

Christine Heitsch

Strings, Trees, and RNA Folding

We highlight some challenges and opportunities at the interface of discrete mathematics and molecular biology, illustrating that this interaction motivates new combinatorial theorems as well as advancing biomedical applications.

High school teaches us that RNA's role is to mediate the production of proteins from DNA. Closer inspection, however, reveals a vast complexity of structure and diversity of function. As illustrated in Figures 1 and 2, RNA molecules are essential to cellular processes ranging from bacterial communication to viral capsid assembly. It goes almost without saying that advancing knowledge of how RNA functions in these diverse roles has the potential for tremendous scientific impact.

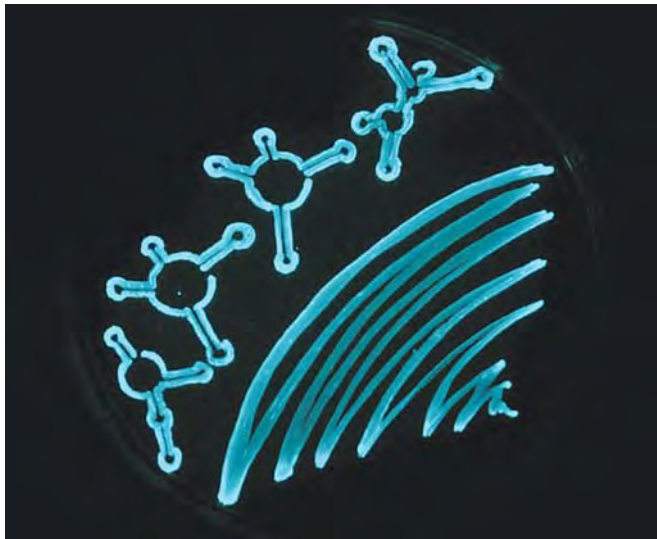


Figure 1. How do bacteria communicate? Four RNA molecules (diagrammed here through marine bioluminescence) are essential to the quorum sensing process by which bacteria regulate collective behavior, ranging from this benign light display to cholera toxicity. Despite high sequence similarity and known functional redundancy [1], the branching of the four possible RNA structures can vary significantly. New results in combinatorics and its applications provide insight into this important, yet difficult-to-determine, molecular characteristic.

A basic biological principle is that “function follows form”; that is, to understand what a molecule *does*, one

Christine Heitsch is professor of mathematics at the Georgia Institute of Technology. Her e-mail address is heitsch@math.gatech.edu.

For permission to reprint this article, please contact: reprint-permission@ams.org.

DOI: <http://dx.doi.org/10.1090/noti1567>

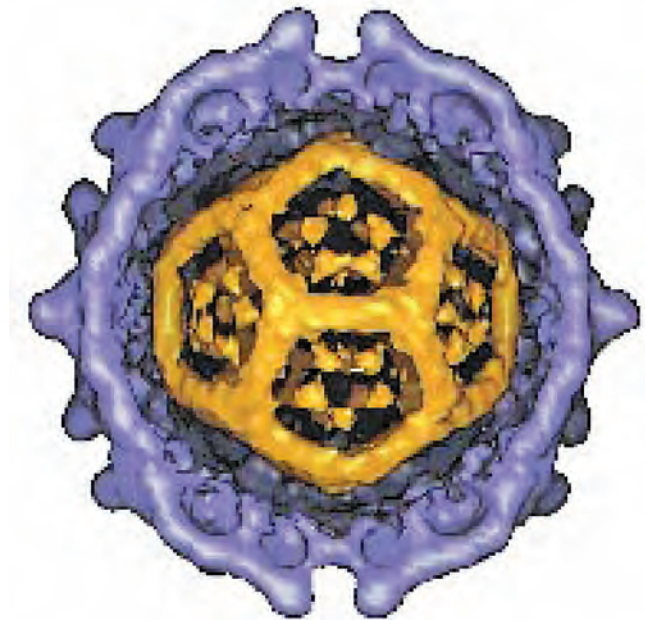


Figure 2. How do viral capsids assemble? The RNA genome (gold) is partially visible inside the quasi-icosahedral protein capsid (purple). The resolution of this detailed crystal structure [2] is such that the number and approximate length of many runs of stacked base pairs are known, but the composition and connectivity of these helices are not. New knowledge of RNA branching configurations is needed to understand how this viral sequence folds into its dodecahedral cage.

must first know how it is structured. Generically, an RNA molecule is a single nucleotide sequence which folds into a 3D conformation via a set of noncrossing, intrasequence base pairings known as a secondary structure. Given that the goal of 3D experimental determination remains inaccessible for most RNA structures, computational predictions of possible base pairing configurations, such as the four outlined in Figure 1, are essential for generating functional hypotheses.

Yet accurate prediction of the branching of these structures remains a fundamental open question. Given the combinatorial nature of the problem, this is an opportunity for discrete models, methods, and analyses to provide insights into a complex biomolecular process.

To begin to appreciate the challenge, first consider the model sequence $G^6A^4C^6 = GGGGGG AAAA CCCCCC$. Under the Watson-Crick association of G and C, it folds into a structure with one helix (run of 6 stacked base pairs) and one loop (single-stranded region) known as a “hairpin.” Thus, the lower left structure in Figure 1 has 3 hairpins, 1 internal loop, and the central loop from which branch 3 helical arms.

To analyze this coding of 2D branching information in a linear biochemical chain, we model an RNA structure

