.

Finite volume nil-manifolds are compact and are circle bundles over T^2 with non-zero Euler number, or are finitely covered by such circle bundles. Finite volume Solv-manifolds fiber over S^1 with T^2-fiber or have a two-sheeted covering of this type. The monodromy of this fibration is given by an element of $SL_2(\mathbb{Z})$ whose eigenvalues are real and have absolute value different from 1. Geometric structures based on $\widetilde{PSL_2(\mathbb{R})}$ are Seifert fibered manifolds over hyperbolic two-dimensional orbifolds of finite area. The non-compact examples are all diffeomorphic to interiors of compact three-manifolds all of whose boundary components are tori.

Examination of all the possibilities shows us the following:

Corollary 2.2.1. *If X^3 is orientable and admits a locally homogeneous riemannian metric of finite volume, then X is diffeomorphic to the interior of a compact three-manifold with boundary, all of whose boundary components are tori. Furthermore, each of these tori has fundamental group which injects into the fundamental group of X. The only geometric structures with finite volume but non-compact examples are those modelled on $\mathbb{H}^3$, $\mathbb{H}^2 \times \mathbb{R}$ and $\widetilde{PSL_2(\mathbb{R})}$. The manifolds of finite volume of these types are either hyperbolic three-manifolds or are Seifert-fibered with hyperbolic two-dimensional base.*

3 Topology of Three-manifolds

There are several reasons that, unlike the case in dimension two, not every compact three-manifold can admit a homogeneous metric of finite volume. The most obvious reason has to do with the fact that three-manifolds are not necessarily prime, whereas, with one exception, all locally homogeneous manifolds are prime.

3.1 The prime decomposition

Definition 3.1.1. A three-manifold X is said to be *prime* if it is not diffeomorphic to S^3 and if every $S^2 \subset X$ that separates X into two pieces has the property that one of these two pieces is a three-ball. Equivalently, anytime we write X as a connected sum of two manifolds one of them must be the three-sphere.

One of the first theorems in the topology of three-manifolds is due to Knesser:

Theorem 3.1.2. *Every three-manifold admits a decomposition as a connected sum of prime three-manifolds, called prime factors. This decomposition is unique up to the order of the prime factors (and diffeomorphism of the factors).*

N.B. To reverse the process and reconstruct the manifold from its prime factors one needs to work with oriented manifolds and oriented connected sums. That is to say two non-diffeomorphic three-manifolds can have the same prime factors.

An essential $S^2 \subset X$ is a two-sphere that does not bound a three-ball in X. If this sphere separates, then cutting X along it and filling in the holes with three-balls implements a connected sum decomposition of X. If on the other hand a $S^2 \subset X$ does not separate, then this procedure has the effect of removing from X a prime factor diffeomorphic to $S^2 \times S^1$.

Because of the prime decomposition, all questions about the topology of three-manifolds can be reduced to questions about prime three-manifolds. For example:

Corollary 3.1.3. *Σ^3 is a homotopy three-sphere if and only if all its prime factors are homotopy three-spheres. Thus, there is a counterexample to the Poincaré conjecture if and only if there is a prime counterexample.*

By the sphere theorem, if $\pi_2(M) \neq 0$, then X contains an essential S^2.

3.2 Thurston's Geometrization Conjecture

Excluding manifolds with locally homogeneous metrics based on $S^2 \times \mathbb{R}$, all manifolds with locally homogeneous metrics have universal coverings which are diffeomorphic either to $\mathbb{R}^3$ or S^3. This means that every S^2 in the universal cover bounds a three-ball. The same is then true on the original manifold. Thus, the only 3-manifolds with locally homogeneous metrics and essential S^2's are $S^2 \times S^1$ and $\mathbb{R}P^3 \# \mathbb{R}P^3$, the latter being the only example of a non-prime manifold admitting a locally homogeneous metric.

Thurston's geometrization conjecture is about prime manifolds. For simplicity we restrict attention to orientable manifolds.

Conjecture 3.2.1. *(Thurston's Geometrization Conjecture) Let M be a compact, orientable, prime three-manifold. Then there is an embedding of a disjoint union of 2-tori and Klein bottles $\coprod_i T_i^2 \subset M$ such that every component of the complement admits a locally homogeneous riemannian metric of finite volume.*

The locally homogeneous metrics on the various components of this cutting process can be modelled on different homogeneous manifolds.

Since we are working with orientable manifolds, the boundary of a neighborhood of a Klein bottle is a two-torus, so that the resulting geometric pieces are interiors of compact manifolds all of whose boundary components are tori.

