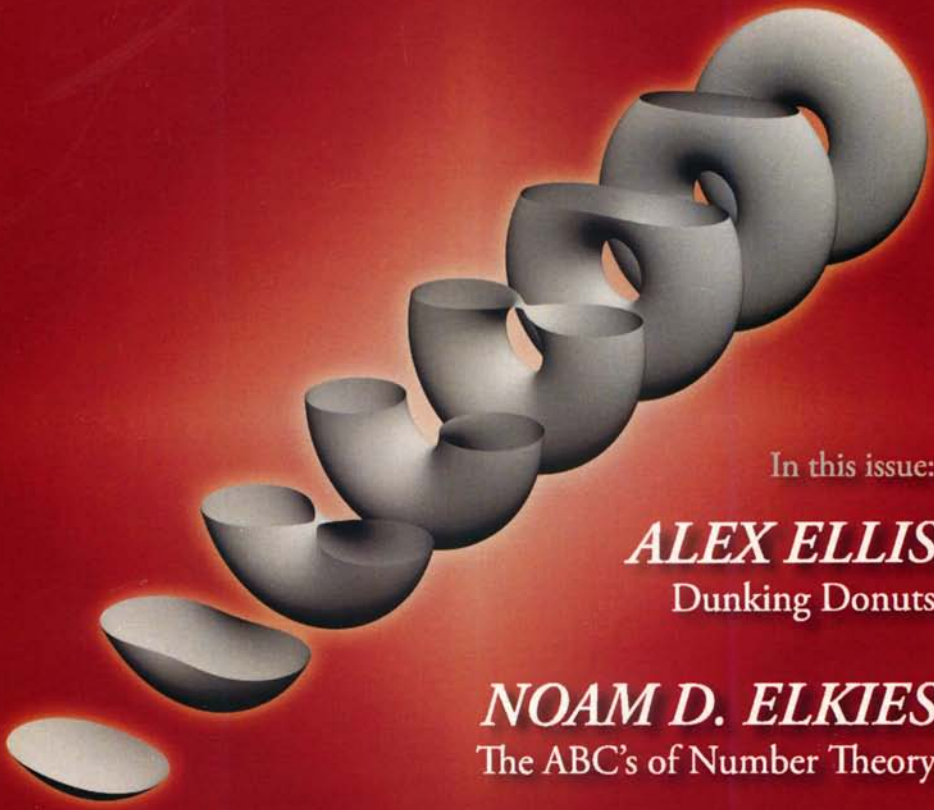
The Harvard College Mathematics Review

Vol. 1, No. 1

Spring 2007



In this issue:

ALEX ELLIS
Dunking Donuts

NOAM D. ELKIES
The ABC's of Number Theory

HCMR

A Student Publication of Harvard College

Instructions for Authors All submissions should include the name(s) of the author(s), institutional affiliations (if any), and both postal and e-mail addresses at which the corresponding author may be reached. General questions should be addressed to Editor-In-Chief Scott Kominers at hcmr@hcs.harvard.edu.

Articles. The Harvard College Mathematics Review invites the submission of quality expository articles from undergraduate students. Articles may highlight any topic in undergraduate mathematics or in related fields, including computer science, physics, applied mathematics, mathematical economics, and statistics.

Authors may submit articles electronically, in .pdf, .ps, or .dvi format, to hcmr@hcs.harvard.edu, or in hard copy to

The Harvard College Mathematics Review
Student Organization Center at Hilles
Box # 360
59 Shepard Street
Cambridge, MA 02138.

Submissions should include an abstract and reference list. Figures, if used, must be of publication quality. If a paper is accepted, high-resolution scans of hand drawn figures and/or scalable digital images (in a format such as .eps or .pdf) will be required.

Problems. The HCMR welcomes the submission of original problems in any field of mathematics, as well as solutions to previously proposed problems.

Proposers should submit problems to Problems Editor Zachary Abel, either at hcmr-problems@hcs.harvard.edu or at the address above. A complete solution or a detailed sketch of the solution should be included, if known.

Solutions to previous problems should be sent to hcmr-solutions@hcs.harvard.edu or to the address above. Solutions should include the problem reference number. All correct solutions will be acknowledged in future issues. The most outstanding solutions received will be published.

Advertising. Advertising inquiries should be sent to hcmr-advertise@hcs.harvard.edu, addressed to Business Manager Charles Nathanson.

Subscriptions. One-year (two issues) subscriptions are available, at rates of $10.00 for students, $15.00 for other individuals, and $30.00 for institutions. All inquiries should be sent to Distribution Manager Nike Sun at hcmr-subscribe@hcs.harvard.edu.

Cover Image. The image on the cover illustrates the application of Alexander P. Ellis's "dunking function" to a torus of genus one, as described in his article "Dunking Donuts: Culinary Calculations of the Euler Characteristic." The image was created in Mathematica™ by Graphic Artist Zachary Abel.

The Harvard College Mathematics Review is produced and edited by a student organization of Harvard College.

Staff 2006–2007

Editor-In-Chief
Scott Kominers '09

Design Director
Brett Harrison '10

Business Manager
Charles Nathanson '09

Articles Editor
Shrenik Shah '09

Distribution Manager
Nike Sun '09

Features Editor
Sam Lichtenstein '09

Graphic Artist
Zachary Abel '10

Problems Editor
Zachary Abel '10

Cover and Logo Design
Hannah Chung '09

Issue Production Directors
Zachary Abel '10
Brett Harrison '10

Board of Reviewers
Zachary Abel '10
Pablo Azar '09
Kelley Harris '09
Brett Harrison '10
Scott Kominers '09
Rosen Kralev '09
Menyoung Lee '10
John Lesieutre '09
Sam Lichtenstein '09
Charles Nathanson '09
Shrenik Shah '09
Nike Sun '09

Board of Copy Editors
Zachary Abel '10
Pablo Azar '09
Jannis R. Brea '10
Charles Chen '09
Kelley Harris '09
Brett Harrison '10
Paul Kominers
Scott Kominers '09
Menyoung Lee '10
Sam Lichtenstein '09
Daniel Litt '10
Charles Nathanson '09
Shrenik Shah '09

Faculty Sponsors
Dean Benedict H. Gross '71
Professor Peter Kronheimer

Contents

0 From the Editor

Scott Kominers '09
Harvard University
Cambridge, MA 02138
kominers@fas.harvard.edu

It is my great pleasure to introduce the inaugural issue of *The Harvard College Mathematics Review* (HCMR). Since I first proposed the journal, I have hoped that The HCMR would help students learn and appreciate advanced mathematics. True to the magazine's mission, this issue contains expository articles on topics drawn from undergraduate-level foundations and surveys of undergraduate research, as well as student-appropriate original problems.

Back in high school, my mathematics teacher gave me her collection of old issues of *The College Mathematics Journal* and *The American Mathematical Monthly*. I dove in, skimming abstract after abstract. Every so often, I would manage to find an article with an introduction I understood.

I would always start to read these articles, but I never quite had the mathematical background to finish them. Nonetheless, it was exciting to see what "mathematics" really is. Reading about a range of fields helped awaken me to the depth and beauty of mathematics.

I go back to that same stack of journals annually. Each time I return, I find that I understand more than I did the last time. This is how math evolves for me. As I learn, I feel myself approach the day when I can open a journal to a random article and comprehend it in its entirety.

I would appreciate any commentary or feedback you have. Please direct your comments and questions to hcmr@hcs.harvard.edu or to me personally at kominers@fas.harvard.edu. I also invite you to submit to future issues. We publish articles, short notes, and problems in any field of pure or applied mathematics at the undergraduate level. Please see the inside cover for submission guidelines.

We at The HCMR are greatly indebted to **Dean Benedict H. Gross**, who has volunteered his time, advice, and expertise throughout the production process and to **Professor Peter Kronheimer**, who has been with The HCMR as an advisor since our earliest days. We also extend our warmest thanks to **Professor Noam D. Elkies** for guidance, commentary, and of course for his fantastic feature article on "The ABC's of Number Theory," and to **Professor Dennis Gaitsgory** for sharing one of his early teaching experiences in our endpaper. We are grateful to **Dean Paul J. McLoughlin II** for his administrative help in establishing The HCMR organization and to **Mr. Christopher C. Mihelich '02** for his LaTeX advice. Finally, we could never have produced this issue without the generous support of **The Harvard Mathematics Department**.

Every one of us currently involved in The HCMR is a founding member; we are proud and excited to have seen our project finally come to fruition. QED.

Scott Kominers '09
Editor-In-Chief, The HCMR

STUDENT ARTICLE

1

Dunking Donuts: Culinary Calculations of the Euler Characteristic

Alexander P. Ellis '07[†]
Harvard University
Cambridge, MA 02138
apellis@gmail.com

Abstract

Motivated by a remarkable 18th-century result about polyhedra known as Euler's formula, we will develop the notion of the Euler characteristic χ in the more modern context of CW complexes. The fact that χ is a homotopy invariant gives an easy (perhaps trivializing) proof of Euler's formula. We then develop two non-elementary methods of computing χ in specific cases: Morse theory and the Poincaré-Hopf Index Theorem. Both will be used to compute the Euler characteristic of closed orientable surfaces, using culinary analogies. In an appendix, the former will also be used to compute the Euler characteristic of real projective space.

Most of this paper requires only an understanding of multivariable calculus and basic point-set topology. While the reader would be aided by a modest background in differential and algebraic topology at a few points, the degree of formality does not require this.[‡]

1.1 The Euler Characteristic and CW Complexes

The **Euler characteristic** $\chi(P)$ of a polyhedron P is defined to be the number F of its faces, minus the number E of its edges, plus the number V of its vertices:

$$\chi(P) = F - E + V.$$

We consider any n-sided polygon to be "filled in," so it has one face. Then we immediately have:

$$\chi(\text{any } n\text{-gon}) = 1 - n + n = 1.$$

We have easily seen that the Euler characteristic of a polygon is independent of the number and arrangement of these sides; less obviously, any convex polyhedron satisfies

$$\chi(\text{any convex polygon}) = 2.$$

This fact, known as **Euler's formula**, was known to Leonhard Euler (1707-1783), the namesake of χ. From Euler's formula, it is not hard to prove the classification of Platonic solids. (The original

[†]Alexander P. Ellis, Harvard '07, is a mathematics concentrator and English minor. Originally from New York City, Alex attended Stuyvesant High School. Starting in the fall, he will spend a year studying at Cambridge University, in Part III of the Mathematical Tripos, after which he plans to return to the United States to pursue a PhD in pure mathematics. His mathematical interests are primarily in geometry and topology, and in their connections with other branches of mathematics, as well as with physics. He also has a knack for counting the number of letters in words quickly.

[‡]Diagrams for this article were created in METAPOST by Graphic Artist Zachary Abel '10, based on drawings submitted by the author.

classification argument, which proceeds by adding up angles at a vertex, appears in Book XIII of Euclid's *Elements*.)

There is a more modern definition of χ which generalizes it to a homotopy invariant of CW complexes. Once we see what a CW complex is, all this means is that stretching, bending, folding, and compressing our space will not change its Euler characteristic; we may not, however, cut or glue.

We will define the notion of a **CW complex** inductively. A zero-dimensional CW complex is just a set of points, also called the 0-**skeleton**. The data of a one-dimensional CW complex X is a 0-skeleton X_0, a set of closed 1-discs (closed intervals) $\{I_\alpha\}_{\alpha \in A}$, and a set of corresponding maps

$$\{\phi_\alpha : \partial I_\alpha \to X_0\}_{\alpha \in A}$$

taking the boundary of each closed 1-disc to the 0-skeleton. The complex X (or its 1-skeleton X_1) is then the quotient space

$$X = \left(X_0 \amalg \coprod_{\alpha \in A} I_\alpha \right) / \{\phi_\alpha\}_{\alpha \in A}.$$

(The symbol $\coprod$ just means a union of disjoint topological spaces, where the open sets are unions of open sets taken from either space.) When we quotient by a family of maps, we are quotienting by the equivalence relation which identifies each point of each ∂I_α with its image under the corresponding map ϕ_α. Geometrically, we are just attaching each closed 1-disc I_α to X_0 by gluing its endpoints to their images under ϕ_α. Inductively, an n-dimensional CW complex is given by an $(n-1)$-skeleton X_{n-1}, a set of $\{D_\beta\}_{\beta \in B}$ of closed n-discs,[1] and attaching maps $\{\psi_\beta : \partial D_\beta \to X_{n-1}\}_{\beta \in B}$. The complex is then the quotient space

$$X = \left(X_{n-1} \amalg \coprod_{\beta \in B} D_\beta \right) / \{\phi_\beta\}_{\beta \in B}.$$

Further details can be found in Chapter 0 of [Ha].

An example which will be useful in just a moment: the n-sphere $S^n = \{v \in \mathbb{R}^{n+1} : |v| = 1\}$ is homeomorphic to the CW complex given by:

one 0-cell, the point p

one n-cell D with attaching map $\phi(x) = p$ for all $x \in \partial D$.

In other words, we start with the closed n-disc D, and glue the entire bounding $(n-1)$-sphere to a point.

Now say we have an n-dimensional CW complex Y whose k-cells are given by the set C_k. Write $\mathrm{Card}(C_k)$ for the cardinality for C_k, that is, the number of k-cells. Furthermore, say that each C_k is a finite set. Then we define the Euler characteristic of Y to be

$$\chi(Y) = \sum_{k=0}^{n} (-1)^k \, \mathrm{Card}(C_k).$$

This generalizes our earlier definition, since vertices, edges, and faces can be taken to be the 0-, 1-, and 2-cells of a two-dimensional CW complex. It turns out (see section 2.2 of [Ha]) that χ is a homotopy invariant in the sense mentioned earlier.[2] In particular, homeomorphic CW complexes have the same χ,

[1] By n-disc, we simply mean a space homeomorphic to the unit ball in $\mathbb{R}^n$, that is, $\{v \in \mathbb{R}^n : |v| < 1\}$. When we add the adjective **closed**, we simply mean the closure in $\mathbb{R}^n$ of such a set.

[2] For those familiar with cellular homology, the proof is not hard. One can show purely algebraically that given a chain complex $C_0 \to C_1 \to C_2 \to \cdots$ of finitely generated abelian groups, $\sum(-1)^k \, \mathrm{rk}(C_k) = \sum(-1)^k \, \mathrm{rk}(H_k)$, where H_k is the k-th homology group of the complex. In the case of the cellular complex, C_k is simply a freely generated $\mathbb{Z}$-module with rank equal to the number of k-cells, so $\chi(X) = \sum(-1)^k \, \mathrm{rk}(C_k(X)) = \sum(-1)^k \, \mathrm{rk}(H_k(X))$. And since the Betti numbers $b_k = \mathrm{rk}(H_k(X))$ are homotopy invariants, so is the Euler characteristic $\chi(X)$.

since every homeomorphism is certainly a homotopy equivalence. Viewed conversely, we can compute χ of a given space by choosing a CW complex on it, and our computation will not depend on our choice of CW structure. (This is tautologous, since by "choosing a CW structure" on a space we merely mean finding a CW complex homeomorphic to our space.)

As a corollary to all of this, we have an immediate proof of Euler's formula, that all convex polyhedra "miraculously" have Euler characteristic equal to 2. Indeed, any convex polyhedron can be "smoothed out" by a homotopy equivalence (in fact a homeomorphism) into a 2-sphere. Then as explained above, the 2-sphere has one 0-cell and one 2-cell, and thus has Euler characteristic

$$\chi(S^2) = 1 - 0 + 1 = 2.$$

Similarly and more generally, we have

$$\chi(S^n) = \begin{cases} 0 & n \text{ is odd} \\ 2 & n \text{ is even.} \end{cases}$$

1.2 A Little Morse Theory

In a landmark 1934 paper [Mo], Marston Morse (1892-1977) initiated the theory which came to bear his name. The basic idea of Morse theory is to study a smooth manifold by a certain class of smooth functions on it, called Morse functions. It turns out that the typical smooth function is a Morse function.

Let M be a smooth (C^∞) manifold, and let $f : M \to \mathbb{R}$ be a smooth function on M. Recall that a **critical point** of f is a point p such that df_p is a degenerate linear map. In this case, this is equivalent to saying that in a local coordinate system $\{x_1, \ldots, x_n\}$ around p, all the first partial derivatives vanish:

$$p \text{ is a critical point of } f \quad \Leftrightarrow \quad \frac{\partial f}{\partial x_1}(p) = \ldots = \frac{\partial f}{\partial x_n}(p) = 0.$$

In single-variable calculus, we measure the behavior of a function at a critical point by looking at the sign of the second derivative, if non-vanishing. If the second derivative vanishes, we need to consider higher derivatives (think of $f_1(x) = x^3$ and $f_2(x) = x^4$ at $x = 0$). Analogously, we want to consider **non-degenerate** critical points, which are defined to be critical points where the matrix of second partial derivatives determines a non-degenerate bilinear form:

$$\text{the critical point } p \text{ of } f \text{ is non-degenerate} \quad \Leftrightarrow \quad \det\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(p)\right) \neq 0,$$

where i and j are the row and column indices. Then the class of functions which we can easily work with are those whose critical points are all non-degenerate; we call these **Morse functions**. The obvious generalization of looking at the sign of the single-variable first derivative is to look at the signs of the eigenvalues of df_p. However, this would force us to worry about existence of real eigenvalues, and this may not even be stable under change of coordinates. Instead we appeal to a famous and convenient result which guarantees a "nice" set of coordinates.

Lemma 1. *(The Morse Lemma.) Let p be a non-degenerate critical point of the smooth function $f : M \to \mathbb{R}$. Then there exists a neighborhood U of p and a coordinate system $\{y_1, \ldots, y_n\}$ on U centered at y such that on U,*

$$f(y) = f(p) \pm y_1^2 \pm y_2^2 \pm \ldots \pm y_n^2.$$

Furthermore, any such coordinate system will give the same numbers of positive and negative terms in the above.

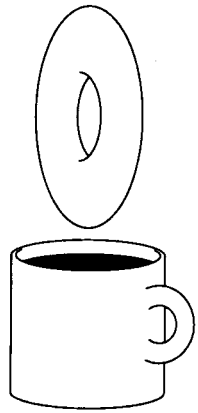


Figure 1.1: Dunking a donut (torus) into coffee

Since our focus is on different tools for computing the Euler characteristic and not on a rigorous development of Morse theory, we refer the reader to section 2 of [Mi] for a proof. We call the number of negative terms $\mathrm{MInd}(f;p)$, the **Morse index** of f at p; intuitively, the Morse index measures the number of independent directions in which f decreases.

For any real number a, let

$$M^a = f^{-1}((-\infty, a]).$$

The intuitive picture is a follows. Say we are dunking a donut into a cup of coffee, as in Figure 1; the manifold in question is the torus T which is the surface of this donut. Define the function $h : T \to \mathbb{R}$ by

$$\begin{aligned} h(p) &= \text{the height of the submerged part of } T \text{ when } p \text{ first touches the coffee} \\ &= \text{the vertical distance from the bottom of the donut to } p. \end{aligned}$$

We will call h, and its later generalizations, the "dunking function." It is not hard to check that h is a Morse function. Figure 2 shows T^a for various values of a. The set of critical points of h is $\{p_0, p_1, p_2, p_3\}$, as pictured. Their indices are:

$$\begin{aligned} \mathrm{MInd}(h; p_0) &= 0 \\ \mathrm{MInd}(h; p_1) &= 1 \\ \mathrm{MInd}(h; p_2) &= 1 \\ \mathrm{MInd}(h; p_3) &= 2. \end{aligned}$$

This is not hard to see: p_0 is a local (in fact, global) minimum, so any direction is a direction of increase, so it has index 0. p_1 decreases if you walk down towards p_0, and increases if you want up the inside of the hole towards p_2, so it has index 1. And so forth.

The first major application of the Morse index, and the one we care about for our purposes, is that it allows you to construct a CW complex homotopy equivalent to M.

Theorem 2. *Let p be a critical point of the Morse function $f : M \to \mathbb{R}$, and set $a = f(p)$. Suppose $f^{-1}([a-\epsilon, a+\epsilon])$ for some $\epsilon > 0$ is compact and contains no critical points other than p. Then $M^{a+\epsilon}$ has the homotopy type of $M^{a-\epsilon}$, with a cell of dimension $\mathrm{MInd}(f;p)$ adjoined.*

(For a proof, see section 3 of [Mi].) So a Morse function gives us a CW structure on M, up to homotopy equivalence. And since χ is a homotopy invariant, this is as good as we need. Combined with

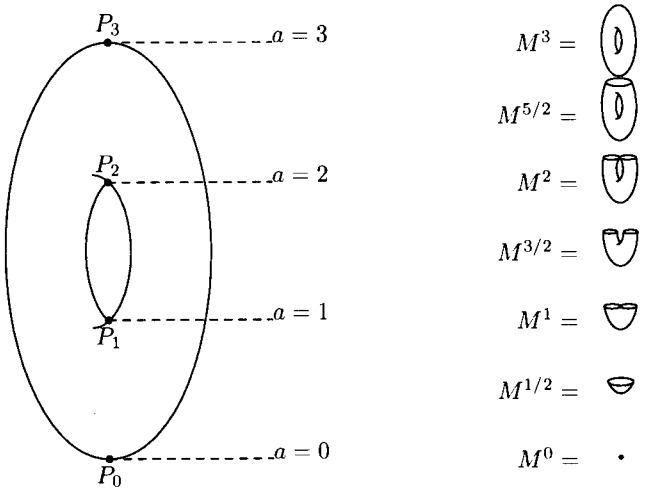


Figure 1.2: The torus, at and between the critical points of its dunking map

the fact that every smooth manifold admits Morse functions (see section 6 of [Mi]), we immediately obtain:

Corollary 3. *Every smooth manifold is homotopy equivalent to a CW complex.*

This implies that the Euler characteristic is defined for all smooth manifolds. If we set

$$A_k(f) = \text{the number of critical points of } f \text{ with Morse index } k$$

and apply Theorem 2, we have

$$\chi(M) = \sum_{k=0}^{n} (-1)^k A_k.$$

Define the surface Σ_g of genus g to be the surface of a g-holed donut; for example, $\Sigma_0 = S^2$ (a "donut hole") and $\Sigma_1 = T$. Consider the "dunking function" h above, but now more generally on any Σ_g; see Figure 3. h always has exactly one maximum and one minimum, and two saddle points (points of Morse index 1) for each hole; we have

$$\begin{aligned} A_0(h) &= 1 \\ A_1(h) &= 2g \\ A_2(h) &= 1 \\ \chi(\Sigma_g) &= 1 - 2g + 1 = 2 - 2g. \end{aligned}$$

1.3 Vector Fields and the Poincaré-Hopf Index Theorem

We now turn to smooth (tangent) vector fields on M. We will think of M as embedded in some $\mathbb{R}^N$ and the vector fields as tangent to $M \subset \mathbb{R}^N$ (if you are aware of the terminology, you may think more abstractly of the vector fields as sections of the tangent bundle TM). For this section only we restrict our attention to the case where M is two-dimensional, but we will indicate the correct generalization to higher dimensions.

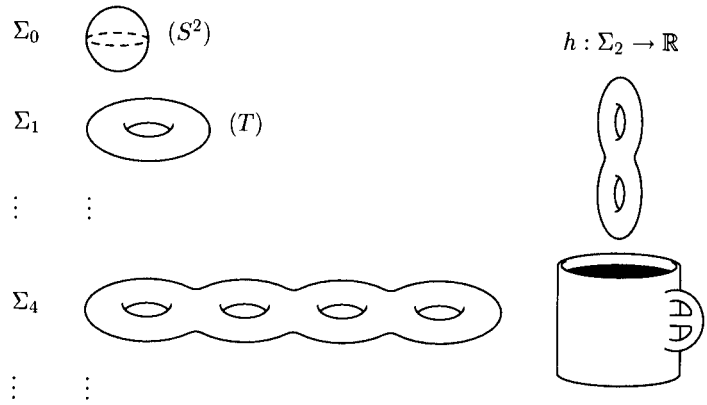


Figure 1.3: Surfaces of higher genus; dunking a two-holed donut into coffee

Let $M \subset \mathbb{R}^N$ be a smooth manifold embedded in Euclidean space. The vector space of vectors tangent to M at a point p, called the **tangent space** T_pM to M at p, is of the same dimension as M. Let $\{x_1, \ldots, x_n\}$ be a smooth coordinate system for M centered at p; that is, p is the point for which $x_j = 0$ for all j. We write $(a_1, \ldots, a_n)$ for the point with coordinate $a_j = x_j$. Then the tangent space can be written as

$$T_pM = \text{span}\{v_1, \ldots, v_n\},$$

where

$$v_j = \frac{d}{dt}|_{t=0}(0, \ldots, 0, t, 0, \ldots, 0)$$

(the only non-zero entry is the j-th). The corresponding picture is that if we were to trace out a curve given by increasing only coordinate x_j, the vector $v_j \in T_pM$ would be the velocity vector of this curve as it passed p. If you are not comfortable or familiar with the language of tangent spaces, you may just picture these vectors as the tangent plane to a surface $M \subset \mathbb{R}^3$. We define a **vector field** on M to be a choice of vector $v(p) \in T_pM$ for each $p \in M$.

Let $v : M \to \mathbb{R}^N$ be a smooth vector field, and let p be an isolated zero of v. Let $y = (y_1, y_2)$ be a local set of coordinates centered at p, and choose a small circle S_ϵ of radius $\epsilon > 0$ centered at p in these coordinates. Then the map

$$\rho_v : S_\epsilon \to S^1$$
$$\rho_v(y) = \frac{v(y)}{|v(y)|}$$

can be defined, and we define the **local index** of v at p to be

$$\text{Ind}(v; p) = w(\rho_v).$$

Here, $w(\rho_v)$ is the winding number of ρ_v around S^1 (the net number of times ρ_v wraps around S^1 when we go around S_ϵ once, with counterclockwise being the positive sense).[3]

To see what local indices look like, consider Figure 4. If we walk around the small dotted-line circle centered at the zero of the vector field, we can see the local index by counting how many counterclockwise revolutions the arrows make. In example (a), the image $\rho_v(x)$ starts pointing to the right,

[3]For the topologically advanced: More generally, for $\dim(M) = n \geq 2$, S_ϵ is an $(n-1)$-sphere, and instead of $w(\rho_v)$, we use the topological degree of the map $\rho_v : S_\epsilon \to S^{n-1}$. One can prove that for ϵ small enough, the local index is well-defined. For more details, see chapter 3 of [Gu].

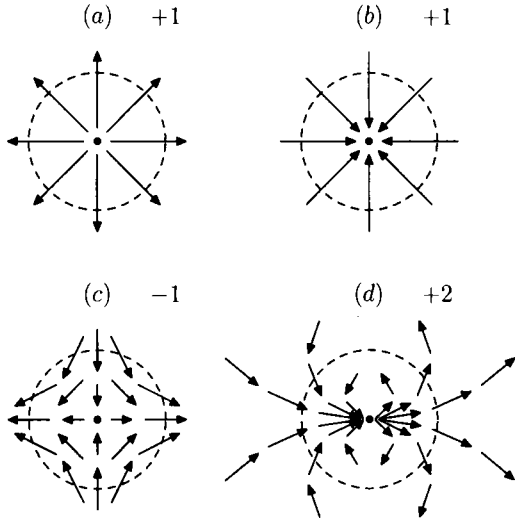


Figure 1.4: Local indices at zeroes of a vector field

then points upwards, then left, then down, and then right again; ρ_v has traversed S^1 once in the counterclockwise direction, so the local index is $+1$. So a "source" has index $+1$. Looking at (b), we see that a "sink" also has index $+1$: starting to the right of the zero, ρ_v starts pointing left, then down, then right, then up, and finally left again. It takes something like the situation in (c) to get a negative local index. Doing the same sort of walk around, ρ_v starts pointing right, then down, then left, then up, and finally right again; we have traversed S^1 once in the clockwise direction. Example (d) shows a local index of $+2$. If v has finitely many zeroes, then we define the **global index** (or simply the **index**) of v to be the global sum of its local indices:

$$\mathrm{Ind}(v) = \sum_{v(x)=0} \mathrm{Ind}(v; x).$$

The remarkable namesake of this section is the following:

Theorem 4. *(The Poincaré-Hopf Index Theorem.) Let v be a smooth vector field on M with finitely many zeroes. Then the global index of v equals the Euler characteristic of M:*

$$\mathrm{Ind}(v) = \chi(M).$$

The two-dimensional case was proved by Jules Henri Poincaré (1854-1912) in 1885; Heinz Hopf (1894-1971) proved the general case in 1927. In particular, the full Poincaré-Hopf Index Theorem predates Morse theory. A proof using Morse theory, however, is popular; see chapter 12 of [Ma]. For a proof using the Lefschetz fixed point theorem, see chapter 3 of [Gu]. The immediate corollary of this theorem is that the global index is the same, regardless of which vector field you choose; this is analogous to the fact that the alternating sum $\sum(-1)^k A_k$ did not depend on the choice of Morse function.

We now use the Poincaré-Hopf Index Theorem to compute again the Euler characteristic of the surface Σ_g. The vector field we will choose is again culinary: the "hot fudge vector field" v_{hf} depicted in Figure 5. Simply stand Σ_g on end as shown, and pour hot fudge over the surface. In an ideal steady state situation, all the fudge enters at one point on top, and all the fudge drips off at one point on the bottom. Then define the value of v_{hf} at a point to be the instantaneous velocity vector of the hot fudge

flow at that point. We have a source at the top and a sink at the bottom (neglecting the inflow and outflow, which are not tangent to the surface), and saddle points (points which look like Figure 4c) at the top and bottom of each hole (you should try to picture this yourself). We saw earlier that sources and sinks have index $+1$ and saddles have index -1, so we conclude

$$\chi(\Sigma_g) = \mathrm{Ind}(v_{hf}) = 1 + (2g)(-1) + 1 = 2 - 2g.$$

If you compare how the computations went here and in the section on Morse theory, in both cases each hole contributed two "negative units" (odd dimensional CW cells or negative index zeroes), and the two ends each contributed one "positive unit." Since the computations are similar in nature, it makes sense that one is able to prove the Poincaré-Hopf Index Theorem using Morse theory.

We conclude this section with a corollary, which contains a famous and amusingly named result as a special case.

Corollary 5. *A smooth manifold M with $\chi(M) \neq 0$ does not admit a smooth, nowhere vanishing vector field.*

Proof. Let v be a smooth vector field on M. Then by the Poincaré-Hopf Index Theorem, $\mathrm{Ind}(v) = \chi(M) \neq 0$. If v were nowhere vanishing, the sum defining $\mathrm{Ind}(v)$ would be empty, forcing $\mathrm{Ind}(v) = 0$; this is impossible. □

Corollary 6. *The surface Σ_g of genus g admits a nowhere vanishing smooth vector field if and only if $g = 1$, that is, if and only if Σ_g is the torus.*

Proof. The "only if" direction is immediate from the previous corollary and our earlier computation, $\chi(\Sigma_g) = 2 - 2g$. Conversely, we can construct a nowhere vanishing vector field on the torus by the process depicted in Figure 6: first take a nowhere vanishing vector field on S^1, and then revolve the entire construction about an axis away from it. □

The special case $g = 0$, that is $\Sigma_0 = S^2$, is known as the "Hairy Ball Theorem." Intuitively, it states that there is always at least one point on the surface of Earth with no wind blowing. Equivalently, if the Earth had hair, it would necessarily have a bald spot.

1.4 An Example: Real Projective Space

Define **real projective space**[4] **of dimension** n to be the quotient space

$$\mathbb{RP}^n = \mathbb{R}^{n+1} - \{0\}/\sim$$
$$v \sim w \quad \Leftrightarrow \quad v = \lambda w \text{ for some } \lambda \in \mathbb{R} - \{0\}.$$

Since the equivalence class of a non-zero vector v is the one-dimensional subspace of $\mathbb{R}^{n+1}$ spanned by v (minus the point 0) and every one-dimensional subspace contains a non-zero vector, $\mathbb{RP}^n$ is just the set of one-dimensional subspaces of $\mathbb{R}^{n+1}$, topologized.

Note that a particular one-dimensional subspace $U \subset \mathbb{R}^{n+1}$ intersects the unit sphere $S^n \subset \mathbb{R}^{n+1}$ in exactly two points, namely the two vectors $v, -v$ of length 1 in U. Thus any even function[5] on S^n determines a function on $\mathbb{RP}^n$; it is easy to check that if such a function is smooth on S^n, it is smooth on $\mathbb{RP}^n$ as well. Let $\{a_0, a_1, \ldots, a_n\}$ be an ordered set of distinct, non-zero real numbers; for simplicity, assume they are in ascending order. Define the function

$$f : S^n \to \mathbb{R}$$
$$f(x) = a_0 x_0^2 + a_1 x_1^2 + \ldots + a_n x_n^2;$$

[4] We borrow greatly from chapter 12 of [Ma] for the first half of this section.

[5] Recall that a function $f : V \to X$ on a vector space V is said to be **even** if $f(v) = f(-v)$ for all $v \in V$.

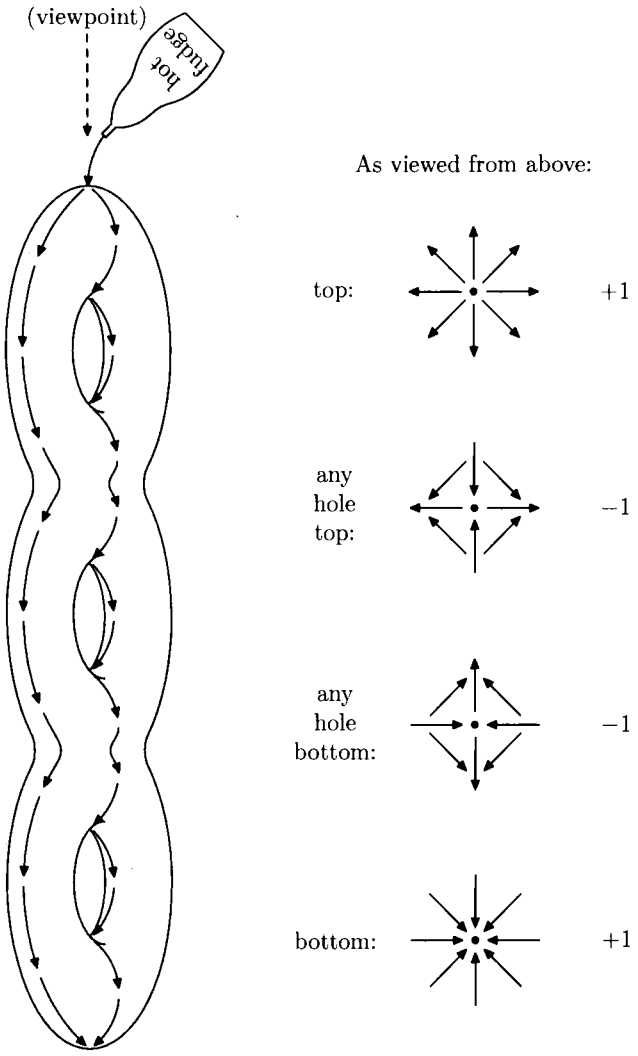


Figure 1.5: A vector field constructed by coating a three-holed donut in hot fudge

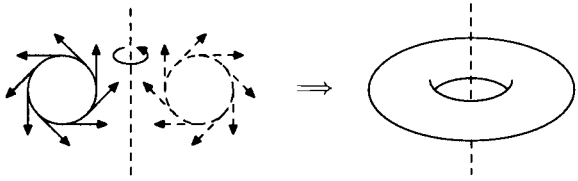


Figure 1.6: Constructing a non-vanishing vector field on the torus

using the standard coordinates $\{x_0, \ldots, x_n\}$ on $\mathbb{R}^{n+1}$. Since f is even, it determines a function on $\mathbb{RP}^n$, which by abuse of notation, we also call f.