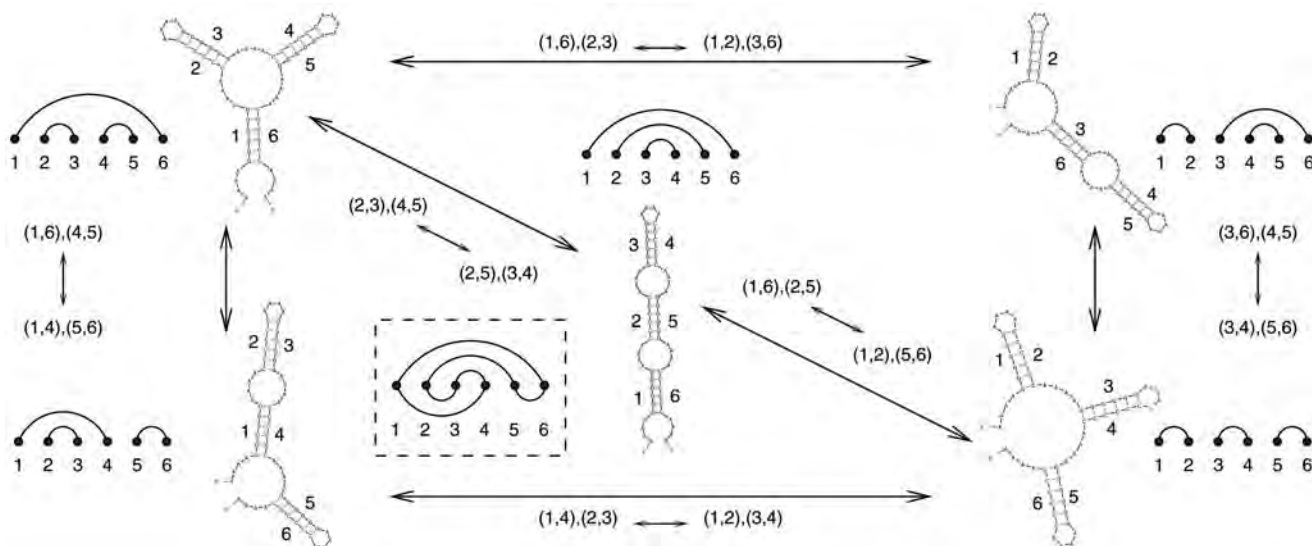


Figure 3. How do branching configurations relate? Five arrangements of loops and helices for the RNA sequence $A^4(G^6 A^4 C^6 A^4)^3$, determined by the pairing of G (labeled 1, 3, 5) and C (labeled 2, 4, 6) segments. Arrows indicate movement between structures under allowable pairing exchanges. Two configurations with a graph geodesic of length 2 comprise a meander; the corresponding noncrossing perfect matchings form a single closed loop when drawn on the same endpoints, one above and the other below, as illustrated in the inset figure. Although these closed meanders arise in various mathematical settings, their exact enumeration problem remains open.

as a plane tree—that is, a rooted tree whose subtrees are linearly ordered, by mapping loops to vertices and helices to edges. In Figures 1 or 3 this would correspond to collapsing each circular region to a point and merging two parallel lines into one. This abstraction preserves the basic structural arrangement, including information about the sequence ordering and the energetic types of different loop structures.

Mathematics is a vital source of new structural insights.

This highlights one of the critical issues: there can be exponentially many different possible low-energy branching configurations for an arbitrary RNA sequence.

Hence, even in a situation, as in Figure 2, where some experimental information about the 3D structure is known, mathematics is a vital source of new structural insights. As we will discuss, by using strings and trees as a combinatorial model of RNA folding, we can analyze different possible branching configurations at viral genome length scales. We prove theorems, using methods from enumerative, probabilistic, and geometric combinatorics, which address questions such as: *What are the trade-offs*

among different types of loop structures? Is there a typical degree of loop branching? What is the dependence on the thermodynamic optimization parameters? Since the accuracy of computational base pairing predictions decays rapidly with sequence length, these theoretical results help separate structural signals from thermodynamic noise, thereby supporting alternative hypotheses in viral capsid assembly.¹

Conversely, under a suitable abstraction, the space of branching configurations reveals significant combinatorial structure. The challenge of understanding the different possible low-energy secondary structures for an RNA sequence motivates a new local move on plane trees/noncrossing perfect matchings. This yields graphs, as in Figure 3, isomorphic to the Hasse diagram for the lattice of noncrossing partitions. By recapitulating this well-known combinatorial structure, we gain insights that now allow us to count and characterize the orbits under the Kreweras complementation operator. This result then naturally leads to considering some new approaches to the challenging open problem of meander enumeration. In this way, we illustrate that the interaction of the two disciplines is fruitful for mathematics as well as beneficial for biology.

¹For more on the mathematics of molecular machines, see “Building Polyhedra by Self-Assembly” by Govind Menon in this issue (page 822).

References

- [1] D. H. LENZ, K. C. MOK, B. N. LILLEY, R. V. KULKARNI, N. S. WINGREEN, and B. L. BASSLER, The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*, *Cell* **117** (2004), no. 1, 69–82.
- [2] L. TANG, K. N. JOHNSON, L. A. BALL, T. LIN, M. YEAGER, and J. E. JOHNSON, The structure of Pariacoto virus reveals a dodecahedral cage of duplex RNA, *Nat. Struct. Biol.* **8** (2001), 77–83.

Image Credits

Figure 1 courtesy of Henke and Bassler, Princeton University.

Figure 2 courtesy of *Nature Structural and Molecular Biology*, Nature Publishing Group, reprinted with permission [2].

Figure 3 courtesy of Christine Heitsch, Georgia Institute of Technology. Secondary structures predicted and drawn by mfold software, available via unafold.rna.albany.edu.

Photo of Christine Heitsch courtesy of James Heitsch, University of Illinois at Chicago.

ABOUT THE AUTHOR

Christine Heitsch is a mathematician with research interests in RNA structure and function. Her work in discrete mathematical biology is supported by a BWF Career Award and an NIH R01. She is also PI on an NSF workforce grant addressing the challenges and opportunities of interdisciplinary career paths.



Christine Heitsch

Jonathan R. Kujawa

ABSTRACT. In joint work with Brian Boe and Daniel Nakano, we provide a concrete description of the spectrum for several tensor triangulated categories which naturally appear in representation theory.

Realizing the Spectrum of Tensor Categories

The usefulness of attaching geometry to algebraic objects goes back at least to Descartes. Using geometry we can obtain qualitative information about our original algebraic object. Given a polynomial with real coefficients, $p(x)$, we teach schoolchildren to look at the graph of $y = p(x)$ in \mathbb{R}^2 . By examining the x-intercepts, y-intercepts, and end behavior they can say things about the degree, leading coefficient, the constant term, and so on.

A more modern example is the prime ideal spectrum, $\text{Spec}(R)$, of a commutative ring R . More generally, given a finitely generated R -module M , we can define the support of M , $\text{supp}(M)$, to be the subset of $\text{Spec}(R)$ consisting of all prime ideals P such that M localized at P does not vanish. The geometry of the spectrum and support again captures algebraic information. For example, for two modules the support of a direct sum is the union of the supports, and the support of the tensor product is the intersection of the supports.

Now suppose C is a category which admits both a direct sum and a tensor product. We also assume the tensor product is “commutative” in that there are canonical isomorphisms $X \otimes Y \cong Y \otimes X$ for all pairs of objects X and Y . For example, C could be $k\text{-vec}$, the category of finite-dimensional k -vector spaces over a fixed ground field k with the usual direct sum and tensor product operations. In this case the canonical isomorphism $V \otimes W \rightarrow W \otimes V$ is given by the “flip” map $v \otimes w \mapsto w \otimes v$. The ground field acts as the identity for the tensor product in that there are canonical isomorphisms $k \otimes V \cong V \cong V \otimes k$. Such tensor categories are common throughout mathematics. Another elementary example is the category of closed, orientable surfaces with direct sum given by disjoint union and tensor product given by connected sum.