N.B. There can in general be more than one such family of tori and Klein bottles up to isotopy for which the conclusion of the geometrization conjecture holds. For example, let Σ_2 be a surface of genus 2. Then $\Sigma_2 \times S^1$ has a geometric structure. Let $T^2 \subset \Sigma_2 \times S^1$ be the torus lying over a loop in Σ_2 separating it into two once-punctured tori. Then each of the two components of $\Sigma_2 \times S^1 \setminus T$ also has a geometric structure.

To obtain uniqueness one takes a family in X satisfying the conclusion with a minimal number of components. With this condition, the family is unique up to isotopy.

Notice by Corollary 2.2.1 each torus boundary component has fundamental group injecting into each of the three-manifolds that it bounds. It follows from Van Kampen's theorem, that the fundamental group of each torus injects into $\pi_1(M)$. Such tori are called *incompressible tori*. More generally, a surface of genus at least one in a three-manifold is said to be *incompressible* if its fundamental group injects into the fundamental group of the three-manifold. It follows immediately that the Klein bottles also are incompressible. Thus, one can formulate the Geometrization Conjecture as saying that there is a decomposition of M along incompressible tori and Klein bottles into pieces whose interiors admit finite volume complete geometric structures.

3.3 Relation to the Poincaré Conjecture

Suppose that we have a prime homotopy three-sphere Σ that satisfies the conclusion of the geometrization conjecture. Since $\pi_1(\Sigma) = \{1\}$, Σ has no incompressible tori and hence the decomposition of Σ referred to in Conjecture 3.2.1 must be trivial. That is to say Σ has a locally homogeneous metric. Again since $\pi_1(\Sigma)$ is trivial, the homogeneous model for Σ must itself be compact and Σ is diffeomorphic to the model space. The only compact three-dimensional model is S^3, and consequently, Σ is diffeomorphic to S^3. Notice that this argument (except for the very last step) applies equally well to prime 3-manifolds of finite fundamental group. The conclusion is that the Geometrization Conjecture for such manifolds implies that they are space-forms, i.e., are quotients of S^3 by a subgroup of $SO(4)$ acting freely. Such groups were classified by Hopf, see [M].

4 The Ricci Flow

Let us now introduce Hamilton's Ricci flow, cf [H1]. This is a parabolic evolution equation a riemannian metric on a manifold. As we shall see it is a non-linear analogue of the heat equation for the metric tensor. Its general form is

$$g'(t) = F(g(t)),$$

where a solution is a one-parameter family of metrics whose time derivative is given by a functional $F(g(t))$. This functional must be of the same tensor type as the metric, i.e., a contravariant symmetric two-tensor and should (we hope) involve no more than second derivatives of the metric. It must also be natural (i.e., independent of the coordinates used to write the equation locally). There are not too many choices for such an F, and the one Hamilton introduced is F equal to a constant multiple of the Ricci curvature on the metric.

4.1 Curvature

Recall that given a riemannian metric, there is a unique symmetric connection ∇ on the tangent bundle of the manifold. Defining

$$R(X,Y) = \nabla_X \circ \nabla_Y - \nabla_Y \circ \nabla_X - \nabla_{[X,Y]}$$

introduces the curvature operator which is a section of the second exterior power of the cotangent bundle with values in the orthogonal endomorphisms of the tangent bundle. One then defines the curvature tensor

$$Rm(X, Y, Z, W) = -\langle R(X, Y)Z, W\rangle.$$

This is a section of the fourth exterior power of the cotangent bundle. It is skew in the first two variables and skew in the last two variable and symmetric under interchange of variables $(1, 2)$ with variables $(3, 4)$. Thus, we can view it as a symmetric tensor on the second exterior power of the cotangent bundle. The Ricci curvature is the trace on Rm on the middle two variables. Thus, if $e_1, \ldots, e_n$ is an orthonormal frame at a point we have

$$Ric(e_i, e_j) = \sum_{k=1}^{n} Rm(e_i, e_k, e_k, e_j).$$

This is a contravariant symmetric two-tensor on the manifold. In local coordinates, its leading term is a linear expression in the second partial derivatives of the entries of the matrix g_{ij} describing the metric.

