Mathematics is Four Dimensional: A Basis for a Professional Master's Degree

William A. Massey
Bell Labs, Lucent Technologies
Murray Hill, NJ 07974
will@research.bell-labs.com

January 10, 2000

Abstract

This paper presents a new unifying vision of mathematics as having four dimensions. We also discuss the significance of CAARMS meetings and the development of a professional masters degree in the context of this new paradigm.

Keywords: Professional Masters Program

1 Personal Mathematical Journey

In 1978, I had my first mathematics publication [1]. My A. B. degree in mathematics from Princeton University was completed a year earlier, specializing in algebraic number theory. Up and coming mathematics and science majors were warned during the 1970's that academic jobs were scarce and it might not be possible to pursue a career in mathematics research. At the time, my plan was to enjoy "pure mathematics" for now as an undergraduate and major in it. Mathematics classes in college were coupled with physics courses in mechanics, electromagnetism, and quantum mechanics. Pragmatic urges to shift over to "applied mathematics" would be put off until graduate school.

During my senior undergraduate year of 1977, I applied for a minority fellowship sponsored by Bell Laboratories called the Cooperative Research Fellowship Program[1]. This was a program initiated in 1972 by prominent African-American Bell Labs scientists. Bell Laboratories was one of the few major research institutions that actually had a history of distinguished African-American scientists going back to the 1940's. In hiring Dr. William Lincoln Hawkins in 1942, an African-American chemist and National Engineering medalist winner, Bell Labs broke the "color barrier" five years before major league baseball did so with the hiring of Jackie Robinson in 1947. For the past 28 years, CRFP has helped Bell Laboratories increase the number of African-, Hispanic-, and Native-American minority Ph.D.'s in the engineering, mathematical, and physical sciences in the United States (over 120 Ph.D. 's and counting).

In 1977, the year of my graduating from college, I was awarded a CRFP fellowship to pursue a Ph.D. in mathematics at Stanford University and work that summer at Bell Laboratories in Murray Hill, New Jersey. One of the successful features on the fellowship was pairing every awardee with a Bell Labs scientist as a mentor. My mentor was James McKenna, head of one of the departments in the Mathematical Sciences Research Center. He introduced my to John Morrison, who in turn introduced me to the applied mathematical field of queueing theory. That summer collaboration led to my first technical paper in 1978.

My experiences from Bell Labs made me eager to learn more about the field of queueing theory. As a graduate student at Stanford University, I took many courses in probability theory, stochastic processes, and applied probability as I was studying for qualifying exams in real analysis, complex analysis and abstract algebra. My Ph.D. dissertation in queueing theory was written under the direction of Joseph B. Keller.

Since 1981, I have worked at Bell Laboratories and currently have over 40 publications in the fields of queueing theory, performance modelling, telecommunications, and applied probability.

The first thing that my undergraduate training gave me was a sense of the difference between "cookbook calculus" and a rigorous axiom-definition-theorem-proof presentation of calculus, as taught in an honors calculus course taught by the late Benard Fox. Reading Michael Spivak's book *Calculus* was an eye-opener and it introduced me to the world of mathematical theory. The second thing learned was the importance of taking courses in physics. Many mathematical concepts such as cross-products, gradient, curl, divergence,

[1]The Cooperative Research Fellowship Program was founded in 1972 and is currently co-sponsored by Lucent Technologies and Bell Laboratories. More information about the program can be found at http://www.bell-labs.com/fellowships/CRFP.

eigenvalues, Hilbert spaces, differential equations, and Fourier transforms, were first encountered in my physics courses. This was especially helpful when I took an honors multivariate calculus course using Michael Spivak's book *Calculus on Manifolds*. When the class lectures covered exterior derivatives of differential forms, you could divide the class into two halves. The students with a background in physics who had seen a curl or a divergence before could understand the motivation behind these exterior derivatives. The ones who had not could do the formal manipulations but had no idea why anyone would want to. Third, what Princeton considered to be core material for mathematical theory was real analysis, complex analysis, and abstract algebra. This conception of mathematical theory would be reinforced at Stanford since their qualifying exams were structured the same way. Finally, my experience in writing an expository undergraduate thesis in algebraic number theory, was acquired under the direction of the late Bernard Dwork.