Such a category can be thought of as a categorical analogue of a commutative ring, with the direct sum as the “addition” and the tensor product as the “multiplication.” With this in mind it is natural to ask for the notion of an ideal. A *tensor ideal* of C is a full subcategory I which has the property that (1) if M is an object of I and N an

object of C , then $M \otimes N$ is an object of I , and (2) $A \oplus B$ is an object of I if and only if both A and B are objects of I . The second condition is the requirement that I be a “thick” subcategory. Thinking of the kernels of ring homomorphisms, the thick condition becomes plausible once we notice that for any functor of tensor categories, $F : C \rightarrow D$, we have $F(A \oplus B) \cong 0$ if and only if $F(A) \cong 0$ and $F(B) \cong 0$. Finite-dimensional k -vector spaces is a categorical version of a field in that it has no proper ideals. Namely, if V is a nonzero vector space in I , then by writing it as a direct sum of one-dimensional subspaces and using the thick condition, it follows that k lies in I . Therefore every vector space $W \cong W \otimes k$ lies in the tensor ideal I .

Things become more interesting when the category is not semisimple. In this case the structure of the category can be quite complicated. Just classifying the objects is already a hopeless task in all but the easiest examples. A more reasonable goal is to describe the tensor ideals. Doing so gives us an idea of the coarse structure of the category and, in particular, gives information about when one object can be obtained from another by direct sums, direct summands, and tensor products.

In easy cases, the tensor ideals can be described by hand. For example, let k be a fixed ground field which is algebraically closed and of characteristic $p > 0$. Let C_p be the cyclic group of order p and let $C_p\text{-mod}$ be the category of finite-dimensional C_p -modules, that is, finite-dimensional k -vector spaces with a linear action by the elements of C_p . Then $C_p\text{-mod}$ again admits a tensor product. Namely, given C_p -modules M and N , define $M \otimes N$ to be the tensor product as vector spaces with C_p action given by the formula $g.(m \otimes n) = (g.m) \otimes (g.n)$ for $g \in C_p$. There are p nonisomorphic C_p -modules, Q_1, \dots, Q_p , which cannot be written as a direct sum of smaller modules. The dimension of Q_d as a k -vector space is d . In particular, Q_1 is the unique simple C_p -module, and Q_p is the unique projective indecomposable module. In this case, direct calculations show that there are two tensor ideals: the entire category and the full subcategory consisting of projective modules. A more general and much more difficult problem is to classify the tensor ideals of $G\text{-mod}$ when G is any finite group in which p divides the order of G .

Now let \mathcal{K} be a tensor triangulated category consisting of compact objects. That is, \mathcal{K} is a triangulated category with a compatible tensor product and subject to suitable finiteness assumptions. Approximately ten years ago Paul Balmer introduced geometry by defining the *spectrum* of \mathcal{K} in the spirit of commutative ring theory. In this setting a tensor ideal is taken to be a full triangulated subcategory with properties (1) and (2) as above. A tensor ideal I is called *prime* if it is a proper ideal and if whenever $A \otimes B$ is an object of I , either A or B is an object of I . The spectrum of \mathcal{K} , $\text{Spc}(\mathcal{K})$, is then the collection of all prime tensor ideals with the Zariski topology. Balmer also defined the support of any object M in \mathcal{K} as the set of all prime tensor ideals which do not contain M .

Jonathan R. Kujawa is professor of mathematics at the University of Oklahoma. His e-mail address is kujawa@math.ou.edu.

The research of the author was partially supported by NSF grant DMS-1160763 and NSA grant H98230-16-0055.

For permission to reprint this article, please contact: reprint-permission@ams.org.

DOI: <http://dx.doi.org/10.1090/noti1565>

Balmer proved that the spectrum and support for \mathcal{K} are universal in a precise sense among support theories for \mathcal{K} . He also proved that it provides a classification of the thick tensor ideals of

Tensor triangular geometry is a beautiful and powerful theory.

\mathcal{K} and, hence, that it provides a geometric answer to our earlier question. Balmer, collaborators, and others have gone on to show that tensor triangular geometry is a rich theory which brings valuable new tools and insights to a variety of settings. Those who are able to attend the November Western Sectional Meeting at the University of California, Riverside, will have the opportunity to hear this story in person at Balmer's Invited Address "An invitation to tensor-triangular geometry."

Tensor triangular geometry is a beautiful and powerful theory. However, for tensor triangular categories of interest it is desirable to have a concrete description of the spectrum. Such a realization is both useful for applications and to connect it to existing theories. For example, when he introduced tensor triangular geometry, Balmer proved that if \mathcal{K} is the homotopy category of bounded complexes of finitely generated projective R -modules, then the spectrum and support recover $\text{Spec}(R)$ and its support. He also showed that if \mathcal{K} is the stable module category for $G\text{-mod}$ when G is a finite group, then the spectrum and support match the long-studied spectrum of the cohomology ring of G and cohomological support varieties for G -modules. In particular, this recovers the classification of thick tensor ideals in this setting first obtained by Benson-Carlson-Rickard twenty years ago. This shows that tensor triangular geometry encompasses known theories. Additional examples have since been computed. Nevertheless, it remains a challenging problem to give an explicit realization of the spectrum and support for tensor triangulated categories of interest.

In joint work with Brian Boe and Daniel Nakano, we provide a description of the spectrum for several tensor triangulated categories which appear in nature. We give explicit, down-to-earth descriptions of the spectrum for the stable category of finite-dimensional modules for the complex Lie superalgebra $\mathfrak{g}(m|n)$ and for the stable category of finite-dimensional modules for quantized enveloping algebras at a root of unity. The goal of the talk will be to describe these results through a gentle introduction involving plenty of examples.

Photo Credit

Photo of Jonathan Kujawa by Anne Dunn, courtesy of Jonathan Kujawa.

ABOUT THE AUTHOR

When not doing math, **Jon Kujawa** can be found riding his bike over the rolling hills of Oklahoma or enjoying a good book and a good scotch.



Jonathan Kujawa

Govind Menon

Building Polyhedra by Self-Assembly

Gromov begins an interesting—and speculative—recent article [2] with the question, “Is there mathematics in biology?” The answer, I think, is yes, but this is not immediately apparent, since the real underlying question is whether modern biology can inspire new forms of mathematics in a way that compares to the deep ties that bind mathematics and physics. If we believe that an essential aspect of mathematics lies in the discovery of abstract principles from empirical knowledge, there is little doubt that biology today presents us with an abundance of the “raw stuff.” What seems much harder is to process this raw stuff into beautiful mathematics, especially if one begins with the genetic code and the theory of evolution.