4.2 The Ricci Flow Equation

Hamilton introduced the Ricci Flow equation:

$$g'(t) = -2Ric(g(t)).$$

To indicate the analogy with non-linear versions of the heat equation let us write the evolution equation for the scalar curvature R (which by definition is the trace of the Ricci curvature). The equation is

$$R'(t) = \Delta R(t) + \frac{2}{3}R^2(t) + |Ric^0(t)|^2,$$

where Ric^0 is the traceless part of the Ricci curvature. It is clear from this expression that the minus sign in the original equation is forced on us – without it we would have non-linear versions of the backwards heat equation which is an ill posed equation. The factor of 2 is for convenience only.

In [H1, H3, H4] Hamilton proved:

- Short-time existence. If g_0 is a smooth metric then there is some $\epsilon > 0$ depending on g_0 and a solution to the Ricci flow equation defined for $t \in [0, \epsilon)$ with $g(0) = g_0$.
- If the solution exists on the time interval $[0, T)$ but does not extend to any strictly larger time interval, then there is a point x in the manifold for which curvature tensor $Rm(x, t)$ of the metric $g(t)$ is unbounded as t approaches T.

In [H1] Hamilton also analyzed manifolds the Ricci flow on 3-manifolds with non-negative Ricci.

Theorem 4.2.1. (Hamilton) *Let X^3 be a compact connected three-manifold with non-negative Ricci curvature. Then one of the following happens:*

- *The Ricci curvature becomes strictly positive for all $t > 0$ sufficiently small. In this case, the Ricci flow develops a singularity in finite time. As the singularity develops the diameter of the manifold is going to zero. Rescaling the evolving family of metrics so that their diameter is 1 leads to a family of metrics converging to a metric of constant positive curvature. In particular, the manifold is diffeomorphic to the quotient of S^3 by a free orthogonal action of a finite group.*
- *The manifold is finitely covered by a metric product of a compact surface of positive curvature and S^1. Again the Ricci flow develops a singularity in finite time, and the manifold in question is diffeomorphic to $S^2 \times S^1$, or is finitely covered by such a manifold.*
- *The metric is flat and the evolution equation is constant. In this case, of course, the manifold is finitely covered by T^3.*

In [H2] Hamilton went on to analyze under certain extra assumptions what happens to the metric as $t \mapsto \infty$.

Theorem 4.2.2. (Hamilton) *Suppose that the Ricci flow on a compact riemannian 3-manifold M exists for all $t \in [0, \infty)$ and that the normalized curvature $Rm(x,t)\cdot t$ is bounded as $t \mapsto \infty$. Then there is a finite set of complete hyperbolic manifolds H_i of finite volume and for each $t >> 1$ an embedding $\varphi_t \colon \coprod_i H_i \to M$ which are converging on each compact subset of $\coprod_i H_i$ to an isometry provided that we rescale the metric at time t by a constant dependent on t. The tori in the H_i near infinity are incompressible tori in M and the complement of the image of φ_t has a decomposition along incompressible tori and Klein bottle into pieces each of which admits a locally homogeneous metric of finite volume.*

Thus, under two addition hypothesis – existence of the Ricci flow for all time and a bound on the normalized curvature – Hamilton established the conclusion of the Geometrization Conjecture.

5 Perelman's Claims

In a sequence of three manuscripts [P2, P3, P4] posted on the math archive in the last 14 months Perelman has given arguments for a sequence of claims about the Ricci flow and its relation to the Geometrization Conjecture. The overriding claim is that Hamilton's program can be completed without the extra hypotheses he imposed and hence Geometrization Conjecture is true and can be proved using the Ricci flow, generalized to allow surgeries.

5.1 Singularities at finite time

Perelman's first result is about the nature of the singularities that develop in finite time under Ricci flow.

Claim 5.1.1. *Let M be a compact, orientable riemannian manifold. Then there is $Q < \infty$ depending only on an upper bound of the curvature of M and a lower bound for the volume of any unit metric ball in M such that the following holds. Suppose that the Ricci flow with these initial conditions exists for $t \in [0, T)$ but does not extend past. Then , the regions where $|Rm|(x, t) \geq Q$ as t approaches T are of one of the following five types:*

- *a closed manifold with a metric collapsing to a point.*
- *a closed manifold diffeomorphic to $S^2 \times S^1$, $\mathbb{R}P^3$, S^3 or $\mathbb{R}P^3 \# \mathbb{R}P^3$.*
- *a tube diffeomorphic to $S^2 \times I$ covered by regions where the metric is very close to a standard round metric of curvature $Q' \geq Q$ on the sphere times a interval whose length is a large multiple of $1/\sqrt{Q'}$.*
- *a region double covered by a tube as in the previous case.*
- *a manifold diffeomorphic to D^3 with a neighborhood of the boundary of the third type.*

Furthermore, in the first case the metric when rescaled to have a fixed diameter is converging to a metric of constant positive curvature.