We will determine and classify the critical points of f, conclude that it is a Morse function, and use this to build a CW structure on $\mathbb{RP}^n$. Afterwards, we will re-construct this CW structure in a more elementary fashion. As a corollary of either approach, we will compute $\chi(\mathbb{RP}^n)$.

Since the properties of f at a point are local in nature, we can continue working with the explicit embedding $S^n \subset \mathbb{R}^{n+1}$. At the point $x = (x_0, \ldots, x_n)$, the tangent space is

$$T_x S^n = \{v = (v_0, \ldots, v_n) \subset \mathbb{R}^{n+1} : \sum x_i v_i = 0\},$$

and the first partial derivatives are given by

$$\frac{\partial f}{\partial x_i} = 2a_i x_i.$$

However, these are the partial derivatives with respect to the coordinates of the ambient space, $\mathbb{R}^{n+1}$. We do not need all of them to vanish; we merely need the gradient vector to be orthogonal to all vectors in the tangent space (for the more advanced: we need the differential to be the zero linear functional). In other words, we need to show that

$$\sum_{i=0}^{n} \frac{\partial f}{\partial x_i} v_i = 0 \text{ for all } v = (v_0, \ldots, v_n) \in T_x S^n.$$

Since $|x| = 1$, it is impossible for the partial derivatives to all simultaneously vanish due to x being zero; instead, we use the relation $x \cdot v = 0$ for all $v \in T_x S^n$. The above equation holds, then, if and only if $x = (x_0, \ldots, x_n)$ and $(a_0 x_0, \ldots, a_n x_n)$ are parallel. But since the a_i are all distinct, this occurs if and only if $x = \pm e_i$, where e_i is the vector of all zeroes, except for a 1 in the i-th place. There are $2(n+1)$ such points on S^n, but only $n+1$ on $\mathbb{RP}^n$, since $e_i \sim -e_i$.

We now check that e_0 is a nondegenerate critical point and compute its Morse index. A local coordinate system $\{y_1, \ldots, y_n\}$ is defined by

$$(y_1, \ldots, y_n) \in \mathbb{R}^n \leftrightarrow \left(\pm\sqrt{1 - \sum y_i^2}, y_1, \ldots, y_n\right) \in S^n.$$

In terms of these coordinates, f looks like

$$f(y_1, \ldots, y_n) = a_0\left(1 - \sum_{i=1}^{n} y_i^2\right) + \sum_{i=1}^{n} a_i y_i^2 = a_0 + \sum_{i=1}^{n} (a_i - a_0) y_i^2.$$

The matrix of second partial derivatives is just

$$\begin{pmatrix} 2(a_1 - a_0) & & & \\ & 2(a_2 - a_0) & & \\ & & \ddots & \\ & & & 2(a_n - a_0) \end{pmatrix}.$$

Since the a_i are all distinct, the matrix is invertible, and $\pm e_0$ is a nondegenerate critical point. Also, the chosen coordinates are evidently of the form the Morse lemma guarantees, so we can read off the Morse index. Since the points were chosen to be in ascending order, each $a_i - a_0$ is positive and $\mathrm{MInd}(f; \pm e_0) = 0$. The same analysis holds for each $\pm e_k$, with the exception that $(a_0 - a_k), (a_1 - a_k), \ldots, (a_{k-1} - a_k)$ will all be negative. Thus in the general case, we have

$$\mathrm{MInd}(f; \pm e_k) = k.$$

Then the resulting CW structure on $\mathbb{RP}^n$ has one cell in each dimension from 0 through n inclusive, and

$$\chi(\mathbb{RP}^n) = \begin{cases} 1 & n \text{ is even} \\ 0 & n \text{ is odd.} \end{cases}$$

Finally, we give an elementary, geometric construction of this same CW structure. We begin by introducing **homogeneous coordinates** on $\mathbb{RP}^n$; while not coordinates in the usual sense, they are a convenient way of working explicitly in $\mathbb{RP}^n$. We will use $(n+1)$-tuple notation for the $\mathbb{R}^{n+1}$ our copy of $\mathbb{RP}^n$ is obtained from. The homogeneous coordinate for the point $p \in \mathbb{RP}^n$ is $[x_0, \ldots, x_n]$, where $x = (x_0, \ldots, x_n)$ is any non-zero vector in the one-dimensional subspace p of $\mathbb{R}^{n+1}$. In other words, in homogeneous coordinates,

$$[x_0, \ldots, x_n] = [y_0, \ldots, y_n] \quad \Leftrightarrow \quad x_i = \lambda y_i \text{ for all } i, \text{ and } \lambda \neq 0.$$

Also, a bracketed $(n+1)$-tuple $[x_0, \ldots, x_n]$ represents a point of $\mathbb{RP}^n$ if and only if not all its entries are zero.

Define an open subset $U_0 \subset \mathbb{RP}^n$ by

$$U_0 = \{[x_0, \ldots, x_n] \in \mathbb{RP}^n : x_0 \neq 0\}.$$

This is well-defined because nonzero scalar multiplication does not depend upon whether or not $x_0 = 0$, and it is open because it is the inverse image of $\mathbb{R} - \{0\}$ under the even, continuous map on S^n taking each point to the absolute value of its e_0 coordinate. The smooth map $\mathbb{R}^n \to U_0$ given by

$$(x_1, \ldots, x_n) \mapsto [1, x_1, \ldots, x_n]$$

has smooth two-sided inverse

$$[x_0, \ldots, x_n] \mapsto \left(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}\right).$$

This is well-defined because $x_0 \neq 0$, and because scaling all entries on the left does not affect the values on the right. Thus U_0 is diffeomorphic to $\mathbb{R}^n$; it is an n-cell. In order to determine what $\mathbb{RP}^n - U_0$ is, note that a bracketed $(n+1)$-tuple $[x_0, \ldots, x_n]$ is in $\mathbb{RP}^n - U_0$ if and only if $x_0 = 0$ but not all entries are zero; equivalently, a point of $\mathbb{RP}^n - U_0$ is just a choice of $x_1, \ldots, x_n$, not all zero. In other words, this complement is nothing other than a copy of $\mathbb{RP}^{n-1}$. We have found that

$$\mathbb{RP}^n = D^n \cup \mathbb{RP}^{n-1},$$

where we write D^k for an open k-dimensional cell. Noting that $\mathbb{RP}^0$ is just a point and inducting downwards,

$$\mathbb{RP}^n = D^0 \cup D^1 \cup \cdots \cup D^n.$$

This is the desired CW structure. Intuitively, $\mathbb{RP}^n$ contains an n-dimensional plane, and a copy of $\mathbb{RP}^{n-1}$ "at infinity"; this $\mathbb{RP}^{n-1}$ represents all possible directions in $\mathbb{R}^n$, up to identifying opposite directions. For instance, the projective plane contains the ordinary plane, as well as a circle's worth ($\mathbb{RP}^1 = S^1$) of infinities, each point on this circle being a direction in which you can go off to infinity from the plane.

1.5 Conclusion

To recap: as early as Euler, the curious observation had been made that the quantity $F - E + V$ corresponding to a convex polyhedron always equals 2. This so-called Euler characteristic was computed for other sorts of shapes, and results about it were proven, but it was not until the machinery of homotopy invariance became available that these results became "trivial" to prove. Indeed, any convex polyhedron can be "smoothed out" into a sphere, which has Euler characteristic 2.

In cases where we cannot immediately see what the Euler characteristic is by such a geometric trick, we can employ more sophisticated methods in our computations. The results of Morse and of Poincaré and Hopf that we have encountered tell us that given *almost any* vector field or smooth function on a manifold, we can compute the Euler characteristic of that manifold; viewed conversely, we can read these theorems as describing a topological constraint on any vector fields (with finitely many zeroes) or smooth (Morse) functions which may appear on a given manifold.

References

[Gu] Victor Guillemin and Alan Pollack: *Differential Topology*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1974.

[Ha] Allen Hatcher: *Algebraic Topology*. Cambridge: Cambridge Univ. Press, 2002.

[Ma] Ib Madsen and Jørgen Tornehave: *From Calculus to Cohomology: De Rham cohomology and characteristic classes*. Cambridge: Cambridge Univ. Press, 1997.

[Mi] John Milnor: *Morse Theory*, Princeton, N.J.: Princeton Univ. Press, 1969 (Annals of Mathematics Studies **51**).

[Mo] Marston Morse: *The Calculus of Variations in the Large*. Providence, R.I.: American Math. Society, 1934. (Colloquium Publications **18**).

STUDENT ARTICLE

2

Dirichlet's Prime Number Theorem: Algebraic and Analytic Aspects

Igor Rapinchuk '07†
Harvard University
Cambridge, MA 02138
rapinch@fas.harvard.edu

Abstract

The focus of this paper is the famous theorem on primes in arithmetic progressions due to Dirichlet: *if a and $m > 0$ are relatively prime integers, then there exist infinitely many primes of the form $a + km$ with k a positive integer.* The proof of this theorem in the general case uses analytic techniques, and in fact some key statements heavily rely on complex analysis. The case $a = 1$, however, can be handled by purely algebraic methods as we will show in Section 2.1 following suggestions given in [La]. In Section 2.2, we will outline the idea of the proof of Dirichlet's theorem as it is presented in [IR] and [Kn] for the case $m = 4$. Finally, in Section 2.3, after a brief discussion of characters of finite abelian groups following [Se], we will present the proof of Dirichlet's theorem (cf. [IR, Kn, Se]).

2.1 There are infinitely many primes $p \equiv 1 (\mathrm{mod}\ m)$: algebraic proof

Let $P = \{2, 3, 5, \cdots\}$ be the set of all primes. For relatively prime integers a and $m > 0$ we let

$$P_{a(m)} = \{p \in P \mid p \equiv a (\mathrm{mod}\ m)\}.$$

Dirichlet's Prime Number Theorem states that $P_{a(m)}$ is always infinite. In this section, we will prove this for $a = 1$ by using purely algebraic techniques. It is interesting that the argument can be traced back to Euclid's proof of the fact that P is infinite: if $P = \{p_1, \ldots, p_r\}$ then for any prime factor p of $p_1 \cdots p_r + 1$ we have $p \notin \{p_1, \ldots, p_r\}$, a contradiction. We will now do a couple of simple examples which demonstrate that suitable modifications of Euclid's method allow one to find infinitely many primes in certain arithmetic progressions.

Proposition 1. *The sets $P_{1(4)}$ and $P_{3(4)}$ are infinite.*

Proof. $P_{1(4)}$: We will use the well-known fact that primes in $P_{1(4)}$ (in other words, primes of the form $4k + 1$) can be characterized as those primes > 2 for which the congruence $x^2 \equiv -1 (\mathrm{mod}\ p)$ has a solution. Assume that $P_{1(4)}$ contains only finitely many primes, say, $p_1 = 5, p_2 = 13, \ldots, p_n$. Consider $a = 4p_1^2 \cdots p_n^2 + 1$, and let p be a prime factor of a. Then, just as in Euclid's proof, $p \notin \{p_1, \ldots, p_n\}$.

†Igor Rapinchuk '07 is a mathematics concentrator living in Kirkland House. He came to Harvard from Charlottesville, VA, where he graduated from Albemarle High School. His main mathematical interests are in algebraic geometry and algebraic number theory, with related interests in algebra and complex analysis. Following graduation, Igor plans to pursue graduate studies in mathematics, and will, in particular, be spending the next academic year in the Math Tripos, Part III program at the University of Cambridge as a Gates Cambridge Scholar.

On the other hand, $p|a$ implies that $-1 \equiv (2p_1 \cdots p_n)^2 (\mathrm{mod}\ p)$, and therefore $p \in P_{1(4)}$ (as obviously $p > 2$). So, p is a "new" prime in $P_{1(4)}$, contradicting our original assumption. Thus $P_{1(4)}$ is infinite.

$P_{3(4)}$: Again, assume that $P_{3(4)}$ contains only finitely many primes: $p_1 = 3, p_2 = 7, \ldots, p_n$. Consider $b = 4p_2 \cdots p_n + 3$. Clearly, b is odd, not divisible by 3, and satisfies $b \equiv 3(\mathrm{mod}\ 4)$. Then all prime factors of b cannot be belong to $P_{1(4)}$ as otherwise we would have $b \equiv 1(\mathrm{mod}\ 4)$. Since $P = \{2\} \cup P_{1(4)} \cup P_{3(4)}$, we conclude that b has a prime factor $p \in P_{3(4)}$. But obviously $p \notin \{p_1, \ldots, p_n\}$, which again yields a contradiction. □

It is important to observe that the above argument for $P_{1(4)}$ already contains the idea that we will use to prove that $P_{1(m)}$ is infinite for any m : show that there exists a polynomial $f(X) \in \mathbb{Z}[X]$ (for $m = 4$ we used $f(X) = X^2 + 1$) such that *any* prime factor $p \nmid m$ of $f(a)$, where $a \in \mathbb{Z}$, belongs to $P_{1(m)}$, and on the other hand, the values $f(a)$ as a runs through $\mathbb{Z}$ have infinitely many prime divisors. We will show that the latter property holds in fact for any nonconstant integer polynomial (Lemma 2), while the former property holds for the m-th cyclotomic polynomial $\Phi_m(X)$ (see the proof of Theorem 4). This approach to proving that $P_{1(m)}$ is infinite is suggested in Problems 20 and 21 in Ch. VI of [La]. We also notice that our argument for $P_{3(4)}$ depends on the fact that an odd prime can get only in one of the two classes, $P_{1(4)}$ or $P_{3(4)}$, mod 4, and therefore may not be generalizable for $m > 4$.

Lemma 2. (Problem 20 in [La], Ch. VI) *Let*

$$f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0 \in \mathbb{Z}[X]$$

be a nonconstant polynomial. Then the nonzero values $f(a)$ with $a \in \mathbb{Z}^+$ are divisible by infinitely many primes.

Proof. We can assume that $a_0 \neq 0$ as otherwise for any prime p, the value $f(pa)$ is divisible by p for any $a \in \mathbb{Z}$, and of course one can pick an a so that $f(pa) \neq 0$. Next, observe that

$$f(a_0 X) = a_0 g(X) \text{ where } g(X) = a_n a_0^{n-1} X^n + \cdots + 1,$$

so it is enough to show that the nonzero values $g(a)$ with $a \in \mathbb{Z}^+$ are divisible by infinitely many primes. In other words, we can assume that $a_0 = 1$. Suppose that the nonzero values $f(a)$ for $a \in \mathbb{Z}$ are divisible only by finitely many primes, say, $p_1, \ldots, p_r$. Consider $F(X) = f(p_1 \cdots p_r X)$. Then $F(X)$ is a nonconstant integer polynomial of degree n, hence assumes each value at not more than n values of the variable. In particular, there exists $a \in \mathbb{Z}^+$ such that $F(a) \neq 0, \pm 1$. Then it follows from our construction that $F(a)$ is divisible by some p_i where $i \in \{1, \ldots, r\}$. But

$$F(a) = a_n (p_1 \cdots p_r a)^n + \cdots + a_1 (p_1 \cdots p_r a) + 1,$$

so the fact that $p_i | F(a)$ implies that $p_i | 1$. This is a contradiction, proving the lemma. □

Obviously, the above proof of Lemma 2 is based on the same idea as Euclid's proof. We will now give another proof of Lemma 2 which gives some additional quantitative information. For a subset $A \subset \mathbb{Z}$ and a natural number N we let $A(N) = \{a \in A \mid |a| \leq N\}$. We will use the following simple idea: given two subsets $A, B \subset \mathbb{Z}$, to show that $A \not\subset B$ it is enough to find N such that $|A(N)| > |B(N)|$. We will apply this idea to the sets

$$A = \{f(a) \mid a \in \mathbb{Z}^+ \text{ and } f(a) \neq 0\}$$

and, assuming that the numbers in A are divisible only by finitely many primes $p_1, \ldots, p_r$,

$$B = \{p_1^{\alpha_1} \cdots p_r^{\alpha_r}\}.$$

Let $M = \max\{|a_n|, \ldots, |a_0|\}$. Then for any $a \in \mathbb{Z} \setminus \{0\}$ we have

$$|f(a)| \leq |a_n||a|^n + \cdots + |a_0| \leq M(n+1)|a|^n.$$

It follows that if $d \in \mathbb{N}$ is such that $M(n+1)d^n \leq N$ then all the nonzero numbers among $f(1), \ldots, f(d)$ belong to $A(N)$. Since f assumes each value at not more than n different values of the variable, we get that

$$|A(N)| \geq \frac{d-n}{n} = \frac{d}{n} - 1 \geq \frac{1}{n}\left(\left(\frac{N}{M(n+1)}\right)^{1/n} - 1\right) - 1$$

because for d one can take

$$d = \left[\left(\frac{N}{M(n+1)}\right)^{1/n}\right] > \left(\frac{N}{M(n+1)}\right)^{1/n} - 1. \tag{2.1}$$

Since $(1+n)^{1/n} \leq 2$, we finally get that

$$|A(N)| \geq \frac{N^{1/n}}{2M^{1/n}n} - 2.$$

On the other hand, since $p_i \geq 2$, we see that $p_1^{\alpha_1} \cdots p_r^{\alpha_r} \leq N$ implies that

$$\alpha_1 + \cdots \alpha_r \leq \log_2 N,$$

and in particular, $\alpha_i \leq \log_2 N$ for all $i = 1, \ldots, r$. It follows that

$$|B(N)| \leq (\log_2 N + 1)^r.$$

Since $N^{1/n}/(\log_2 N)^r \to \infty$ as $N \to \infty$, we find that

$$|A(N)| > |B(N)|$$

for all sufficiently large N. Thus, $A \not\subset B$, which yields another proof of Lemma 2. In fact, we proved the following.

Proposition 3. *Fix a natural number r and pick N so that*

$$\frac{N^{1/n}}{2M^{1/n}n} - 2 > (\log_2 N + 1)^r.$$

If d is defined by (2.1) then the nonzero numbers among $f(1), f(2), \ldots, f(d)$ have at least $(r+1)$ distinct prime divisors.

We are now ready to prove the main result of this section.

Theorem 4. *For any $m > 0$, the set $P_{1(m)}$ is infinite.*

Let $\Phi_m(X)$ denote the m-th cyclotomic polynomial (cf. [Co], Sec. 9.1, or [La], Ch. VI, Sec. 3).

Lemma 5. (Problem 21(a) in [La], Ch. VI) *Let p be a prime, a and $m > 0$ be integers prime to p. Then $p|\Phi_m(a)$ if and only if the image $\bar{a}$ of a in $(\mathbb{Z}/p\mathbb{Z})^*$ has order (exactly) m.*

Proof. First, suppose $\bar{a}$ has order m in $(\mathbb{Z}/p\mathbb{Z})^*$. Then $\bar{a}^m = \bar{1}$, or equivalently $p|(a^m - 1)$. On the other hand, for any d such that $0 < d < m$, we have $\bar{a}^d \neq \bar{1}$, and therefore $p \nmid (a^d - 1)$. By Proposition 9.1.5 in [Co], we have

$$X^m - 1 = \prod_{d|m} \Phi_d(X) \tag{2.2}$$

and therefore

$$a^m - 1 = \prod_{d|m} \Phi_d(a). \tag{2.3}$$

Let d be a proper divisor of m. Since $\Phi_d(a)|(a^d-1)$, it follows from the above that $p \nmid \Phi_d(a)$. On the other hand, $p|(a^m-1)$, so we conclude from (2.3) that $p|\Phi_m(a)$.

Conversely, suppose $p|\Phi_m(a)$. Then it follows from (2.3) that $p|(a^m-1)$, i.e. $\bar{a}^m = \bar{1}$. This means that the order of $\bar{a}$ divides m. Suppose the exact order of $\bar{a}$ is $m' < m$ (clearly, $m'|m$). Then using a factorization similar to (2.3) in which m is replaced with m' we see that there exist $d|m'$ such that $p|\Phi_d(a)$ (of course, $d < m'$). Then $\bar{a}$ is a root of both reductions $\overline{\Phi_n(X)}$ and $\overline{\Phi_d(X)}$ mod p. It follows from (2.2) that $\bar{a}$ is a multiple root of $\overline{X^m-1}$. But since $p \nmid m$, the latter has no multiple roots. A contradiction, proving that the order of $\bar{a}$ is exactly m. □

Proof of Theorem 4. First, let us show that for a prime $p \nmid m$, the conditions $p|\Phi_m(a)$ and $p \equiv 1 (\mathrm{mod}\ m)$ are equivalent (Problem 21(b) in [La], Ch. VI). Indeed, if $p|\Phi_m(a)$ then by Lemma 5, the order of $\bar{a}$ is m. Thus, $(\mathbb{Z}/p\mathbb{Z})^*$ contains an element of order m, and therefore its order $p-1$ is divisible by m, i.e. $p \equiv 1 (\mathrm{mod}\ m)$. Conversely, suppose $p \equiv 1 (\mathrm{mod}\ m)$. Since the group $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic of order $p-1$, it contains an element $\bar{a}$ of order m. Then by Lemma 5, $p|\Phi_m(a)$.

Now, by Lemma 2, the values $\Phi_m(a)$ with $a \in \mathbb{Z}$ are divisible by infinitely many primes. As we have seen, all these primes belong to $P_{1(m)}$, implying that $P_{1(m)}$ itself is infinite. □

Since cyclotomic polynomials can be described explicitly (see [La], pg. 280), one can use Proposition 3 to find, for given m and r, a natural number d such that among prime divisors of the integers $\Phi_m(1), \ldots, \Phi_m(d)$ there are at least r distinct primes $\equiv 1 (\mathrm{mod}\ m)$. For example, if m is a prime then the cyclotomic polynomial $\Phi_m(X)$ has degree $n = m-1$ and the maximum of its coefficients is $M = 1$. So, if we choose N so that

$$\frac{N^{1/(m-1)}}{2(m-1)} - 2 > (\log_2 N + 1)^r$$

and define d by (2.1) then the prime divisors $\neq m$ of the numbers $\Phi_m(1), \ldots, \Phi_m(d)$ yield at least r distinct primes in $P_{1(m)}$.

2.2 The idea of the proof of Dirichlet's Theorem

The idea of Dirichlet's proof of the Prime Number Theorem can be traced back to Euler's proof of the fact that there exist infinitely many primes. Euler considered the generalized harmonic series

$$\sum \frac{1}{n^s}. \tag{2.4}$$

For $s \in \mathbb{C}$, we have $|n^s| = n^{\mathrm{Re}\, s}$, so it follows that (2.4) converges whenever $\mathrm{Re}\, s > 1$. (In fact, it converges absolutely, implying in particular that the series obtained by any permutation of the terms of (2.4) converges to the same number, see [Ru], Theorem 3.55.) The sum of (2.4) for $s \in \mathbb{C}$ such that $\mathrm{Re}\, s > 1$ is denoted $\zeta(s)$, and the correspondence $s \mapsto \zeta(s)$ is called the **(Riemann) zeta function**. The key step in Euler's proof is the following.

Lemma 6. *For $s \in \mathbb{C}$, $\mathrm{Re}\, s > 1$, we have*

$$\zeta(s) = \prod_{p \in P} \frac{1}{1-p^{-s}}, \tag{2.5}$$

where P is the set of all primes.

Proof. We recall that we write $a = \prod_{n=1}^{\infty} a_n$ if $\lim_{d\to\infty} \prod_{n=1}^{d} a_n = a$. In (2.5), we consider the natural order on $P = \{p_1, \ldots, p_d, \ldots\}$, so that

$$\prod_{p \in P} \frac{1}{1-p^{-s}} = \prod_{i=1}^{|P|} \frac{1}{1-p_i^{-s}}$$

where the cardinality $|P|$ is either a finite (natural) number or infinity (in fact, the order on P doesn't matter). Fix $d \geq 1$, and let $\mathbb{N}_d$ denote the set of natural numbers whose prime factors belong to $\{p_1, \ldots, p_d\}$. Since

$$\frac{1}{1-p^{-s}} = \sum_{n=0}^{\infty} \frac{1}{p^{ns}}$$

and the geometric series in the right-hand side is absolutely convergent, we have

$$\prod_{i=1}^{d} \frac{1}{1-p_i^{-s}} = \sum_{n \in \mathbb{N}_d} \frac{1}{n^s}, \tag{2.6}$$

as absolutely convergent series can be multiplied term-by-term (cf. [Ru], Theorem 3.50). Notice that the order of summation in the right-hand side of (2.6) doesn't matter as the series converges absolutely. Now, we have

$$\zeta(s) - \prod_{i=1}^{d} \frac{1}{1-p_i^{-s}} = \sum_{n \in \mathbb{N} - \mathbb{N}_d} \frac{1}{n^s}.$$

Clearly, any number in $\mathbb{N} - \mathbb{N}_d$ is strictly greater than $p_d \geq d$, so

$$\left| \sum_{n \in \mathbb{N} - \mathbb{N}_d} \frac{1}{n^s} \right| \leq \sum_{n=d+1}^{\infty} \frac{1}{n^{\operatorname{Re} s}} \longrightarrow 0 \text{ as } d \to \infty,$$

and (2.5) follows. □

Now, suppose that P is finite. Then $\prod_{p \in P} \frac{1}{1-p^{-1}}$ is a finite number, say A. For any $s \in \mathbb{R}$, $s > 1$, we have

$$\zeta(s) = \prod_{p \in P} \frac{1}{1-p^{-s}} \leq \prod_{p \in P} \frac{1}{1-p^{-1}} = A;$$

i.e. $\zeta(s)$ is bounded above by A as $s \to 1^+$. Let us show that this is not the case. For any $d \in \mathbb{N}$, we have

$$\sum_{n=1}^{d} \frac{1}{n^s} \leq \zeta(s) \leq A.$$

Taking the limit as $s \to 1^+$, we get $\sum_{n=1}^{d} 1/n \leq A$ for all d. This implies that the harmonic series $\sum_{n=1}^{\infty} 1/n$ converges, a contradiction. Thus, P is infinite. Using a bit more analysis, we can derive the following stronger statement, which is crucial for Dirichlet's Theorem.

Proposition 7. *For $s \in \mathbb{C}$, $\operatorname{Re} s > 1$, let*

$$\lambda(s) = \sum_{p \in P} \frac{1}{p^s}.$$

Then $\lambda(s)$ is unbounded as $s \to 1^+$ in $\mathbb{R}$, and consequently the series $\sum_{p \in P} 1/p$ diverges.

Proof. Since $\zeta(s) > 0$ for $s > 1$, we derive from (2.5) that

$$\ln \zeta(s) = \sum_{p \in P} -\ln\left(1 - p^{-s}\right). \tag{2.7}$$

Using the expansion

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \quad \text{for } |x| < 1,$$

we get

$$-\ln\left(1 - p^{-s}\right) = \frac{1}{p^s} + \frac{1}{2p^{2s}} + \frac{1}{3p^{3s}} + \cdots. \tag{2.8}$$

Let

$$g_p(s) = \frac{1}{2p^{2s}} + \frac{1}{3p^{3s}} + \cdots.$$

Clearly, for any $s > 1$ we have

$$0 < g_p(s) \leq \frac{1}{2p^{2s}}\left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots\right) = \frac{1}{p^{2s}} \cdot \frac{1}{2(1-p^{-s})} \leq \frac{1}{p^{2s}}.$$

It follows that for any d,

$$\sum_{i=1}^{d} g_{p_i}(s) \leq \sum_{i=1}^{d} \frac{1}{p_i^{2s}} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} = \zeta(2).$$

So, for any $s > 1$, the series $\sum_{p \in P} g_p(s)$ converges and its sum $g(s)$ satisfies $0 \leq g(s) \leq \zeta(2)$; in particular $g(s)$ remains bounded as $s \to 1^+$. On the other hand, by combining (2.7) and (2.8), we obtain

$$\ln \zeta(s) = \lambda(s) + g(s).$$

Since $\zeta(s)$ is unbounded and $g(s)$ is bounded as $s \to 1^+$, we conclude that $\lambda(s)$ is unbounded.

Now, suppose the series $\sum_{p \in P} 1/p$ converges, say to B. Then for any $s > 1$ and any $m \in \mathbb{N}$ we have

$$\sum_{i=1}^{m} \frac{1}{p_i^s} \leq \sum_{i=1}^{m} \frac{1}{p_i} \leq B.$$

Taking the limit as $m \to \infty$, we get $\lambda(s) \leq B$, a contradiction. □

The idea of the proof of Dirichlet's Theorem is to establish an analog of Proposition 7 for the function which is defined just like λ, but using, instead of all primes, only those primes that occur in a given arithmetic progression. More precisely, for $s \in \mathbb{C}$, Re $s > 1$, define

$$\nu_{a(m)}(s) = \sum_{p \in P_{a(m)}} \frac{1}{p^s}.$$

Then to prove that $P_{a(m)}$ is infinite (which is what Dirichlet's theorem claims) it is enough to show that $\nu_{a(m)}(s)$ is unbounded as $s \to 1^+$. In the remaining part of this section we will show (following [IR], Ch. 16, Sec. 2 and [Kn], Ch. VII, Sec. 1) how this idea can be implemented for $m = 4$; in other words, we will show that $P_{1(4)}$ and $P_{3(4)}$ are infinite.

We obviously have

$$\lambda(s) = 2^{-s} + \nu_{1(4)}(s) + \nu_{3(4)}(s),$$

so it follows from Proposition 7 that the function

$$\lambda_+(s) = \nu_{1(4)}(s) + \nu_{3(4)}(s)$$

is unbounded as $s \to 1^+$, and therefore at least one of the functions $\nu_{1(4)}(s)$ or $\nu_{3(4)}(s)$ has this property. What we want to show is that *both* functions have this property. For this we need to identify the contributions of $\nu_{1(4)}(s)$ and $\nu_{3(4)}(s)$ to $\lambda_+(s)$ separately. The sets $P_{1(4)}$ and $P_{3(4)}$ can be separated by the following function χ defined on $\mathbb{Z}$:

$$\chi(n) = \begin{cases} 0 & n \equiv 0 \pmod 2, \\ 1 & n \equiv 1 \pmod 4, \\ -1 & n \equiv 3 \pmod 4. \end{cases}$$

Consider

$$\lambda_-(s) = \sum_{p \in P} \frac{\chi(p)}{p^s}.$$

(Notice that this series absolutely converges for all $s \in \mathbb{C}$, Re $s > 1$.) Clearly,

$$\nu_{1(4)}(s) = \frac{1}{2}(\lambda_+(s) + \lambda_-(s)) \text{ and } \nu_{3(4)}(s) = \frac{1}{2}(\lambda_+(s) - \lambda_-(s)).$$

So, since $\lambda_+(s)$ is unbounded as $s \to 1^+$, to prove that both $\nu_{1(4)}(s)$ and $\nu_{3(4)}(s)$ have this property, it is enough to show that $\lambda_-(s)$ remains bounded.

Proposition 8. *The function $\lambda_-(s)$ remains bounded as $s \to 1^+$.*

Proof. Consider the series

$$L_-(s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

This series converges absolutely for all $s \in \mathbb{C}$, Re $s > 1$, but its real advantage over $\lambda_-(s)$ is that it is alternating, and therefore its sum can be easily estimated (notice that $\lambda_-(s) = -3^{-s} + 5^{-s} - 7^{-s} - 11^{-s} + \cdots$ is not alternating). We have

$$L_-(s) = 1 - 3^{-s} + 5^{-s} - 7^{-s} + \cdots = (1 - 3^{-s}) + (5^{-s} - 7^{-s}) + \cdots$$

from which it follows that $L_-(s) > (1 - 3^{-s}) > 2/3$ for all $s > 1$. Similarly, from

$$L_-(s) = 1 - (3^{-s} - 5^{-s}) - (7^{-s} - 9^{-s}) - \cdots$$

we conclude that $L_-(s) < 1$ for all $s > 1$. To connect $L_-(s)$ and $\lambda_-(s)$, we observe that the function χ is a multiplicative homomorphism, using which and repeating the proof of Lemma 6 word-for-word, one proves that

$$L_-(s) = \prod_{p \in P} \frac{1}{1 - \chi(p)p^{-s}}$$

(see Proposition 16(i) for a general statement). Then proceeding as in the proof of Proposition 7, we see that

$$\ln L_-(s) = \sum_{p \in P} -\ln(1 - \chi(p)p^{-s})$$

and

$$-\ln(1 - \chi(p)p^{-s}) = \frac{\chi(p)}{p^s} + \frac{\chi(p)^2}{2p^{2s}} + \frac{\chi(p)^3}{3p^{3s}} + \cdots$$

It follows that

$$\ln L_-(s) = \lambda_-(s) + h(s)$$

where $h(s)$ is a function that remains bounded as $s \to 1^+$. We showed above that $2/3 < L_-(s) < 1$ for all $s > 1$, so the boundedness of $\lambda_-(s)$ as $s \to 1^+$ follows. □

2.3 The proof of Dirichlet's Theorem

The function χ used in Section 2.2 to separate $P_{1(4)}$ and $P_{3(4)}$ can be viewed as a character of $(\mathbb{Z}/4\mathbb{Z})^*$ extended by 0 on the numbers (or classes of numbers mod 4) that are not relatively prime to 4. So, it is not surprising that the proof of Dirichlet's theorem for arbitrary m uses characters of $(\mathbb{Z}/m\mathbb{Z})^*$ extended to $\mathbb{Z}/m\mathbb{Z}$ by 0 on the classes that are not relatively prime to m. For this reason, we begin with a brief discussion of characters of finite abelian groups, following [Se], Ch. VI, Sec. 1.