The topic of my talk is not true biology, but an instance of “synthetic biology.” All biological organisms build themselves or “self-assemble.” This is, of course, familiar to us from our everyday experience, but my talk will be about much smaller organisms. For the past twenty years, nanotechnologists have been trying to manufacture devices by mimicking biological self-assembly and the exquisite design of molecular machines. The goal of my talk is to advertise one aspect of this rapidly growing field and to explain how an important biological example—the self-assembly of viruses with icosahedral symmetry—can inspire and guide the development of self-assembly in technology.

Viruses are biological organisms that lack the cellular machinery necessary for independent existence. The simplest viruses consist of genomes contained within a protein shield (the capsid). The capsid disassembles when the virus attacks a host cell; the virus genome then hijacks the host cell and uses it to make many more copies of virus genome and proteins, which then rapidly reassemble into new copies of the virus. The natural design of viruses has two elegant features that should appeal to all mathematicians: genetic economy and structural symmetry. The genetic sequences of primitive viruses are very short. For example, the genome of MS2, a well-studied virus, has only 3,569 nucleotides that code for four proteins (lysis, replicase, maturation, and coat protein), each of which has a very specific function. The lysis enzyme degrades the cell wall of the host, and the replicase catalyzes the reproduction of the virus. The other two proteins are used to build the MS2 capsid: it consists of 180 copies of the coat protein, pinned at one end by the maturation protein, in a beautiful arrangement of dimers with icosahedral symmetry (Figure 1). While the genome of MS2 has been

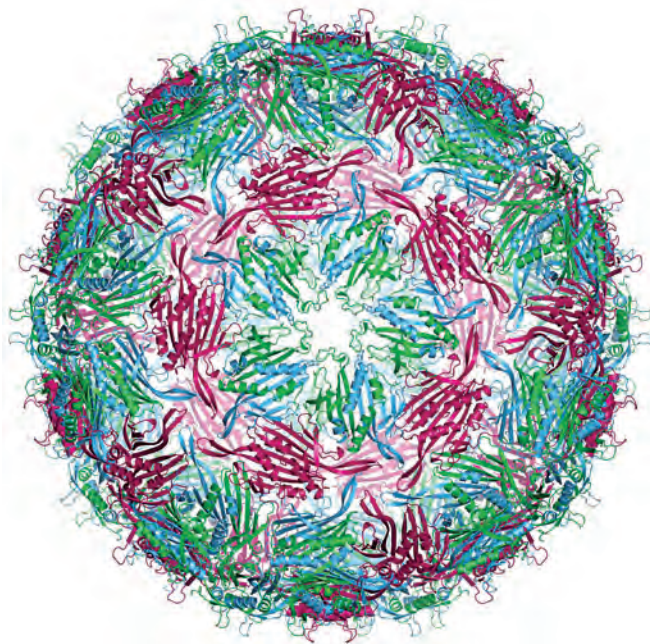


Figure 1. A ribbon-diagram showing the structure of the bacteriophage MS2. The coat protein exists in three distinct conformations (A, B, and C), which merge in pairs into A/B dimers (blue/green) and C/C dimers (maroon). A/B dimers cluster into pentamers around the 5-fold axes of an icosahedron, three alternating A/B and C/C clusters form at the 3-fold axes, and the C/C dimers sit as axes of 2-fold symmetry.

known since the mid-1970s, it is only recently that the intricate combinatorial structure of the co-assembly of the capsid with RNA folding was deciphered by Reidun Twarock and her colleagues [1].¹

The self-assembly of viruses has inspired many examples of synthetic self-assembly. My work has mainly been in collaboration with David Gracias, an experimentalist at Johns Hopkins University. Over the past fifteen years, David has used photolithography to design many devices and containers that fold themselves into a final shape once they are released from a substrate. The devices built in his lab are small (a hair’s width and smaller), but much larger than viruses such as MS2. This allows us to observe the pathways of self-folding, unlike the process of self-assembly of viruses, which must be inferred indirectly (Figure 2).

The unfolding of a polyhedron into a planar net is a classical problem in discrete geometry, and our collaboration began when David asked me what the best net should be for a self-folding dodecahedron. The issue here is a combinatorial explosion. The cube has only 11 nets, each

Govind Menon is professor of applied mathematics at Brown University. His e-mail address is govind_menon@brown.edu.

For permission to reprint this article, please contact: reprint-permission@ams.org.

DOI: <http://dx.doi.org/10.1090/noti1564>

¹For more on connections between combinatorics and molecular biology, see “Strings, Trees, and RNA Folding” by Christine Heitsch in this issue (page 817).

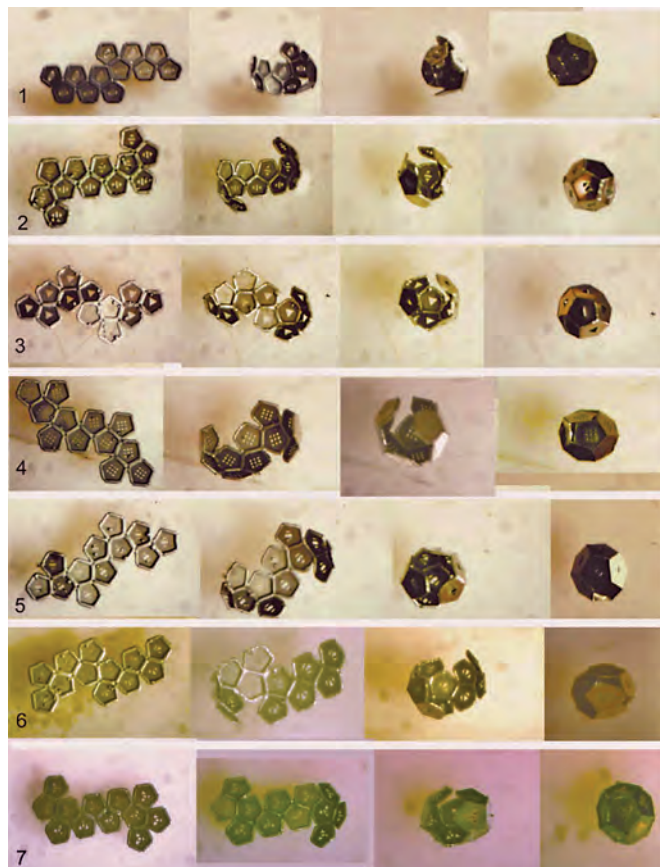


Figure 2. Optical microscope images of surface-tension-driven self-assembly of a dodecahedron from a net. The sides of each face of the dodecahedron are $300\ \mu\text{m}$.

of which may be tested in the lab. However, the dodecahedron has 43,380 nets, and, to my surprise and delight, simple heuristics along with our computations revealed the best nets in the lab [4]. Since then our work has evolved into a study of the pathways of self-assembly [3]. This has required some surprisingly sophisticated mathematics. My current goal is to understand the conformational diffusion of polyhedral linkages. More formally, this involves a rigorous formulation for Brownian motion on algebraic varieties defined by polyhedral linkages, along with effective algorithms for simulation.