Thus, in the first two cases the manifold becomes completely singular at the limiting time and admits a metric homogeneous metric of non-negative sectional curvatures. That is to say the manifold satisfies the conclusion of the Geometrization Conjecture. Thus, we can assume that manifolds of these types do occur we can remove them at the singular time, knowing that these satisfy the conclusion of the Geometrization Conjecture. We are left to consider tubes, regions double covered by tubes and regions diffeomorphic to D^3 with a neighborhood of the boundary being a tube. We assume (for simplicity of the exposition only) that the manifold in question admits no $\mathbb{R}P^2$, then the only cases that we are left to consider are tubes and three-balls with tubes as neighborhoods of the boundary. We call the latter *caps*.

5.2 The Surgery Process

To treat singularities occurring inside of tubes and caps, Perelman introduces the notion of Ricci flow with surgery (Hamilton had earlier introduced in [H5] a similar idea in studying certain four-dimensional Ricci flows). The idea is to proceed to the first singular time T. There the metric outside the tubes and caps converges to a smooth (non-complete) metric. What is happening to the metric is that inside the tubes and caps is that one or more singularities are developing, but these are fairly deep within the tubes and caps, far from the boundary. To do surgery on a tube one fixes a parameter $Q' >> Q$. One starts near each end of the tube, where the metric remains smooth and the norm of the curvature is close to Q and proceeds into the tube until one finds a long piece of tube with metric close to the product metric of a round metric of curvature Q' on S^2 with the usual metric on an interval $[-a, a]$ with $Q' >> Q$ and $a\sqrt{Q'} >> 1$. We cut this subtube open along $S^2 \times \{0\}$.

By damping the metric on $S^2 \times [-a/2, 0]$ down to the product metric near the 0 end we can then glue in a standard metric on the three-ball. We do this process near each end of the tube and throw away the rest of the limit in between these two surgery regions.

For a cap, the process is similar except that since there is only one boundary component there is only one surgery. Again, after doing the cutting we throw away all of the cap past the surgery S^2. As we mentioned above, if the singularity is of the other type then we remove the entire manifold and the flow is empty from then on. (More precisely, if the manifold is disconnected then we throw away all components that are becoming singular of one of the other types.)

The result of this process, called Ricci flow with surgery, is to produce a closed manifold M_T at the first singular time T. Most of the manifold is the result of the time T Ricci flow but we have gone in and removed a neighborhood of the singularities that develop at this singular time and sewed in fixed riemannian metrics on 3-disks that we create by hand. Having made the closed riemannian manifold M_T, we use it as the initial conditions of a Ricci flow at time $t = T$. In this way we continue the Ricci flow with surgery to the next singular time. At the next singular time T_1 we do exactly the same surgery process, and then re-start the Ricci flow.

Because the process of doing a surgery removes a fixed amount of volume from the three-manifold (how much depends on the choice of the parameters Q' and a) and because under Ricci flow volume increases at no more than a fixed exponential rate, it follows that there can be only finitely many surgery times in any compact time interval, though there they may well be infinitely surgeries as we allow time to go all the way to $+\infty$. (There is a technical issue here. For reasons having to do with controlling some of the important quantities in these arguments, Perelman has to allow the surgery parameter Q' to grow as function of t, but always remaining finite in finite time. This means that the estimate of how much volume is removed is a decreasing function of the surgery time. But still it is uniformly bounded below on any compact time interval, giving an upper bound estimate on how many surgery times can occur in this interval.) In this way Perelman creates a Ricci flow with surgery with any compact riemannian 3-manifold as initial conditions. This flow with surgery exists for all $t \in [0, \infty)$.

In this process the manifold changes topological type by the surgery process, but the topological changes are threefold:

- (Tube Surgery) Remove a tube $S^2 \times I$ from the manifold and glue in disks onto each end.
- (Cap Surgery) Remove a three-ball from the manifold and glue in another one.
- (Collapsing Component Surgery) A component of the manifold is entirely removed.

The first type of operation is a usual topological surgery. This is how the prime decomposition required in the Geometrization Conjecture is implemented in the Ricci flow with surgery. (Of course, there is no reason to expect that every one of the surgeries produces a step in the prime decomposition. Some may simply split off S^3's, but this is harmless.)