My years at Stanford were the beginning of my transition from "pure mathematician" to "applied mathematician". My training through courses and qualifying examination preparation were in pure mathematics but my Ph.D. dissertation topic was in queueing theory, under the direction of an outstanding applied mathematician Joseph Keller. I was also spending my summers working at Bell Laboratories. What I learned during this time was that you do not attack an applied mathematics problem by starting with some well known theory and superimpose it on the model. Instead, you start with understanding the model and see what type of mathematics grows organically from it. If the resulting mathematics is standard, then look for the right theorem in the right book. However, if the resulting mathematics in non-standard, then you need develop the appropriate mathematical theory yourself. Through Joseph Keller, I was fortunate enough to be exposed to a family of queueing models (ones with time-varying rates) that were more realistic as models but lacking in a properly developed mathematical theory. I still find this to be the best of all worlds for an applied mathematician.

My years as a researcher at Bell Labs has given me the opportunity to acquire a deeper understanding about both telecommunications and mathematics (in particular the theory of stochastic processes). This happy synergy has come about through understanding the history of queueing theory and comparing the developments between advances in telecommunication systems and advances in queueing models that capture the behavior of these new systems.

I also met many African-American scientists at Bell Labs who were simultaneously prominent researchers in their fields but also had a sense of obligation to encourage more African-American students to pursue Ph.D. 's and excel in the world of science. During the 1980's I noticed that there were research conferences held by and for African-American chemists and similar conferences were held by and for African-American physicists. This motivated me to help bring about a similar conference in 1995 called the Conference for African American Researchers in the Mathematical Sciences (CAARMS). Since then I have taken on the role of "perpetual organizer" and have helped to shape the last five CAARMS events.

2 Four Dimensional Mathematics

Lay to rest the "pure" versus "applied" paradigm and replace it by a new model. Use this four dimensional perspective to discuss the creation of a "professional Masters program" in

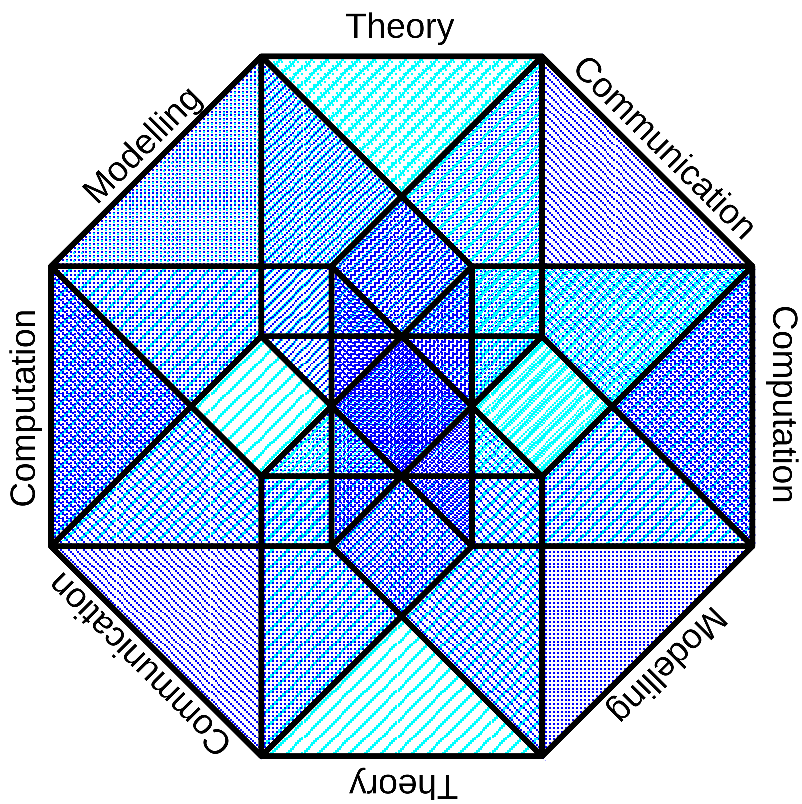

Figure 1: The four dimensions of mathematics

mathematics.

We define the four dimensions of mathematics to be (in alphabetical order): communication, computation, modelling, and theory.

We define *communication* to be the methodology through which mathematical ideas are conveyed to other mathematicians or the world at large. Examples of media that facilitate this type of communication are blackboard and chalk, computer languages, Latex, and pencil and paper. The products of the communication dimension are books, journal articles, lectures, mathematical notation, and software. It is not surprising that communication is an important dimension since math has a long distinguished role as the language of science. An underappreciated part of communication is notation. Its importance is easily conveyed through the following example:

Problem 1 *Multiply* MMMCCVI *and* LXXIII.

We define *computation* to be the collection of systematic and efficient methodologies for making mathematical calculations. The products of the computation dimension are algorithms with the goal of producing the simplest and fastest algorithms possible.