References

- [1] E. C. DYKEMAN, P. G. STOCKLEY, and R. TWAROCK, Solving a Levinthal's paradox for virus assembly identifies a unique antiviral strategy, *Proceedings of the National Academy of Sciences* (2014).
- [2] M. GROMOV, Crystals, proteins, stability and isoperimetry, *Bull. Amer. Math. Soc. (N.S.)* **48** (2011), 229–257.
- [3] RYAN KAPLAN, JOSEPH KLOBUŠICKÝ, SHIVENDRA PANDEY, DAVID H. GRACIAS, and GOVIND MENON, Building polyhe-

dra by self-assembly: Theory and experiment, *Artificial Life* (2014). MR 2774091

- [4] S. PANDEY, M. EWING, A. KUNAS, N. NGUYEN, D. H. GRACIAS, and G. MENON, Algorithmic design of self-folding polyhedra, *Proceedings of the National Academy of Sciences* **108** (2011), 19885–19890.

Image Credits

Figure 1 ©Dr Neil Ranson, University of Leeds, UK. Used under the Creative Commons Attribution-Share Alike 3.0 Unported License.

Figure 2 courtesy of Shivendra Pandey.

Photo of Govind Menon courtesy of Govind Menon.



Govind Menon

ABOUT THE AUTHOR

Govind Menon is an applied mathematician at Brown University. He primarily works on the dynamics of disordered systems arising in materials science and physics. This includes models of turbulence, kinetics of phase transitions, and, most recently, random matrix theory and numerical linear algebra.

Kevin M. Pilgrim

Semigroups of Branched Mapping Classes: Dynamics and Geometry

A rational function of a single complex variable defines a continuous map of the Riemann sphere to itself. In the early 1980s, W. Thurston gave a topological characterization of certain rational functions among the much larger set of self-branched coverings of the sphere. Implicit in his development are generalizations of mapping class groups. These generalizations will be the focus of my talk.

Mapping Class Groups

The simplest mapping class group is that of the torus. The mapping class group of the torus T^2 is the group $\text{Mod}(T^2)$ of orientation-preserving self-homeomorphisms of the torus, where two such maps are identified if they are isotopic, i.e. connected by a continuous path of homeomorphisms.

Matrices provide lots of examples. Let $T^2 = \mathbb{R}^2/\mathbb{Z}^2$ be the usual presentation of the torus as a quotient of the plane. Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The linear map $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by $\begin{bmatrix} x \\ y \end{bmatrix} \mapsto A \begin{bmatrix} x \\ y \end{bmatrix}$ descends to a homeomorphism $f : T^2 \rightarrow T^2$ that preserves orientation if and only if it sends \mathbb{Z}^2 onto itself and $\det(A) > 0$. Equivalently, $a, b, c, d \in \mathbb{Z}$ and $\det(A) = 1$, i.e. $A \in \text{SL}_2(\mathbb{Z})$. For example:

- $A = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$. A calculation shows A has order 6. The map f is *periodic*.
- $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Since $A^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$, f has infinite order. The x -axis is an eigenspace of A . Its image under projection to T^2 gives a simple closed curve which is preserved by f . Infinite order maps for which some iterate preserves a curve are called *aperiodic reducible*.
- $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. Again A has infinite order. The eigenspaces have irrational slopes and project to dense subsets of T^2 . If we use suitable local coordinates (u, v) given by the eigenspaces, f looks like $(u, v) \mapsto (\lambda u, \lambda^{-1} v)$ where $\lambda = \frac{3+\sqrt{5}}{2}$ is the larger eigenvalue. No iterate of f fixes a curve. Such f are called *irreducible*.

It turns out that every element of $\text{Mod}(T^2)$ arises in this way: the mapping class group $\text{Mod}(T^2)$ is naturally isomorphic to $\text{SL}_2(\mathbb{Z})$. Put another way, mapping classes of the torus are determined by their induced action on the fundamental group \mathbb{Z}^2 of T^2 .

Kevin Pilgrim is professor of mathematics at Indiana University. His research is partially supported by the Simons Foundation. His e-mail address is pilgrim@indiana.edu.

For permission to reprint this article, please contact: reprint-permission@ams.org.

DOI: <http://dx.doi.org/10.1090/noti1566>

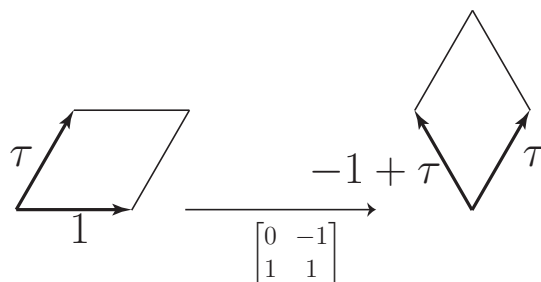


Figure 1. Shown are two fundamental domains for the torus $\mathbb{C}/\langle 1, \tau \rangle$ where $\tau = e^{2\pi i/6}$. The \mathbb{R} -linear map induced by $1 \mapsto \tau$ and $\tau \mapsto -1 + \tau$ is a rotation of order 6 and descends to an isometry on the torus.

Conjugacy

Two elements $f, g \in \text{Mod}(T^2)$ are *conjugate* if $g = h^{-1}fh$ for some $h \in \text{Mod}(T^2)$. When does this happen? Thinking dynamically, two maps f, g are conjugate if they coincide after a change of coordinates via an element h of $\text{Mod}(T^2)$. Properties like being periodic, reducible, or irreducible are thus invariant under conjugacy. The *conjugacy problem* asks, given f, g , can you tell if f and g are conjugate? And if the answer is “yes,” can you produce such an element h ? For $\text{Mod}(T^2)$ the answer is “yes”; the classification of conjugacy classes goes back to Gauss.