Let G be a finite abelian group. By a **character** of G we mean a group homomorphism $\chi\colon G \to \mathbb{C}^*$. All characters of G form a group under the operation $(\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g)$, which will be denoted $\widehat{G}$ and called the **dual** of G.

Example 2.3.1. Let $G = \mathbb{Z}/n\mathbb{Z}$. Then any $\chi \in \widehat{G}$ is completely determined by its value $\chi(\bar{1})$. Since $\bar{1}$ has order n, we get $\chi(\bar{1})^n = 1$, i.e. $\chi(\bar{1})$ belongs to the group μ_n of n-th roots of unity. Conversely, given any $\zeta \in \mu_n$, the correspondence $\chi\colon \bar{a} \to \zeta^a$ is a character of G such that $\chi(\bar{1}) = \zeta$. Thus, the map

$$\widehat{G} \ni \chi \mapsto \chi(\bar{1}) \in \mu_n$$

is a bijection. Moreover, the equation $(\chi_1\chi_2)(\bar{1}) = \chi_1(\bar{1})\chi_2(\bar{1})$ tells us that this map is a group homomorphism, hence in fact a group isomorphism. Thus, in this example $\widehat{G} \simeq \mu_n$ (noncanonically), which means that a finite cyclic group is isomorphic to its group of characters. Furthermore, if $\zeta_n = \cos(2\pi/n) + i\sin(2\pi/n)$ then the corresponding character $\chi(\bar{a}) = \zeta_n^a$ has the property $\chi(\bar{a}) \neq 1$ whenever $\bar{a} \neq \bar{0}$, so for any nontrivial element of a cyclic group there is a character that does not vanish on this element.

We will now extend these observations to arbitrary finite abelian groups.

Proposition 9. *Let G be a finite abelian group. Then*

(i) $G \simeq \widehat{G}$ (noncanonically), in particular, $|G| = |\widehat{G}|$;

(ii) for any $g \in G$, $g \neq e$, there exists $\chi \in \widehat{G}$ such that $\chi(g) \neq 1$.

Proof. We first observe that if $G = G_1 \times G_2$ then the correspondence

$$\widehat{G} \xrightarrow{\theta} \widehat{G_1} \times \widehat{G_2}, \quad \chi \mapsto (\chi|G_1, \chi|G_2),$$

is an isomorphism of groups. Indeed, it follows from the definition of multiplication on the character group that θ is a group homomorphism. Since G_1 and G_2 generate G, θ is injective. Finally, given $(\chi_1, \chi_2) \in \widehat{G_1} \times \widehat{G_2}$, the map $\chi\colon G \to \mathbb{C}^*$ defined by $\chi(g) = \chi_1(g_1)\chi_2(g_2)$ if $g = (g_1, g_2)$ is a character of G which restricts to χ_1 and χ_2 on G_1 and G_2 respectively, proving that θ is surjective.

By the structure theorem for finite abelian groups (see [Ar], Theorem 12.6.4), $G \simeq G_1 \times \cdots \times G_r$, where G_i are cyclic groups. Then it follows by induction from the above remark that the correspondence

$$\widehat{G} \xrightarrow{\iota} \widehat{G_1} \times \cdots \times \widehat{G_r}, \quad \chi \mapsto (\chi|G_1, \ldots, \chi|G_r),$$

is a group isomorphism. According to the example, $\widehat{G_i} \simeq G_i$ for all $i = 1, \ldots, r$, yielding (i). If now $g \in G$ is a nontrivial element then $g = (g_1, \ldots, g_r)$ and there exists an i such that g_i is nontrivial. As we observed in the example, there exists $\chi_i \in \widehat{G_i}$ such that $\chi_i(g_i) \neq 1$. Then the character $\chi \in \widehat{G}$ corresponding under ι to the r-tuple $(\chi_{01}, \ldots, \chi_i, \ldots, \chi_{0r})$, where χ_{0j} is the trivial character of G_j, has the property $\chi(g) \neq 1$. □

Corollary 10. *Let H be a subgroup of G, and let $\widehat{G} \xrightarrow{\rho} \widehat{H}$ be the homomorphism given by restriction. Then ρ is surjective.*

Proof. Assume the contrary. Since $|\widehat{G}| = |G|$ and $|\widehat{H}| = |H|$, this means that $|\ker \rho| > [G : H]$. But any $\chi \in \ker \rho$, having trivial restriction to H, induces a character of $\bar{\chi} \in \widehat{G/H}$ defined by $\bar{\chi}(gH) = \chi(g)$. Clearly, the map $\ker \rho \to \widehat{G/H}$, $\chi \mapsto \bar{\chi}$, is injective, so we obtain $|G/H| = |\widehat{G/H}| > [G : H]$, a contradiction. □

For a fixed $g \in G$, the map $\delta_g \colon \widehat{G} \to \mathbb{C}^*$, $\delta_g(\chi) = \chi(g)$, is a character of $\widehat{G}$. Moreover, the map $\varepsilon \colon G \to \widehat{\widehat{G}}$, $g \mapsto \delta_g$ is a group homomorphism.

Corollary 11. *ε is a group isomorphism. Thus, G is (canonically) isomorphic to its second dual $\widehat{\widehat{G}}$.*

Indeed, it follows from (ii) that ε is injective. On the other hand, by (i), $|G| = |\widehat{G}| = |\widehat{\widehat{G}}|$, whence ε is an isomorphism.

The following proposition and especially its corollaries play a crucial role in the proof of Dirichlet's theorem.

Proposition 12. *(i) Let $\chi \in \widehat{G}$. Then*

$$\sum_{x \in G} \chi(x) = \begin{cases} |G| & \chi \text{ is trivial,} \\ 0 & \text{otherwise.} \end{cases}$$

(ii) Let $x \in G$. Then

$$\sum_{\chi \in \widehat{G}} \chi(x) = \begin{cases} |G| & x = e, \\ 0 & x \neq e. \end{cases}$$

Proof. (i): The first assertion is clear. To prove the second, pick $y \in G$ so that $\chi(y) \neq 1$. Then

$$\sum_{x \in G} \chi(x) = \sum_{x \in G} \chi(xy) = \left(\sum_{x \in G} \chi(x) \right) \chi(y)$$

It follows that

$$(\chi(y) - 1) \sum_{x \in G} \chi(x) = 0,$$

and therefore $\sum_{x \in G} \chi(x) = 0$.

(ii): In the notations introduced prior to Corollary 11,

$$\sum_{\chi \in \widehat{G}} \chi(x) = \sum_{\chi \in \widehat{G}} \delta_x(\chi).$$

Since $\delta_x = 1 \Leftrightarrow x = 1$, our claim follows from part (i) applied to $\widehat{G}$. □

Corollary 13. *For $x, y \in G$ we have*

$$\sum_{\chi \in \widehat{G}} \chi(x)^{-1} \chi(y) = \begin{cases} |G| & x = y, \\ 0 & x \neq y. \end{cases}$$

Indeed, $\sum_{\chi \in \widehat{G}} \chi(x)^{-1} \chi(y) = \sum_{\chi \in \widehat{G}} \chi(x^{-1}y)$, so we can apply (ii).

Now, fix $m \geq 1$ and let $G_m = (\mathbb{Z}/m\mathbb{Z})^*$; clearly, $|G_m| = \varphi(m)$. Given $\chi \in \widehat{G_m}$, we extend it to a function on all of $\mathbb{Z}/m\mathbb{Z}$ by defining its value to be 0 on classes mod m that are not relatively prime to

m. Composing this function with the canonical homomorphism $\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ we obtain a function on $\mathbb{Z}$ that will be denoted by the same letter χ. Notice that $\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in \mathbb{Z}$. A special role in the proof is played by (the function on $\mathbb{Z}$ obtained from) the trivial character χ_0 which in this context is called the **principal character.** Thus, $\chi_0(a) = 1$ if a is relatively prime to m, and 0 otherwise. For each $\chi \in \widehat{G_m}$, we define

$$\lambda(s, \chi) = \sum_{p \in P} \frac{\chi(p)}{p^s}.$$

Since $|\chi(a)| \leq 1$ for all $a \in \mathbb{Z}$, the series in the right-hand side absolutely converges for all $s \in \mathbb{C}$, $\operatorname{Re} s > 1$.

Corollary 14. *In the notations introduced in Section 2.2, for any integer a prime to m we have*

$$\nu_{a(m)}(s) = \frac{1}{\varphi(m)} \sum_{\chi \in \widehat{G_m}} \chi(a)^{-1} \lambda(s, \chi)$$

for any $s \in \mathbb{C}$, $\operatorname{Re} s > 1$.

Indeed, using the definition of $\lambda(s, \chi)$ we obtain

$$\begin{aligned} \sum_{\chi \in \widehat{G_m}} \chi(a)^{-1} \lambda(s, \chi) &= \sum_{\chi \in \widehat{G_m}} \chi(a)^{-1} \sum_{p \in P} \frac{\chi(p)}{p^s} \\ &= \sum_{p \in P} \frac{\sum_{\chi \in \widehat{G_m}} \chi(a)^{-1} \chi(p)}{p^s} \\ &= \sum_{p \in P_{a(m)}} \frac{\varphi(m)}{p^s} = \varphi(m) \cdot \nu_{a(m)}(s) \end{aligned}$$

as

$$\sum_{\chi \in \widehat{G_m}} \chi(a)^{-1} \chi(p) = \begin{cases} \varphi(m) & p \equiv a \pmod{m}, \\ 0 & x \neq \text{otherwise.} \end{cases}$$

according to Corollary 13.

The following theorem comprises the most technically complicated part of the proof of Dirichlet's theorem.

Theorem 15. *(i) The function $\lambda(s, \chi_0)$ is unbounded as $s \to 1^+$.*
(ii) For $\chi \neq \chi_0$, the function $\lambda(s, \chi)$ remains bounded as $s \to 1^+$.

Theorem 15 in conjunction with Corollary 14 immediately implies Dirichlet's theorem. Indeed, Theorem 15 implies that the function

$$\nu_{a(m)}(s) = \frac{1}{\varphi(m)} \sum_{\chi \in \widehat{G_m}} \chi(a)^{-1} \lambda(s, \chi)$$

is unbounded as $s \to 1^+$. Since

$$\nu_{a(m)}(s) = \sum_{p \in P_{a(m)}} \frac{1}{p^s},$$

this implies that the set $P_{a(m)}$ is infinite.

The remaining part of this section is devoted to proving Theorem 15. Assertion (i) is easy: we obviously have

$$\lambda(s) = \sum_{p | m} \frac{1}{p^s} + \lambda(s, \chi_0),$$

so the required fact immediately follows from Proposition 7. On the contrary, assertion (ii) is very difficult. First, as we have already seen in the proof of Proposition 8, it may be easier to work instead of $\lambda(s,\chi)$ with a similar expression in which the summation runs over all natural numbers instead of just primes:

$$L(s,\chi)=\sum_{n=1}^{\infty}\frac{\chi(n)}{n^s}.$$

This series absolutely converges for $s\in\mathbb{C}$, $\operatorname{Re} s>1$ and defines a function in this domain which is called the **Dirichlet L-function** corresponding to the character χ. The following proposition relates $L(s,\chi)$ and $\lambda(s,\chi)$.

Proposition 16. *For any character χ mod m and any $s\in\mathbb{C}$, $\operatorname{Re} s>1$, we have the following:*

(i) $L(s,\chi)=\prod_{p\in P}\dfrac{1}{1-\chi(p)p^{-s}}$;

(ii) $\ln L(s,\chi)=\lambda(s,\chi)+g(s,\chi)$ *where $g(s,\chi)$ is bounded as $s\to 1^+$.*

Proof. (i): We will imitate the proof of Lemma 6. Again, let $\mathbb{N}_d$ denote the set of natural numbers whose prime factors are among the first d primes $p_1,\dots,p_d$. For a fixed prime p we have

$$\begin{aligned}\frac{1}{1-\chi(p)p^{-s}}&=1+\frac{\chi(p)}{p^s}+\left(\frac{\chi(p)}{p^s}\right)^2+\cdots\\&=1+\frac{\chi(p)}{p^s}+\frac{\chi(p^2)}{p^{2s}}+\cdots.\end{aligned}$$

It follows that

$$\begin{aligned}\prod_{i=1}^{d}\frac{1}{1-\chi(p_i)p_i^{-s}}&=\prod_{i=1}^{d}\left(1+\frac{\chi(p_i)}{p_i^s}+\frac{\chi(p_i^2)}{p_i^{2s}}+\cdots\right)\\&=\sum_{n\in\mathbb{N}_d}\frac{\chi(n)}{n^s}\end{aligned}$$

because

$$\frac{\chi(p_1)^{\alpha_1}}{p_1^{\alpha_1}}\cdots\frac{\chi(p_d)^{\alpha_d}}{p_d^{\alpha_d}}=\frac{\chi(p_1^{\alpha_1}\cdots p_d^{\alpha_d})}{p_1^{\alpha_1}\cdots p_d^{\alpha_d}}.$$

Now,

$$L(s,\chi)-\prod_{i=1}^{d}\frac{1}{1-\chi(p_i)p_i^{-s}}=\sum_{n\in\mathbb{N}-\mathbb{N}_d}\frac{\chi(n)}{n^s}.$$

Since any $n\in\mathbb{N}-\mathbb{N}_d$ is $\geq d$, we have

$$\left|\sum_{n\in\mathbb{N}-\mathbb{N}_d}\frac{\chi(n)}{n^s}\right|\leq\sum_{n=d+1}^{\infty}\frac{1}{n^{\operatorname{Re} s}}\longrightarrow 0\ \text{ as }\ d\to\infty,$$

proving (i).

(ii): Here the argument is similar to the proof of Proposition 7. From (i) we derive that

$$\ln L(s,\chi)=\sum_{p\in P}-\ln(1-\chi(p)p^{-s})$$

On the other hand,

$$-\ln(1-\chi(p)p^{-s})=\frac{\chi(p)}{p^s}+\frac{\chi(p)^2}{2p^{2s}}+\frac{\chi(p)^3}{3p^{3s}}+\cdots=\frac{\chi(p)}{p^s}+g_p(s,\chi)$$

where

$$g_p(s,\chi) := \frac{\chi(p)^2}{2p^{2s}} + \frac{\chi(p)^3}{3p^{3s}} + \cdots .$$

Then

$$|g_p(s,\chi)| \le \frac{1}{2p^{2s}} + \frac{1}{3p^{3s}} + \cdots \le \frac{1}{p^{2s}}$$

as we have seen in the proof of Proposition 7. Then for any d,

$$\sum_{i=1}^{d} |g_{p_i}(s,\chi)| \le \sum_{i=1}^{d} \frac{1}{p_i^{2s}} \le \sum_{n=1}^{\infty} \frac{1}{n^2} = \zeta(2).$$

This means that for any $s > 1$, the series $\sum_{p\in P} g_p(s,\chi)$ absolutely converges, and its sum $g(s,\chi)$ satisfies $|g(s,\chi)| \le \zeta(2)$, hence remains bounded as $s \to 1^+$. Since $\ln L(s,\chi) = \lambda(s,\chi) + g(s,\chi)$, (ii) is proven. □

It follows from Proposition 16(ii) that to complete the proof of Theorem 15 one needs to show that if $\chi \neq \chi_0$, $L(s,\chi)$ approaches some *nonzero* number as $s \to 1^+$. This part of the argument heavily relies on complex analysis. Let

$$\zeta_m(s) = \prod_{\chi\in\widehat{G_m}} L(s,\chi).$$

Proposition 17. *(i) $L(s,\chi_0)$ extends meromorphically to the domain $D = \{s \in \mathbb{C} \mid \operatorname{Re} s > 0\}$ with the only pole at $s = 1$, and this pole is simple.*

(ii) For $\chi \neq \chi_0$, $L(s,\chi)$ extends holomorphically to D.

(iii) $\zeta_m(s)$ extends meromorphically to D with a pole at $s = 1$.

Assume for now Proposition 17. Then for $\chi \neq \chi_0$,

$$\lim_{s\to 1^+} L(s,\chi) = L(1,\chi),$$

which is a finite number. Suppose $L(1,\chi) = 0$ for at least one character $\chi \neq \chi_0$. Then in the product $L(s,\chi_0)L(s,\chi)$ the zero of $L(s,\chi)$ would annihilate the pole of $L(s,\chi_0)$ at $s = 1$, implying that the product is actually holomorphic at $s = 1$. Since the L-functions for all other characters are also holomorphic at $s = 1$, we would get that $\zeta_m(s)$ is holomorphic at $s = 1$, which contradicts Proposition 17(iii).

Analyticity in parts (i) and (ii) is derived from the following general statement.

Lemma 18. *Let U be an open set of $\mathbb{C}$ and let $\{f_n\}$ be a sequence of holomorphic functions on U which converges uniformly on every compact subset of U to a function f. Then f is holomorphic in U.*

Proof. See [Se], pg. 64-65. □

Proof of Proposition 17(i). First, we will show $\zeta(s)$ extends to a meromorphic function on D with a simple pole at $s = 1$. For $s > 1$ we have

$$\frac{1}{s-1} = \int_1^{\infty} t^{-s}dt = \sum_{n=1}^{\infty} \int_n^{n+1} t^{-s}dt.$$

Hence we can write

$$\zeta(s) = \frac{1}{s-1} + \sum_{n=1}^{\infty} \left(\frac{1}{n^s} - \int_n^{n+1} t^{-s}dt \right) = \frac{1}{s-1} + \sum_{n=1}^{\infty} \int_n^{n+1} (n^{-s} - t^{-s})dt.$$

Set now

$$\phi_n(s) = \int_n^{n+1} (n^{-s} - t^{-s})dt \text{ and } \phi(s) = \sum_{n=1}^{\infty} \phi_n(s).$$

Our goal is to show that $\phi(s)$ is defined and analytic in D; then $1/(s-1) + \phi(s)$ will be the required meromorphic extension of $\zeta(s)$. Since each of the functions $\phi_n(s)$ is analytic in D, the analyticity of ϕ will follow from Lemma 18 if we can show that the series $\sum \phi_n(s)$ converges uniformly on every compact subset of D. But any compact subset of D is contained in

$$K_{\sigma,c} = \{s \in \mathbb{C} \mid \operatorname{Re} s \geq \sigma,\ |s| \leq c\}$$

for some $c, \sigma > 0$. Let $\psi_{n,s}(t) = n^{-s} - t^{-s}$. Then for any $t_0 \in [n, n+1]$ we have

$$\begin{aligned} |\psi_{n,s}(t_0)| &= |\psi_{n,s}(t_0) - \psi_{n,s}(n)| \\ &\leq \max_{t\in[n,n+1]} |\psi'_{n,s}(t)| \cdot |t_0 - n| \\ &\leq \max_{t\in[n,n+1]} \left|\frac{s}{t^{s+1}}\right| = \frac{|s|}{n^{\operatorname{Re} s+1}}. \end{aligned}$$

So, for $s \in K_{\sigma,c}$, we have

$$|\phi_n(s)| \leq \max_{t\in[n,n+1]} |\psi_{n,s}(t)| \leq \frac{|s|}{n^{\operatorname{Re} s+1}} \leq \frac{c}{n^{\sigma+1}}.$$

Since the series $\sum \frac{c}{n^{\sigma+1}}$ converges, the series $\sum \phi_n(s)$ uniformly converges on $K_{\sigma,c}$ by the Weierstrass M-test (cf. [Ru], Theorem 7.10).

Now, it remains to relate $\zeta(s)$ and $L(s, \chi_0)$. Suppose $m = q_1^{\alpha_1} \cdots q_r^{\alpha_r}$. Let $\mathbb{N}'$ be the set of all natural numbers of the form $q_1^{\beta_1} \cdots q_r^{\beta_r}$, and let $\mathbb{N}''$ be the set of all natural numbers that are relatively prime to m. Then any $n \in \mathbb{N}$ can be uniquely written in the form $n = n'n''$ with $n' \in \mathbb{N}'$, $n'' \in \mathbb{N}''$. It follows that

$$\zeta(s) = \sum_{n\in\mathbb{N}} \frac{1}{n^s} = \left(\sum_{n\in\mathbb{N}'} \frac{1}{n^s}\right)\left(\sum_{n\in\mathbb{N}''} \frac{1}{n^s}\right)$$

But

$$\sum_{n\in\mathbb{N}''} \frac{1}{n^s} = L(s, \chi_0)$$

and

$$\sum_{n\in\mathbb{N}'} \frac{1}{n^s} = \left(1 + \frac{1}{q_1^s} + \frac{1}{q_1^{2s}} + \cdots\right) \cdots \left(1 + \frac{1}{q_r^s} + \frac{1}{q_r^{2s}} + \cdots\right) = \prod_{i=1}^{r} \frac{1}{1 - q_i^{-s}}.$$

So, $L(s, \chi_0) = \zeta(s)F(s)$, where $F(s) = \prod_{i=1}^{r}(1 - q_i^{-s})$. Since $F(s)$ is holomorphic and has no zeroes in D, we obtain our claim. □

Proof of Proposition 17(ii). We will prove analyticity of $L(s, \chi)$ in D for $\chi \neq \chi_0$ by showing that the series

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

converges uniformly on compact subsets of D. The proof imitates the proof of Abel's and Dirichlet's test for convergence of series of the form $\sum a_n b_n$ (cf. [Ru], Theorem 3.41). Let $a_n = \chi(n)$, $b_n = n^{-s}$. To apply the Cauchy criterion, we need to show that $|\sum_{n=M}^{N} a_n b_n|$ becomes arbitrarily small uniformly on $K_{\sigma,c}$ for $M < N$ if M is large enough. Let $A_n = \sum_{k=1}^{n} a_n$. The crucial thing is that the assumption

$\chi \neq \chi_0$ implies that $|A_n| \leq C$ for some constant C independent of n (which, of course, is false for $\chi = \chi_0$!). Indeed, for any $a \in \mathbb{Z}$ we have $\chi(a) = \chi(a+m)$, and besides it follows from Proposition 12(i) that

$$\sum_{n=1}^{m} \chi(n) = 0.$$

Thus, if $n = dm + r$ where $0 \leq r < m$ then

$$A_n = \sum_{k=1}^{n} \chi(k) = \sum_{k=dm+1}^{dm+r} \chi(k) = \sum_{k=1}^{r} \chi(k) = A_r$$

where by convention $A_0 = 0$. So, $C = \max\{|A_1|, \ldots, |A_{m-1}|\}$ will work.

Substituting $a_n = A_n - A_{n-1}$, we get

$$\sum_{n=M}^{N} a_n b_n = \sum_{n=M}^{N-1} A_n(b_n - b_{n+1}) + A_N b_N - A_{M-1} b_M. \tag{2.9}$$

We have seen in the proof of part (i) that

$$|n^{-s} - (n+1)^{-s}| \leq \frac{|s|}{n^{\operatorname{Re} s+1}}.$$

So it follows from (2.9) that

$$\left| \sum_{n=M}^{N} a_n b_n \right| \leq \sum_{n=M}^{N-1} \frac{C|s|}{n^{\operatorname{Re} s+1}} + \frac{C}{M^{\operatorname{Re} s}} + \frac{C}{N^{\operatorname{Re} s}}.$$

Thus, if $s \in K_{\sigma,c}$ then

$$\left| \sum_{n=M}^{N} a_n b_n \right| \leq Cc \sum_{n=M}^{N-1} \frac{1}{n^{\sigma+1}} + \frac{2C}{M^{\sigma}}.$$

Since the series $\sum \frac{1}{n^{\sigma+1}}$ converges, we see that $|\sum_{n=M}^{N} a_n b_n|$ becomes arbitrarily small uniformly on $K_{\sigma,c}$ if M is large enough, completing the proof. □

Proof of Proposition 17(iii). We only need to show that $\zeta_m(s)$ cannot be holomorphic at $s = 1$.

Lemma 19. *For an integer a prime to m, let $f(a)$ denote the order of $\bar{a}$ in G_m, and let $g(a) = \varphi(m)/f(a)$. If T is a variable then*

$$\prod_{\chi \in \widehat{G_m}} (1 - \chi(a)T) = (1 - T^{f(a)})^{g(a)}.$$

Proof. Let H be the cyclic subgroup of G_m generated by $\bar{a}$; $|H| = f(a)$. Then the set $\{\chi(\bar{a}) \mid \chi \in \widehat{H}\}$ is precisely the set of all $f(a)$-th roots of unity. It follows that

$$\prod_{\chi \in \widehat{H}} (X - \chi(a)) = X^{f(a)} - 1.$$

Substituting $X = T^{-1}$ and multiplying by $T^{f(a)}$, we get

$$\prod_{\chi \in \widehat{H}} (1 - \chi(a)T) = 1 - T^{f(a)}.$$

Now, the homomorphism of restriction $\widehat{G_m} \to \widehat{H}$ is surjective (Corollary 11) and its kernel has order $g(a)$. It follows that

$$\prod_{\chi \in \widehat{G_m}} (1 - \chi(a)T) = \left(\prod_{\chi \in \widehat{H}} (1 - \chi(a)T) \right)^{g(a)} = (1 - T^{f(a)})^{g(a)}.$$

□

Using Lemma 19, we can transform the expression for $\zeta_m(s)$:

$$\zeta_m(s) = \prod_{\chi \in \widehat{G_m}} L(s, \chi) = \prod_{p \in P} \left(\prod_{\chi \in \widehat{G_m}} \frac{1}{1 - \chi(p)p^{-s}} \right) = \prod_{(p,m)=1} \frac{1}{(1 - p^{-f(p)s})^{g(p)}}. \tag{2.10}$$

Since

$$\frac{1}{(1 - p^{-f(p)s})} = 1 + \frac{1}{p^{f(p)s}} + \frac{1}{p^{2f(p)s}} + \cdots,$$

it follows from (2.10) that $\zeta_m(s)$ can be written in the form

$$\zeta_m(s) = \sum_{n=1}^{\infty} \frac{c_n}{n^s} \tag{2.11}$$

(a **Dirichlet series**) with $c_n \geq 0$, and the series converges for $|s| > 1$. Assume now that $\zeta_m(s)$ is holomorphic at $s = 1$. Then $\zeta_m(s)$ is holomorphic everywhere in D. By applying [Se], Prop. 7, Ch. VI, we conclude that the series in (2.11) converges everywhere in D. To see that this is false, we observe that

$$\begin{aligned} \frac{1}{(1 - p^{-f(p)s})^{g(p)}} &= (1 + p^{-f(p)s} + p^{-2f(p)s} + \cdots)^{g(p)} \\ &\geq 1 + p^{-\varphi(m)s} + p^{-2\varphi(m)s} + \cdots = \frac{1}{1 - p^{-\varphi(m)s}}. \end{aligned}$$

So,

$$\sum_{n=1}^{\infty} \frac{c_n}{n^s} \geq \prod_{(p,m)=1} \frac{1}{1 - p^{-\varphi(m)s}} = \sum_{(n,m)=1} n^{-\varphi(m)s} = L(\varphi(m)s, \chi_0).$$

But we already know that $L(\varphi(m)s, \chi_0)$ diverges for $s = \varphi(m)^{-1}$, so $\sum \frac{c_n}{n^s}$ cannot converge for the same value of s. A contradiction. □

This concludes the proof of Dirichlet's theorem.

References

[Ar] Michael Artin: *Algebra.* Englewood Cliffs, N.J.: Prentice Hall, 1991.

[Co] David A. Cox: *Galois Theory.* Hoboken, N.J.: Wiley-Interscience, 2004.

[IR] Kenneth Ireland and Michael Rosen: *A Classical Introduction to Modern Number Theory.* New York: Springer, 1990 (Graduate Texts in Math. **84**).

[Kn] Anthony W. Knapp: *Elliptic Curves.* Princeton, N.J.: Princeton Univ. Press, 1992 (Mathematical Notes **40**).

[La] Serge Lang: *Algebra.* New York: Springer, 2002 (Graduate Texts in Math. **211**).

[Ru] Walter Rudin: *Principles of Mathematical Analysis.* New York: McGraw-Hill Book Co., 1976.

[Se] Jean-Pierre Serre: *A Course in Arithmetic.* New York: Springer, 1973 (Graduate Texts in Math. **7**).

3 Quivers

Virginia Fisher '08[†]
Harvard University
Cambridge, MA

Eloy Lopez[‡]
California State Northridge
Northridge, CA

Tiago Macedo and Lonardo Rabelo[*]
IMECC - UNICAMP
Caixa Postal
Campinas-SP, Brazil

Abstract

This project is based on the study of two kinds of representation theory: quiver representation theory and Lie algebra representation theory. By looking at some simple examples, we'll show how the two are connected. Indeed, we'll identify the isomophism classes of simple and indecomposable representations of a particular quiver with relation with the equivalence classes of simple and indecomposable representations of $\mathfrak{sl}_2(\mathbf{k})$. Throughout this paper, $\mathbf{k}$ will indicate an algebraically closed field of characteristic 0.[§]

3.1 Quivers

3.1.1 Definitions

A **quiver** is directed graph $Q = (Q_0, Q_1)$ where Q_0 is the set of vertices (which is assumed to be finite) and Q_1 the set of arrows, with maps $h, t : Q_1 \to Q_0$ which assign to each arrow its head and tail, respectively. Every vertex $i \in Q_0$ has an associated edge e_i such that $h(e_i) = t(e_i) = i$.

A **path** is a sequence of arrows $p = a_1 a_2 \cdots a_n$ such that $h(a_{k+1}) = t(a_k)$ for $k = 1, \ldots, n-1$. The **head** of the path is $h(a_1)$, and the **tail** of the path is $t(a_n)$. Each e_i (defined above) is a trivial path which starts and ends at the vertex i.

An **oriented cycle** is a path p such that $h(p) = t(p)$, and $h(a_i) \neq t(a_j)$ for any other $i \neq j + 1$.

It is easy to see the following property:

Proposition 1. *A quiver with an oriented cycle has an infinite set of paths.*

[†] Virginia A. Fisher, Harvard '08, hails from New York's capital district and resides at the Dudley Coop. Besides mathematics, she studies philosohy and Persian language.

[‡] Eloy Lopez is a first year masters student in math at California State University at Northridge.

[*] Tiago Macedo and Lonardo Rabelo are both juniors at Unicamp in Campinas, Brazil.

[§] This work was developed during the International Research Experiences for Students in Mathematics (IRES) hosted by the Universidade Estadual de Campinas (UNICAMP), SP, Brazil, in July of 2006, and was funded by the National Science Foundation, CNPq (grant 451.154/2006-1) and FAEPEX-UNICAMP (grant 163/2006).

Given a quiver Q, the **path algebra** $\mathbf{k}Q$ is the $\mathbf{k}$-vector space generated by all paths in Q with multiplication rule:

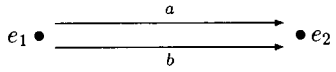
$$p * q = \begin{cases} pq & \text{if } h(q) = t(p) \\ 0 & \text{otherwise.} \end{cases}$$

Given a point $i \in Q_0$, we have that e_i is the null path beginning and ending at that point, so $a * e_i = a$ whenever $t(a) = i$ and $e_i * b = b$ whenever $h(b) = i$. Note that the path algebra has a unit given by $\sum_{i \in Q_0} e_i$.

Example 3.1.1. The **Jordan quiver** J has one vertex $J_0 = \{1\}$ and one nontrivial arrow $J_1 = \{e_1, a\}$, such that $t(a) = h(a) = 1$.

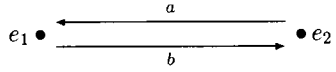
Its path algebra has basis $\{e_1, a, a^2, \ldots\}$ and is thus infinite-dimensional.

Example 3.1.2. The **2-Kronecker** quiver K_2



has finite-dimensional path algebra with basis $\{e_1, e_2, a, b\}$.

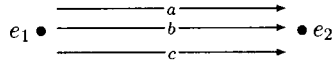
However, the **cyclic 2-Kronecker** quiver C_2 presented by



has an oriented cycle, and its path algebra is infinite-dimensional with basis

$$\{e_1, e_2, a, b, ba, ab, aba, bab, \ldots, (ba)^k, (ab)^k, a(ba)^k, b(ab)^k, \ldots\}.$$

Example 3.1.3. The **3-Kronecker** quiver K_3 presented by



has finite-dimensional path algebra with basis $\{e_1, e_2, a, b, c\}$.

3.1.2 Quivers with relations

We can impose further relations on the composition of arrows. This is equivalent to quotienting the path algebra by the apropriate ideal.

Example 3.1.4. For the Jordan quiver defined above, we can impose the relation $a^k = e_1$ for some $k \in \mathbb{N}$. The resulting path algebra has basis $\{e_1, a, a^2, \ldots, a^{k-1}\}$.

Example 3.1.5. For the quiver C_2 defined above, we can impose the relation $ab = e_2$, to obtain a quiver with path algebra basis given by $\{e_1, e_2, a, b, ba\}$.

3.2 Quiver Representations

3.2.1 Definitions

Given a quiver Q, a **quiver representation** of Q is a collection $\{V_i | i \in Q_0\}$ of finite dimensional $\mathbf{k}$-vector spaces together with a collection $\{\phi_a : V_{t(a)} \to V_{h(a)} | a \in Q_1\}$ of $\mathbf{k}$-linear maps such that $\phi_a \phi_b = \phi_{ab}$.

From now on, we will denote a representation by $\mathcal{R} = (\{V_i\}, \{\phi_a\})$.

Example 3.2.1. The representations $(\{V\}, \{\phi\})$ of the Jordan quiver are given by all $n \times n$ matrices A_ϕ, where $n = \dim V$.

Suppose $\mathcal{R} = (\{V_i\}, \{\phi_a\})$ and $\mathcal{R}' = (\{W_i\}, \{\psi_a\})$ are representations of Q. Then $\mathcal{R}'$ is a **subrepresentation** of $\mathcal{R}$ if

- for every $i \in Q_0$, W_i is a subspace of V_i and
- for every $a \in Q_1$, the restriction of $\phi_a : V_{t(a)} \to V_{h(a)}$ to $W_{t(a)}$ is equal to $\psi_a : W_{t(a)} \to W_{h(a)}$.

The **zero representation** of Q is given by $(\{V_i\}, \{\phi_a\})$ such that $V_i = 0$ for all $i \in Q_0$ and ϕ_a is the zero map for all $a \in Q_1$. A non-zero representation $\mathcal{R}$ is called **simple representation** if the only subrepresentations of $\mathcal{R}$ are the zero representation and $\mathcal{R}$ itself.