The second type of operation is meaningful metrically since the balls in question have different metrics, but topologically it does nothing. The last type of operation may change the topology significantly since entire components are removed, but, as we have seen, each of these components satisfies the conclusion of the Geometrization Conjecture, so that we are removing pieces of the prime decomposition of the original three-manifold that do satisfy the conjecture and keeping those about which we have not yet learned anything.

5.3 Limits at infinity

The next step in Perelman's argument is to analyze the limits of the Ricci flow with surgery as $t \mapsto \infty$. Once again in this part of the argument he is following the path initiated by Hamilton in [H2], though his situation is more complicated for two reasons – (i) he has Ricci flow with surgeries instead of a Ricci flow and (ii) he does not make the curvature bound assumption as Hamilton did. In spite of these complications Perelman claims that much of Hamilton's analysis can be adapted. He claims:

Claim 5.3.1.

- *There is a finite collection of complete hyperbolic manifolds H_i of finite volume and for all sufficiently large t an embedding*

$$\varphi_t \colon \coprod_i H_i \to M_t$$

which for sufficiently large t is arbitrarily close to an isometry on an arbitrarily large compact subset of $\coprod_i H_i$ provided that we rescale the metric on M_t by $\sqrt{c/t}$ for an appropriate constant c independent of t.

- *For all $t >> 1$ the boundary tori of the H_i are incompressible tori in M_t.*
- *For all $t >> 1$, the complement of the image of a sufficiently large compact subset of $\coprod_i H_i$ has a metric which is arbitrarily collapsed with lower curvature bounds.*

For a point x in a complete riemannian manifold define $r(x)$ to be the supremum of $r \geq 0$ such that the $Rm(y) \geq -r^2$ on the metric ball $B_r(x)$ of radius r centered at x. To say that a metric is w-collapsed at a point x with lower curvature bounds means that $\mathrm{Vol}(B_{r(x)}(x)) < wr^3(x)$. To say the complement of the image of φ_t is arbitrarily collapsed with lower curvature bounds as $t \mapsto \infty$ means that given any $w > 0$ there is $T(w) < \infty$ such that this region is w-collapsed with lower curvature bounds for all $t >> T(w)$.

The fact that we have pieces of the M_t isometrically approaching fixed hyperbolic pieces implies that all the surgeries done at sufficiently large times are done in the collapsed regions of the manifold M_t.

It follows from this claim that the for any $t >> 1$ we can separate M_t along incompressible tori into pieces that are diffeomorphic to complete hyperbolic manifolds of finite volume and pieces where the metric is w-collapsed with lower curvature bounds where w depends on t and goes to zero as t goes to infinity. It remains to study the collapsed pieces. If we can show that these satisfy the Geometrization Conjecture, then we will have completed the proof.

5.4 The Collapsed Regions

At this point we use no more information from the Ricci flow. The argument now turns to results from the theory of collapsed manifolds. In our situation, we have a sequence M_{t_n} of 3-dimensional manifolds that are collapsing with (local) lower curvature bounds. The idea is to pass to a limit X (of a subsequence) which is a more general metric space than a riemannian manifold, but one in which the notion of curvature being bounded below still makes sense. The natural category for such objects is the category of Alexandrov spaces, see [B] and [P1]. Such spaces have a uniform dimension. In the case of limits the dimension of the limit is most that of the manifolds in the sequence. The analysis the depends on the dimension of the limiting Alexandrov space X. Let us first analyze the case of a sequence of 3-manifolds converges with (global) curvature bounds to an Alexandrov space X, cf [ST1, ST2].

5.4.1 X has dimension 0

This is the case when the manifold (or one of its components) is collapsing to a point. In this case, the manifolds in the sequence eventually have positive curvature. These satisfy the Geometrization Conjecture.

5.4.2 X has dimension 1

In this case X is either a circle or an interval and the manifolds M locally fibers over X with two-torus fibers. Hence, these manifolds are Solv-manifolds or are $T^2 \times I$. Of course, the Solv-manifolds satisfy the Geometrization Conjecture.

5.4.3 X has dimension 2

In this case the manifolds M 'fibers' over X with circle fibers. The quotation marks reflect the fact that the structure is that of a Seifert fibration over a two-dimensional orbifold.

5.4.4 X has dimension 3

This case can be quite delicate since three-dimensional Alexandrov spaces are quite complicated. But in this case the Ricci flow gives us more. The parabolic nature of the equation tells us that curvature bounds imply bounds on all derivatives of curvature. Hence, in this case the limit is actually a smooth riemannian manifold not a more general Alexandrov space. This manifold is automatically flat, and hence satisfies the Geometrization Conjecture.