Geometrization

There is another way to think about the classification of conjugacy classes, via *geometrization*. Roughly, geometrization is the problem of finding a “nice” or “optimal” geometric structure for some topological object. Think about the (surface of the) bagel you ate this morning. It inherits a metric from usual Euclidean 3-space. That metric has variable curvature. In contrast, the Euclidean metric on the plane \mathbb{R}^2 descends to a metric on T^2 which is *flat*: the curvature is zero everywhere. The map f induced by $A = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$ seems simple enough, but it distorts this Euclidean metric! There are, however, lots of flat metrics on T^2 . They can be fruitfully organized by points in the upper-half-plane $\mathbb{H} \subset \mathbb{C}$. Given $\tau \in \mathbb{H}$, we can identify \mathbb{R}^2 as a real vector space with \mathbb{C} via $(1, 0) \mapsto 1$ and $(0, 1) \mapsto \tau$, and transport the Euclidean metric on \mathbb{C} to one on \mathbb{R}^2 via this identification. If we take the Euclidean metric on \mathbb{C} and take $\tau = \exp(2\pi i/6)$, f becomes an isometric rotation of order 6; see Figure 1.

In general, if $f \in \text{Mod}(T^2)$ and if we equip T^2 with a Euclidean metric corresponding to a point $\tau \in \mathbb{H}$ in this way, then we can transport this metric via f to get another metric on T^2 . This leads to an action of $\text{Mod}(T^2)$ on the space of metrics \mathbb{H} . The action is given by $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \tau \mapsto \frac{a\tau + b}{c\tau + d}$.

Given $\tau_1, \tau_2 \in \mathbb{H}$, there is a natural notion of distance $d(\tau_1, \tau_2)$ between the corresponding flat metrics on T^2 that turns out to be the *hyperbolic metric* on \mathbb{H} ; its

infinitesimal form is given by $ds = \frac{2|d\tau|}{\mathfrak{J}(\tau)}$. This yields a natural invariant of $f \in \text{Mod}(T^2)$ via its *minimum displacement* $\inf\{d(\tau, A.\tau) : \tau \in \mathbb{H}\}$. This can be (1) zero and realized by some fixed-point τ , (2) zero and not realized by any point, or (3) positive and realized by points τ lying along some geodesic that is preserved by the action of f . These correspond to the three cases above. The first and third cases are called *geometrizable*: an optimal metric exists. For reducible cases, we can get close to optimal, but we cannot achieve it.

General Mapping Class Groups

Now suppose S is an arbitrary closed orientable surface and $P \subset S$ is finite. The mapping class group $\text{Mod}(S, P)$ is the group of orientation-preserving homeomorphisms $f : S \rightarrow S$ for which $f(P) = P$, where again two are identified if they are isotopic. Roughly summarizing thousands of pages of hard work by many mathematicians: the flavor of the results for the torus outlined above extend to arbitrary mapping class groups. Elements of $\text{Mod}(S, P)$ are determined by their action on the fundamental group. A geometrizable element is either periodic and represented by an isometry or irreducible and represented by a map which in (mildly singular local) Euclidean coordinates looks again like $(x, y) \mapsto (\lambda x, \lambda^{-1}y)$. A class that is not geometrizable reduces canonically into geometrizable pieces by cutting along some finite invariant collection of pairwise disjoint curves. The conjugacy problem is significantly harder but solvable.

Branched Mappings

We now drop the assumption that f is a homeomorphism and require only that it is a *finite branched covering* unramified away from P . Equivalently, $f : S \rightarrow S - P$ is a covering map of some degree $d \geq 1$; we require still that $f(P) \subset P$. The Riemann-Hurwitz formula implies that if $d \geq 2$, then the surface S is either the torus T^2 or the sphere S^2 and that if S is the torus, the map f is unramified. Composition descends to a well-defined map on isotopy classes, and we obtain a countable semigroup $\text{BrMod}(S, P)$.

Again the case of the torus is instructive. The conjugacy problem in this semigroup boils down to the question: Given a pair of 2 by 2 integral matrices A, B of common determinant larger than one, when is $B = P^{-1}AP$ for some $P \in \text{SL}_2(\mathbb{Z})$? This question is equivalent to the classification of real quadratic fields and is still unsolved.

So I'll focus on the case $S = S^2$. It turns out the case $\#P = 4$ is already interesting.

Complex Dynamics

Complex dynamics provides a natural source of geometrizable elements of $\text{BrMod}(S^2, P)$. Let's identify the Riemann sphere $\hat{\mathbb{C}}$ with S^2 in the usual way via stereographic projection. For a parameter $c \in \mathbb{C}$ let $p_c : \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$ be given by $p_c(z) = z^2 + c$. Dynamics is concerned with *iteration*. We study sequences like $\{z, p_c(z), p_c(p_c(z)), \dots\}$, called the *orbit* of z . On the one hand, an easy exercise shows

that if $|z| > \max\{2, |c|\}$, then this sequence converges to infinity. On the other hand, the quadratic formula shows that there are fixed points. So there is a nonempty compact subset K_c consisting of points z for which the orbit of z is bounded. The boundary J_c of this locus is the typically fractal *Julia set* and is the locus of chaotic behavior. We have $z \in J_c$ if and only if the orbit of z is bounded, but there are arbitrarily small perturbations of z for which the orbit becomes unbounded.

Looking at the polynomials $p_c^n(0) - p_c^m(0)$, $n, m = 0, 1, 2, \dots$, we find there is a countably infinite set of parameter values c for which the forward orbit of the branch point at the origin is finite. These *critically finite* parameters play a crucial role in complex dynamics and provide a rich source of examples of semigroups $\text{BrMod}(S^2, P)$.

An Example

Let $g(z) = z^2 + i$. There are branch points at the origin and the point at infinity where g is locally 2-to-1. The point at infinity ∞ is fixed by g , and $0 \mapsto i \mapsto i - 1 \mapsto -i \mapsto i - 1$. Let's take $P := \{i, i - 1, -i, \infty\}$; note that $g(P) \subset P$. We have just constructed an element of $\text{BrMod}(S^2, P)$.

How might we get other elements? We can do so by *twisting* via pre- and post-composition with elements of $\text{Mod}(S^2, P)$. Let's look at all such elements:

$$\mathcal{H}_g := \{[h_0 \circ g \circ h_1] : h_0, h_1 \in \text{Mod}(S^2, P)\} \subset \text{BrMod}(S^2, P).$$

This set decomposes into conjugacy classes. How many are there? As shown in a groundbreaking paper by L. Bartholdi and V. Nekrashevych, the set of conjugacy classes in \mathcal{H}_g consists of two geometrizable elements classified by polynomials $z^2 \pm i$ and a \mathbb{Z} 's worth of classes represented by reducible elements that fix a curve in a way that provides an obstruction to geometrization.