We define *modelling* to be the fundamental aspects of the real world that are captured and articulated through the language of mathematics. The premier example of the modelling dimension is physics. It is has a vast collection of mathematical tools that were created to describe some physical phenomenon. Einstein's theory of general relativity can be viewed as the "killer application" for the theory of differential geometry.

Finally, *theory* is well known to be the encoding, description, and summary of mathematical experiences through the narrative of proofs. The theory dimension is one of establishing axioms and producing theorems. One important aspect of theory is developing different types of *intuition*. A given problem is easier to prove in the context of a algebraic, geometric, physical, or probabilistic intuition. To show the advantages of algebraic intuition, consider the following example:

Problem 2 *Picture two planes in Euclidean space that intersect at exactly one point.*

This problem is difficult to view geometrically since this is an impossibility within Euclidean three space. At a minimum we must be in Euclidean four space to have this type of intersection. Shifting to an algebraic perspective, construct one of the planes as the set of four-vectors $(x, y, 0, 0)$ for arbitrary scalars x and y. If the second plane is the set of four-vectors $(0, 0, z, w)$ for arbitrary scalars z and w, then it is immediate that these two planes intersect at the single point $(0, 0, 0, 0)$.

Now consider the far reaching theorem below whose proof is immediate when viewed through the lens of geometric intuition.

Theorem 1 *Let A be a closed convex subset of a Hilbert space. For all x not belonging to A, there exists a unique y in A that realizes the minimal distance from some point in A to x.*

Proof: The proof is Figure 2. Assume that y and z are both points in A that minimize the distance from x to any point in A. The points x, y, and z form an isoceles triangle whose two equal sides meet at the vertex x. Since A is a convex set containing both points y and z, then the triangle side opposite the vertex x is a subset of A. Being in a Hilbert space means that we have the notion of orthogonality and so this problem can be solved by high school geometry. Dropping a perpendicular line segment from x to the opposite triangle side constructs a line segment from x to a point in A that is shorter that the equal sides of the triangle, if y and z are distinct points. Since this contradicts our premise, we must have $y = z$ and the point that minimizes the distance to A is unique. ∎

We can extend this multidimensional metaphor by showing that these four dimensions of mathematics are a "linearly independent" set. No three dimensions are a substitute for the remaining one. Discussing the interplay between using software and proving theorems can illustrate these ideas.

> Using software is not a substitute for designing software.

The use of software can be thought of as an aspect of mathematical communication. By contrast the act of designing software can be a mixture of theory, modelling, computation, and communication.

> Designing software is not an effective substitute for proving theorems.

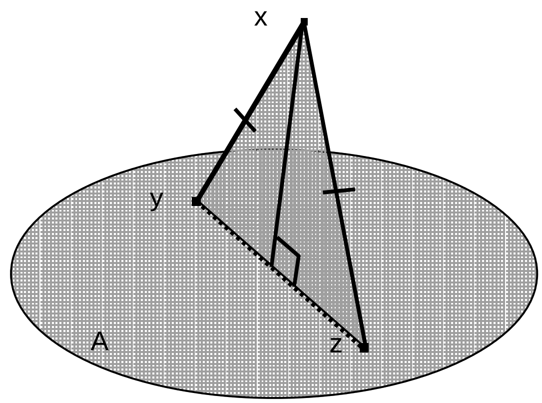

Figure 2: Counterproof of uniqueness for minimal distance to convex set

Creating computational tools to construct a set of computational examples may lead to new mathematical insights and suggest new theorems to prove, but the ultimate theorem must still be proved. Given the computer proof of the Four Color Theorem and the growing computer science field of theorem checking this view may be challenged in the future. For the present however, the case can still be made for human beings proving theorems.

> Proving theorems is not a substitute for using and designing computational software.

This is easily verified by asking why a numerical analyst would not solve a set of linear equations numerically by using Cramer's rule.

The mathematical typesetting, computer language Latex is a tool primarily used by mathematicians but it is not mathematical theory, computation, or modelling.

While these dimensions are mutually independent, they complement each other. We give some examples of what theory contributes to the other dimensions.