If $\mathcal{R} = (\{V_i\}, \{\phi_a\})$ and $\mathcal{S} = (\{W_i\}, \{\psi_a\})$ are representations of Q then we can define the **direct sum representation** $\mathcal{R} \oplus \mathcal{S} = (\{U_i\}, \{\rho_a\})$ by taking:

- $U_i = V_i \oplus W_i$ for every $i \in Q_0$, and
- $\rho_a : V_{t(a)} \oplus W_{t(a)} \to V_{h(a)} \oplus W_{h(a)}$, given by the matrix $\begin{pmatrix} \phi_a & 0 \\ 0 & \psi_a \end{pmatrix}$.

If $\mathcal{R}$ and $\mathcal{S}$ are two representations of Q, then a **representation morphism** $\Phi : \mathcal{R} \to \mathcal{S}$ is a collection of $\mathbf{k}$-linear maps $\{\varphi_i : V_i \to W_i | i \in Q_0\}$ such that the diagram

$$\begin{array}{ccc} V_{t(a)} & \xrightarrow{\phi_a} & V_{h(a)} \\ \downarrow{\scriptstyle \varphi_{t(a)}} & & \downarrow{\scriptstyle \varphi_{h(a)}} \\ W_{t(a)} & \xrightarrow[\psi_a]{} & W_{h(a)} \end{array}$$

commutes for all $a \in Q_1$. If φ_i is invertible for every $i \in Q_0$, then the morphism Φ is called an **isomorphism** and $\mathcal{R}$ and $\mathcal{S}$ are called **isomorphic representations**.

A representation $\mathcal{R}$ of a quiver Q is called **decomposable** if $\mathcal{R} \simeq \mathcal{S} \oplus \mathcal{T}$ where $\mathcal{S}$ and $\mathcal{T}$ are nonzero subrepresentations of Q. A nonzero representation is called **indecomposable** if it cannot be written as such a direct sum. For any quiver Q, the simple representations of Q form a subclass of its indecomposable representations.

3.2.2 Isomorphism Classes of Representations

The study of quiver representations is significantly simplified if we consider isomorphism classes of quiver representations rather than representations themselves. To find a representative element of each isomorphism class, we apply representation isomorphisms to change the basis of the vector space at each vertex in order to simplify the matrices for the maps at each arrow. For representations over $\mathbb{C}$ with equidimensional vector spaces at each vertex, this process is the same as that of the Jordan normal form.

Example 3.2.2. For the Jordan quiver with one vertex and one arrow, every isomorphism class of representations has a representative element of the form $\mathcal{R} = (\{V_1\}, \{J_1\})$ where J_1 is a matrix in Jordan normal form and V_1 is a vector space with the associated basis.

This is a direct consequence of the theorem that every square matrix $M = P^{-1}JP$, where J is in Jordan form and P is an invertible matrix corresponding to the change of basis required to isolate the eigenspaces of the operator; see [Ha].

If we restrict ourselves to representations with invertible maps at each arrow, we may simultaneously describe the isomorphism classes of representations of quivers which differ from each other only in the orientation of their arrows. Note that the invertibility condition implies that the representation must have equidimensional vector spaces at all vertices. The isomorphism classes of these representations can often be described neatly, by analogy to the case of the Jordan quiver.

Example 3.2.3. Given any representation $\mathcal{R}$ of the cyclic 2-Kronecker quiver C_2 of the form

$$\mathcal{R} = (\{V_1, V_2\}, \{A, B\}),$$

where A and B are both invertible, we can find an isomorphic representation of the form:

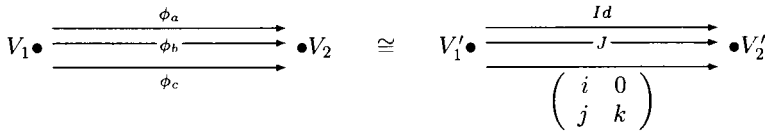
with J in Jordan form. To find this isomorphic representation, let $\mathcal{B}_1$ and $\mathcal{B}_2$ be bases for V_1 and V_2, respectively. Take P_0 to be the change-of-basis matrix taking $\mathcal{B}_2$ to $A\mathcal{B}_1$. This is possible because invertibility implies equidimensionality. Then the representation isomorphism $\Phi_0 = (Id, P_0)$ yields the isomorphic representation

$$\mathcal{R}' = (\{V_1', V_2'\}, \{Id, BA\}).$$

We can find an invertible matrix P_1 such that $BA = P_1^{-1} J P_1$, where J is a Jordan-form matrix. Applying the representation isomorphism $\Phi_1 = (P_1, P_1)$ yields the desired isomorphic representation.

Since we are considering only representations $(\{V_i\}, \{\phi_a\})$ with ϕ_a invertible for all $a \in Q_1$, quivers that differ only in the direction of their arrows (such as K_2 and C_2) have the same sets of representations. However, these quivers still have different representation theories (for example, different classes of simple representations), because the definition of a subrepresentation depends on the direction of arrows.

Example 3.2.4. The case of the 3-Kronecker quiver K_3 is more complicated than that of the 2-Kronecker, because we may not be able to simultaneously put the maps on the second and third arrows in Jordan normal form. However, in the case $\dim(V_1) = \dim(V_2) = 2$, we will always be able to conjugate bases and obtain an isomorphic representation of the form:



3.2.3 Simple Representations

Definition 2. An **i-th canonical representation** $\mathcal{R}_i$ for the quiver $Q = (Q_0, Q_1)$ is a representation of the form

$$\mathcal{R} = (\{V_j = \delta_{ij}\mathbf{k}\}, \{\phi_a = 0 \text{ for all } a \in Q_1\}).$$

where $\delta_{ij} = 1$ when $i = j$, $\delta_{ij} = 0$ when $i \neq j$.

Proposition 3. *Let Q be a quiver with no oriented cycles. A representation $\mathcal{R}$ of Q is simple if and only if it is canonical.*

Proof. ($\Leftarrow$) A canonical representation $\mathcal{R}$ must be simple, because its only proper subrepresentation is the zero representation.

($\Rightarrow$) We will show that every non-canonical representation has a canonical subrepresentation.

Lemma 4. *If $Q = (Q_0, Q_1)$ is a quiver with no oriented cycles, then there is some vertex $i \in Q_0$ such that $i \neq t(a)$ for all arrows $a \in Q_1$. Such an arrow is called a* ***sink.***

Proof. Suppose for every $v_i \in Q_0$, $v_i = t(a)$ for some $a \in Q_1$. Choose some $v_1 \in Q_0$ and form a path as follows: Choose a_n such that $t(a_n) = v_n$. Write $v_{n+1} = h(a_n)$, and repeat. As Q_0 is a finite set, eventually we will get $v_{n+1} = v_i$ for some $i \leq n$. Then $p = a_i \cdots a_n$ is an oriented cycle in Q. But by assumption, Q has no oriented cycles, so some vertex in Q must be a sink. □

So let Q be a quiver with no oriented cycle, let $x_1 \in Q_0$ be a vertex such that $t(a) \neq x_1$ for all $a \in Q_1$. Given an arbitrary representation $R = (V_i, \rho_a)$, if $V_{x_1} \neq \{0\}$, then write $x_n = x_1$ and proceed to the construction of $\mathcal{S}$ below.

If $V_{x_1} = \{0\}$, define $Q' = (Q'_0 = Q_0 \backslash \{x_1\}, Q'_1 = Q_1 \backslash \{a \in Q_1 | h(a) \neq x_1\})$. As Q contained no oriented cycles, and $Q'_0 \subset Q_0$, $Q'_1 \subset Q_1$, Q' contains no oriented cycle, so we may apply the lemma.

So we may let $x_2 \in Q'_0$ be a vertex such that $t(a) \neq x_2$ for all $a \in Q'_1$. Define the representation $\mathcal{R}'$ of Q' by restricting the representation $\mathcal{R}$, and repeat the process described above.

If $\mathcal{R}$ is a non-trivial representation of Q, we will eventually find $x_n \in Q_0$ such that $V_{x_n} \neq \{0\}$ but $V_{h(a)} = \{0\}$ for all $a \in Q_1$ such that $t(a) = x_n$.

Construct a representation $\mathcal{S}$ of Q by taking

$$\mathcal{S} = (\{W_i = \delta_{ni}\mathbf{k}\}, \{\phi_a = 0 \text{ for all } a \in Q_1\}).$$

Then $\mathcal{S}$ is a proper canonical subrepresentation of $\mathcal{R}$. To see this, observe that $W_i \subseteq V_i$ for all $i \in Q_0$ and define the inclusion morphism from $\mathcal{S}$ into $\mathcal{R}$ by $P = \{P_i : W_i \hookrightarrow V_i | i \in Q_0\}$.

To check that all maps commute, first note that for $a \in Q_1$ such that $t(a) \neq x$, $W_{t(a)} = \{0\}$. So $\psi_a : W_{t(a)} \to W_{h(a)}$ and $P_{t(a)} : W_{t(a)} \to V_{t(a)}$ must both be the zero map. Hence, for all $a \in Q_1$ such that $t(a) \neq x$ we have: $P_{h(a)} \circ \psi_a = \varphi_a \circ P_{t(a)} = 0$ so the morphism commutes.

Now, for all a such that $t(a) = x$, we know that $V_{h(a)} = W_{h(a)} = \{0\}$. So $\varphi_a : V_{t(a)} \to V_{h(a)}$, $\psi_a : W_{t(a)} \to W_{h(a)}$ and $P_{h(a)} : W_{h(a)} \to V_{h(a)}$ must all be the zero map. So for all $a \in Q_1$ such that $t(a) = x$, we have $P_{h(a)} \circ \psi_a = \varphi_a \circ P_{t(a)} = 0$ and the morphism commutes.

Therefore, $\mathcal{S}$ is a subrepresentation of $\mathcal{R}$ of the desired form. □

3.2.4 Indecomposable Representations

Here we will work with the examples we have given above. The invertibility of maps and the dimension vectors will play an important role in giving all the indecomposable representations for some given quiver.

Example 3.2.5. A representation $\mathcal{R}$ of the Jordan quiver J is indecomposable if it is isomorphic to a representation with matrix for ϕ_a in Jordan form with a single eigenblock, as such a matrix cannot be rewritten as a direct sum of two smaller matrices.

Example 3.2.6. For the oriented 2-Kronecker quiver C_2 in Example 3.1.2, we have the following classification:

Proposition 5. *A representation $\mathcal{R} = (\{V_1 = \mathbb{C}^m, V_2 = \mathbb{C}^n\}, \{\phi_a, \phi_b\})$ of C_2 is indecomposable if and only if one of the following holds:*

- *$\mathcal{R} \cong \mathcal{R}' = (\{V_1, V_2\}, \{Id, J_\lambda\})$ where J_λ is a matrix in Jordan normal form with only one eigenblock.*
- *$(\phi_b \circ \phi_a)^k = 0$ for some $k \in \mathbb{Z}^+$ and $\dim \ker \phi_b \circ \phi_a = 1$.*

Proof. Without loss of generality, $m \geq n$.

We describe the possible cases, and prove decomposability or indecomposability for each case.

1. If the composite map $\phi_b \circ \phi_a : V_1 \to V_1$ is invertible, then we must have $m = n$, and ϕ_a, ϕ_b both invertible. Thus, by changing bases, we can find an isomorphic representation with $\phi'_a = I$, ϕ'_b represented by a matrix in Jordan form.

   Then, as shown above, $\mathcal{R}$ is indecomposable if and only if the matrix for ϕ'_b has only one Jordan block.

2. If the composite map $\phi_b \circ \phi_a : V_1 \to V_1$ is not invertible, we have two cases:

   (a) $\phi_b \circ \phi_a$ is nilpotent, i.e. $(\phi_b \circ \phi_a)^k = 0$ for some $k \in \mathbb{Z}^+$

      i. Suppose $\dim \ker \phi_b \circ \phi_a = 1$. Then take $x \in \ker \phi_b \circ \phi_a$. Then any y in the kernel of $\phi_b \circ \phi_a$ must be a scalar multiple of x. Suppose $\mathcal{R}$ is not indecomposable, so that $\mathcal{R} = \mathcal{R}' \oplus \mathcal{R}''$ where

      $$\mathcal{R}' = (\{W_1, W_2\}, \{A|_{W_1}, B|_{W_2}\}) \ , \ \mathcal{R}'' = (\{U_1, U_2\}, \{A|_{U_1}, B|_{U_2}\})$$

      are both non-trivial. Without loss of generality, $x \in W_1$.

      First suppose $y \in U_1, y \neq 0$. Then by definition of decomposability, $(\phi_b \circ \phi_a)^i \in U_1$ for all $i \in \mathbb{Z}_+$. But $(\phi_b \circ \phi_a)^k = 0$, so pick the least $j \in \mathbb{Z}^+$ such that $(\phi_b \circ \phi_a)^j = 0$. Then $(\phi_b \circ \phi_a)^{j-1} \in \ker \phi_b \circ \phi_a$ so for the y chosen above, $(\phi_b \circ \phi_a)^{j-1} y = \lambda x \in W_1$. But by assumption, $y \in U_1$. Thus, $U_1 = \{0\}$ so $V_1 = W_1$.

      Now suppose $y \in U_2$, $y \neq 0$. Then, by a dimension counting argument, either $y = \phi_a x$ for some $x \in V_1$ or $\phi_b y = x$ for some nonzero $x \in V_1$. In either case, $y \in W_2$ by the invariance of subrepresentations. But by assumption, $y \in U_2$, so $U_2 = \{0\}$ so $V_2 = W_2$. Therefore, $\mathcal{R}''$ in the direct sum is the trivial subrepresentation, so $\mathcal{R}$ is indecomposable.

      ii. Suppose $\dim \ker \phi_b \circ \phi_a > 1$. Write $\ker \phi_b \circ \phi_a = W_0 \oplus U_0$, both of which are non-zero. For $x \in V_1$, write $j_x \in \mathbb{Z}^+$ is the minimal integer such that $(BA)^{j_x} = 0$, and define:

      $$W_1 = \{x \in V_1 | (\phi_b \circ \phi_a) j_x - 1x \in W_0\} \ , \ U_1 = \{x \in V_1 | (\phi_b \circ \phi_a)^{j_x - 1} \in U_0)\}$$

      These two sets define a decomposition of $\mathcal{R}$, so $\mathcal{R}$ is decomposable.

   (b) $\phi_b \circ \phi_a$ is not nilpotent; i.e. $(\phi_b \circ \phi_a)^k \neq 0$ for all $k \in \mathbb{Z}^+$. Then there is some integer j such that $V_1 = \ker(\phi_b \circ \phi_a)^j \oplus W_1$ and $(\phi_b \circ \phi_a)^j|_{W_1}$ is invertible. These sets define a decomposition of $\mathcal{R}$, so $\mathcal{R}$ is decomposable. □

Corollary 6. *The vector spaces V_1, V_2 of an indecomposable representation of C_2 can only have dimensions* $\dim V_1 = m = n = \dim V_2$ *or* $\dim V_1 = m$, *where* $m \pm 1 = n = \dim V_2$.

Proof. In the first case, V_1 and V_2 must be equidimensional. In the second case, $\dim \ker \phi_b \circ \phi_a = 1$ implies $|\dim V_1 - \dim V_2| \leq 1$. □

3.3 Lie Algebras and Their Representations

Definition 7. A **Lie Algebra** $\mathfrak{g}$ is a (non-associative) algebra with the multiplication rule given by a bilinear map $[\ ,\ ]$ which satisfies

- $[x, x] = 0$ for all $x \in \mathfrak{g}$,
- $[x, [y, x]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in \mathfrak{g}$.

These two properties imply that the $[\ ,\ ]$ operation is anti-symmetric, i.e. $[x,y] = -[y,x]$ for all $x, y \in \mathfrak{g}$.

We can construct a Lie algebra from any associative algebra by defining the bracket operation as the commutator $[a,b] = ab - ba$.

3.3.1 Representations of $\mathfrak{sl}_2(\mathbf{k})$

The simple linear algebra $\mathfrak{sl}_2(\mathbf{k}) = \{A \in M_2(\mathbf{k}) | \operatorname{tr} A = 0)\}$ of traceless 2×2 matrices is a Lie Algebra with bracket operation defined by the commutator $[A,B] = AB - BA$ and basis:

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \ h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We will describe the isomorphism classes of certain subclasses of the simple and indecomposable representations of $\mathfrak{sl}_2(\mathbf{k})$ and show that these correspond to simple and indecomposable representations of the 2-Kroneker quiver with an oriented cycle under the relation $ab = 0$, as described in Example 3.1.5.

We will restrict ourselves to the category $\mathcal{O}(\mathfrak{sl}_2)$ of representations of $\mathfrak{sl}_2(\mathbf{k})$ such that

- $V = \bigoplus_{k \in Z} V_k$, where $V_k = \{v \in V | hv = kv\}$ is the eigenspace with eigenvalue k for the action of h on V,
- $V_k = 0$ for $k \geq 0$,
- Each V_k is finite dimensional.

Given $v \in V_k$, using the bracket properties, calculation gives:

- $h(v) = kv$,
- $h(f(v)) = (k-2)f(v)$,
- $h(e(v)) = (k+2)e(v)$.

In other words, the action of f takes the eigenspace V_k with eigenvalue k to V_{k-2} with eigenvalue $k-2$, and the action of e takes V_k to V_{k+2}.

By the given properties, we have that each representation has a maximal eigenvalue $m \in \mathbb{Z}$ and:

$$V_k = \begin{cases} \mathbf{k} & \text{if } k = m - 2i \text{ for } i \in \mathbb{Z}_+, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, if we take $v_0 \in V_m$, the set $\mathcal{B}=\{v_i | i \in \mathbb{Z}^+\}$, where $v_i = f^i(v_0)$, defines a basis for

$$V(m) = \bigoplus_{k \leq m} V_k.$$

From the equations above, we can calculate that $e(v_i) = i(m-i+1)v_{i+1}$.

Now, we will describe a chain of examples of such representations; for this we assume that $\mathbf{k} = \mathbb{C}$.

Example 3.3.1. (The Verma Module)

Let $M(m)$ be the $\mathfrak{sl}_2$-module with underlying vector space

$$M(m) = \bigoplus_{i \geq 0} \mathbf{k} v_i$$

and the action given by

$$h(v_0) = mv_0, \qquad v_i = f^i(v_0), \qquad e(v_i) = i(m-i+1)v_{i+1}.$$

It is easy to check that this is in fact a representation and we have a diagramatic picture as in Figure 3.1.

We have defined m as the greatest eigenvalue of $M(m)$. If m is negative, then the map e does not annihilate any of the other eigenspaces, and we have an infinite-dimensional simple representation.

If the greatest eigenvalue, m, is nonnegative, the action of e will annihilate the eigenspace $M(m)_{-m-2}$ since $e(v_{m+1}) = 0$. In this case, the representation will not be simple; in fact, the $\mathbf{k}$-subspace

$$\bigoplus_{j \geq m+1} \mathbf{k}v_j$$

is a subrepresentation isomorphic to $M(-m-2)$. It will, however, be indecomposable, because the subspace

$$\bigoplus_{-m \leq i \leq m} M(m)_i$$

is not invariant under the action of f.

Now, for $m \geq 0$ $m \in \mathbb{Z}$, taking the quotient representation

$$V(m) = M(m)/M(-m-2)$$

gives a second example of simple representation, the only one with nonnegative integer maximal eigenvalue, $V(m)$. Its structure is shown in Figure 3.2(a).

These representations can be related by the following non-split short exact sequence:

$$0 \to V(-m-2) \to M(m) \to V(m) \to 0.$$

Example 3.3.2. ($P(-m-2)$ and $M^*(m)$)

Let $M(m)$ be the Verma module as defined above, and define another linearly independent eigenvector w_0 with eigenvalue $-m-2$ such that $e(w_0) = v_m$. From w_0, we can derive another set $\{w_i\}_{i \in \mathbb{N}}$ of eigenvectors by the rule $w_i = f^i(w_0)$ for each eigenvalue $\lambda = -m-2(i+1)$. Take the direct sum of the Verma Module with the eigenspaces spanned by these w_i's together with the action of e given by $e(w_i) = i(-m-i-1)w_{i-1} + v_{m+i}$.

Now we can consider the $\mathbf{k}$-vector space

$$M(m) \oplus \bigoplus_{j \geq 0} \mathbf{k}w_j$$

with action given by $h(w_0) = (-m-2)w_0$, $w_i = f^i(w_0)$ and $e(w_i) = i(-m-i-1)w_{i-1} + v_{m+i}$. It is easy to verify that these formulas turn the vector space $\bigoplus_{j \geq 0} \mathbf{k}v_j \oplus \bigoplus_{i \geq 0} \mathbf{k}w_i$ into a module belonging to category $\mathcal{O}(\mathfrak{sl}_2)$. We denote this module by $P(-m-2)$ and it has the diagram shown in Figure 3.2(b).

Notice that the $\mathbf{k}$-subspace $\bigoplus_{j \geq m+1} \mathbf{k}v_j$ is a subrepresentation isomorphic to $M(-m-2)$. Then, taking the quotient of these two representations, we define $M^*(m) = P(-m-2)/M(-m-2)$. This gives another non-split short exact sequence:

$$0 \to M(-m-2) \to P(-m-2) \to M^*(m) \to 0.$$

Furthermore, $M^*(m)$ has the diagram shown in Figure 3.2(c).

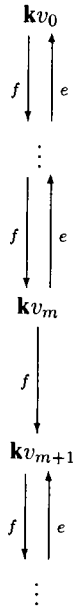


Figure 3.1: The Verma module $M(m)$.

Also, we can see that $M^*(m)$ has $V(m)$ as a subrepresentation which gives directly the next non-split short exact sequence:

$$0 \to V(m) \to M^*(m) \to M(-m-2) \to 0.$$

Finally, as $P(-m-2)$ has $M(m)$ as a subrepresentation, we get another short exact sequence:

$$0 \to M(m) \to P(-m-2) \to M(-m-2) \to 0.$$

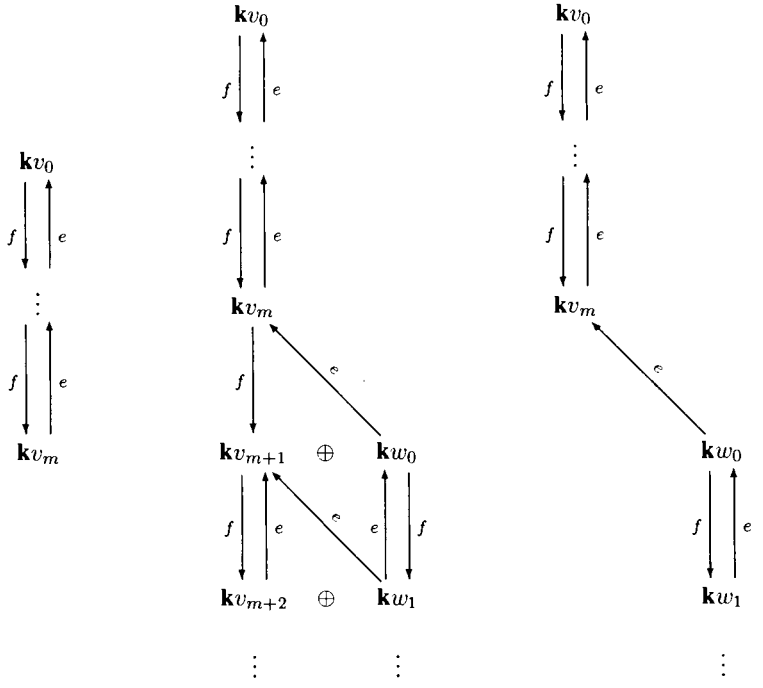
The following proposition tell us that the above examples are actually all of the examples of indecomposable modules in $\mathcal{O}(\mathfrak{sl}_2)$.

Proposition 8. *The following short exact sequences*

$$\begin{aligned} 0 &\to M(-m-2) \to M(m) \to V(m) \to 0 \\ 0 &\to V(m) \to M^*(m) \to M(-m-2) \to 0 \\ 0 &\to M(-m-2) \to P(-m-2) \to M^*(m) \to 0 \\ 0 &\to M(m) \to P(-m-2) \to M(-m-2) \to 0 \end{aligned}$$

are a complete set of equivalence class representatives of non-split short exact sequences of representations in the category $\mathcal{O}(\mathfrak{sl}_2)$.

A proof can be found in [FH]. In particular, every indecomposable representation in the category $\mathcal{O}(\mathfrak{sl}_2)$ is isomorphic to one of the examples given above.



(a) $V(m)$ (b) $P(-m-2)$ (c) $M^*(m)$

Figure 3.2: More $\mathfrak{sl}_2$-modules.

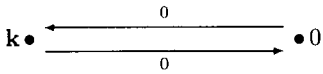
3.4 A matching example

At this point, for each nonnegative integer m, we have found two simple representations ($M(-m-2)$ and $V(m)$) and three indecomposable representations ($M(m)$, $M^*(m)$, $P(-m-2)$) of $\mathfrak{sl}_2(\mathbf{k})$.

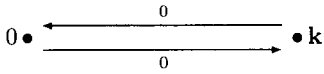
The next step is to match these representations with representations of some quiver. In order to do that, we consider the quiver with relation given in Example 3.1.5. The following proposition gives the classification of the simple and indecomposable representations of this quiver.

Proposition 9. *For the quiver with relation given in Example 3.1.5, there are two simple representations given by*

1.

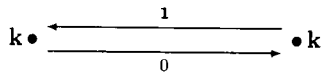


2.

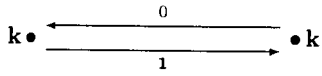


and there are three indecomposable representations given by:

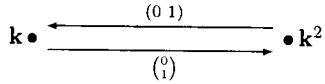
3.



4.



5.



To prove this proposition, we use the previous results described above.

Notice that we have exactly the same number of simple and indecomposable representation. This suggests the following correspondence between each simple and indecomposable representations:

Proposition 10. *There is a bijective correspondence, which preserves inclusions and quotients, between equivalence classes of simple and indecomposable modules of the 2-Kronecker quiver with relation from Example 3.1.5 and the simple and indecomposable modules of $\mathcal{O}(\mathfrak{sl}_2)$, given as follows:*

- $1 \leftrightarrow V(m)$
- $2 \leftrightarrow M(-m-2)$
- $3 \leftrightarrow M^*(m)$
- $4 \leftrightarrow M(m)$
- $5 \leftrightarrow P(-m-2)$

Proof. By Proposition 8, we have short exact sequences

$$0 \to V(m) \to M^*(m) \to M(-m-2) \to 0,$$
$$0 \to M(-m-2) \to M(m) \to V(m) \to 0,$$
$$0 \to M(-m-2) \to P(-m-2) \to M^*(m) \to 0,$$
$$0 \to M(m) \to P(-m-2) \to M(-m-2) \to 0.$$

Quiver representation #5 above has representations #2 and #3 as subrepresentations; #4 has #2 as a subrepresentation; and #3 has #1 as subrepresentation.

The result follows by observing that $P(-m-2)$ has two lie algebra subrepresentations corresponding to the two quiver subrepresentations of #5. Then by a dimension analysis for the last exact sequence, we get the correspondence between the two simple representations. Similarly, $M(m)$ has $M(-m-2)$ as a subrepresentation, corresponding to the quiver subrepresentation #2 in #4, and $M^*(m)$ has $V(m)$ as a subrepresentation, corresponding to the quiver subrepresentation #1 in #3. □

This correspondence is not an accident. In fact, the 2-Kronecker quiver corresponds to the Lie algebra $\mathfrak{sl}_2$ under a correspondence developed by Kac and Moody. In this more general matching, a quiver corresponds to a matrix representing (t_{ij}) where t_{ij} is the number of arrows between vertices i and j. This matrix then is used to formulate a set of relations which describe the corresponding Lie algebra.

Acknowledgments

This is the final report of work developed during the International Research Experiences for Students in Mathematics (IRES) hosted by the Universidade Estadual de Campinas (UNICAMP), SP, Brazil, in July of 2006. The IRES was funded by the National Science Foundation, CNPq (grant 451.154/2006-1) and FAEPEX-UNICAMP (grant 163/2006). The authors would like to thank the Department of Mathematics at UNICAMP for their hospitality, their advisors Professor Marcos Jardim and Adriano Moura, for suggesting the problem and for useful discussions and Professors Helena Lopes and M. Helena Noronha for organizing the event.

References

[De] Harm Derksen: Lecture notes for Math 711, Winter 2001. `http://www.math.lsa.umich.edu/~hderksen/math711.w01/math711.html`

[FS] Jürgen Fuchs and Christoph Schweigert: *Symmetries, Lie Algebras and Representations: A Graduate Course for Physics*. Cambridge: Cambridge Univ. Press, 1997.

[FH] William Fulton, and Joe Harris: *Representation Theory: a First Course*. New York: Springer, 1991 (Graduate Texts in Math. **129**).

[Ha] Paul R. Halmos: *Finite-Dimensional Vector Spaces*, 2nd ed. Princeton, N.J.: Van Nostrand, 1958.

[Hu] James E. Humphreys: *Introduction to Lie Algebras and Representation Theory*. New York: Springer, 1987 (Graduate Texts in Math. **9**).

[Ka] Victor G. Kac: *Infinite Dimensional Lie Algebras*, 3rd ed. Cambridge: Cambridge Univ. Press, 1990.

[Sa] Alistair Savage: Finite-dimensional algebras and quivers, preprint. `arXiv:math/0505082v1` [math.RA].

STUDENT ARTICLE

4

A Fitness-Based Model for Complex Networks

Zhou Fan '10[†]
Harvard University
Cambridge, MA 02138
zhoufan@fas.harvard.edu

Abstract

Complex networks such as the World Wide Web and social relationship networks are prevalent in the real world, and many exhibit similar structural properties. In this paper, a fitness-based model is developed for these complex networks. This model employs a purely "better-get-richer" method of network construction that is believed to realistically simulate the growth process of most real-world networks. Both computer-simulated results and theoretical analysis show that the degree distribution of networks created with this model depends on the distribution of vertex fitnesses; a power-law fitness distribution results in the commonly observed scale-free network structure. In addition, results indicate a small average path length and large clustering coefficient, in accordance with real-world phenomena. It is proposed that this model may serve as a possible explanation of the prevalence of scale-free networks in the real world.[‡]

4.1 Introduction

There are many examples of complex networks in the world, from the more common World Wide Web and social relationship networks to the more obscure power grid of the Western United States and network of scientific paper citations. Over the past decades, researchers have noted that many such real-world networks exhibit similar properties in structure and have studied and modeled them together under the term **complex networks**. A greater understanding of the structure of these abstract complex networks will undoubtedly heighten our understanding of the behavior of their real-world counterparts. Indeed, the study of complex networks has already led to advances in areas such as immunization and Internet simulation [BB1]. In this paper, we will provide a model of network growth similar to an existing model, but we will incorporate a fitness concept, and we will examine the structural properties of our model in comparison to real-world phenomena.

4.2 Background

In the field of complex networks, the individual network components are represented by vertices of a graph and the connections between them are represented by the edges. For instance, the vertices of a network representing the World Wide Web would be the web pages, with two vertices connected by an

[†]Zhou Fan, Harvard '10, is a prospective concentrator in mathematics or applied mathematics. He was born in Hangzhou, China and grew up in Parsippany, New Jersey, where he graduated from Parsippany Hills High School.

[‡]Part of the research for this paper was conducted at the 2005 Research Science Institute under the guidance of King Y. Yick, sponsored by a grant from the Center for Excellence in Education.

edge when there is a link from one page to the other. (For the purposes of this paper, we consider only undirected and unweighted edges.) It has been observed that the vertex degrees of a large majority of complex networks satisfy a power-law distribution, and such networks are called **scale-free** [AB].

The Barabási-Albert model (BA model), one of the most basic and widely-accepted models of complex networks, captures their scale-free structure [AB]. The BA model constructs networks based on the two ideas of network growth and preferential attachment: more popular vertices of a network attract more new vertices. In addition to being scale-free, networks constructed using this model have a small average path length between vertices and display a relatively high tendency for a vertex's neighbors to connect to each other; this tendency is known as **clustering**. Both of these properties are also observed in real-world networks [AB]. One should note, however, that the BA model always predicts a power-law degree distribution where the probability density function of the vertex degrees, k, scales according to k^{-3}, while the degree distributions of real-world networks have varying powers of k. Also, a few real-world networks have an exponential degree distribution [St].

4.3 Fitness

The BA model relies on preferential attachment, the idea that a more popular website or scientific paper will attract more links or citations. A fundamentally different concept is that a more helpful, useful, ingenious, or simply "better" vertex will attract more such edges. This second concept is **fitness**-based, and the "better" vertices are deemed to be more **fit**. A weakness of the BA model is that it does not address fitness; for example, it does not allow a newer but very good scientific paper to become more frequently cited than an older but less significant one. Thus, a modification of the BA model has been developed that uses both preferential attachment and fitness [BB2]. This modified model, in essence, assumes that preferential attachment and fitness are separate and parallel causes of network structure.

In our paper, we examine whether a model based on fitness alone, without preferential attachment, can produce results similar to those produced by the BA model. This is intuitively reasonable; for example, a popular scientific paper probably becomes more frequently cited because it is better than other papers. We thus hypothesize that a model based solely on the fitness concept may produce results similar to those of the BA model. It should be noted that a network model based solely on the fitness concept has already been developed by Caldarelli et al., but it uses an approach to network construction different from that used in the BA model [CCRM, SC]. In this study, we instead examine a network construction algorithm based on the BA construction algorithm, but we employ the fitness concept instead of preferential attachment.

Specifically, our algorithm is as follows: Fix a probability distribution of fitnesses, $\rho(\eta)$, and the number of edges m with which a newly formed vertex starts. When the network grows sufficiently large so that the initial vertices do not matter, m becomes the average vertex degree. Begin with N_0 vertices, where N_0 is small. Randomly assign to each vertex a fitness value η from the fitness distribution $\rho(\eta)$, where a high value of η corresponds to a vertex that is more fit. Once a fitness value is assigned to a vertex, it does not change. At each time step $t = 1, 2, 3, \ldots$, add one vertex to the network, connect it to m existing vertices, and assign to it a fitness value based on $\rho(\eta)$. For each of these m new connections, the probability of connecting to an existing vertex i with fitness η_i is proportional to η_i, i.e.,

$$P = \frac{\eta_i}{\sum_{j=1}^{N} \eta_j}$$

with N being the size of the network prior to the addition of this new point. We connect the m edges so that no two edges connect to the same vertex.