What if we use a different polynomial? Now let $f(z) = z^2 + c$ where c is the unique parameter for which $\mathfrak{J}(c) > 0$ and the origin is periodic of least period 3; now put $P = \{0, c, c^2 + c, \infty\}$. The polynomial f is called the *Douady Rabbit* polynomial, since its Julia set (shown in the left-hand side of Figure 4) looks a bit like a rabbit. Now the corresponding set of conjugacy classes consists of just three geometrizable elements, classified by the three values of c for which the origin has period 3 under iteration of $z^2 + c$.

G. Kelsey and R. Lodge have just announced an extension of these results to all quadratics with $\#P = 4$.

Highlights

There are several results about the semigroup $\text{BrMod}(S^2, P)$ that mirror results for the mapping class group $\text{Mod}(S, P)$.

A first point—one exploited by L. Bartholdi and V. Nekrashevych—is that branched mapping classes are faithfully encoded by algebraic data known as *wreath recursions*; like induced maps on fundamental groups, they are well defined up to conjugacy. As an illustration, equation (1) gives wreath recursion data for the Rabbit

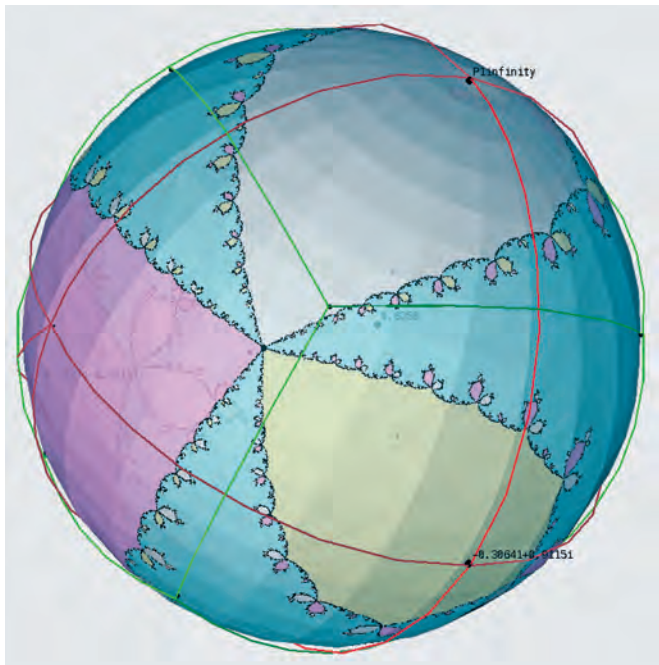


Figure 2. A computer program reads the group-theoretic data faithfully encoding Douady's Rabbit. It produces a numerical approximation for the polynomial map. It then draws an approximation to its fractal Julia set on the sphere.

polynomial $f(z)$; here a, b, c are free generators for the fundamental group of $S^2 - P$. The upshot is that we can represent and compute with elements of $\text{BrMod}(S^2, P)$. A program by L. Bartholdi takes this algebraic data as input and returns a numerical approximation for the rational geometrization (if it exists) and an approximation of its associated Julia set; see Figure 2:

$$(1) \quad \begin{aligned} a &= \langle a^{-1}b^{-1}, cba \rangle(12), \\ b &= \langle a, 1 \rangle, \\ c &= \langle b, 1 \rangle. \end{aligned}$$

A second point is that our understanding of the corresponding geometrization questions is advancing rapidly. In our setting, typical geometrizable elements of degree larger than one are represented by maps which are expanding in a suitable sense. Critically finite polynomials and rational maps in this sense are expanding. However, there do exist nonrational expanding maps. L. Bartholdi and D. Dudko give a characterization of expanding classes. Recently N. Selinger and M. Yampolsky and independently L. Bartholdi and D. Dudko have shown that a general map decomposes algorithmically into geometrizable pieces.

A third point is that the semigroup $\text{BrMod}(S^2, P)$ acts naturally on the space of conformal structures on (S^2, P) . When $\#P = 4$ this is again the hyperbolic plane \mathbb{H} . Figure 3 shows the limit set of the action of the subsemigroup of

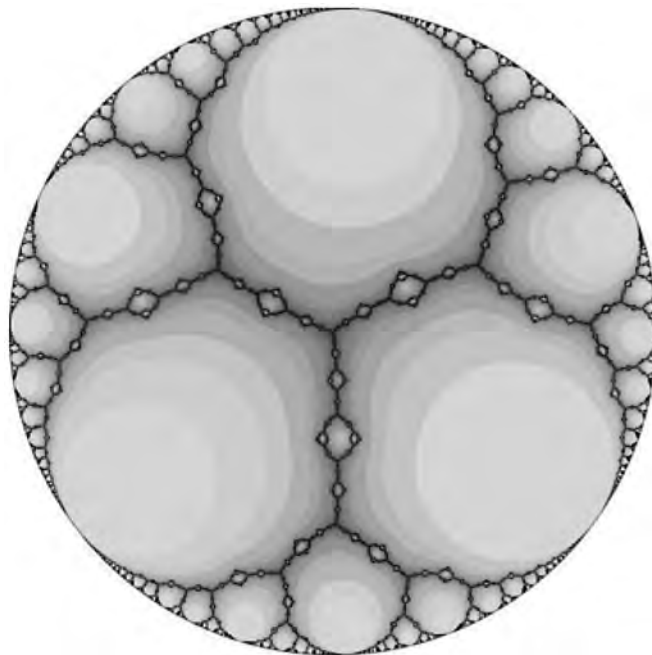


Figure 3. The limit set of the subsemigroup of $\text{BrMod}(S^2, P)$ generated by $\text{Mod}(S^2, P)$ and the Douady Rabbit polynomial $f(z) = z^2 + c$, drawn in the disk model.

$\text{BrMod}(S^2, P)$ generated by the Rabbit polynomial f and the group $\text{Mod}(S^2, P)$.