We define the random process $\{ N(t) \mid t \geq 0 \}$ to be a *simple point process* if the sample paths of N are increasing step functions with unit jumps where $N(0) = 0$. For a semi-open time interval $(s, t]$ we define the increment of N to be $N(t) - N(s)$. The process N is counting its own jumps so the increment of N for the time interval $(s, t]$ equals the number of jumps that N makes during that interval. The following theorem is due to Prekopa:

Theorem 2 (Prekopa) *A simple point process is non-homogeneous Poisson if and only if it has independent increments for all mutually disjoint time intervals.*

Even though this theorem was proved in the 1950's, there is still an on-going discussion of the appropriateness of the Poisson process for telecommunication traffic modelling. This theorem clarifies the situation. If we are discussing the arrivals of large number of individual people making calls (or connections in the parlance of data traffic), then it is reasonable to make the modelling assumption of independent increments. However, if we are discussing file transfers and breaking them down to transmit a stream of data packets, then there will be correlations between the number of packet arrivals in disjoint time intervals. When a statistical analysis of these arrival streams are made, one actually discovers that the nonhomogeneous Poisson model is not appropriate for modelling the arrival of data packets but it still is an appropriate model for the calling or connection traffic.

(How theory informs computation)

Every proof of an asymptotic expansion points to a computational shortcut that approximates some formula or solution to an equation. A classic example of this is Stirling's approximation of the factorial function.

Theorem 3 (Stirling's Approximation) *If* $n! = 1 \cdot 2 \cdot \cdots \cdot n$ *for all positive integers* n*, then*

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}(n/e)^n} = 1. \tag{2.1}$$

3 Professional Masters Program

There is a bright employment future for mathematics but is there one for mathematicians?

Many young mathematicians express concern about future job opportunities. The recent shortage of academic jobs has led to some strange reactions. The reality is that Ph.D. level mathematics is needed in government and industrial fields such as telecommunications, encryption and national security, as well as finance. Despite this demand, some academicians have made statements such as: If I cannot secure an academic position for a graduate student who is narrowly trained in my field, then perhaps we should train fewer graduate students. This point of view makes sense only if the following premise is allowed, that a Ph.D. training in mathematics is totally useless for anything except an academic research position in that field. I use this reducto ad absurdum argument to suggest that there is an extrinsic value in having a deep understanding of mathematics.

How to help the future of mathematics?

Radical changes in the undergraduate and Ph.D. programs are not necessary. A simple way to improve the training of students in the mathematical sciences is to upgrade the masters program in mathematics. In fields such as engineering or statistics, receiving a masters is an honorable terminal degree. For fields such as mathematics and physics however, the masters degree has the connotation of a "consolation prize" for people who did not obtain a Ph.D. This unfortunate perception needs to be eliminated.

A modest proposal.

Below is a sketch of a "professional masters program" in mathematics. It is a curriculum that reflects the view of mathematics as an organic four dimensional whole. This program also attempts to provide future math students with skills that will prepare them either for a job or a Ph.D. program in the mathematical sciences. If the Ph.D. path is taken, this may lead to some academic, business, government, or industrial job in the mathematical sciences where having mathematical skills in all four dimensions will be critical.

This most significant part of the curriculum is the theoretical dimension. We divide it into two core areas. The first core area is analysis:

- Theoretical Calculus (epsilon-delta proofs or non-standard analysis)

- Theoretical Multivariate Calculus (generalized Stokes theorem)
- Real Analysis (measure theory)
- Complex Analysis

All students entering a Masters program will have had courses in calculus, multivariate calculus and differential equations. The first two courses of the analysis core gives a consolidated review of this undergraduate material from a rigorous theorem-proof perspective. The two canonical books that cover this type of material are [4] and [5]. They provide a solid mathematical foundation for learning real and complex analysis. It should be noted that by "real analysis" we mean a course that presents measure theory and the Lebseque integral.

The second theory core area is algebra:

- Theoretical Linear Algebra
- Abstract Algebra

This combination allows students to consolidate an understanding of linear algebra from a theoretical perspective and also to see the relationship between this field and abstract algebra. For example, when discussing eigenvalues for matrices, there is a connection between the canonical Jordan normal form for matrices and the decomposition of finitely generated modules over principal ideal domains.

The computational courses will be as follows:

- Numerical Linear Algebra
- Computational Programming Language (Examples: C, C++, Fortran)

The course on numerical linear algebra reveals the fundamental role that linear algebra plays in the computational dimension of mathematics. Moreover, this field also reveals how differing goals for the same field affects the ultimate theory that is developed. In the computational realm, returning to the example of matrix eigenvalues, Jordan normal form is viewed as "numerically unstable". However an insightful theorem, which is rarely found in linear algebra books that emphasize the "algebraic", is found in books on numerical or matrix analysis. This result is the Gershgorin circle theorem, which we now state:

Theorem 4 (Gershgorin) *Let $\mathbf{M} = \{ m_{ij} \mid i, j = 1, \ldots, n \}$ be a matrix where all its entries are complex numbers. Every eigenvalue of $\mathbf{M}$ belongs to some circle in the complex plane, centered at some diagonal entry m_{ii}, with radius $\sum_{j \neq i} |m_{ij}|$.*

The theorem is a simple inequality, but it communicates significant geometric insight into where the eigenvalues of a matrix are in the complex plane. Moreover, this theorem also communicates insight into the limiting behavior of discrete time Markov chains. Recall that a matrix $\mathbf{P} = \{ p_{ij} \mid i, j = 1, \ldots, n \}$ is defined to be stochastic if all its entries are non-negative real and all the row sums equal one.