4.4 Degree Distribution

The degree distribution of networks created using this fitness-based algorithm can be examined using the **continuum theory**, a method developed by Barbási and Albert in which network growth is treated as a continuous process to allow simplification of the model using calculus [AB]. Such an approximation should match closely with discrete network growth, provided that we consider networks of sufficiently large scale, i.e., networks that undergo a large number of timesteps. Consider a vertex V with fitness η, and assume that its degree k_V is a continuous function of time. Because during each unit of time m new edges are formed, we expect that

$$\frac{dk_V}{dt} \approx m\frac{\eta}{\sum_{j=1}^{N}\eta_j}.$$

For large enough N we can make the approximation

$$\sum_{j=1}^{N}\eta_j \approx N\bar{\eta} = (N_0+t)\bar{\eta},$$

where $\bar{\eta}$ is the expected value of η. So

$$\frac{dk_V}{dt} = m\frac{\eta}{(N_0+t)\bar{\eta}}.$$

Integration yields

$$k_V = \int \frac{m\eta}{(N_0+t)\bar{\eta}}dt = \frac{m\eta}{\bar{\eta}}\ln(N_0+t) + C.$$

Let t_0 be the time that this vertex was added to the network. Since $k_V = m$ at time $t = t_0$,

$$C = m - \frac{m\eta}{\bar{\eta}}\ln(N_0+t_0)$$

$$k_V = m + \frac{m\eta}{\bar{\eta}}\ln\frac{N_0+t}{N_0+t_0}.$$

We can now calculate the cumulative distribution function (CDF) of k_V as

$$\begin{aligned}\mathbb{P}(k_V \le k) &= \mathbb{P}\left(m + \frac{m\eta}{\bar{\eta}}\ln\frac{N_0+t}{N_0+t_0} \le k\right)\\ &= \mathbb{P}\left(\ln\frac{N_0+t}{N_0+t_0} \le \frac{\bar{\eta}(k-m)}{m\eta}\right)\\ &= \mathbb{P}\left(\frac{N_0+t}{e^{\frac{\bar{\eta}(k-m)}{m\eta}}} - N_0 \le t_0\right).\end{aligned}$$

There are $N_0 + t$ total vertices in the network, so for any particular τ, $1 \le \tau \le t$, the probability that $t_0 = \tau$ is $\frac{1}{N_0+1}$ and the probability that $t_0 = 0$ (the vertex is a starting vertex) is $\frac{N_0}{N_0+t}$. Thus in the continuous analogue, $\mathbb{P}(t' \le t_0) = \frac{N_0+t'}{N_0+t}$, so So

$$\begin{aligned}\mathbb{P}(k_V \le k) &= \mathbb{P}\left(\frac{N_0+t}{e^{\frac{\bar{\eta}(k-m)}{m\eta}}} - N_0 \le t_0\right)\\ &= \frac{1}{N_0+t}\left[t - \left(\frac{N_0+t}{e^{\frac{\bar{\eta}(k-m)}{m\eta}}} - N_0\right)\right]\\ &= 1 - e^{\frac{\bar{\eta}(m-k)}{m\eta}}.\end{aligned}$$

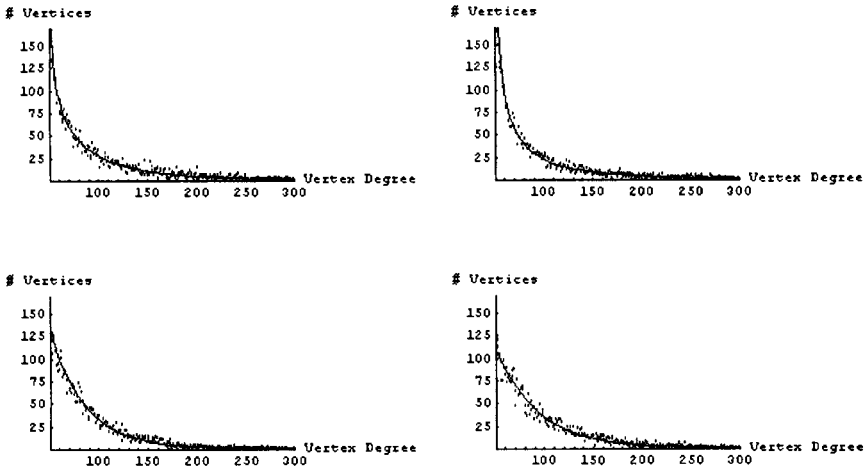


Figure 4.1: Predicted and simulated degree distributions. Solid lines represent predictions of the continuum theory and scatter plots represent simulated results for (a) uniform $\rho(\eta)$, $0 < \eta < 1$; (b) exponential $\rho(\eta) = e^{-\eta}$, $0 < \eta < \infty$; (c) power-law $\rho(\eta) \sim \eta^{-3}$, $1 < \eta < \infty$; (d) power-law $\rho(\eta) \sim \eta^{-4}$, $1 < \eta < \infty$.

We obtain the probability density function (PDF) of the vertex degree by differentiating the CDF with respect to k:

$$\frac{d}{dk}\mathbb{P}(k_V \leq k) = \frac{\bar{\eta}}{m\eta} e^{\frac{\bar{\eta}(m-k)}{m\eta}}.$$

This is the PDF for the degree of a vertex of fitness η, which we will denote as $P(k_\eta)$. To obtain the overall PDF, we take a weighted average of these fitness-based PDFs with the weights being the probabilities of having a fitness η. In other words,

$$P(k) = \int_{\eta_{\min}}^{\eta_{\max}} \rho(\eta) P(k_\eta) d\eta,$$

or

$$P(k) = \int_{\eta_{\min}}^{\eta_{\max}} \rho(\eta) \frac{\bar{\eta}}{m\eta} e^{\frac{\bar{\eta}(m-k)}{m\eta}} d\eta. \tag{4.1}$$

In this overall PDF, k is the continuous random variable for vertex degree, m is the constant for the average vertex degree, $\rho(\eta)$ is the PDF of fitnesses η, $\bar{\eta}$ is the expected value of η as determined by $\rho(\eta)$, and $\eta_{\min}$ and $\eta_{\max}$ are the bounds of the fitness values. It is important to note that this PDF does not depend on the present time t or the network size N. That is, as long as the size of the network is large enough so that the initial approximations are true, the PDF for the vertex degrees is constant over time as new vertices are added to the network.

We can scale equation (4.1) by multiplying by the total number of vertices N to predict the degree distribution of the network. To verify the predictions of the continuum theory, we numerically simulated this network construction algorithm for $m = 50$, $N = 5000$ and a variety of distributions $\rho(\eta)$ and calculated the degree distributions. The data from the simulations matches our theoretical result (Figure 4.1).

We also note that, as in previously developed fitness-based modifications of the BA model, $P(k)$ depends on the fitness distribution $\rho(\eta)$, and that for our model $P(k)$ is very versatile and varies greatly

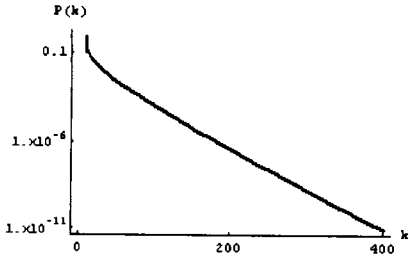


Figure 4.2: Semilog plot of $P(k)$ for uniform $\rho(\eta)$, $0 < \eta < 1$ and $m = 10$.

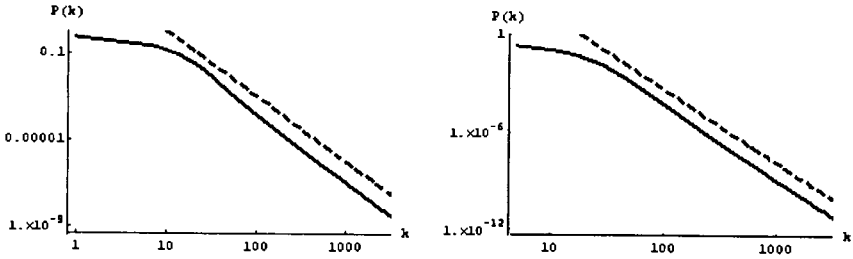


Figure 4.3: Log-log plot of $P(k)$ for $m = 10$ and (a) $\rho(\eta) \sim \eta^{-3}$, (b) $\rho(\eta) \sim \eta^{-4.5}$. Solid lines are plots of $P(k)$; dashed lines are plots of k^{-3} and $k^{-4.5}$ for (a) and (b) respectively.

with different fitness distributions. Evaluating this integral for a uniform fitness distribution over varying bounds and average vertex degree m results in varying exponential-tailed distributions for $P(k)$; one distribution is shown in Figure 4.2. Evaluating this integral for a power-law fitness distribution $\rho(\eta) \propto \eta^{-b}$ over varying m yields distributions of $P(k)$ with power-law tails of the same power $-b$; two such distributions are shown in Figure 4.3. Thus, with different power-law fitness distributions, we can obtain scale-free networks with degree distributions of various powers.

4.5 Path Length

Two other empirical properties observed in real-world networks are a small average path length between vertices and a high tendency for small clusters of highly connected nodes to form. We examined path length and clustering of networks produced by our model using computer simulations, and we draw comparisons both to empirical data and to results of the BA model. All data for path length and clustering coefficients are average values over 50 network constructions. We find through our simulations that our fitness-based algorithm does generate networks with small average path lengths. Using a power-law distribution with power $-b$, we find that for fixed values of N and m, the average path length of a network quickly increases to a low asymptotic limit as b increases (Figure 4.4a). Fixing m and b, we observe that the average path length increases logarithmically with N, a phenomenon also observed both in the original BA model and in random graphs (Figure 4.4b) [AB]. However, as in the BA model, the path lengths of our networks are of the same order of magnitude but consistently lower than those of real-world networks of the same size and average vertex degree, indicating that our algorithm may be overly effective, as compared to real-world processes, in bringing the vertices of the network closer together.

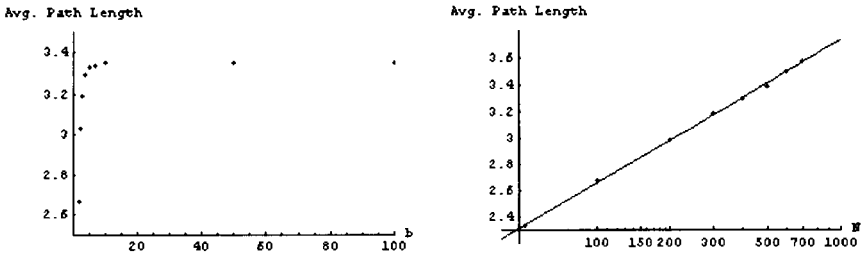


Figure 4.4: (a) Linear plot of path length versus b for $N = 300$ and $m = 3$. (b) Log-linear plot of path length versus N for $b = 3$ and $m = 3$. Solid line is the exponential regression curve.

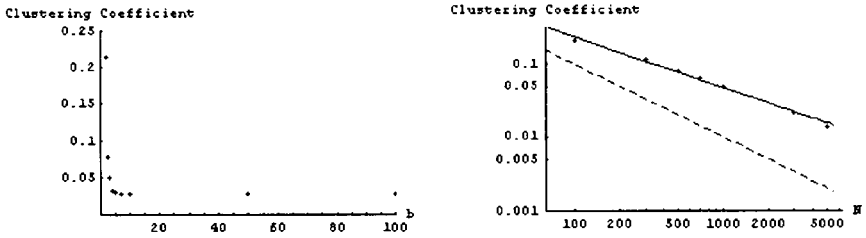


Figure 4.5: (a) Linear plot of clustering coefficient C versus b for $N = 1000$ and $m = 10$. (b) Log-log plot of C versus N for $b = 3$ and $m = 10$. Solid line corresponds to $N^{-0.7}$, and dashed line corresponds to C for a random graph with $m = 10$.

4.6 Clustering Coefficient

To quantify the concept of clustering, we use the clustering coefficient C developed by D. J. Watts [Wa]. C is the average of $\frac{2E_i}{k_i(k_i-1)}$ for all vertices i in the network, where k_i is the degree of vertex i and E_i is the number of edges in the subgraph of its k_i neighboring vertices. As in the case path length, if we fix the network size N and the average vertex degree m, then the clustering coefficient rapidly decreases to an asymptotic limit as b increases (Figure 4.5a). To obtain an idea of how large or small these clustering coefficients are, we fix m and b and compare the clustering coefficients of our networks to those of random graphs for different values of N (Figure 4.5b). We first note that the clustering coefficients of our networks are consistently higher than those of random graphs of the same size (whose clustering coefficients are given by $\frac{m}{N}$), and this difference increases with the size of the network. Secondly, C decreases with N as a power-law, as is observed for both random graphs and BA networks. Finally, the power of this relationship between C and N is -1 for random graphs, -0.75 for BA networks, and -0.70 for our fitness-based networks, while for real-world networks, this power is 0 and network size does not seem to affect the value of C [AB].

4.7 Conclusion

We have created a network model that parallels a simple and accepted existing model, the BA model, but that uses a "better-get-richer" instead of "richer-get-richer" growth algorithm. Our study indicates several important facts about our fitness-based network model. The first is that through a power-law fitness distribution, we can obtain scale-free networks. It may seem that having a power-law fitness

distribution is an arbitrary criterion, but in many real-world situations where individuals such as people or cities are ranked according to wealth or some other measure of "fitness," these fitnesses fall under a power law distribution, as is stated in the empirical Zipf's law [CCRM]. Thus, it may be a reasonable hypothesis that real-world networks have power-law fitness distributions. If this were true, our model would indicate that Zipf's law and the ubiquitous nature of scale-free networks in the real world might be related phenomena. The varying powers of the degree distributions of real networks can be explained by varying powers of fitness distributions; the analysis of our model shows that these two powers are equal.

A second observation is that in our model, non-power-law distributions of fitness result in other network structures. Specifically, a uniform fitness distribution results in an exponential degree distribution. This may be related to certain real-world networks that are indeed not scale-free but follow such an exponential degree distribution. The Western United States power grid and the network of neurons in a human brain are notable instances of such exponential distributions [AB]. The structures of these two networks in particular are heavily influenced by the physical location of their vertices, and thus the vertex fitness values may be more indicative of the number of other vertices that physically surround them and thus may fall under a relatively more uniform probability distribution than the fitnesses of networks without this distance restriction.

A final observation is that our fitness-based networks with power-law fitness distributions very closely resemble the BA network, particularly with respect to how path length and clustering scale with network size. Along with a scale-free degree distribution, this is evidence that our models are very similar in structure to BA networks. Thus, we have shown that newly added vertices of a network do not need knowledge of the popularity of the current vertices in order to maintain a scale-free network structure, and that knowledge of the vertex popularity values (as in the BA model) does not alter three of the most significant structural properties. It should be noted, though, that this result is dependent on the hypothesis that fitness distributions are power laws.

Important work needs to be done in studying on a microscopic level the growth patterns of particular real-world networks to determine their underlying fitness distributions. Further work in this area can also be done by examining models with vertex fitnesses that vary over time, as well as by adding complications such as directed and weighted edges. Overall, we have shown that a fitness-based variation of the BA model can produce some of the important trends observed in the structure of real-world complex networks.

4.8 Acknowledgements

The bulk of this research was performed in the summer of 2005 under the mentorship of King Y. Yick, graduate student of mathematics at MIT. It was conducted as part of the Research Science Institute (RSI), sponsored by the Center for Excellence in Education. Staff of RSI 2005, in particular Dr. Jenny Sendova, contributed to the original drafting of this paper.

References

[AB] Reka Albert and Albert-László Barabási: Statistical mechanics of complex networks. *Reviews of Modern Physics* **74** #1 (2002), 47–97. See also references therein.

[BB1] Albert-László Barabási and Eric Bonabeau: Scale-free networks. *Scientific American* (May 2003), 50–59.

[BB2] Ginestra Bianconi and Albert-László Barabási: Competition and multiscaling in evolving networks. *Europhysics Letters* **54** #4 (2001), 436–442.

[CCRM] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Muñoz: Scale-free networks from varying vertex intrinsic fitness. *Physical Review Letters* **89** #25 (2002), 8702.

[Ev] David Everitt: Generating random variables. Available at `http://www.it.usyd.edu.au/~deveritt/networksimulation/rv.pdf` (2005/08/01).

[SC] Vito D. P. Servedio and Guido Caldarelli: Vertex intrinsic fitness: How to produce arbitrary scale-free networks. *Physical Review E* **70** (2004), 056126.

[St] Steven H. Strogatz: Exploring complex networks. *Nature* **410** (2001), 268–276.

[Wa] Duncan J. Watts: Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology* **105** #2 (1999), 493–527.

STUDENT ARTICLE

5

Does Every Polynomial Root Have a Simple Approximation?

Bryan Gin–ge Chen '07†
Harvard University
Cambridge, MA 02138
bryangingechen@gmail.com

Abstract

The practice of neglecting small terms of an equation is analyzed in the case of polynomial root approximations. Our discussion centers on the following new result: The roots of a polynomial can be approximated self-consistently by roots of much simpler equations consisting of pairs of terms from the polynomial.

5.1 Introduction

The essence of mathematical modeling is to take a real–world question and translate it into a mathematical problem which can then be solved, yielding insight into the original question. In the course of such modeling, approximations must invariably be made.

It is not an exaggeration to say that all important equations take the schematic form

$$t_1(p_1, p_2, \dots) + t_2(p_1, p_2, \dots) + \cdots = 0,$$

where t_n are arbitrary terms and p_m are arbitrary parameters.

One commonly used approximation simplifies these equations by choosing some subset of terms deemed to be the most important and then neglecting all the others. If we choose only the two largest, we end up with

$$t_i(p_1, p_2, \dots) + t_j(p_1, p_2, \dots) \approx 0.$$

We will refer to such an approximation as a **dominant balance approximation**, since it seeks a solution which "balances" the two dominant terms against each other. The question we are interested in is: How often can solutions to an equation be approximated by the behavior of a few dominant terms?

We investigated this question for the case of polynomial equations of arbitrary order in a single variable.

†Bryan Gin-ge Chen, Harvard '07, is a physics concentrator in Adams House. He hails from Centerville, Ohio, where he attended Centerville High School. His mathematical interests extend freely to all that is unreasonably effective in the natural sciences, including calculus and linear algebra, scaling and similarity solutions, topology, and symmetry.

5.2 Example: $a_5x^5 + a_1x + a_0 = 0$

When are the roots of this quintic trinomial dominated by the behavior of just two of the terms? In other words, when are the roots of the equations

$$a_5x^5 + a_1x = 0$$
$$a_5x^5 + a_0 = 0$$
$$a_1x + a_0 = 0$$

close to the actual roots?

Note that if the roots of these equations are relatively close to the true roots, then we have indeed found simple solutions to the trinomial equation — these approximations depend on only two coefficients!

We say that a choice of terms yields a **self-consistent** root when those terms are larger (in absolute value) than any other terms when we set x equal to the root of the simplified equation. For instance, the roots of the equation for which we say that a_5x^5 and a_1x are dominant are $x = 0$ or $x = \left(-\frac{a_1}{a_5}\right)^{1/4}$ (all four of these 4th roots). We ignore the $x = 0$ root as this is obviously inconsistent with our assumption that a_5x^5 and a_1x are largest near the root. Therefore the self-consistency condition is

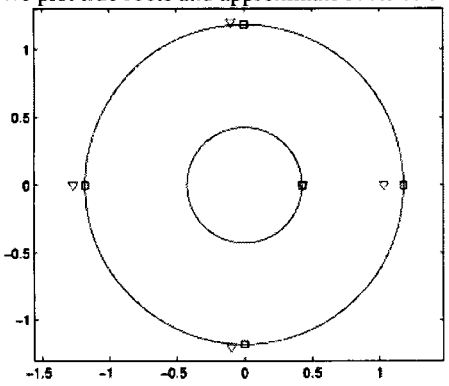
$$\left|a_5\left(-\frac{a_1}{a_5}\right)^{5/4}\right| = \left|a_1\left(-\frac{a_1}{a_5}\right)^{1/4}\right| > |a_0|,$$

or equivalently $\left|\frac{a_0^4 a_5}{a_1^5}\right| < 1$.

Working out the other cases yields the following result:

$$\left|\frac{a_0^4 a_5}{a_1^5}\right| < 1 \iff (a_1x, a_5x^5) \text{ and } (a_0, a_1x) \text{ yield self-consistent roots,}$$
$$\left|\frac{a_0^4 a_5}{a_1^5}\right| > 1 \iff (a_0, a_5x^5) \text{ yields self-consistent roots.}$$

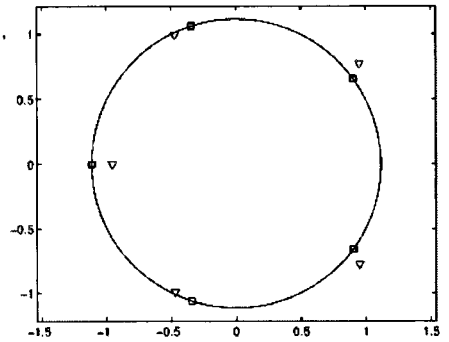
We plot true roots and approximate roots of one from each "class" of trinomial below:



$.832x^5 - 1.62x + .692 = 0$, $\left|\frac{a_0^4 a_5}{a_1^5}\right| = .0169$

true roots: .434, 1.04, -1.27, $-.0993 \pm 1.20i$;

approx. roots: .426, ± 1.18, $\pm 1.18i$.



$.690x^5 + .669x + 1.19 = 0$, $\left|\frac{a_0^4 a_5}{a_1^5}\right| = 10.4$

true roots: $-.956$, $-.471 \pm -.991i$, $.949 \pm .773i$;

approx. roots: -1.12, $-.345 \pm 1.06i$, $.902 \pm .656i$.

These examples were generated using MATLAB with coefficients drawn from a normal distribution centered at zero with unit variance.

5.3 Generalization to Arbitrary Polynomials

It is possible to generalize the ideas in the previous column and prove the following result:

Theorem 1. *Given a nondegenerate*[1] *polynomial* $a_n x^n + a_{n_1} x^{n_1} + a_{n_2} x^{n_2} + \cdots + a_{n_p} x^{n_p} + a_0$ *where* $n = n_0 > n_1 > n_2 > \cdots > n_p > n_{p+1} = 0$, *all self-consistent approximations to the roots of this polynomial will come from pairs* $\{(a_n x^n, a_{n_{j_1}} x^{n_{j_1}}), (a_{n_{j_1}} x^{n_{j_1}}, a_{n_{j_2}} x^{n_{j_2}}), \cdots, (a_{n_{j_s}} x^{n_{j_s}}, a_0)\}$, *where* $\{n_{j_k}\}$ *is a subsequence of the* $\{n_l\}$ *above.*

The pairs of terms which give the self-consistent approximations can be bracketed together as below:

$$a_n \overbrace{x^n + a_{n_1} x^{n_1} + \ldots + a_{n_{j_1}}} \overbrace{x^{n_{j_1}} + \ldots + a_{n_{j_s}}} \overbrace{x^{n_{j_s}} + \ldots + a_{n_p} x^{n_p} + a_0}$$

Theorem 1 essentially states that this series of brackets will not cross itself, and will reach from the $a_n x^n$ term to a_0. Note that since a pair $(a_j x^j, a_k x^k)$ yields $k - j$ different approximate roots (via the $\frac{1}{k-j}$th roots of $-\frac{a_k}{a_j}$), the total number of self-consistent approximate roots is guaranteed to be $(n_{j_s} - 0) + (n_{j_{s-1}} - n_{j_s}) + \cdots + (n_{j_2} - n_{j_1}) + (n - n_{j_1})$ which telescopes to n.

The nondegeneracy condition in Theorem 1 excludes polynomials which have pairs of terms which are not completely dominant at the approximate roots. In our quintic trinomial example, the degenerate polynomials would be those for which $\left|\frac{a_0^4 a_5}{a_1^5}\right| = 1$. This condition gives us the set of polynomials for which all three terms are equally large when evaluated at the approximate roots.

Note that Theorem 1 does not say anything about the accuracy of these self-consistent approximations — it merely states that they exist.

Sketch of proof. The key 'trick' is to transform the self-consistency inequalities by taking logarithms. In the quintic trinomial example, letting $A_j = \log |a_j|$, we have:

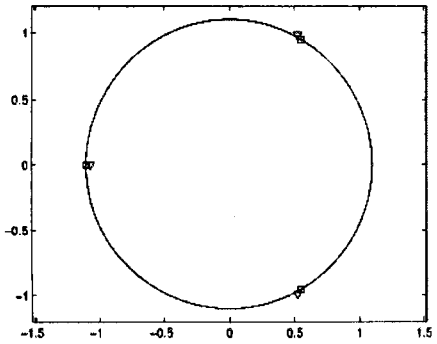
$$A_0 + A_2 - 2A_1 < 0 \Rightarrow (A_0, A_1, A_2) \cdot (1, -2, 1) < 0,$$
$$A_0 + A_2 - 2A_1 > 0 \Rightarrow (A_0, A_1, A_2) \cdot (1, -2, 1) > 0.$$

Here $\cdot$ is the ordinary dot product. The two classes have become half–spaces in $\mathbb{R}^3$.In the general case, the classes are cones in $\mathbb{R}^{j-2}$ (where j is the number of terms) defined by a set of dot–product inequalities. Using a 1958 result due to Samelson, Thrall and Wesler [STW], we can show that these cones partition all of $\mathbb{R}^{j-2}$.

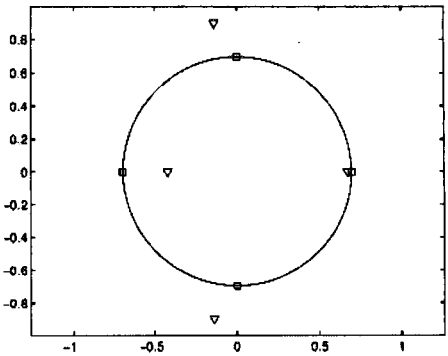
5.4 Self-Consistent Approximation Picture Gallery

The following polynomials were generated in MATLAB by choosing coefficients from a normal distribution with unit variance. The approximate roots were found by an algorithm based on Theorem 1. The values ϵ are quantities analogous to $\left|\frac{a_0^4 a_5}{a_1^5}\right|$ in the example, but are now chosen so that a root is self-consistent if $\epsilon < 1$ for all ϵ.

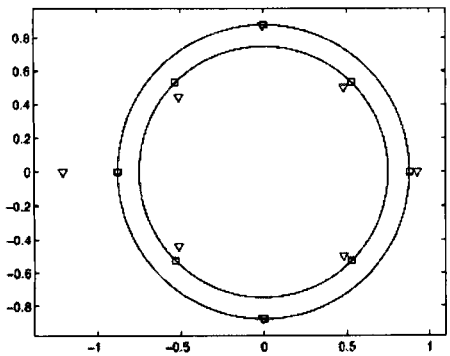
[1]For the purposes of Theorem 1, a polynomial $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ is **degenerate** if the vector $(\log |a_n|, \log |a_{n-1}|, \ldots, \log |a_0|)$ is in the linear space spanned by $(1, 1, \ldots, 1)$ and $(n, n-1, \ldots, 0)$.



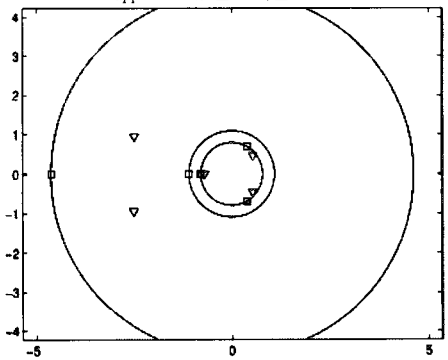
$1.20x^3 + .0198x^2 + .157x + 1.60 = 0$
$\epsilon_j = \{1.24 \times 10^{-3}, 3.34 \times 10^{-6}\}$
true roots: -1.07, .525±.987i;
approx. roots: -1.10, .550±.953i.



$-2.17x^4 + .0592x^3 - 1.01x^2 + .614x + .508 = 0$
$\epsilon_j = \{.501, .859, 2.36 \times 10^{-6}\}$
true roots: -.424, .673, -.138±.895i;
approx. roots: ±.695, ±.695i



8th order polynomial equation
$\max \epsilon_j = .0779, \min \epsilon_j = 1.07 \times 10^{-16}$
selected true roots: .927, -1.21, -.509±.442i; closest approx. roots: .879, -.879, -.530±.530i



$.262x^5 + 1.21x^4 + 1.32x^3 - .931x^2 - .0112x + .645 = 0$
$\epsilon_j = \{2.59 \times 10^{-6}, .719, .38, .235\}$
true roots: -.695, .534±.459i, -2.50±.948i; approx. roots: -1.09, .394±.682i, -4.62, -7.88.

The approximate roots we chose in this way are in general quite close to the actual roots. However, in the fourth plot there is a pair of complex roots that is approximated by a pair of real roots. It appears that partitioning using Theorem 1 does not place some polynomials correctly.

5.5 Self-Consistency Is Not Enough

Though the self-consistency condition gives us a simple criterion for choosing dominant terms, the choices do not always yield good approximations. The reason is that the self-consistency condition completely ignores the possibility of multiple roots.

If we apply Theorem 1 to the quadratic equation $a_2x^2 + a_1x + a_0 = 0$, we find:

$$\left|\frac{a_0a_2}{a_1^2}\right| < 1 \iff (a_1x, a_2x^2) \text{ and } (a_0, a_1x) \text{ yield self-consistent roots,}$$

$$\left|\frac{a_0a_2}{a_1^2}\right| > 1 \iff (a_0, a_2x^2) \text{ yields self-consistent roots.}$$

Thus the degenerate, borderline case is when $\left|\frac{a_0a_2}{a_1^2}\right| = 1$. However, by the quadratic formula

$x = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_0a_2}}{2a_2}$, there is a multiple root when $\frac{a_0a_2}{a_1^2} = \frac{1}{4}$, so this should be the borderline case.

5.6 Series Expansions for Roots

Let us take a different approach now. Let our polynomial be $\sum_{j=0}^{n} a_j x^j = 0$. Choose two terms $a_k x^k$ and $a_j x^j$. We now rescale by $x \mapsto y\left(-\frac{a_j}{a_k}\right)^{\frac{1}{k-j}}$, which yields (after division by a simplifying factor):

$$\frac{a_n\left(-\frac{a_j}{a_k}\right)^{\frac{n}{k-j}}}{\left(-\frac{a_j^k}{a_k^j}\right)^{\frac{1}{k-j}}} y^n + \cdots + y^k + \cdots - y^j + \cdots + \frac{a_0}{\left(-\frac{a_j^k}{a_k^j}\right)^{\frac{1}{k-j}}} = 0.$$

We can rewrite this as

$$c_n y^n + c_{n-1} y^{n-1} + \cdots + y^k + \cdots - y^j + \cdots + c_0 = 0,$$

where for all $l \neq k, j$, $c_l = \left((-a_j)^{l-k} a_k^{j-l} a_l^{k-j}\right)^{\frac{1}{k-j}}$. It is clear that if the $n-1$ coefficients c_l are all sufficiently small, then $y \approx 1^{\frac{1}{k-j}}$ and $x \approx \left(-\frac{a_j}{a_k}\right)^{\frac{1}{k-j}}$. But what do we mean by sufficiently small? We can write a series solution for y by assuming an ansatz of the form

$$y_{(j,k)} = \sum_{s_0, s_2, \ldots, s_n = 0}^{\infty} A_{s_0, s_1, \ldots, s_n} c_n^{s_n} c_{n-1}^{s_{n-1}} \cdots c_0^{s_0},$$

where we have $n-1$ quantities c_l and s_l (no c_j, c_k or s_j, s_k). It makes sense to say that the terms are sufficiently small if this series converges. It is thus more natural to say that the pair $a_k x^k$, $a_j x^j$ is dominant at the root if the series for $y_{(j,k)}$ converges (rather than using the self-consistency conditions)!

The self-consistency conditions are equivalent to requiring that all $|c_l| < 1$; however, the domain of convergence of this series is in general a more complicated object.

Series solutions of polynomials can be written in terms of hypergeometric functions, but the domains of convergence are only known in some cases; see [St, PT].

5.7 What If We Iterate?

Suppose that we have some method of choosing pairs of terms which gives us approximate roots from dominant balances. Note that

$$f(z) = \frac{f^{(n)}(0)}{n!} z^n + \frac{f^{(n-1)}(0)}{(n-1)!} z^{n-1} + \cdots + f(0),$$

where $f^{(m)}$ is the m-th derivative of f.

The approximate root we get by assuming that the j-th and k-th terms are dominant is then

$$\left(-\frac{f^{(j)}(0)}{f^{(k)}(0)}(k-j)!\right)^{\frac{1}{k-j}}.$$

To improve on this root, instead of deriving the next term of a series (as in the previous section), consider $f(u - z_1)$ where $z_1 = \left(-\frac{f^{(j)}(0)}{f^{(k)}(0)}(k-j)!\right)^{\frac{1}{k-j}}$. Note that for any z_1, the roots of $f(z)$ are precisely $u + z_1$ for the roots u of the polynomial $f(u + z_1)$. Using Taylor's theorem, we have

$$f(u + z_1) = \frac{f^{(n)}(z_1)}{n!} u^n + \frac{f^{(n-1)}(z_1)}{(n-1)!} u^{n-1} + \cdots + f(z_1).$$

If the terms proportional to u^p and u^q (with $q > p$) are dominant terms of this polynomial, an approximate root of this polynomial will be $u_2 = \left(-\frac{f^{(p)}(z_1)}{f^{(q)}(z_1)}(q-p)!\right)^{\frac{1}{q-p}}$, which gives us

$$z_2 = u_2 + z_1 = z_1 + \left(-\frac{f^{(p)}(z_1)}{f^{(q)}(z_1)}(q-p)!\right)^{\frac{1}{q-p}}.$$

In general, there are multiple choices of pairs of terms which will give us dominant terms. Furthermore, each of these $(k-j)$th root expressions above will have $k-j$ different solutions — this shows that this process will **branch**. If we keep iterating, we have the recursive function:

$$z_{m+1} = z_m + \left(-\frac{f^{(p_m)}(z_m)}{f^{(q_m)}(z_m)}(q_m - p_m)!\right)^{\frac{1}{q_m - p_m}}.$$

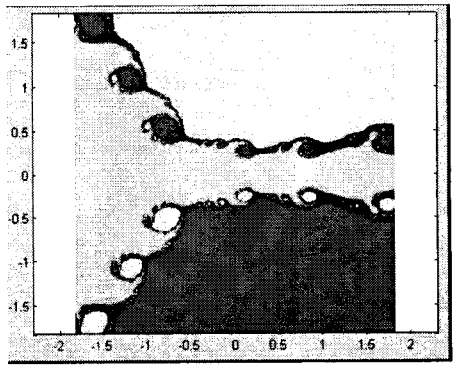
Note that if $q_m = 1$ and $p_m = 0$ for all m, this process no longer branches. In fact, we now have

$$z_{m+1} = z_m - \frac{f(z_m)}{f'(z_m)},$$

which is precisely Newton's Method! Thus we may interpret Newton's method as an iterated dominant balance method which always assumes that the 1st order and 0th order terms dominate, or rather, we might interpret the iteration of a dominant balance method as a branching version of Newton's method.