Finally, the subject is immensely rich. In joint work with W. Floyd, G. Kelsey, S. Koch, R. Lodge, W. Parry, and E. Saenz, we systematically study elements of $\text{BrMod}(S^2, P)$, $\#P = 4$, for which the branch points are simple. Such maps turn out to be closely related to affine torus maps, and this observation leads to them being computationally tractable. A database of tens of thousands of examples is now online at www.math.vt.edu/netmaps/index.php. Figure 4 illustrates the circle of ideas. The first two figures (top left) show respectively the fractal Julia set for the

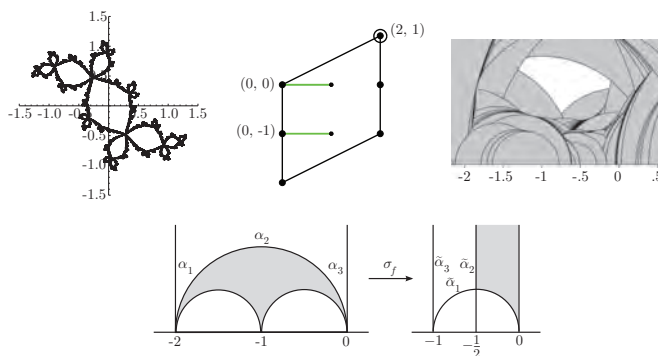


Figure 4. Follow the Rabbit from its fractal Julia set to its induced action on the upper-half-plane.

Rabbit polynomial f and a diagram encoding the relation between f and an affine torus map. The (right top) figure deals with the geometry of the induced map $\sigma_f : \mathbb{H} \rightarrow \mathbb{H}$: there is a unique fixed point in the white region. The final pair (bottom) illustrates that σ_f maps an ideal triangle to a Schwarz triangle and is extended by reflection.

Image Credits

Figure 1 courtesy of Kevin M. Pilgrim.

Figure 2 courtesy of L. Bartholdi.

Figure 3 courtesy of S. Koch.

Figure 4 courtesy of W. Floyd et al.

Photo of Kevin M. Pilgrim courtesy of Indiana University College of Arts and Sciences.

ABOUT THE AUTHOR

Kevin Pilgrim studies conformal dynamical systems, broadly interpreted. He is grateful to the Simons Foundation and to the AIM SQuaRE program for support.



Kevin M. Pilgrim

Bruce Sagan

The Protean Chromatic Polynomial

I am very excited to have the opportunity to share some of the ideas surrounding one of my favorite objects in combinatorics, the chromatic polynomial, during my Invited Address. Let me start by defining each of the terms in my title.

The Merriam-Webster Dictionary defines “protean” as “of or resembling Proteus in having a varied nature or ability to assume different forms.” In Greek mythology, Proteus was one of the gods of the sea and thus was associated with its constantly changing nature. In a similar manner, the chromatic polynomial gives one information about many things which, a priori, have nothing to do with its original purpose, as described below.

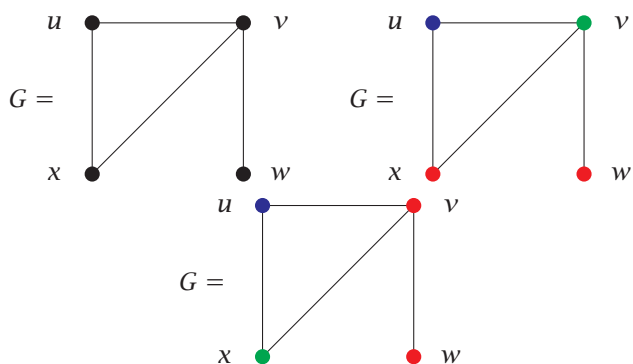


Figure 1. A graph and two colorings

“Chromatic” refers to color, and our general topic is the coloring of the vertices of a graph. A (*combinatorial*) graph, G , consists of a set of vertices V and a set of edges E which connect pairs of vertices. For example, the graph in the upper left in Figure 1 has vertex set $V = \{u, v, w, x\}$ and edge set $E = \{uv, ux, vx, vw\}$. A *coloring* of G is a function $c : V \rightarrow S$ where S is called the *color set*. The coloring is *proper* if the endpoints of every edge have different colors. The coloring in the upper right of Figure 1 is proper, while the one on the bottom is not because the edge $e = vw$ has the same color on both endpoints. The *chromatic number*, $\chi(G)$, is the smallest number of colors needed to properly color G . The graph in Figure 1 has $\chi(G) = 3$ since the upper right image exhibits a proper coloring with three colors, and the triangle uvx cannot be colored with fewer colors. Maybe the most famous theorem in graph theory is the Four Color Theorem, which states that if a graph is planar (can be drawn in the plane without edge crossings), then $\chi(G) \leq 4$. This statement was a conjecture for over a hundred years until

Bruce Sagan is professor of mathematics at Michigan State University. His e-mail address is sagan@math.msu.edu.

For permission to reprint this article, please contact: reprint-permission@ams.org.

DOI: <http://dx.doi.org/10.1090/noti1568>

it was finally proved by Wolfgang Haken and Kenneth Appel in 1976. Their proof caused quite a stir in the mathematical community, because it was the first to use a substantial amount of computing time, and the large number of cases could not all be checked by hand.

The chromatic polynomial was introduced in 1912 by George Birkhoff as a possible tool for proving the then Four Color Conjecture. Although it did not turn out to be useful for the eventual proof, it has more than justified its existence through its many other applications. Let t be a nonnegative integer. The *chromatic polynomial*, $P(G; t)$, is the number of proper colorings $c : V \rightarrow \{1, 2, \dots, t\}$. It is not apparent at first blush why this cardinality should be called a polynomial. However, this will become clearer if we compute $P(G; t)$ for the graph in Figure 1. Suppose we color the vertices in the order u, v, w, x . Then there are t choices for the color of u since it is the first vertex to be colored. After that, there will be $t - 1$ choices for the color of v , since it cannot be the same color as u . Similar reasoning shows that there are $t - 1$ choices for w . Finally, x is adjacent to both u and v , and these two vertices have different colors, so we can color x in $t - 2$ ways. The net result is that

$$P(G; t) = t(t - 1)^2(t - 2) = t^4 - 4t^3 + 5t^2 - 2t,$$

which is a polynomial in t , the number of colors!

One can show that $P(G; t)$ is always a polynomial in t and give nice characterizations of its degree, coefficients, and other properties. Furthermore, it has connections with many other objects of study, including acyclic orientations of graphs, hyperplane arrangements, and even Chern classes in algebraic geometry. I will explain these during my lecture, as well as present some recent work with Joshua Hallam and Jeremy Martin relating $P(G; t)$ to yet another graphical concept, increasing spanning forests. If you are at the Buffalo AMS Sectional Meeting in September, I hope to see you at my talk.

Image Credits

Figure 1 courtesy of Bruce Sagan.

Photo of Bruce Sagan by Robert Chandler, courtesy of Bruce Sagan.

ABOUT THE AUTHOR

Bruce Sagan does research in enumerative, algebraic, and geometric combinatorics. He is best known for his book *The Symmetric Group*. He has twice received the MSU mathematics department’s Frame Teaching Excellence Award. When not doing mathematics, Bruce plays folk music. The photograph shows him playing the Swedish nyckelharpa.



Bruce Sagan