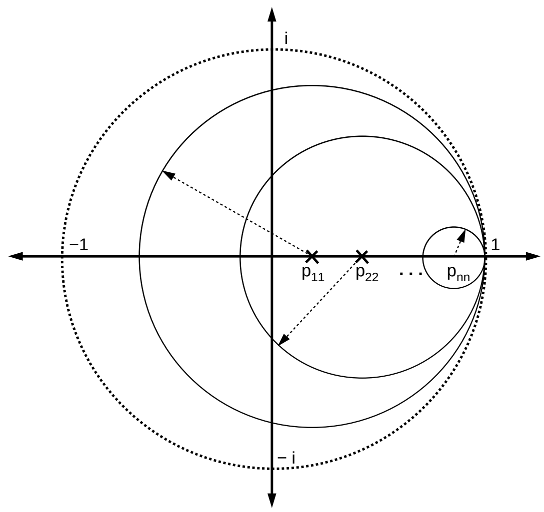

Figure 3: Gergorin circles for stochastic matrices

The matrix **P** has one as an eigenvalue which is the spectral radius for all the eigenvalues. Moreover, when every diagonal entry is strictly positive, then all the other eigenvalues for **P** have a modulus strictly less than one.

The proof for this result is geometric (see Figure 3).

A programming course allows mathematicians to see how they can take ideas from the theoretical dimension and translate them into computational algorithms.

The communication topics to be covered are:

- Mathematical Prototyping Language (Examples: Mathematica, Matlab, Maple)
- Mathematical Documentation Language (Example: Latex)
- Webpage Markup Language (Example: HTML)
- Graphical User Interface Language (Examples: Java, Visual Basic, Visual C++)

The first three topics can all be covered in one all purpose course on "computer numeracy" and the last topics should be a full semester course by itself. The first of these computer languages can be used by mathematicians as a natural extension of pencil and paper for working out examples. The last three computer languages prepare mathematicians for a future where they may be marketing their own work. Computer languages such as Latex allow mathematicians to typeset their own papers and publish their own books. Languages such as HTML allow mathematicians to distribute and promote their paper and books on personal websites that they can create themselves for the internet. Graphical user interface languages like Java, Visual Basic, or C++ allow mathematicians to transform computational algorithms into "user friendly" software that can be sold to the technical community at large.

Finally, the modelling courses can be taken in one of the following subjects:

- Physics

- Economics
- Operations Research.

By grounding the mathematical training in a specific context, we see how to abstract phenomena in the "real world" into mathematical principles. In turn we can analyze these principles in the world of mathematics to acquire a deeper insight into the original phenomena.

Now we shape these courses into a two year curriculum. To the courses below, we also add an expository Masters thesis and a summer industrial or government internship.

- Year 1 : Semester 1
  - Theoretical Calculus
  - Theoretical Linear Algebra
  - Computational Programming Language
  - Computer Numeracy (Prototyping and Documentation, Webpage Creation)
- Year 1 : Semester 2
  - Abstract Algebra
  - Theoretical Multivariate Calculus
  - Graphical User Interface Language
- Summer Session
  - Industrial/Government Internship or Two Summer Elective Courses
- Year 2 : Semester 1
  - Real Analysis
  - Numerical Linear Algebra
  - Modelling
- Year 2 : Semester 2
  - Complex Analysis
  - Masters Thesis
  - Modelling

References

[1] Massey, W. A. and Morrison, J. A., *Calculation of Steady State Probabilities for Content of Buffer with Correlated Inputs*, Bell Sys. Tech. J., 57:9 (November 1978), pp. 3097-3117.

[2] Prekopa.

[3] Horn, R. A. and Johnson, C. R. *Matrix Analysis.* Cambridge University Press, 1985.

[4] Spivak, M. *Calculus* (Third Edition). Publish or Perish, Inc., 1994.

[5] Spivak, M. *Calculus on Manifolds.* Addison-Wesley, Publishing Company, 1965.

[6] Willinger, W.