If we color each point in $\mathbb{C}$ according to which point it converges to upon iterating Newton's method, we produce the Newton fractal. Similarly, with a branching algorithm, we can color the points of $\mathbb{C}$ according to which set of points we obtain. Below, we compare the "Newton fractals" of Newton's method, and iterating the self-consistent root algorithm based on Theorem 1.

This image shades each point in the plane according to which roots it goes to on iteration of the self-consistent roots method.



This image shades each point in the plane according to which roots it goes to after iterating Newton's method.

5.8 Conclusion

The results described in this paper give evidence in one case for a fact which equation–solvers have known intuitively for a long time: the solutions to equations are often determined in a large part by the behavior of the largest terms in the equation. Are there similar results for other equations — in particular, ordinary differential equations or partial differential equations?

Acknowledgments

I thank Professor Michael Brenner for introducing me to this problem and advising me throughout my work. Thanks to Takuya Kitagawa '08 for helpful discussions as well.

References

[PT] Mikael Passare and August Tsikh: Algebraic equations and Hypergeometric Series, *The Legacy of Niels Henrik Abel* (New York: Springer, 2004), 563–582.

[STW] Hans Samelson, R.M. Thrall and Oscar Wesler: A partition theorem for Euclidean n–space, *Proc. of the AMS* **9** (1958), 805–807.

[St] Bernd Sturmfels: Solving algebraic equations in terms of $\mathcal{A}$–hypergeometric series, *Discrete Math.* **210** (2000), 171–181.

FACULTY FEATURE ARTICLE

6

The ABC's of Number Theory

Prof. Noam D. Elkies[†]
Harvard University
Cambridge, MA 02138
elkies@math.harvard.edu

Abstract

The ABC conjecture is a central open problem in modern number theory, connecting results, techniques and questions ranging from elementary number theory and algebra to the arithmetic of elliptic curves to algebraic geometry and even to entire functions of a complex variable. The conjecture asserts that, in a precise sense that we specify later, if A, B, C are relatively prime integers such that $A + B = C$ then A, B, C cannot all have many repeated prime factors. This expository article outlines some of the connections between this assertion and more familiar Diophantine questions, following (with the occasional scenic detour) the historical route from Pythagorean triples via Fermat's Last Theorem to the formulation of the ABC conjecture by Masser and Oesterlé. We then state the conjecture and give a sample of its many consequences and the few very partial results available. Next we recite Mason's proof of an analogous assertion for polynomials $\mathsf{A}(t), \mathsf{B}(t), \mathsf{C}(t)$ that implies, among other things, that one cannot hope to *dis*prove the ABC conjecture using a polynomial identity such as the one that solves the Diophantine equation $x^2 + y^2 = z^2$. We conclude by solving a Putnam problem that predates Mason's theorem but is solved using the same method, and outlining some further open questions and fragmentary results beyond the ABC conjecture.[‡]

6.1 Pythagorean triples: $x^2 + y^2 = z^2$

An ordered triple (x, y, z) of integers is called a **Pythagorean triple** if and only if it solves the Diophantine equation $x^2 + y^2 = z^2$; that is, if and only if $|x|$ and $|y|$ are the lengths of the sides, and $|z|$ the length of the hypotenuse, of a right triangle. (We allow degenerate triangles with a "side" of length zero.) It is well-known that every such triple is proportional to

$$(x, y, z) = (m^2 - n^2, 2mn, m^2 + n^2) \tag{6.1}$$

for some integers m, n. Equivalently (dividing by n^2 to obtain polynomials in the single rational variable $t = m/n$), the solution (x, y, z) is proportional to $(t^2 - 1, 2t, t^2 + 1)$ for some $t \in \mathbb{Q}$, or to $(1, 0, 1)$ which arises for "$t = \infty$" (corresponding to $(m, n) = (1, 0)$). That is, all Pythagorean triples are accounted for by the single polynomial identity

$$(t^2 - 1)^2 + (2t)^2 = (t^2 + 1). \tag{6.2}$$

[†]Noam D. Elkies earned his doctorate in mathematics in 1987 at Harvard, where his advisors where Professors Barry Mazur and Benedict H. Gross. After three years in Harvard's Society of Fellows he joined the Mathematics faculty and has remained at Harvard since. Most of his research is in number theory, usually Diophantine geometry (the combination of algebraic geometry and Diophantine equations) and/or computational number theory. Other interests include some combinatorial mathematics (lattices and codes, incidence geometry, and combinatorial games) and, outside of mathematics, classical music (mostly composition and piano) and chess (usually chess problems and endgames).

[‡]Supported in part by NSF grant DMS-0501029.

This classical fact can be profitably approached from many points of view.[1] In one familiar approach, illustrating an important method in algebraic geometry, we first divide by z^2 to obtain the equivalent $(x/z)^2 + (y/z)^2 = 1$, so we now seek rational solutions of $X^2 + Y^2 = 1$, or geometrically a **rational point** (a point with both coordinates rational) on the unit circle. Note that two nonzero solutions $(x : y : z)$ in integers yield the same solution (X, Y) in rationals if and only if they are proportional, so that by going from $x^2 + y^2 = z^2$ to $X^2 + Y^2 = 1$ we have automatically identified proportional Pythagorean triples (corresponding to similar right triangles). The unit vector $(1, 0)$ is an obvious rational point on the circle. This point yields only a degenerate Pythagorean triple, but we can use it to find any other rational point (X, Y) using the straight line through (X, Y) and $(1, 0)$. The general such line is $Y = -t(X - 1)$, where the slope $-t$ must be rational if X and Y are. (We choose $-t$ rather than t for consistency with equation (6.2).) Substituting $-t(X - 1)$ for Y in $X^2 + Y^2 = 1$ we get the quadratic equation $X^2 + t^2(X - 1)^2 = 1$, one of whose solutions must be $X = 1$. The other solution is then the root of

$$\frac{X^2 + t^2(X-1)^2 - 1}{X - 1} = (t^2 + 1)X - (t^2 - 1),$$

that is, $X = (t^2 - 1)/(t^2 + 1)$. Then $Y = -t(X - 1) = 2t/(t^2 + 1)$, so we have recovered the rational point corresponding to the solution $(t^2 - 1, 2t, t^2 + 1)$ of $x^2 + y^2 = z^2$. See Figure 1, which shows this construction for $t = 2$.

This procedure readily generalizes: instead of $X^2 + Y^2 - 1$ we can use any irreducible polynomial $P(X, Y)$ of degree 2, and instead of the initial point $(1, 0)$ we can use any rational solution (X_0, Y_0) of $P(X, Y) = 0$; the lines through (X_0, Y_0) not tangent to the curve $P(X, Y) = 0$ at that point then parametrize all other rational points on the curve. [Try $X^2 + Y^2 = 2$ and $X_0 = Y_0 = 1$. What goes wrong if we attempt this for $P(X, Y) = X^2 + Y^2$ and $X_0 = Y_0 = 0$? Note that $X^2 + Y^2$ is irreducible over the rationals, but not over $\mathbb{C}$ where it factors as $(X + iY)(X - iY)$.] The technique even works in some settings beyond plane curves of degree 2, including notably degree-3 plane curves with a double point; see Figure 2 for the example of the double point $(0, 0)$ on the curve $(X + Y)^3 = XY$. In our special case of $X^2 + Y^2 = 1$ and $(X_0, Y_0) = (1, 0)$ we can make yet another connection: if $(X, Y) = (\cos\theta, \sin\theta)$ then our line $Y = -t(X - 1)$ makes an angle of $\theta/2$ with the vertical. This can be seen by elementary plane geometry for $0 < \theta < \pi$, starting from the fact that $(0, 0)$, $(1, 0)$ and (X, Y) are vertices of an isosceles triangle (this too is shown in Figure 1); in general one must remember that θ is defined only up to integer multiples of 2π. In any case, this gives $t = \cot(\theta/2)$, so our parametrization is equivalent to the trigonometric half-angle formulas that give $\cot(\theta/2)$ as a rational function of $(\sin\theta, \cos\theta)$ and vice versa:

$$\cot\frac{\theta}{2} = \frac{\sin\theta}{1 - \cos\theta}; \qquad \cos\theta = \frac{\cot^2(\theta/2) - 1}{\cot^2(\theta/2) + 1}, \quad \sin\theta = \frac{2\cot(\theta/2)}{\cot^2(\theta/2) + 1}. \tag{6.3}$$

These formulas reappear in integral calculus in the guise of the universal substitution that converts $\int f(\sin\theta, \cos\theta)\, d\theta$ (where f is any rational function) into $\int F(t)\, dt$ for some rational function $F \in \mathbb{R}(t)$, which can then be expanded in partial fractions to obtain an elementary antiderivative. Equivalently this lets us integrate any rational function of X and $\sqrt{1 - X^2}$ with respect to X, and the generalization to quadratic $P(X, Y) = 0$ lets us replace $\sqrt{1 - X^2}$ by the square root of any quadratic polynomial.

[1] Besides the algebro-geometric method we follow, at least four others come to mind, which suggest various perspectives on and generalizations of the result. The most elementary may be to begin with the trigonometric identities (6.3), or with an equivalent geometric calculation with isosceles and right triangles. An elementary derivation from unique factorization in $\mathbb{Z}$ is obtained by removing common factors from (x, y, z), switching x, y if necessary to make x odd, and using the factorization $x^2 = z^2 - y^2 = (z - y)(z + y)$ and the fact that $\gcd(z - y, z + y) = 1$ to write $z \pm y = (m \pm n)^2$ for some coprime integers m, n. See for instance [IR, p.23, Exercise 12]. Alternatively, factor $z^2 = (x + iy)(x - iy)$ in the ring $\mathbb{Z}[i]$ of Gaussian integers, and use unique factorization in $\mathbb{Z}[i]$; this explains why x and y are the real and imaginary parts of $(m + in)^2$. Finally, for $X, Y \in \mathbb{Q}$ we have $X^2 + Y^2 = 1$ if and only if the element $X + iY$ of $\mathbb{Q}(i)$ has norm 1, which by Hilbert's Theorem 90 is equivalent to $X + iY = w/\overline{w}$ for some nonzero $w \in \mathbb{Q}(i)$. Taking $t = \mathrm{Re}(w)/\mathrm{Im}(w)$ we recover $X + iY = (t^2 - 1 + 2it)/(t^2 + 1)$. See [Ta].

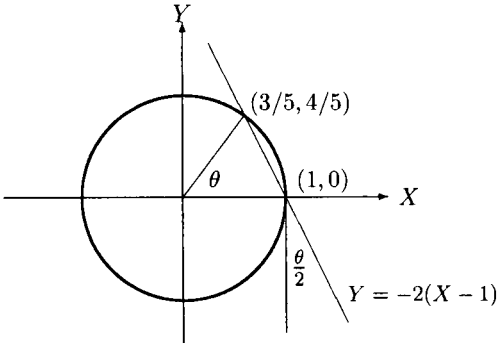


Figure 1: $X^2 + Y^2 = 1$

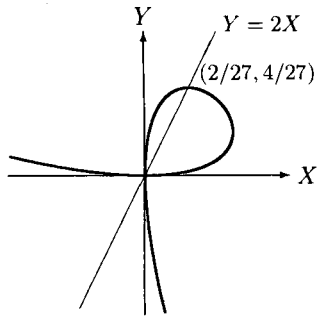


Figure 2: $(X + Y)^3 = XY$

But we have digressed from our main plot, to which we now return by looking at $x^2 + y^2 = z^2$ and the parametrization (6.1) or (6.2) from another point of view. We ask: *How many solutions does the Diophantine equation $x^2 + y^2 = z^2$ have in integer triples (x, y, z)?* Our parametrizations provide infinitely many (x, y, z) even when we identify proportional solutions, but we can still ask how common these solutions are. To make this vague question more precise, for all $N > 0$ define $C(N)$ to be the number of solutions of $x^2 + y^2 = z^2$ in integers such that x^2, y^2, z^2 are relatively prime and of absolute value at most N. (We give the condition on x, y, z in this form because of the way we intend to generalize it to other Diophantine equations, though of course for $x^2 + y^2 = z^2$ the absolute value condition is equivalent to the single inequality $z^2 \leq N$.) Then the existence of infinitely many non-proportional Pythagorean triples is equivalent to the fact that $C(N) \to \infty$ as $N \to \infty$, and we ask: *How quickly does $C(N)$ grow?*

Using either of the forms (6.1) and (6.2) of our parametrization of Pythagorean triples we see that $C(N)$ should grow as some multiple of $N^{1/2}$. For instance, (6.1) gives points (m, n) in the circle $m^2 + n^2 \leq N^{1/2}$, whose number is asymptotic to the area $\pi N^{1/2}$ of the circle. This is not quite right because we must count only relatively prime (m, n), and if both m and n are odd then we must remove a common factor of 2; but each of these corrections changes the asymptotic formula only by a constant factor. As it happens this factor is $2/(3\zeta(2)) = 4/\pi^2$, making $C(N) \sim (4/\pi)N^{1/2}$. But it is the exponent $1/2$ that concerns us here, and we could have guessed this exponent much more easily as follows. Let $A = x^2$, $B = y^2$, and $C = z^2$. Then

$$A + B = C,$$

and the number of solutions of $A + B = C$ in relatively prime integers in $[-N, N]$ is asymptotically proportional to N^2. Of the $2N + 1$ integers in $[-N, N]$, approximately $N^{1/2}$ are squares (and all but one are squares in two different ways, but this will not affect the exponent of N, only the coefficient of that power). So, if we pick A, B, C independently and uniformly at random from the integers in $[-N, N]$, the probability that all three will be squares is asymptotically proportional to $N^{-3/2}$. While we actually choose A, B, C not at random but subject to $A + B = C$, it seems a reasonable guess that the fraction of such (A, B, C) all of which are squares is still roughly $N^{-3/2}$, giving a total of roughly $N^{2-\frac{3}{2}} = N^{1/2}$ such triples in that range.

If you think this seems suspiciously easy, you are right: we are only guessing the correct answer (up to a constant factor), not proving it. This kind of heuristic is quite naïve, and can easily fail. For instance, for the equations $x^2 + y^2 + z^2 = 0$ or $x^2 + y^2 = 3z^2$ we might similarly expect the number of solutions with all three terms in $[-N, N]$ to grow at the same $N^{1/2}$ rate. But neither of these equations has any solution other than the trivial $(0, 0, 0)$: the first obviously so, because the terms x^2, y^2, z^2 are all nonnegative; and the second because after removing common factors from (x, y, z) we get a contradiction mod 3.[2] In the other direction, the heuristic might grossly underestimate the

[2] In fact these two obstructions are more similar than they might seem: $x^2 + y^2 + z^2 = 0$ has no nontrivial solution in the

number of solutions. Consider for example solutions in relatively prime integers of $(x+y)^3 = xyz$ (the homogeneous form of the curve $(X+Y)^3 = XY$ shown in Figure 2). We might expect very few solutions, on the grounds that there are about $8H^3$ triples (x,y,z) of integers in $[-H,H]$, and in that range $(x+y)^3 - xyz$ can be as large as a multiple of H^3, so should vanish with probability only c/H^3 for some $c > 0$, leaving a constant expected number of solutions no matter how large H is. Somewhat more reasonably, we could start with the number of solutions in $\max(|x|,|y|,|z|) \in (2^{h-1}, 2^h]$ and then sum over $h \le \log_2 H$; but even then we would guess that the number of solutions with $\max(|x|,|y|,|z|) \le H$ grows only logarithmically. But in fact the rational parametrization by lines through the origin shows that the correct order of growth is $H^{2/3}$. Here the failure of the naïve heuristic can be attributed to the singularity of our curve at the origin. In higher dimensions, examples are known where our heuristic fails for other, subtler reasons.

Still, such failures are not surprising. What is remarkable is how often such a naïve heuristic gives the correct answer when this answer can be established, and an answer consistent with or close to the predictions of more refined conjectures and heuristics when the correct answer is not known but the problem fits into a suitable mathematical framework. In the next few sections we illustrate this by successively generalizing the problem of solving $x^2 + y^2 = z^2$ until we reach the ABC conjecture.

6.2 Fermat's "Last Theorem" (FLT): $x^n + y^n = z^n$

Of the many fruitful generalizations of $x^2 + y^2 = z^2$, one of the most natural and by far the best known is the Fermat equation $x^n + y^n = z^n$ for $n \ge 2$. Again we seek solutions in nonzero integers, or equivalently solutions of $X^n + Y^n = 1$ in rational numbers $X = x/z$, $Y = y/z$. The locus of $X^n + Y^n = 1$ is known as the **n-th Fermat curve**; Figures 3 and 4 show part of the real locus for $n = 3$ and the entire real locus for $n = 4$, and are typical of the visual appearance (albeit not necessarily of the arithmetic or algebraic geometry) of Fermat curves with $n \ge 3$ odd or even respectively.

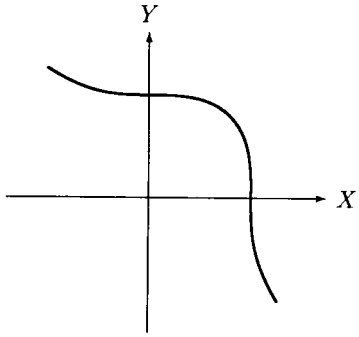


Figure 3: $X^3 + Y^3 = 1$

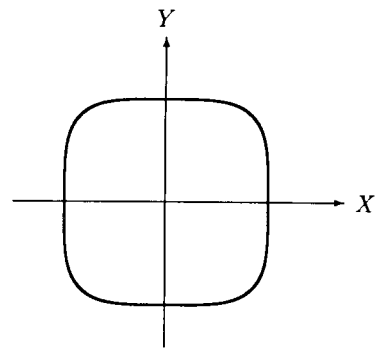


Figure 4: $X^4 + Y^4 = 1$

Fermat's "Last Theorem" (FLT) is the assertion, recorded by Fermat in 1637 and proved by him at least for $n = 4$, that for $n \ge 3$ there are no solutions of $x^n + y^n = z^n$ in nonzero integers; equivalently,

real field $\mathbb{R}$, and $x^2 + y^2 = 3z^2$ has no nontrivial solution in the field $\mathbb{Q}_3$ of 3-adic numbers. Since we live in the real world rather than the 3-adic world, the former obstruction is more intuitive to us, but both $\mathbb{R}$ and $\mathbb{Q}_3$ (and more generally $\mathbb{Q}_p$ for any prime p) are completions of $\mathbb{Q}$ with respect to the corresponding valuations on $\mathbb{Q}$, and decades of experience have shown the advantage of regarding the real and p-adic valuations of $\mathbb{Q}$ on as equal a footing as possible.

At this point we cannot resist another digression. Both $x^2 + y^2 + z^2 = 0$ and $x^2 + y^2 = 3z^2$ are obstructed not just over $\mathbb{R}$ and $\mathbb{Q}_3$ respectively, but also over $\mathbb{Q}_2$. It turns out that for *any* irreducible homogeneous quadratic $P(x,y,z)$ there are at most finitely many completions of $\mathbb{Q}$ in which there are no nonzero solutions of $P(x,y,z) = 0$, and that the number — call it ν — of such completions (either real or p-adic) is always even; this is equivalent to Quadratic Reciprocity. Conversely, any finite subset of $\{\mathbb{R}, \mathbb{Q}_2, \mathbb{Q}_3, \mathbb{Q}_5, \mathbb{Q}_7, \ldots\}$ of even size can arise this way, a fact that ultimately amounts to the determination of the 2-torsion of the Brauer group of $\mathbb{Q}$. Finally, if $\nu = 0$ then $P(x,y,z) = 0$ does in fact have nontrivial rational solutions; that is, the Hasse principle holds for homogeneous quadratics in three variables over $\mathbb{Q}$.

that the n-th Fermat curve has no rational points other than $(\pm 1, 0)$ and $(0, \pm 1)$ (with minus signs allowed only when n is even). Why should $n \geq 3$ behave so differently from $n = 2$? Let us consult our heuristic for estimating the expected number of solutions of $x^n + y^n = z^n$ with $\max(|x^n|, |y^n|, |z^n|) \in (N/2, N]$. (Every solution (x, y, z) will satisfy this condition with $N = 2^h$ for a unique nonnegative integer h.) As before we write $(A, B, C) = (x^n, y^n, z^n)$, and observe that $A + B = C$, and that the number of triples (A, B, C) of integers with $A + B = C$ and $\max(|A|, |B|, |C|) \in (N/2, N]$ is asymptotically proportional to N^2. But now we want each of them to be not a square but an n-th power for some $n \geq 3$, and n-th powers get rarer as n increases. Indeed the number of n-th powers in $[-N, N]$ grows only as $N^{1/n}$, so the probability that three integers A, B, C chosen independently and uniformly at random in that range are all n-th powers is asymptotically proportional to $N^{3((1/n)-1)}$. We thus expect roughly $N^{2+3((1/n)-1)} = N^{(3-n)/n}$ such triples with $A + B = C$. The exponent $(3 - n)/n$ is positive, zero, or negative according as $n < 3$, $n = 3$, or $n > 3$. Taking $N = 2^h$ and summing over h, we thus expect the solutions to be plentiful for $n < 3$ (the number of solutions up to N growing as a positive power of N), sparse for $n = 3$, and finite in number for $n > 3$. The same should be true of primitive[3] integral solutions of $A_0 x^n + B_0 y^n = C_0 z^n$ for any fixed choice of A_0, B_0, C_0, corresponding to rational points on the curve $A_0 X^n + B_0 Y^n = C_0$.

It turns out that each of these predictions is essentially correct. For $n = 1$ the result is almost trivial. For $n = 2$ we saw that, once the curve $A_0 X^2 + B_0 Y^2 = C_0$ has a rational point P, the lines through P yield the expected plenty of rational points on the curve. For $n \geq 3$ we must appeal to more advanced and recent results on Diophantine equations. When $n = 3$, the curve $E : A_0 X^3 + B_0 Y^3 = C_0$ is a nonsingular cubic plane curve, and thus an **elliptic curve** assuming it has a rational point P.[4] Here it is not so easy to get new rational points, because a typical line through P meets E at two more points, which in general are not rational. To obtain a new rational point we must use the line joining two rational points on E, or tangent to one rational point. This is shown in Figure 5 for the curve with $(A_0, B_0, C_0) = (1, 1, 91)$: the line through the rational points[5] $(3, 4)$ and $(6, -5)$ meets E again at $(9/2, -1/2)$, and the tangent at $(6, -5)$ meets E again at $(-204/341, 1535/341)$.

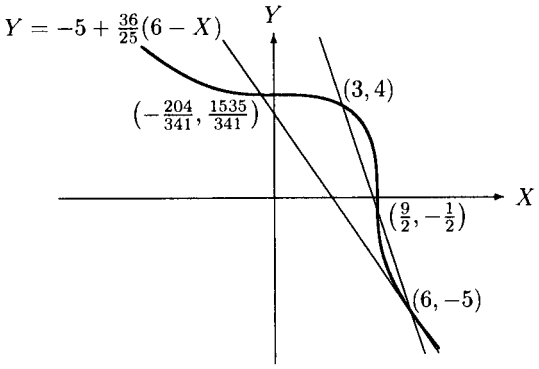


Figure 5: some rational points on $X^3 + Y^3 = 91$

[3] An integer solution (x, y, z) of a homogeneous polynomial equation $p(x, y, z) = 0$ is said to be **primitive** if $\gcd(x, y, z) = 1$. Every integer solution other than $(0, 0, 0)$ can be written uniquely as (kx, ky, kz) for some primitive solution (x, y, z) and some positive integer k.

[4] It is known that in characteristic zero such a curve is always isomorphic to one of the more familiar form $Y^2 = P_3(X)$ for some polynomial P_3 with distinct roots. See [Sil, Chapter III, §3] for such isomorphisms, and [Sil, Chapter III, §1] for standard formulas for elliptic curves.

[5] The value $C_0 = 91$ was chosen so that our curve has two simple rational points $(3, 4)$ and $(6, -5)$. This required a simple but nontrivial solution of $X^3 + Y^3 = X'^3 + Y'^3$. It would have been nice to use the famous "Ramanujan taxicab" example $C_0 = 1729 = 1^3 + 12^3 = 9^3 + 10^3$; but this would make it hard to draw a clear and accurate Figure 5, because $(1, 12)$ is too close to an inflection point of E and $(10, 9)$ too close to the middle of the curve. Our example with $(3, 4)$ and $(6, -5)$ relies instead on another famous identity $3^3 + 4^3 + 5^3 = 6^3$, which is tantalizingly reminiscent of $3^2 + 4^2 = 5^2$ but alas does not generalize further: $\sum_{m=3}^{n+2} m^n \neq (n + 3)^n$ once $n > 3$.

By drawing more lines and tangents we can generate infinitely many rational points on $X^3 + Y^3 = 91$, and it can be shown that every rational point can be obtained this way. As one might guess from the case of $(-204/341, 1535/341)$, the resulting primitive solutions of $x^3 + y^3 = 91z^3$ grow rapidly, and it turns out that the number of primitive solutions with all variables in $[-N, N]$ is asymptotic only to $R \log N$ for some $R > 0$. There are similar results for any elliptic curve E. By a famous theorem of Mordell [Mo] there is a finite list of rational points on E from which all other points can be recovered by repeatedly drawing chords and tangents through points already known or constructed. More precisely, Mordell uses the chords-and-tangents construction to give the set $E(\mathbb{Q})$ of rational points on E the structure of an abelian group,[6] and proves that this group is finitely generated. It then follows from the Néron-Tate theory of canonical heights that the number of rational points $(x/z, y/z)$ with each of x, y, z in $[-N, N]$ is asymptotic to $R(\log N)^{\rho/2}$, where ρ is the rank of the abelian group $E(\mathbb{Q})$ and R is a positive constant depending on E. The curve has finitely many rational points if and only if $\rho = 0$. It is known that this happens for the cubic Fermat curve $X^3 + Y^3 = 1$, whose only rational points are the obvious $(1, 0)$, $(0, 1)$, and the point at infinity $(X : Y : 1) = (1 : -1 : 0)$.

Finally, for $n > 3$ the curve $A_0X^n + B_0Y^n = C_0$ is a smooth plane curve of degree at least 4. Mordell conjectured that (as our heuristics suggest) every such curve has only finitely many rational points.

At any rate there is no longer a general method for constructing new points out of known ones; even the line through two known points, or tangent to one known point, meets the curve in $n - 2$ more points (allowing points with complex coordinates), and those points need not be rational once $n - 2 > 1$. For example, the line $X + Y = 1$ through the rational points $(X, Y) = (0, 1)$ and $(1, 0)$ on the Fermat quartic $X^4 + Y^4 = 1$ meets the curve again in a pair of Galois-conjugate points, each defined only over $\mathbb{Q}(\sqrt{-7})$, namely $\left(\frac{1}{2}(1 \pm \sqrt{-7}), \frac{1}{2}(1 \mp \sqrt{-7})\right)$. More generally, Mordell conjectured that any algebraic curve of genus at least 2 has only finitely many rational points. (The genus of a curve is a measure of its complexity[7]; an irreducible plane curve of degree d has genus $(d-1)(d-2)/2$ at most, with equality if and only if the curve is smooth; an elliptic curve has genus 1, and rationally parametrized curves have genus 0.) Mordell's conjecture was finally proved by Faltings, who gave two entirely different proofs [F1, F2]. Like Mordell's proof of the finite generation of $E(\mathbb{Q})$ for an elliptic curve E, both of Faltings' proofs are "ineffective": Mordell's proof yields an upper bound on the rank, and either of Faltings' proofs yields an upper bound on the number of rational points, but in general there may be no way to find a list of points and prove that it accounts for all the rational points on the curve. While much more is known now than at the time of Mordell's or even Faltings' proof, the general problems of making those theorems effective remain open.

A final note on Mordell's and Faltings' theorems: while they share the mystery of ineffectivity, the proofs are of quite a different flavor. Mordell's proof for elliptic curves can be traced back to Fermat's proof of the case $n = 4$ of FLT (showing in effect that the elliptic curves $Y^2 = X^4 \pm 1$ associated to the Diophantine equations $x^4 \pm y^4 = z^2$ have rank zero), and can be regarded as the culmination of Fermat's work in this direction. On the other hand, Faltings' proofs, together with the proof of FLT by Wiles and Taylor [Wil, TW], depend heavily on some of the most abstract and difficult results and techniques of late twentieth-century number theory; it would take an expository paper at least as long as this one to even give a sense of these methods to a reader not already acquainted with them.

6.3 The Darmon-Granville theorem: $x^p + y^q = z^r$

Another natural way to generalize the Fermat equation is to allow different exponents, changing $x^n + y^n = z^n$ to $x^p + y^q = z^r$. Here p, q, r are fixed positive integers that are not necessarily equal, and x, y, z are integer unknowns. Solving this equation is equivalent to solving $A + B = C$ under the

[6] While the chord-and-tangent method has been known at least since the time of Fermat, the construction of an abelian group law from it is not obvious. See [Si1, Chapter III, §2] for the details.

[7] At this point it is almost obligatory for an expository paper to cite the fact that an algebraic curve of genus g is one whose graph over $\mathbb{C}$ is an orientable surface with g holes; if nothing else, that is one indication that g measures the curve's complexity.

condition that A be a p-th power, B be a q-th power, and C be a r-th power. The Fermat equation with exponent n is the special case $p = q = r = n$. Applying our heuristic to general (x, y, z), we find that if A, B, C are random integers with $\max(|A|, |B|, |C|) \in (N/2, N]$ then they are respectively p-th, q-th, and r-th powers with probability asymptotically proportional to $N^{((1/p)-1)+((1/q)-1)+((1/r)-1)}$, and thus that of the roughly N^2 solutions of $A + B = C$ in that range we might expect about

$$N^{((1/p)-1)+((1/q)-1)+((1/r)-1)} N^2 = N^{((1/p)+(1/q)+(1/r)-1)}$$

to yield solutions of $x^p + y^q = z^r$. As before, the same analysis applies (to the extent we believe it) to the equation

$$A_0 x^p + B_0 y^q = C_0 z^r \tag{6.4}$$

for fixed nonzero A_0, B_0, C_0. This leads us to introduce

$$\delta = \delta(p, q, r) := 1 - \frac{1}{p} - \frac{1}{q} - \frac{1}{r}. \tag{6.5}$$

Our expected number of solutions with $\max(|A|, |B|, |C|) \in (N/2, N]$ is now roughly $N^{-\delta}$, and as before we vary N and expect the solutions to be plentiful, sparse, or bounded according as $\delta < 0$, $\delta = 0$, or $\delta > 0$. The corresponding values of (p, q, r) are as follows.

Exercise 6.3.1. We have $\delta(p, q, r) < 0$ if and only if one of the following conditions holds: the smallest of p, q, r equals 1; the two smallest of p, q, r both equal 2; or (p, q, r) is a permutation of $(2, 3, 3)$, $(2, 3, 4)$, or $(2, 3, 5)$. In this case, if $\min(p, q, r) = 2$ then $1/\delta$ is a negative integer. We have $\delta(p, q, r) = 0$ if and only if (p, q, r) is a permutation of $(3, 3, 3)$, $(2, 4, 4)$, or $(2, 3, 6)$. Otherwise $\delta(p, q, r) \geq 1/42$, with equality if and only if (p, q, r) is a permutation of $(2, 3, 7)$.

The new borderline cases $(2, 4, 4)$ and $(2, 3, 6)$ again yield elliptic curves, with equations $Y^2 = X^4 \pm 1$ and $Y^2 = X^3 \pm 1$ in the simplest case $A_0 = B_0 = C_0 = 1$. It so happens that again each of these elliptic curves has rank zero, and thus only finitely many rational points. For $Y^2 = X^4 \pm 1$ the only rational points not at infinity are obvious ones with $XY = 0$; this is equivalent to Fermat's result that there are no solutions of $x^4 \pm y^4 = z^2$ in nonzero integers. For $Y^2 = X^3 \pm 1$ there is one additional solution[8] $3^2 = 2^3 + 1$, giving rise to a single set of equivalent solutions of $x^2 + y^3 = z^6$ in nonzero integers, namely $(x, y) = (3z^3, -2z^2)$ for nonzero $z \in \mathbb{Z}$. For general A_0, B_0, C_0 there may be infinitely many such equivalence classes, but again their minimal representatives will be quite sparse, with the number of representatives in the range $\max(|A|, |B|, |C|) \leq N$ growing only as a multiple of $(\log(N)^{\rho/2})$ (where as before ρ is the rank of the corresponding elliptic curve).

But for general p, q, r our prediction can be very wide of the mark: there are cases where $\delta > 0$ but solutions are plentiful. For example, the equation $x^3 + y^4 = z^5$ has the solution

$$(x, y, z) = (209952, 11664, 1944) = (2^5 3^8, 2^4 3^6, 2^3 3^5), \tag{6.6}$$

with (A, B, C) proportional to $(1, 2, 3)$ — and indeed every integer solution of $A + B = C$ is proportional to (x^3, y^4, z^5) for some (and thus for infinitely many) integer triples (x, y, z). More generally we have:

Exercise 6.3.2. Suppose the natural numbers p, q, r are pairwise relatively prime, and A_0, B_0, C_0 are any nonzero integers. Then every integer solution of $A+B = C$ is proportional to $(A_0 x^p, B_0 y^q, C_0 z^r)$ for some (and thus for infinitely many) integer triples (x, y, z), and given the initial A, B, C (not all zero) the number of such (x, y, z) with $\max(|A_0 x^p|, |B_0 y^q|, |C_0 z^r|) \leq N$ is asymptotically proportional to $N^{1/(pqr)}$ as $N \to \infty$. Moreover there are triples (p, q, r) of relatively prime numbers for which δ is arbitrarily close to 1.

[8] The elliptic curve $Y^2 = X^3 + 1$ still has rank zero, but with six rational points: one at infinity, one with $X = -1$, and two each with $X = 0$ and $X = 2$. The reader can check that no further points are obtained by intersecting the curve with the tangent line at any of these points, or the line through any two of them. For instance, $(X, Y) = (2, 3)$ is the third point of intersection of $Y^2 = X^3 + 1$ with the line $Y = X + 1$ through the obvious points $(-1, 0)$ and $(0, 1)$.