Now we must adapt this limiting process where we have a global lower bound to curvature to our situation where the curvature lower bound is only a local one. Perelman's claim is that one obtains a further decomposition of the manifolds in the sequence along incompressible tori into pieces of the above types, i.e., into pieces satisfying the conclusion of the Geometrization Conjecture. Notice that one will never see pieces of the first type

(dimension X is 0) as limits as $t \mapsto \infty$ since these pieces would eventually have positive curvature and hence disappear at finite time.

6 Concluding Remarks

Perelman's picture of the effect of Ricci flow with surgery fits perfectly with Thurston's Geometrization Conjecture. The surgeries at finite time implement the connected sum decomposition into prime factors. This decomposition is necessary before one can hope to find homogeneous metrics. Then the decomposition along tori and Klein bottles occurs in analyzing the limits as $t \mapsto \infty$. The first decomposition is along incompressible tori into inflating hyperbolic pieces and collapsing pieces. The second decomposition involves only the collapsing pieces. It is a decomposition into collapsing pieces whose collapsing limit is 1-dimension and those whose limit is of 2-dimension. The first are either T^2-bundles over S^1 (these are components of the manifolds in the sequence) or $T^2 \times I$. The latter separate different pieces whose collapsing limit is 2-dimensional, these being Seifert fibered three-manifolds that are collapsing to two-dimensional orbifolds. Thus, we get components that are decomposed along tori into Seifert fibered pieces, components that are Seifert fibered, and components that are Solv-manifolds. There is one other possibility which is a collapsing component whose limit is a flat three-manifold. Since these are smoothly collapsing to a smooth limit, the manifolds in the sequence are eventually diffeomorphic to the limit.

Notice that the hyperbolic metrics are produced by the Ricci flow, but the other locally homogeneous metrics are imposed by hand. On the collapsing pieces we obtain topological rather than metric conclusions from the Ricci flow and Alexandrov space theory, but these conclusions are strong enough to allow classification of the topological types. Fortunately, these are all known to possess locally homogeneous metrics (essentially by induction on dimension).

The three-sphere and other manifolds of finite fundamental group are Seifert fibered. Thus, they have collapsing metrics with two-dimensional limits. One can ask whether these manifolds occur in the limit as $t \mapsto \infty$. In [P3] Perelman argues that in fact this does not happen. Any manifold whose fundamental group is a free product of finite groups and infinite cyclic groups must disappear in finite time under the Ricci flow with surgery process.

7 References

[B].Y. Burago, M. Gromov and G. Perelman, "A. D. Aleksandrov spaces with cruvature bounded below," Russian Math. Surveys **47** (1992), 1-58.

[H1]. R. Hamilton, " Three-manifolds with positive Ricci curvature," J. Differential Geom. **17** (1982), 255-306.

[H2]. R. Hamilton, "Non-singular solutions of the Ricci flow on three-manifolds," Comm. Analysis and Geometry **7** (1999), 695-729.

[H3]. R. Hamilton, " Formation of singularities in the Ricci flow," Surveys in Differential Geometry **2** (1995), 7-136, International Press.

[H4]. R. Hamilton, "The Harnack estimate for the Ricci flow," Jour. DIff. Geom. **37** (1993), 225-243.

[H5]. R. Hamilton, "Four-manifolds with positive isotropic curvature," Comm. Anal. Geom. **5** (1997), 1-92.

[M]. J. Milnor, "Groups which act on S^n without Fixed Points," Amer. Journal of Math. 1957, 623- 630.

[P1]. G. Perelman, 'Spaces with curvature bounded below," Proceedings of ICM 1994, 517-525.

[P2]. G. Perelman, "The entropy formula for the Ricci flow and its geometric applications," arXiv.math.DG/0211159.

[P3]. G. Perelman, "Ricci flow with surgery on three-manifolds," arXiv.math.DG/0303109.

[P4]. G. Perelman, " Finite extinction time for the solutions to the Ricci flow on certain three-manifolds," arXiv.math.DG/0307245.

[ST1]. T. Shioya and T. Tamaguchi, "Volume collapsed three-manifolds with a lower curvature bound," arXiv.math.DG/0304472.

[ST2]. T. Shioya nad T. Tamaguchi, "Collapsing three-manifolds under a lower curvature bound," J. Differential Geom., **56** (2000), 1-66.