The exponent $1/(pqr)$, though usually small, is positive for all p, q, r; hence if p, q, r are pairwise relatively prime our equation $A_0x^p + B_0y^q = C_0z^r$ has "plentiful solutions" by our standards, even when the value of δ is almost as positive as it can be. This seems to utterly demolish our heuristic, which suggests that when $\delta > 0$ there should be only finitely many solutions, and moreover that this tendency should be more pronounced the larger δ gets. But even in favorable cases like the "twisted Fermat curves" $A_0x^n + B_0y^n = C_0z^n$ our heuristic holds only for primitive solutions, those with x, y, z pairwise relatively prime. Indeed we should not expect the heuristic to hold when x and y have a large common factor, say d, because then $A = A_0x^n$ and $B = B_0y^n$ are both multiples of d^n, which makes $A + B$ much likelier to be of the form C_0z^n than a random number of the same size. Our construction of plentiful solutions such as (6.6) likewise exploits large common factors. We thus restrict attention to solutions with $(A, B, C) = (A_0x^p, B_0y^q, C_0z^r)$ *relatively prime.*[9,10] In this case our heuristic agrees precisely with the remarkable theorem of Darmon and Granville (1995):

Theorem 1. *[DG]: Let p, q, r be natural numbers such that $\delta(p, q, r) > 0$, and let A_0, B_0, C_0 be any nonzero integers. Then there are finitely many triples (x, y, z) of integers with $\gcd(x, y, z) = 1$ satisfying the equation (6.4).*

As with FLT and Faltings' theorem, the proof is alas much too advanced for us to be able to even outline the main ingredients here — though we do note that one key step is an application of Faltings' theorem itself!

Exercise 6.3.3. The Darmon-Granville theorem may seem a bit stronger than what we suggested, because (A, B, C) might still have a common factor coming from the coefficients A_0, B_0, C_0. But given those coefficients there are only finitely many possible values of $d := \gcd(A, B, C)$. Use this to show that there are also only finitely many equations $A_1x_1^p + B_1y_1^q = C_1z_1^r$ whose integer solutions satisfying $\gcd(A_1x_1^p, B_1y_1^q, C_1z_1^r) = 1$ account for all solutions of (6.4) with $\gcd(x, y, z) = 1$. Therefore if we knew Darmon-Granville only under the more restrictive hypothesis that A_0x^p, B_0y^q, C_0z^r be relatively prime, we could deduce the result in the form quoted above.

Seeing that the Darmon-Granville theorem for equation (6.4) generalizes Faltings' finiteness result for the case $p = q = r$ of twisted Fermat curves, can we also generalize FLT to the special case $A_0 = B_0 = C_0 = 1$ of (6.4), finding all solutions of $x^p + y^q = z^r$ with $\delta(p, q, r) > 0$ and $\gcd(x, y, z) = 1$? Our heuristic analysis suggests that there should be only finitely many such triples (x^p, y^q, z^r), but we have no reason to expect that there should be none at all — and we would not be surprised if some of them are quite large, especially for those choices of (p, q, r) that make δ positive but small. Note that the Darmon-Granville theorem gives finiteness for any particular choice of (p, q, r) but (like Faltings' theorem vis-a-vis FLT) leaves open the possibility of infinitely many solutions with (p, q, r) varying as well.

The full answer is still beyond reach, so we report on the known partial results and conjectures. The simplest example is the identity $1 + 8 = 9$ already noted in connection with $(p, q, r) = (2, 3, 6)$; it yields a solution $1^r + 2^3 = 3^2$ for all r, satisfying $\delta(2, 3, r) > 0$ for all $r > 6$. Computer searches reveal 9 more solutions: $13^2 + 7^3 = 2^9$ with $\delta(2, 3, 9) = 1/18$; two solutions

$$2^5 + 7^2 = 3^4, \quad 3^5 + 11^4 = 122^2 \tag{6.7}$$

with $\{p, q, r\} = \{2, 4, 5\}$ and $\delta = 1/20$; two solutions

$$33^8 + 1549034^2 = 15613^3, \quad 43^8 + 96222^3 = 30042907^2 \tag{6.8}$$

[9]We need not specify *pairwise* relatively prime, because the relation $A + B = C$ forces any factor of two of A, B, C to divide the third.

[10]The failure of our naïve heuristic when A, B, C can have large common factors is related to the failure we noted earlier for a singular cubic curve. Here the *surface* $A_0x^p + B_0y^q = C_0z^r$ is highly singular at the origin, and a solution with A, B, C all divisible by a high power of p yields a point (x, y, z) on that surface that is close to that singularity in the p-adic metric.

with $\{p,q,r\} = \{2,3,8\}$ and $\delta = 1/24$; and four solutions

$$\begin{aligned} 2^7 + 17^3 &= 71^2, \quad 17^7 + 76271^3 = 21063928^2, \\ 1414^3 + 2213459^2 &= 65^7, \quad 9262^3 + 15312283^2 = 113^7 \end{aligned} \tag{6.9}$$

with $\{p,q,r\} = \{2,3,7\}$ and the minimal δ value of $1/42$. These computations are reported in [DG], with the discovery of the five large solutions credited to Beukers and Zagier. This list is conjectured to be complete, based both on further computer searches that revealed no other solutions and on various partial results that prove special cases of the conjecture. In particular it would follow from this conjecture (plus FLT for $n = 3$) that $x^p + y^q = z^r$ has no solution in integers $p,q,r \geq 3$ and relatively prime integers x,y,z; this is the **Tijdeman-Zagier conjecture**, for whose solution Andrew Beal offers a $50,000 prize [Mau].

The most recent of the partial results in the direction of the conjecture that there are no further solutions with $\delta(p,q,r) > 0$ is [PSS], a *tour de force* proving that there are no further solutions for $\{p,q,r\} = \{2,3,7\}$. This paper also gives an extensive list (Table 1 at the end of the Introduction) of triples (p,q,r) for which the corresponding result had been proved earlier, including the triples with $\{p,q,r\} = \{2,4,5\}$ and $\{2,3,8\}$ seen in the other known solutions (6.7, 6.8). Another special case is Catalan's conjecture that 8 and 9 are the only consecutive powers of integers, recently proved by Mihăilescu [Mi]; this shows that there are no other solutions with $x = 1$. The proofs of these partial results call on a vast range of number-theoretical techniques, including classical methods of elementary, algebraic, and analytic number theory, Galois representations and modularity as in the proof of FLT, and algebraic geometry of curves. This huge theoretical arsenal is complemented by sophisticated computational and algorithmic tools that are often essential for carrying out algebraic manipulations or for completing a proof that has been reduced to a finite but nontrivial calculation.

What about $\delta(p,q,r) < 0$, when we expect that the number of relatively prime solutions of (6.4) with $\max(|A|,|B|,|C|) \leq N$ can grow as a multiple of $N^{-\delta}$ as $N \to \infty$? We easily dispose of the case where at least one of p,q,r is 1, when we can simply solve (6.4) for the corresponding variable x, y, or z in terms of the other two. So we assume that each of p,q,r is at least 2. In Exercise 6.3.1, we saw that then $-\delta = 1/d$ for some integer $d > 0$. There are choices of the coefficients A_0, B_0, C_0 for which (6.4) has no solutions at all — we already saw the examples $x^2 + y^2 + z^2 = 0$ and $x^2 + y^2 = 3z^2$ with $p = q = r = 2$. But if we assume that there is at least one solution of (6.4) in relatively prime integers then Beukers showed [Beu] that the $N^{1/d}$ guess is correct. Moreover, for each A_0, B_0, C_0 there are finitely many polynomial identities in degree $2d$ that together account for all the relatively prime solutions, in the same way that the single identity (6.2) accounts for all Pythagorean triples. (In fact the Pythagorean parametrization illustrates the special case $A_0 = B_0 = C_0 = 1$, $p = q = r = 2$ of Beukers' result; note that here $\delta = -1/2$ and both sides of the identity are polynomials of degree 4.)

Unlike the Faltings and Darmon-Granville finiteness results, Beukers' is effective: at least in principle one can find all the parametrizations by carrying out a computation whose length is bounded by an explicit function of p,q,r,A_0,B_0,C_0. Doing this in practice still requires some work. For the three exceptional cases where only one of p,q,r equals 2, this work was recently completed by Edwards [Ed]. In particular he gave for the first time the complete solution for $\{p,q,r\} = \{2,3,5\}$ in the case $A_0 = B_0 = C_0 = 1$. There are 27 inequivalent identities, of which the simplest (which Beukers had already obtained) is $X(t)^2 + Y(t)^3 = Z(t)^5$ where

$$\begin{aligned} X(t) &= (t^{10} + 12^4)\,(t^{20} - 12^2 522\,t^{15} - 12^4 10006\,t^{10} + 12^6 522\,t^5 + 12^8), \\ Y(t) &= -t^{20} - 12^2 228\,t^{15} - 12^4 494\,t^{10} + 12^6 228\,t^5 - 12^8, \\ Z(t) &= 12(-t^{11} + 12^2 11\,t^6 + 12^4\,t). \end{aligned} \tag{6.10}$$

For any $m,n \in \mathbb{Z}$ we recover an integer solution of $x^2 + y^3 = z^5$ by taking $x = n^{30}X(m/n)$, $y = n^{20}Y(m/n)$, and $z = n^{12}Z(m/n)$, and these x,y,z are relatively prime if and only if

$$\gcd(m,6n) = 1 \quad \text{and} \quad m \not\equiv 2n \bmod 5. \tag{6.11}$$

For example, $m = n = 1$ yields $36934790165857^2 + 240546239^3 = 267828^5$. To make it such that $\max(|x^2|, |y^3|, |z^5|)$ less than N it is enough to make both $|m|$ and $|n|$ less than some multiple of $N^{1/60}$; the number of such (m, n) satisfying (6.11) is asymptotically proportional to $N^{1/30} = N^{-\delta(2,3,5)}$ as expected.

We conclude this section with another scenic detour: a view of two surprisingly pertinent alternative descriptions of the triples (p, q, r) of integers greater than 1 for which $\delta(p, q, r) < 0$. First, p, q, r satisfy this condition if and only if there exists a spherical triangle Δ with angles π/p, π/q, π/r on the unit sphere Σ, in which case the triangle has area $\pi \cdot (-\delta)$. Second, we have $\delta(p, q, r) < 0$ if and only if the group $\Gamma = \Gamma_{p,q,r}$ with the presentation

$$\Gamma_{p,q,r} := \langle \alpha, \beta, \gamma \mid \alpha^p = \beta^q = \gamma^r = \alpha\beta\gamma = 1 \rangle \tag{6.12}$$

is finite, in which case it has $2d$ elements, where $d = -1/\delta$ as before. The first equivalence follows from the well-known fact that the sum of the angles of Δ exceeds π by an amount equal to the area of Δ. In this case we can take the generators α, β, γ of Γ to be rotations about the vertices of Δ through angles $2\pi/p$, $2\pi/q$, $2\pi/r$, or equivalently the products of pairs of reflections in the edges of Δ. If we identify Σ with the Riemann sphere $\mathbb{CP}^1$ and let t be a complex coordinate on Σ then Γ becomes a finite group of automorphisms of $\mathbb{CP}^1$, which is to say a finite group of fractional linear transformations $t \mapsto (at + b)/(a't + b')$. Then for each of our identities $X(t)^p + Y(t)^q = Z(t)^r$ in degree $2d$ the ratios X^p/Z^r, Y^q/Z^r, etc. are invariant under Γ for a suitable choice of spherical triangle Δ! Moreover, by Galois theory any such ratio T actually *generates* the field of Γ-invariant rational functions of t; that is, $\mathbb{C}(T) = (\mathbb{C}(t))^\Gamma$. For example, when $p = q = r = 2$ our Pythagorean parametrization (6.2) yields functions such as $(t^2 - 1)^2/(2t)^2$ and $(t^2 + 1)^2/(2t)^2$ invariant under the 4-element group isomorphic with $\Gamma_{2,2,2}$ and generated by $t \leftrightarrow -t$ and $t \leftrightarrow 1/t$. For $(p, q, r) = (2, 3, 5)$, we have $\Gamma \cong A_5$, the group of rotational symmetries of a regular icosahedron inscribed in Σ, and the roots of the polynomials[11] X, Y, Z of (6.10) are precisely the 30 edge centers, 20 face centers, and 12 vertices of that icosahedron!

When $\delta(p, q, r) = 0$ or $\delta(p, q, r) > 0$ the triangle Δ is respectively planar or hyperbolic rather than spherical, and the group $\Gamma = \Gamma_{p,q,r}$ generated by pairs of reflections in its edges is no longer finite. But Γ is still intimately connected with $x^p + y^q = z^r$ via automorphisms of Riemann surfaces. When $\delta(p, q, r) = 0$, we can regard Γ as a group of affine linear transformations $t \mapsto at + b$ of $\mathbb{C}$; its finite-index subgroup of translations (with $a = 1$) is then a lattice, and the quotient of $\mathbb{C}$ by this lattice is the elliptic curve we obtained from $x^p + y^q = z^r$. When $\delta = \delta(p, q, r)$ is positive, Δ is a hyperbolic triangle of area $\pi\delta$ and Γ is a discrete group of isometries of the hyperbolic plane $\mathcal{H}$; the quotient $\mathcal{H}/\Gamma$ can be identified with $\mathbb{CP}^1$ via a Γ-invariant meromorphic function on $\mathcal{H}$ analogous to the functions T of the previous paragraph, and quotients of $\mathcal{H}$ by subgroups of finite index in Γ yield finite extensions of $\mathbb{C}(T)$ that are used in the proof of the Darmon-Granville theorem and in the solution of some special cases such as $x^2 + y^3 = z^7$.

6.4 The ABC conjecture: $A + B = C$

Masser and Oesterlé noted that a solution of the Fermat equation, or of a natural generalization such as the equation (6.4) addressed by Darmon and Granville, yields relatively prime numbers A, B, C (such as x^n, y^n, z^n for a primitive Fermat solution) such that $A + B = C$ and each of A, B, C has many repeated prime factors. This inspired them to guess a vastly more general constraint on repeated prime factors in $A, B, A + B$ for coprime integers A, B, and to formulate a precise conjecture on the nature of this constraint, now known as the **ABC conjecture**. This conjecture is stated in terms of an arithmetic function called (for reasons whose explanation would take us too far afield here) the "conductor", defined as follows:

[11] Note that X, Y, Z are regarded as homogeneous polynomials of degrees 30, 20, and 12 respectively, so we count $t = \infty$ among the roots of Z. The other roots of Z are 0 and the ten values of $\varrho\phi$ where ϕ is $(1 \pm \sqrt{5})/2$ (the golden ratio or its algebraic conjugate) and ϱ is one of the five fifth roots of 12^2 in $\mathbb{C}$.

Definition 2. The **conductor** $N(D)$ of a nonzero integer D is the product of the (positive) primes dividing D, counted *without* multiplicity. Equivalently, $N(D)$ is the largest squarefree factor of N.

Example 6.4.1. $N(D_1D_2) \le N(D_1)N(D_2)$ for all nonzero integers D_1, D_2, with equality if and only if they are relatively prime; $N(D^n) = N(D)$ for all nonzero integers D and $n \ge 1$. The following brief table gives $N(D)$ for $24 \le D \le 32$:

D	24	25	26	27	28	29	30	31	32
$N(D)$	6	5	26	3	14	29	30	31	2

The size of the integer $|D|/N(D)$ should be regarded a measure of how far D is from being squarefree, that is, of how rich D is in repeated prime factors.

Conjecture 3. *(Masser-Oesterlé [Oe]): For every real $\epsilon > 0$ there exists $c_\epsilon > 0$ such that*

$$N(ABC) > c_\epsilon C^{1-\epsilon} \tag{6.13}$$

holds for all relatively prime natural numbers A, B, C such that $A + B = C$; equivalently, for every real $\epsilon > 0$ there exists $c_\epsilon > 0$ such that

$$N(ABC) > c_\epsilon \max(|A|, |B|, |C|)^{1-\epsilon} \tag{6.14}$$

holds for all relatively prime integers A, B, C such that $\pm A \pm B \pm C = 0$.

The equivalence is elementary, and the more symmetrical form $\pm A \pm B \pm C = 0$ will let us avoid repeating some arguments twice or thrice according to the signs of A, B, C.

In the following exercises, we detail how the ABC conjecture implies "asymptotic FLT" (that is, FLT for sufficiently large n) as well as its generalizations by Darmon-Granville and Tijdeman-Zagier, and then give an equivalent formulation in terms of the "ABC exponent", and explain why the ϵ in (6.13,6.14) cannot be removed.

Exercise 6.4.1. The ABC conjecture applied to $(A, B, C) = (A_0x^p, B_0y^q, C_0z^r)$ implies the Darmon-Granville theorem; moreover, for any p, q, r such that $\delta = \delta(p, q, r) > 0$ and any positive $\epsilon < \delta$, the inequality (6.13) with an explicit value of c_ϵ yields an explicit upper bound on relatively prime integers x, y, z such that $A_0x^p + B_0y^q = C_0z^r$.

Exercise 6.4.2. The ABC conjecture implies the Tijdeman-Zagier conjecture with at most finitely many exceptions; moreover, for any positive $\epsilon < 1/12$ the inequality (6.13) with an explicit value of c_ϵ yields an explicit upper bound on x^p, y^q, z^r in any counterexample to the conjecture.[12]

Exercise 6.4.3. The ABC conjecture for *any* $\epsilon < 1$ implies that Fermat's Last Theorem holds for all but finitely many exponents n. Again, an explicit value of c_ϵ yields an explicit n_0 such that FLT holds for all $n \ge n_0$.

Exercise 6.4.4. The ABC conjecture for *any* $\epsilon < 1$ implies that any finitely generated multiplicative subgroup G of $\mathbb{Q}^*$ contains only finitely many solutions (s, s') of $s + s' = 1$. [Choose generators for G, and let S be the set of primes that divide the numerator or denominator of at least one generator; then $s + s' = 1$ yields $A + B = C$ with $N(ABC) \mid \prod_{p \in S} p$.]

Remark. For this problem, as with the first exercise in this list, the finitude of solutions is already a theorem, without assuming ABC or any other unproved conjecture. Better yet, explicit upper bounds have been given on C as a function of $N(ABC)$ — whereas no such bound is known for the Darmon-Granville theorem without an ABC hypothesis. Still, the proved bounds are much worse than what would follow from (6.13); see below.

[12] The bound $1/12$ can be raised to $1/6$ because Bruin showed [Br] that there are no solutions of $x^3+y^3 = z^4$ or $x^3+y^3 = z^5$ in relatively prime integers.

Exercise 6.4.5. For relatively prime natural numbers A, B, C such that $A + B = C$, define the **ABC exponent** $\theta(A, B, C)$ by

$$\theta(A, B, C) := (\log C)/(\log N(ABC))$$

(so that $C = N(ABC)^{\theta(A,B,C)}$); for example $\theta(1, 8, 9) = \log 9/\log 6 = 1.226+$. Set

$$\Theta := \limsup_{(A,B,C)} \theta(A, B, C),$$

the limsup running over all triples $(A, B, A + B)$ of natural numbers. Then the ABC conjecture is equivalent to $\Theta \leq 1$. In fact the ABC conjecture is equivalent to $\Theta = 1$, because it is elementary that $\Theta \geq 1$ (for instance we may take $(A, B) = (1, 2^r - 1)$ with $r \to \infty$).

Remark. If it is true that $\limsup \theta(A, B, C) = 1$ then the convergence must be very slow: it is known that there are infinitely many examples of $\theta(A, B, C) > 1 + c/\sqrt{\log C}$ for some universal constant $c > 0$; and it is expected, based on probabilistic heuristics such as applied earlier to $A_0x^p + B_0y^q = C_0z^r$, that in fact $\theta(A, B, C) - 1 > (\log C)^{-\vartheta}$ holds infinitely often for all $\vartheta > 1/3$, but only finitely often for each $\vartheta < 1/3$. In particular, the ABC conjecture is consistent with those heuristics. The largest numerical value known for $\theta(A, B, C)$ is 1.6299+, for $2 + 3^{10}109 = 23^5$ (found by Eric Reyssat in 1987). See [Ni] for other large $\theta(A, B, C)$.

Exercise 6.4.6. The inequality (6.13) cannot hold for $\epsilon = 0$ and any positive value of c_0. (One way to prove this is to find for each $\alpha > 0$ a natural number r such that $3^\alpha | 2^r - 1$.)

The ABC conjecture, like FLT, is formulated over $\mathbb{Z}$ but has an equivalent statement over $\mathbb{Q}$ obtained by considering ratios of the variables. If $A + B = C$, consider $F = A/C$, so $1 - F = B/C$. Both fractions are in lowest terms because A, B, C are assumed relatively prime. The conductor $N(A)$ is the product of the primes p such that $F \equiv 0 \bmod p$, and likewise $N(B)$ is the product of the primes p such that $F \equiv 1 \bmod p$. As for $N(C)$, that is the product of primes p for which $F \bmod p$ cannot be found in $\mathbb{Z}/p\mathbb{Z}$ because the denominator C vanishes mod p. Since in this case $p \nmid A$, we say that these are the primes such that "$F \equiv \infty \bmod p$". Hence $N(ABC)$, the LHS of the ABC conjecture (6.13), is the product of primes p such that $F \bmod p$ is one of $0, 1, \infty$. The RHS is $c_\epsilon C^{1-\epsilon}$, in which C is simply the denominator of F. This assumes that A, B, C are positive, that is, that $0 < F < 1$; in the general case we replaced C by $\max(|A|, |B|, |C|)$ (see (6.14)), so now we replace the denominator of F by the **height** $h(F)$. By definition, the height of a rational number m/n with $\gcd(m, n) = 1$ is $\max(|m|, |n|)$. This need not exactly equal $\max(|A|, |B|, |C|)$, but is within a factor of 2, which can be accommodated by changing the constant c_ϵ of (6.14). Thus the ABC conjecture is equivalent to the assertion that for every $\epsilon > 0$ there exists $c_\epsilon > 0$ such that, for all $F \in \mathbb{Q}$, the product of the primes at which F reduces to 0, 1, or ∞ is at least $c_\epsilon h(F)^{1-\epsilon}$.

Geometrically, the reduction of FLT to ABC in Exercise 6.4.3 amounts to applying the ABC conjecture to the value of the rational function $F = (x/z)^n = X^n$ on the n-th Fermat curve. This succeeds because F and $1 - F$ have multiple poles and zeros (some defined only over an algebraic closure $\bar{\mathbb{Q}}$) — that is, the preimages of $0, 1, \infty$ under F have large multiplicities, which makes the total number of preimages counted *without* multiplicity small compared to the degree of F as a rational function on the curve. It turns out that here the degree is n^2, and the number of preimages is $3n$, which is less than n^2 once $n > 3$, and indeed less than δn^2 once $n > 3/\delta$. When we try to generalize this argument to rational points on a general algebraic curve $\mathcal{X}$, we find that it is rare for there to be a rational function F on $\mathcal{X}$ whose degree exceeds the size of $F^{-1}(\{0, 1, \infty\})$ by a large factor, so we cannot usually expect to deduce Mordell's conjecture (finiteness of rational points) for $\mathcal{X}$ from an ABC inequality with ϵ near 1. But Belyi [Bel] shows how to construct a function F satisfying $\deg(F) > \#\left(F^{-1}(\{0, 1, \infty\})\right)$ whenever $\mathcal{X}$ is a curve of genus at least 2 defined by an equation with coefficients in $\bar{\mathbb{Q}}$, and then Mordell's conjecture follows from ABC with ϵ sufficiently small [El1]. Recall that Faltings already proved this conjecture twice without any unproved hypothesis, but the proofs are ineffective; the argument in [El1]

shows that the ABC conjecture with effective constants c_ϵ would yield a completely effective finiteness result for rational points on $\mathcal{X}$.

Many other consequences of the ABC conjecture are known, ranging from elementary special cases (there are only finitely many integers N such as $N = 4, 5, 7$ for which $N! + 1$ is a perfect square) to applications that give unexpected connections with other problems in number theory. A striking example is Silverman's application to Wieferich primes, that is, primes p for which $2^{p-1} \equiv 1 \bmod p^2$, such as 1093 and 3511. (Note that the congruence always holds mod p by Fermat's little theorem. In 1909 Wieferich proved [Wie] that a FLT counterexample $x^p + y^p = z^p$ with $p \nmid xyz$ for some prime p would imply $2^{p-1} \equiv 1 \bmod p^2$.) Such primes are expected to be very rare; indeed none is known other than 1093 and 3511, and any further such prime must exceed $1.25 \cdot 10^{15}$ according to computations reported by Richard McIntosh (`http://www.loria.fr/~zimmerma/records/Wieferich.status`). But it is not even known that the set of *non*-Wieferich primes is infinite! Silverman [Si2] proves the infinitude of non-Wieferich primes under the hypothesis of the ABC conjecture, and shows further that this conjecture implies that for every integer $\alpha \neq 0, \pm 1$ there exist constants c_α, x_α such that for all $x > x_\alpha$ there are at least $c_\alpha \log x_\alpha$ primes $p < x$ satisfying $\alpha^{p-1} \not\equiv 1 \bmod p^2$.

Unfortunately a proof of the ABC conjecture still seems a very distant prospect; it is even much too hard to prove the existence of any $\epsilon < 1$ for which the inequality (6.13) holds for some $c_\epsilon > 0$. To show just how far we are, consider the situation suggested by Exercise 6.4.4: we know $N = N(ABC)$, and want all possible (A, B, C). Let S be the set of primes dividing N. Then the inequality (6.13) for any $\epsilon < 1$ gives an upper bound on solutions of $A + B = C$ in relatively prime integers all of whose prime factors are contained in S. (This is often called the "S-unit equation", because it is equivalent to solving $a + b = 1$ in rational numbers $(a, b) = (A/C, B/C)$ that are units in the ring $\mathbb{Z}[1/N]$ obtained from $\mathbb{Z}$ by inverting all the primes in S.) In particular, there should be only a finite number of solutions. This result is known [La1], but already far from trivial. It was not much harder to give an explicit bound on the *number* of solutions [LM], and by now there are bounds depending only on the size of S, as in [Ev]. But that still gives no control over the *size* of the largest solution, which is what the ABC conjecture addresses. Stewart and Tijdeman gave such a bound in [ST], and the bound was recently improved by Stewart and Yu [SY]. But even the best bounds remain exponential: the *logarithm* of C is only known to be bounded by a multiple of $N^{1/3}(\log(N))^3$. Even these results can be useful; for instance the Stewart-Tijdeman bound $\log C = O(N^{15})$ is already enough to compute in practice the full solution of the S-unit equation when S is not too large (see for instance [dW]). But the known results are still very weak compared with the inequalities that the ABC conjecture predicts and that we need for applications such as the Tijdeman-Zagier conjecture and explicit upper bounds in the Darmon-Granville theorem.

6.5 Mason's theorem: $\mathsf{A}(t) + \mathsf{B}(t) = \mathsf{C}(t)$

A curious feature of the ABC conjecture is that not only does it seem very hard to prove but it is not at all obvious how one might try to *dis*prove the conjecture. If FLT were false, a single counterexample would expose the falsity; likewise for the Catalan and Tijdeman-Zagier conjectures, or the Riemann hypothesis and its variants. But there can be no single counterexample for the ABC conjecture, even for a specific value of ϵ, because the inequality (6.13) can accommodate any given triple (A, B, C) by simply decreasing c_ϵ. Likewise for the formulation of the conjecture in terms of ABC exponents $\theta(A, B, C)$: a single example may break the record for the maximal θ, but has no bearing on Θ which is defined as a lim sup of $\theta(A, B, C)$. Proving that the conjecture is false would require the existence of an *infinite family* of (A, B, C)'s whose ABC exponents approach a limit greater than 1 (or approach ∞), just as we had to construct an infinite family such as $\{(1, 2^r - 1, 2^r)\}_{r=1}^{\infty}$ to prove $\Theta \geq 1$.

The most natural families to try arise from identities $\mathsf{A}(t) + \mathsf{B}(t) = \mathsf{C}(t)$ relating polynomials $\mathsf{A}, \mathsf{B}, \mathsf{C} \in \mathbb{Z}[t]$, not all constant. Recall that we already used such polynomials to construct infinitely many Pythagorean triples, or relatively prime solutions of $x^2 + y^3 = z^5$; in effect we solved these

Diophantine equations in $\mathbb{Z}[t]$, then specialized to $t \in \mathbb{Q}$ and multiplied by powers of the denominator of t to recover integer solutions. Similarly, from polynomials $\mathsf{A}(t), \mathsf{B}(t), \mathsf{C}(t)$ satisfying $\mathsf{A}+\mathsf{B}=\mathsf{C}$ for which $D := \max(\deg(\mathsf{A}), \deg(\mathsf{B}), \deg(\mathsf{C}))$ is positive we get a family of integer solutions A, B, C as follows: for any pair (m, n) of relatively prime integers we take

$$(A, B, C) = n^D(\mathsf{A}(m/n), \mathsf{B}(m/n), \mathsf{C}(m/n)). \tag{6.15}$$

Thus A, B, C are homogeneous polynomials of degree D in (m, n). If $\mathsf{A}, \mathsf{B}, \mathsf{C}$ have repeated factors then so do A, B, C, and with enough repeated factors we can hope to get a sequence with

$$\limsup \theta(A, B, C) > 1.$$

We must assume that $\mathsf{A}(t), \mathsf{B}(t), \mathsf{C}(t)$ are relatively prime as polynomials, else A, B, C will have a common factor for most choices of (m, n). This also means that D is the degree of the quotient $\mathsf{F} = \mathsf{A}/\mathsf{C} \in \mathbb{Q}(t)$ as a rational function of t. Conversely, if the polynomials have no common factors then $\gcd(A, B)$ is bounded above,[13] so dividing each of our triples (A, B, C) of (6.15) by its greatest common divisor yields relatively prime solutions of $A + B = C$ with asymptotically the same ABC exponent as the ratio

$$\frac{\log \max(|A|, |B|, |C|)}{\log N(ABC)} = \frac{\log \max(|A|, |B|, |C|)}{\log(N(A)N(B)N(C))} \tag{6.16}$$

that we would compute if A, B, C were relatively prime.

The numerator of this ratio is easy to estimate: it is $D \log h(m, n) + e$, where

$$D = \max(\deg(\mathsf{A}), \deg(\mathsf{B}), \deg(\mathsf{C}))$$

as above, $h(m, n)$ is the height $|\max(m, n)|$ of (m, n) (or of the rational number m/n as before), and e is an error of bounded absolute value. What of the denominator? Let us try some examples using polynomial identities that we have already encountered. If

$$(\mathsf{A}, \mathsf{B}, \mathsf{C}) = \big((t^2-1)^2, (2t)^2, (t^2+1)^2\big)$$

as in (6.2), then $D = 4$ and we get $(A, B, C) = ((m^2-n^2)^2, (2mn)^2, (m^2+n^2)^2)$ (the squares of the entries of the Pythagorean triple (6.2)), and then $N(ABC)$ is a factor of $(m^2-n^2)2mn(m^2+n^2)$. Hence $N(ABC)$ is bounded above by a multiple of $h(m, n)^6$. We can save two factors of $h(m, n)$ in various ways, for instance by making $(m, n) = (1, 2^r)$ as in Exercise 6.4.5; but that still leaves both numerator and denominator of (6.16) asymptotic to $4 \log h(m, n)$, giving the same lower bound of 1 on Θ that we obtained in that Exercise. Might we do better with the more complicated example $(A, B, C) = (X(t)^2, Y(t)^3, Z(t)^5)$, where X, Y, Z are the polynomials of (6.10)? Now $D = 60$ and A, B, C are respectively a square, a cube, and a fifth power, so $N(ABC)$ is bounded by a multiple of $h(m, n)^{30+20+12} = h(m, n)^{62}$. Again we can save a factor $h(m, n)^2$ thanks to the factor mn of C, but that still brings our bound on $N(ABC)$ only down to a multiple of $h(m, n)^{60} = h(m, n)^D$, and again we fail to improve on $\Theta \geq 1$.

In general, suppose A factors as $\mathsf{A}_0 \prod_i \mathsf{x}_i^{e_i}$ where A_0 is a scalar and the x_i are distinct irreducible polynomials. Let $x_i = n^{\deg \mathsf{x}_i} \mathsf{x}_i(m/n)$. Then $A = n^D \mathsf{A}(m/n) = \mathsf{A}_0 n^{e_\infty} \prod_i x_i^{e_i}$, where $e_\infty := D - \sum_i e_i \deg(\mathsf{x}_i)$ is the multiplicity of n as a factor of the homogeneous polynomial $n^D \mathsf{A}(m/n)$ (which may also be regarded as the "order of vanishing at $t = \infty$" of A when A is regarded as a polynomial of degree D). Hence $N(A)$ is bounded by a constant multiple of $\prod_i x_i$ or $n \prod_i |x_i|$ according as $e_\infty = 0$ or $e_\infty > 0$. Each $|x_i|$ is in turn bounded by a constant multiple of $(h(m, n))^{\deg \mathsf{x}_i}$, and of course $|n| \leq h(m, n)$. It follows that $N(A) \leq h(m, n)^{\nu_D(A)}$ where $\nu_D(A) = \nu_{D,\infty}(A) + \sum_i \deg \mathsf{x}_i$

[13] By the Euclidean algorithm for polynomials there exist $\mathsf{X}, \mathsf{Y} \in \mathbb{Z}[t]$ such that $\mathsf{AX} - \mathsf{BY} = d$ for some nonzero $d \in \mathbb{Z}$, and then $\gcd(A, B) \mid n^D d$ for all $m, n \in \mathbb{Z}$. Repeating this argument with A, B replaced by the relatively prime polynomials $t^D \mathsf{A}(1/t), t^D \mathsf{B}(1/t)$ yields a nonzero integer d' such that $\gcd(A, B) \mid m^D d'$. Thus if $\gcd(m, n) = 1$ then $\gcd(A, B) \mid dd'$.

and $\nu_{D,\infty}(A) = 0$ or 1 according as $e_\infty = 0$ or $e_\infty > 0$. More succinctly, $\nu_D(A)$ is the number of solutions of $\mathsf{F}(t) = 0$ in $\mathbb{CP}^1$, counted *without multiplicity* (note in particular that $e_\infty > 0$ if and only if $\mathsf{F}(\infty) = 0$). We define $\nu_D(B)$ and $\nu_D(C)$ likewise, and observe that they are the numbers of solutions in $\mathbb{CP}^1$ of $\mathsf{F}(t) = 1$ and $\mathsf{F}(t) = \infty$. Putting these together we find that $N(A, B, C)$ is bounded by a constant multiple of $h(m, n)^\nu$ where $\nu = \nu_D(A) + \nu_D(B) + \nu_D(C)$ is the size of $\mathsf{F}^{-1}(\{0, 1, \infty\})$. Moreover, if at least two points in $\mathsf{F}^{-1}(\{0, 1, \infty\})$ are rational then we can save an extra factor of $h(m, n)^2$ as we did before; in fact we expect to save this factor in any case, because there are about H^2 choices of (m, n) with $h(m, n) \in (H/2, H]$, and it is not too hard to show that in fact this $h(m, n)^2$ saving is available for all nonconstant rational functions F. In other words, we can make the denominator of (6.16) no larger than $(\nu - 2) \log h(m, n) + e'$, where e' is another bounded error.

Combining our estimates and letting $h(m, n) \to \infty$, we find that the polynomial identity $\mathsf{A}+\mathsf{B} = \mathsf{C}$ will yield a disproof the ABC conjecture if $\nu < D + 2$. We have already given several examples of $\nu = D + 2$, and there are many others, some of which are very easy to construct (try $(\mathsf{A}, \mathsf{B}, \mathsf{C}) = (1, t^D - 1, t^D)$ for instance). Might we attain $\nu < D + 2$ if we are just a little more clever, or look harder? This is where Mason's theorem enters:

Theorem 4. *[Mas]: If $\mathsf{F} \in \mathbb{C}(t)$ is a rational function of degree $D > 0$ on $\mathbb{CP}^1$ then $\mathsf{F}^{-1}(\{0, 1, \infty\})$ has cardinality at least $D + 2$.*

This ruins our hope for an easy refutation of the ABC conjecture. Viewed more positively, it is evidence for the truth of the conjecture, and indeed can be viewed as an "ABC theorem" for polynomials or rational functions. To make the comparison explicit, we again take logarithms in the conjectured inequality (6.13) to write it as $\log N(ABC) > (1 - \epsilon) \log C - \log(1/c_\epsilon)$. We saw that for polynomials ν and D play the roles of $\log N(ABC)$ and $\log C$ respectively. Thus Mason's theorem is an even stronger statement, because the troublesome terms $-\epsilon \log C$ and $-\log(1/c_\epsilon)$ in the lower bound for $\log N(ABC)$ have been replaced by the helpful $+2$ in the lower bound on ν.

Moreover, while the ABC conjecture seems intractable at present, Mason's theorem can be proved easily. There are several related routes, all exploiting the idea of detecting multiple roots of a polynomial or rational function using its *derivative* — a tool not available for integers or rational numbers. The route we choose uses the logarithmic derivative, for which it will be convenient to assume that ∞ is not a preimage of 0, 1, or ∞. We ensure this by applying to t a fractional linear transformation that moves all the preimages of $\{0, 1, \infty\}$ away from infinity.

Proof. Fix a number t_0 not in $\mathsf{F}^{-1}(\{0, 1, \infty\})$, and let $\mathsf{F}_1(t) = \mathsf{F}(t_0 + (1/t))$, a rational function also of degree D and with the same number of preimages of $\{0, 1, \infty\}$ as F, none of which are at infinity. Let ν_0, ν_1, ν_∞ be the number of preimages of $0, 1, \infty$ respectively. Let Λ be the logarithmic derivative $\mathsf{F}_1'/\mathsf{F}_1$. Then Λ is not identically zero because F_1 is nonconstant, and Λ has a simple pole (that is, has a denominator with a simple root) at each preimage of 0 or ∞, regardless of its multiplicity. Hence the denominator of Λ has degree $\nu_0 + \nu_\infty$. Any root of $\mathsf{F}_1' - 1$ of multiplicity e is a root of Λ of multiplicity $e - 1$. Summing over the roots, we find the the numerator of Λ has at least $D - \nu_1$ roots counted *with* multiplicity, and therefore has degree at least $D - \nu_1$. But the difference between the denominator's and numerator's degrees is the order of vanishing of Λ at infinity, which is at least 2 (to see this, expand F_1 at infinity as $\sum_{i=0}^{\infty} a_i t^{-i} = a_0 + a_1 t^{-1} + a_2 t^{-2} + a_3 t^{-3} + \cdots$ with $a_0 \neq 0$, and calculate $\mathsf{F}_1' = -a_1 t^{-2} - 2a_2 t^{-3} - 3a_3 t^{-4} - \cdots$). Hence $D - \nu_1 \leq \nu_0 + \nu_\infty - 2$, which is equivalent to the desired inequality $\nu_0 + \nu_1 + \nu_\infty \geq D + 2$. □

Since the numerator of the derivative or the logarithmic derivative of A/C is (up to sign) the Wronskian

$$W_2(\mathsf{A}, \mathsf{C}) = \det \begin{vmatrix} \mathsf{A} & \mathsf{C} \\ \mathsf{A}' & \mathsf{C}' \end{vmatrix} = \mathsf{A}\mathsf{C}' - \mathsf{A}'\mathsf{C},$$

the proof can also be formulated in terms of Wronskians. The key fact that F and $\mathsf{F} - 1$ have the same derivative then corresponds to the identity $W_2(\mathsf{A}, \mathsf{C}) = W_2(\mathsf{A} - \mathsf{C}, \mathsf{C})$, which holds because $W_2(\cdot, \cdot)$ is

bilinear and alternating, and forces $W_2(\mathsf{A}, \mathsf{C})$ to vanish at multiple zeros of B. Also equivalent, though not as transparently so, is the proof obtained by applying the Riemann-Hurwitz formula to F. This approach explains the "+2" in Mason's inequality as the Euler characteristic of $\mathbb{CP}^1$, and generalizes to rational functions F of degree $D > 0$ on other compact Riemann surfaces, for which Mason finds the inequality $\#\left(\mathsf{F}^{-1}(\{0,1,\infty\})\right) \geq D + \chi = D + 2 - 2g$, where g is the genus and χ the Euler characteristic of the surface. This is why the rational functions F constructed by Belyi cannot satisfy $\deg(F) > \#\left(F^{-1}(\{0,1,\infty\})\right)$ unless $g \geq 2$. For an elliptic curve we have $g = 1$, so $\deg(F) = \#F^{-1}\left(\{0,1,\infty\}\right)$ is possible, and if the elliptic curve has positive rank then its rational points yield another kind of infinite family of (A, B, C) triples with $\limsup \theta(A,B,C) \geq 1$ (such as $(x^3, y^3, 91z^3)$ for primitive solutions of $x^3 + y^3 = 91z^3$); but the points are too sparse for us to prove that the limsup strictly exceeds 1, and again we come just short of a disproof of the ABC conjecture.

6.6 A Putnam problem: minding our P's and Q's

The last problem of the 1956 William Lowell Putnam Mathematical Competition asks [GGK, p.47]:

> The polynomials $P(z)$ and $Q(z)$ with complex coefficients have the same set of numbers for their zeros but possibly different multiplicities. The same is true of the polynomials $P(z) + 1$ and $Q(z) + 1$. Prove that $P(z) \equiv Q(z)$.

As noted in [GGK, p.431], it must be assumed that at least one of P and Q is not constant, else the claim is false. We thus assume $\max(\deg(P), \deg(Q)) > 0$, and by symmetry may take $m = \deg(P) \geq \deg(Q) = n$. The claim is clearly true if P has distinct roots, because then $Q = cP$ for some $c \in \mathbb{C}$, and if λ is any root of $P + 1$ then $0 = Q(\lambda) + 1 = cP(\lambda) + 1 = -c + 1$ implies $c = 1$. Likewise if $P + 1$ has distinct roots. We must then contend with the case that P and $P + 1$ both have multiple roots — and we know already that the derivative $P' = (P + 1)'$ detects multiple roots of either P or $P + 1$. We proceed as in [GGK, p.431–432]. Let $\lambda_1, \ldots, \lambda_r$ be the distinct roots of P (and thus also of Q), and $\mu_1, \ldots, \mu_s$ the distinct roots of $P + 1$ (and thus also of $Q + 1$). By an argument we can now recognize as the special case of Mason's theorem in which F is a polynomial — and which would fail if $m = 0$ were allowed — we have $m - 1 = \deg(P') \geq 2m - r - s$, whence $r + s \geq m + 1$. But each root of P or $P + 1$ is also a root of $P - Q$, a polynomial of degree at most m. Therefore $P - Q$ is the zero polynomial, and we are done.

The corresponding statement for integers instead of polynomials would be that a positive integer n is determined uniquely by the sets (without the multiplicities) of prime factors of n and of $n + 1$, that is, by the conductors $N(n)$ and $N(n+1)$. We might expect that this should be false, because the proof in the polynomial case hinges on an inequality stronger than can be true for integers. Indeed there are infinitely many counterexamples, the smallest with natural numbers being $n = 2$ and $n' = 8$ (this is yet another appearance of $1 + 8 = 9$), which begins the infinite family $\{n, n'\} = \{2^m - 2, 2^m(2^m - 2)\}$ $(m = 2, 3, 4, \ldots)$. Still, such examples seem quite rare; an exhaustive search finds that the only case with $0 < n, n' < 10^8$ not of the form $\{2^m - 2, 2^m(2^m - 2)\}$ is $\{75, 1215\}$ (with $N(75) = N(1215) = 15$ and $N(76) = N(1216) = 38$). When we allow also negative integers, the identity $N(-n) = N(n)$ gives an involution $\{n, n'\} \leftrightarrow \{-1 - n, -1 - n'\}$ on the set of solutions. Modulo this involution, we find one more infinite family $\{2^m + 1, -(2^m + 1)^2\}$ $(m = 1, 2, 3, \ldots)$, and one more sporadic pair in $(-10^8, 10^8)$, namely $\{35, -4375\}$. The infinite families intersect at $\{2, -4, 8\}$ and $\{-3, 3, -9\}$, which may be the only three-element subsets of $\mathbb{Z}$ mapped to a single point under $n \mapsto (N(n), N(n + 1))$.

Might we generalize the Putnam problem to rational functions F? Since a polynomial is just a rational function with $F^{-1}(\{\infty\}) = \{\infty\}$, we might guess that more generally if F and G are non-constant rational functions with complex coefficients that, when considered as maps from the Riemann sphere $\mathbb{CP}^1$ to itself, satisfy $F^{-1}(\{w\}) = G^{-1}(\{w\})$ for each of $w = 0, 1, \infty$, then $F = G$. (In the Putnam problem, F and G would be the polynomials $P + 1$ and $Q + 1$.) Alas this natural guess is false.

An explicit counterexample is[14]

$$F(z) = \frac{(z-1)^3(z+3)}{16z}, \quad G(z) = h(-3/z) = \frac{(z-1)(z+3)^3}{16z^3},$$

with $F(z) - 1 = (z-3)(z+1)^3/16z$ and $G(z) - 1 = (z-3)^3(z+1)/16z^3$. Here F and G are rational functions of degree 4. Is this the smallest possible? It is probably much harder to completely describe all counterexamples, or even to decide whether there any with $\deg(F) \neq \deg(G)$.

6.7 Further problems and results

In number theory most things that can be done in $\mathbb{Q}$ or $\mathbb{Z}$ generalize, with some additional effort, to number fields K (finite-degree field extensions of $\mathbb{Q}$) and their rings O_K of algebraic integers. This is true of the ABC conjecture, which can be naturally formulated over any K or O_K, and has much the same consequences there as we saw over $\mathbb{Q}$ or $\mathbb{Z}$. Much of the extra effort in making this generalization arises because O_K need not have unique factorization, so some solutions in K of $A + B = C$ may not be proportional to any solution in relatively prime elements of O_K. Thus it is more natural to formulate the conjecture in terms of the ratio $F = A/C$, which is invariant under scaling (A, B, C). Briefly, we replace $N(ABC)$ in the LHS of (6.13) or (6.14) by the product of the *norms* of all prime ideals of O_K at which F is congruent to one of $0, 1, \infty$, and in the RHS we take the $(1-\epsilon)$-th power of the height of F, appropriately defined, rather than of C or of $\max(|A|, |B|, |C|)$. See [Vo, p.84] for the details. Mason's theorem still defeats attempts at easy disproofs — recall that the coefficients of the rational function F were allowed to be arbitrary complex numbers.

More subtle is the question of how the constant c_ϵ in the ABC conjecture should depend on K. In the context of Mason's theorem, if we replace $\mathbb{C}(t)$ by a finite-degree extension we get the function field of a compact Riemann surface of some genus g, and then the lower bound $D + 2$ on the size of $\mathsf{F}^{-1}(\{0, 1, \infty\})$ is lowered by $2g$. Granville and Stark [GS] propose an anologous "uniform ABC conjecture", in which the LHS of (6.13) or (6.14) is multiplied by $|\mathrm{disc}(K/\mathbb{Q})|^{1/[K:\mathbb{Q}]}$ and then the constant c_ϵ in the RHS is independent of K. Remarkably, they then show that this uniform ABC conjecture implies the long-standing conjecture that the class number of an imaginary quadratic field $\mathbb{Q}(\sqrt{-d})$ (with $d > 0$ a squarefree integer) is bounded below by a constant multiple of $d^{1/2}/\log d$, and thus that the Dirichlet L-function attached to an odd character has no "Siegel-Landau zero" (a zero s with $1 - s \ll 1/\log(d)$; the nonexistence of such zeros is an important special case of the Riemann Hypothesis for such L-functions). The proof uses special values of modular functions arising from elliptic curves with complex multiplication by the ring of algebraic integers in $\mathbb{Q}(\sqrt{-d})$.

Finally we consider the generalization to more than three variables, to integers satisfying $\pm A \pm B \pm C \pm D = 0$ and beyond. In each case we ask: Given $\max(|A|, |B|, |C|, \ldots)$, how small can the product $N(A)N(B)N(C)\cdots$ get? As before we must assume that the integers have no common factor. With more than three variables, it no longer follows that they are relatively prime in pairs, but we must at least assume that no proper sub-sum of $\pm A \pm B \pm C \pm \cdots$ vanishes, to avoid such trivialities as $2^r + 1 - 2^r - 1 = 0$. It is then known that an upper bound on $N(A)N(B)N(C)\cdots$ implies an upper bound on $\max(|A|, |B|, |C|, \ldots)$, but again this known bound is much too large for our purpose. Even in the special case $A = A_0 w^n$, $B = B_0 x^n$, etc. we have a difficult question: How

[14]The reader who got this far may well wonder where this counterexample comes from. It arises naturally in the theory of elliptic modular functions. For τ in the upper half-plane $\mathcal{H}$, let $\eta(\tau)$ be the Dedekind eta function $e^{\pi i/12}\prod_{n=1}^{\infty}(1-e^{2\pi i n\tau})^{24}$, and define $\lambda(\tau) = 16(\eta_2^2\,\eta_{1/2}/\eta_1^3)^8 = 16q\prod_{n=1}^{\infty}(1+q^{2n})/(1+q^{2n-1})$ where $\eta_k = \eta(k\tau)$ and $q = e^{\pi i\tau}$. Then λ generates the field of modular functions invariant under the ideal hyperbolic triangle group $\Gamma(2)$, and takes the values $0, 1, \infty$ at the cusps of that group. The function F expresses λ in terms of the generator $-3(\eta_3/\eta_1)^{10}(\eta_{1/2}\eta_2/\eta_{3/2}\eta_6)^4$ of the modular functions for $\Gamma(2)\cap\Gamma_0(3)$, and thus gives explicitly the map from the corresponding modular curve to the modular curve $X(2)$ corresponding to $\Gamma_0(2)$. The coordinate λ of $X(2)$ parametrizes elliptic curves $E : Y^2 = X(X-1)(X-\lambda)$ with all their 2-torsion points rational; z parametrizes 3-isogenies $E \to E'$ between pairs of such curves; and the involution $z \leftrightarrow -3/z$ takes the isogeny $E \to E'$ to the dual isogeny $E' \to E$. See [El2].

are the nontrivial primitive solutions of $A_0w^n + B_0x^n + C_0y^n = D_0z^n$ distributed? Our heuristics suggest that solutions should be plentiful for $n < 4$ (if there is a nonzero solution to begin with), sparse for $n = 4$, and bounded for $n > 4$. Likewise for N variables, with critical exponent $n = N$.

Unfortunately this guess is at best close to the truth. Euler already found a polynomial solution for $w^4+x^4 = y^4+z^4$, giving plentiful solutions for that equation, starting with $133^4+134^4 = 59^4+158^4$. There is even a polynomial family of solutions of $w^5 + x^5 = y^5 + z^5$, though sadly not over $\mathbb{Q}$:

$$w, x = 2t \pm (t^2 - 2), \quad y, z = 2t \pm i(t^2 + 2). \tag{6.17}$$

For $n = 6$ one can still obtain infinitely many primitive solutions for some choices of (A_0, B_0, C_0, D_0), using the polynomial identity

$$(t^2 + t - 1)^3 + (t^2 - t - 1)^3 = 2t^6 - 2.$$

Indeed let $(A_0, B_0, C_0, D_0) = (\alpha^3, \beta^3, 2, 2)$. Then if there are infinitely many rational solutions (t, u, v) of

$$t^2 + t - 1 = \alpha u^2, \quad t^2 - t - 1 = \beta v^2 \tag{6.18}$$

then each yields a rational solution $(u, v, 1, t)$ of $A_0w^6 + B_0x^6 + C_0y^6 = D_0z^6$, and thus a primitive integer solution by clearing common factors. Now it can be shown that (6.18) is an elliptic curve, which has positive rank if it has a single rational point with $t \notin \{0, \pm 1, \infty\}$. The simplest such (α, β) is $(5, 1)$ with $t = 2$, giving $125 + 1 + 2 = 2 \cdot 2^6$. The next few t values for $(\alpha, \beta) = (5, 1)$ are $-82/19$, $-148402/91339$, and $-10458011042/1213480199$, giving the solutions[15]

$$(31, 19, 89, 82), \quad (5009, 91339, 165031, 148402),$$

$$(4363642319, 1213480199, 10981259039, 10458011042).$$

Note that, unlike the ABC conjecture, our naïve guess for $A_0w^n + B_0x^n + C_0y^n = D_0z^n$ was disproved by polynomial identities. Thus even Mason's theorem has no good analogue here. One can use a 3×3 Wronskian to get an "ABCD theorem", and likewise for more variables, but these inequalities are no longer sharp. For example, if (w, x, y, z) is a nontrivial solution in $\mathbb{C}[t]$ of $w^n + x^n = y^n + z^n$ then one can show that $n < 8$ by counting roots of $W_3(w^n, x^n, y^n)$, but it is not known whether $n = 6$ or $n = 7$ can occur, nor whether all nontrivial solutions for $n = 5$ are equivalent with (6.17).

Can we salvage from our predicament a conjecture that is both plausible and sharp? Lang [La2] suggested that such conjectures should still be true "on a nonempty Zariski-open set", that is, when we exclude variables that satisfy some algebraic condition. This may well be true, though the possibility of an unpredictable exceptional set makes Lang's conjectures even harder to test. As an indication of the power of these conjectures, we conclude by citing one striking application. Recall that Mordell conjectured, and Faltings proved, that an algebraic curve of genus $g > 1$ over $\mathbb{Q}$ has only finitely many rational points. The conjecture and proofs are silent on how the number of points can vary with the curve. But Caporaso, Harris, and Mazur showed [CHM] that Lang's conjectures imply a uniform upper bound $B(g)$, depending only on g, on the number of rational points of any genus-g curve over $\mathbb{Q}$!

References

[Bel] G[ennadii] V[ladimirovich] Belyi: On the Galois extensions of the maximal cyclotomic field (in Russian), *Izv. Akad. Nauk. SSSR* **43** (1979), 267–276.

[Beu] Frits Beukers: The Diophantine Equation $Ax^p + By^q = Cz^r$, *Duke Math. J.* **91** (1998), 61–88.

[Br] Nils Bruin: On powers as sums of two cubes, pages 169–184 in *Algorithmic Number Theory (Leiden, 2000)*, Berlin: Springer, 2000 (Wieb Bosma, ed.; *Lecture Notes in Computer Science* **1838**).

[15] I do not know where this construction originated. I must have noticed it by 1988, because my computer files include a listing of these solutions dated May 1988.

[CHM] Lucia Caporaso, Joe Harris, and Barry Mazur: Uniformity of rational points, *J. Amer. Math. Soc.* **10** (1997) #1, 1–35.

[DG] Henri Darmon and Andrew Granville: On the equations $x^p + y^q = z^r$ and $z^m = f(x, y)$, *Bull. London Math. Soc.* #129 (**27** part 6, Nov.1995), 513–544.

[Ed] Johnny Edwards: A Complete Solution to $X^2 + Y^3 + Z^5 = 0$, *J. f. d. reine u. angew. Math.* **571** (2004), 213–236 (also online at `http://www.math.uu.nl/people/edwards/icosahedron.pdf`).

[El1] Noam D. Elkies: ABC implies Mordell, *International Math. Research Notices* **1991** #7, 99–109 [bound with *Duke Math. J.* **64** (1991)].

[El2] Noam D. Elkies: Wiles minus epsilon implies Fermat, pages 38–40 in *Elliptic Curves, Modular forms, and Fermat's Last Theorem* (J. Coates and S.-T. Yau, eds.; Boston: International Press, 1995; proceedings of the 12/93 conference on elliptic curves and modular forms at the Chinese University of Hong Kong).

[Ev] Jan-Hendrik Evertse: On equations in S-units and the Thue-Mahler equation, *Invent. Math.* **75** (1984), 561–584 (1994).

[F1] Gerd Faltings: Endlichkeitssätze für abelsche Varietäten über Zahlkörpern, *Invent. Math.* **73** (1983), 349–366.

[F2] Gerd Faltings: Diophantine Approximation on Abelian Varieties, *Ann. Math.* (2) **133** (1991), 549–576.

[GGK] Andrew M. Gleason, R.E. Greenwood, and Leon M. Kelly: *The William Lowell Putnam Mathematical Competition — Problems and Solutions: 1938–1964.* Washington, D.C.: Math. Assoc. of America, 1980.

[GS] Andrew Granville and Harold M. Stark: ABC implies no 'Siegel zero' for L-functions of characters with negative discriminant, *Invent. Math.* **139** #3 (2000), 509–523.

[IR] Kenneth Ireland and Michael Rosen, *A Classical Introduction to Modern Number Theory*, 2nd ed. New York: Springer, 1990 (Graduate Texts in Math. **84**).

[La1] Serge Lang: Integral points on curves, *Publ. Math. IHES* **6** (1960), 27–43.

[La2] Serge Lang: Hyperbolic and diophantine analysis, *Bull. Amer. Math. Soc.* **14** #2 (1986), 159–205.

[LM] D[onald] J. Lewis and Kurt Mahler: Representation of integers by binary forms, *Acta Arith.* **6** (1960/61), 333–363.

[Mas] R[ichard] C. Mason: *Diophantine Equations over Function Fields*, London Mat. Soc. Lect. Notes Ser. **96**, Cambridge Univ. Press 1984. See also pp.149–157 in Springer LNM **1068** (1984) [=proceedings of Journées Arithmétiques 1983, Noordwijkerhout].

[Mau] R. Daniel Mauldin: A Generalization of Fermat's Last Theorem: The Beal Conjecture and Prize Problem, *Notices of the Amer. Math. Soc.* **44** #11 (1997), 1436–1437. `http://www.ams.org/notices/199711/beal.pdf`

[Mi] Preda Mihăilescu: Primary Cyclotomic Units and a Proof of Catalan's Conjecture, *J. reine angew. Math.* **572** (2004), 167–195.

[Mo] Louis J. Mordell: On the rational solutions of the indeterminate equations of the third and fourth degrees, *Proc. Cambridge Phil. Soc.* **21** (1922), 179–192.

[Ni] Abderrahmane Nitaj: The ABC Conjecture Home Page. `http://www.math.unicaen.fr/kernlmm/~nitaj/abc.html`

[Oe] Joseph Oesterlé: Nouvelles approches du "théorème" de Fermat, *Sém. Bourbaki* 2/1988, exposé #694.

[PSS] Bjorn Poonen, Edward F. Schaefer, and Michael Stoll: Twists of $X(7)$ and primitive solutions to $x^2+y^3 = z^7$. Preprint, 2005 (online at `http://arxiv.org/math.NT/0508174`).

[Si1] Joseph H. Silverman: *The Arithmetic of Elliptic Curves*. New York: Springer 1986 (GTM **106**).

[Si2] Joseph H. Silverman: Wieferich's criterion and the *abc*-conjecture, *J. Number Theory* **30** #2 (1988), 226–237.

[ST] Cameron L. Stewart and Robert Tijdeman: On the Oesterlé-Masser conjecture, *Monatsh. Math.* **102** (1986), 251–257.

[SY] Cameron L. Stewart and Kunrui Yu: On the *abc* conjecture, II, *Duke Math. J.* **108** (2001), 169–181.

[Ta] Olga Taussky: Sums of squares, *Amer. Math. Monthly* **77** #8 (Oct.1970), 805–830.

[TW] Richard Taylor and Andrew Wiles: Ring-theoretic properties of certain Hecke algebras, *Ann. Math.* **141** (1995), 553–572.

[Vo] Paul Vojta: *Diophantine Approximations and Value Distribution Theory*. Berlin: Springer 1987 (Lect. Notes Math. **1239**).

[dW] Benne de Weger: *Algorithms for Diophantine equations*. Amsterdam: Centrum voor Wiskunde en Informatica, 1989 (CWI tract **65**).

[Wie] Arthur Wieferich: Zum letzten Fermat'schen Theorem, *J. f. d. reine u. angew. Math.* **136** (1909), 293–302.

[Wil] Andrew Wiles: Modular elliptic curves and Fermat's Last Theorem, *Ann. Math.* **141** (1995), 443–551.

FEATURE

7

Mathematical Minutiae: Differentiation as a Functor

Athanasios Papaioannou '07†
Harvard University
Cambridge, MA 02138
apap@fas.harvard.edu

Unlike any other article in this journal, this one begins with a warning: Categories, beautiful and powerful as they may be, are not panacea and should be used with great prudence. This short note presents a fun, but silly use of categories.

7.1 The Chain Rule

In what follows, $\mathbb{R}$ denotes the set of real numbers. By $\tau_{\mathbb{R}}$ we mean the category whose objects are pairs (U, u) of open subsets $U \subseteq \mathbb{R}$ together with a point $u \in U$, and whose morphisms $(U, u) \to (U', u')$ are differentiable functions f preserving base points, in the sense that $f(u) = u'$. By $\mathscr{R}$ we mean the category whose unique object is $\mathbb{R}$, and whose morphisms are given by

$$\mathrm{Hom}_{\mathscr{R}}(\mathbb{R}, \mathbb{R}) = \{\phi_a : x \mapsto ax \mid a \in \mathbb{R}\};$$

the composition of ϕ_a and ϕ_b is defined to be ϕ_{ab}. We now claim that the assignation $D : \tau_{\mathbb{R}} \to \mathscr{R}$ given by

$$\begin{aligned} (U, u) &\longmapsto \mathbb{R} \\ (U, u) \xrightarrow{f} (U', u') &\longmapsto \left.\frac{df}{dx}\right|_{x=u} \end{aligned}$$

is a functor.

Indeed, we need to check that, given a diagram of the form

$$(U, u) \xrightarrow{f} (U', u') \xrightarrow{g} (U'', u''),$$

the following relation holds:

$$D(g \circ f) = D(g) \circ D(f).$$

But this last expression can be rewritten as $(g \circ f)'(u) = g'(u')f'(u)$, which is exactly the chain rule at u! Moreover, to say that D preserves the identity is precisely to say that the derivative of $f(x) = x$ is 1, which is clearly true.

† Athanasios Papaioannou, Harvard '07, is a mathematics concentrator.

7.2 Getting more serious

A rather more fruitful way to think about derivations in terms of functors is that of modern geometry. We pursue this with extreme economy, at the expense of using many undefined words. Let's think of smooth manifolds as ringed spaces, i.e., pairs $(M, \mathcal{O}_M)$ consisting of a topological space together with a sheaf of functions, such that $(M, \mathcal{O}_M)$ is locally isomorphic to $(\mathbb{R}^n, \mathcal{O}_{\text{sm}})$, the ringed space of $\mathbb{R}^n$ together with the sheaf of smooth functions on it. To every point of M we may attach a ring, that of the derivations from the stalk of the structure sheaf $\mathcal{O}_{\text{sm},m}$ to $\mathbb{R}$—this is a well-known gadget, the **tangent space** at m. Now, there is a way of compiling all these tangent spaces into the **tangent sheaf** on M, which is the dual to the better-known sheaf of differential forms $\Omega_{M/\mathbb{R}}$. And that these sheaves, like all sheaves, are functors of some sort, should please any rabid categorialist.

FEATURE

8

Problems

The HCMR welcomes submissions of original problems in any field of mathematics, as well as solutions to previously proposed problems. Proposers should direct problems to Problems Editor Zachary Abel at `hcmr-problems@hcs.harvard.edu` or at the address on the inside front cover. A complete solution or a detailed sketch of the solution should be included, if known. Solutions to previous problems should also be directed to the Problems Editor at `hcmr-solutions@hcs.harvard.edu` or at the address on the inside front cover. Solutions should include the problem reference number, as well as the solver's name, contact information, and affiliated institution. Additional information, such as generalizations or relevant references, is also welcome. All correct solutions will be acknowledged in future issues, and the most outstanding solutions received will be published. To be considered for publication, solutions to the problems below should be postmarked no later than *November 1, 2007*. An asterisk beside a problem or part of a problem indicates that no solution is currently available.

S07 – 1. How many hyperplane cuts are necessary to divide a $3 \times 5 \times 7 \times 9 \times 11$ rectangular solid into $3 \cdot 5 \cdot 7 \cdot 9 \cdot 11$ distinct $1 \times 1 \times 1 \times 1 \times 1$ hypercubes, if previously separated pieces can be rearranged between cuts?

Proposed by Joel Lewis '07.

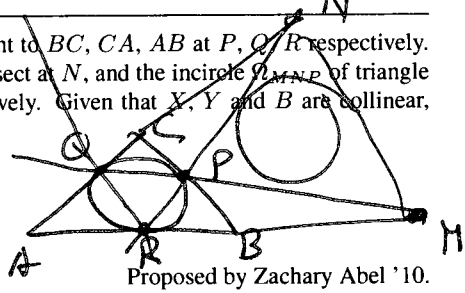
S07 – 2. Suppose $f : [0,1] \to \mathbb{R}$ is an integrable function such that $y \cdot f(x) + x \cdot f(y) \leq x^2 + y^2$. Show that $\int_0^1 f(x)\,dx \leq \frac{\pi}{4}$. (One example of such a function is $f(x) = x$.)

Proposed by Scott Kominers '09.

S07 – 3. The incircle Ω_{ABC} of a triangle ABC is tangent to BC, CA, AB at P, Q, R respectively. Rays PQ and BA intersect at M, rays PR and CA intersect at N, and the incircle Ω_{MNP} of triangle MNP is tangent to MN and NP at X and Y respectively. Given that X, Y and B are collinear, prove:

(a) Circles Ω_{ABC} and Ω_{MNP} are congruent, and

(b) these circles intersect each other in 60° arcs.



Proposed by Zachary Abel '10.

S07 – 4. For a prime p, let $\mathbb{Z}_{(p)} \subset \mathbb{Q}$ denote the localization of the integral domain $\mathbb{Z}$ at the prime ideal (p); that is, the subring of $\mathbb{Q}$ consisting of the rational numbers with denominators prime to p. The canonical homomorphism $\mathbb{Z} \to \mathbb{F}_p$ induces a canonical homomorphism $\phi_p : \mathbb{Z}_{(p)} \to \mathbb{F}_p$, the reduction modulo p homomorphism with kernel the maximal ideal $p\mathbb{Z}_{(p)}$ of the local ring $\mathbb{Z}_{(p)}$. (For example, $\phi_5(1/2) = 3 \in \mathbb{F}_5$.)

Let V be the set of primes p for which $\{\frac{3^n-1}{2^n-1} \mid n \in \mathbb{N}\} \subset \mathbb{Z}_{(p)}$.

(a) Characterize the set V.

(b) Show that V and $P \setminus V$ are both infinite sets, where P is the set of primes. (In other words, show that V is neither finite nor cofinite in the set of primes.)

(c) Show that, for every $p \in V$, the map $\mathbb{N} \to \mathbb{F}_p$ given by $n \mapsto \phi_p\big((3^n - 1)/(2^n - 1)\big)$ is periodic. (For example, $5 \in V$, and the corresponding map $\mathbb{N} \to \mathbb{F}_5$ is $2, 1, 3, 2, 2, 1, 3, 2, 2, 1, 3, 2, \ldots$.)

Proposed by Vesselin Dimitrov '09.

S07 – 5. (a) Prove that, for distinct positive real numbers a and b, the following inequality holds:

$$\frac{a+b}{2} \geq \frac{a^{\frac{a}{a-b}} b^{\frac{b}{b-a}}}{e} \geq \frac{a-b}{\ln a - \ln b}.$$

(b*) Show that both inequalities are strict.

Proposed by Shrenik Shah '09.

FEATURE

9

Endpaper: How to Compute Determinants

Prof. Dennis Gaitsgory†
Harvard University
Cambridge, MA 02138
gaitsgde@math.harvard.edu

During one of my years in graduate school in Israel, I was a teaching fellow for a class on linear algebra. I found the job annoying for two reasons: On one hand, the students were primarily non-math majors. But more importantly, my class started at eight in the morning, which did not rhyme well with my lifestyle at the time. As a result, I could not bring myself to prepare my section in advance. Instead I improvised each time....

One day I found myself explaining determinants. "You know, for a generic matrix a determinant is never zero. Somebody, give me an example of a matrix!" The class produced no reply. They were no less sleepy than I was. In fact, not only were they asleep but they were suspicious as well. They did not want to risk giving a matrix which by misfortune would have a zero determinant, with the gloomy title of "degenerate" attached to it.

So I proceeded: "OK, let's take the first matrix that comes to mind."

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

I set about computing the determinant by the usual formula. I was never good with computations and, once again, I was especially sleepy:

$$1 \cdot 5 \cdot 9 - 2 \cdot 4 \cdot 9 \pm 3 \cdot 4 \cdot 8 + \ldots.$$

It took me a good 10 minutes. And what a shock, the determinant was zero! "I must have made a mistake," I told the class. I ran through the calculations once more, checking every step. Another 10 minutes passed. Zero again!

I tried to save myself. "OK, but sometimes the determinant is zero. Sorry. But now let's take a *really* generic matrix."

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix}$$

Another lengthy computation....

At the end of that semester I was forced to enroll in a special seminar for delinquent instructors.

†Prof. Dennis Gaitsgory is a faculty member of the Harvard Mathematics Department.

Did you enjoy the HCMR?

Subscribe, and submit

your own articles
and problems!

E-mail us at
hcmr@hcs.harvard.edu