Research statement of David Galvin, July 13, 2019

Department of Mathematics, University of Notre Dame

My research is in discrete probability, combinatorics and graph theory, focusing on enumerative problems, set partitions, and applying combinatorial ideas to understanding phase transitions in statistical physics models and long-range correlations in large random structures.

"Enumerative problems" here means "counting": I study families of sets, often coming from fundamental problems in statistical physics and theoretical computer science, and ask "how large is each set?" and "how does this answer change, as the underlying parameters that define the family change?" A solution to the latter question often gives valuable information. The rate at which sets grow as their underlying parameters grow can, for example, be a measure of the complexity of the problem giving rise to the sets — consider how the running time of a program increases as the length of the input does, with linear or quadratic growth leading to a program that may be effective even for large inputs, while larger growth rates may result in a running time that quickly becomes prohibitive.

Moreover, solutions to enumerative problems often give important insight into the structures being counted. Indeed, to count a set it is often necessary first to understand what a "typical" element looks like, or to break the set into clusters with all elements in a cluster having a similar appearance. Such information about the structure of a set can be extremely helpful when developing efficient schemes for generating a random element, as one might want to do when, for example, testing the plausibility of some hypothesis about the set, when it is computationally unfeasible to examine all the elements. More generally enumeration is intimately related to discrete probability — understanding probabilities in large discrete structures often amounts to solving (suitably weighted) enumeration problems.

Enumerative problems are sometimes solved not by direct counting, but via a bijection: a one-to-one correspondence between a set A of unknown size, and another set B of known size. This establishes that the size of A is the same as that of B; but moreover, the fine details of the bijection often provide valuable insights into the structures of both sets. Closely related to bijections are combinatorial proofs. When an answer to an algebraic or analytic problem turns out to be a whole number, it is natural to ask whether this is co-incidental, or whether the answer can be understood as the size of a certain set, whose existence and description may shed new light on the original problem. The search for combinatorial interpretations of solutions to seemingly non-combinatorial problems is both challenging and rewarding.

My recent (post 2012) work in these areas has fallen broadly into three overlapping areas — generalized Stirling numbers, enumeration of H-colorings of graphs, and understanding phase transitions in statistical physics problems. I'll discuss these in some more detail now.

Stirling numbers. Stirling numbers are among the most studied combinatorial sequences. Their importance stems in part from the fact that they can be described combinatorially (the Stirling number of the second kind ${n \brace k}$, for example, counts partitions of $\{1,\dots,n\}$ into k non-empty blocks) as well as admitting algebraic descriptions (for example, ${n \brace k}$ appears in the change-of-basis matrix between the standard basis for polynomials and the falling powers basis). This makes Stirling numbers a natural bridge between combinatorics and algebra.

Via its combinatorial description, ${n \brace k}$ generalizes naturally to graphs: for a graph G the graph Stirling number ${G \brace k}$ counts partitions of the vertex set of G into k non-empty independent sets (sets of mutually non-adjacent vertices). Think of G as encoding partition restrictions: vertices adjacent in G may not lie in the same block of a partition. This notion goes back to Tomescu (1971) and has been taken up by many researchers. If G has no edges — no partition restrictions — then ordinary Stirling numbers are recovered.

In seminal work Harper (1967) found that the sequence $\left(\left\{{n \atop k}\right\}\right)_{k\geq 0}$ is asymptotically normal (its histogram quantifiably approaches the standard normal density as n grows). Asymptotic normality has long played an important role in the theory of enumeration. With undergraduate Do Trong Thanh [24] and later in [31], we substantially extended Harper's result, finding that when G_n is any n-vertex acyclic graph then $(\left\{{G_n \atop k}\right\})_{k\geq 0}$ is asymptotically normal. This is not yet the end of the story, and characterizing those (families of) graphs G_n for which $(\left\{{G_n \atop k}\right\})_{k\geq 0}$ is asymptotically normal is an active research goal.

Scherk (1823) found that Stirling numbers arise naturally in the Weyl algebra on alphabet $\{x, D\}$ with relation $Dx = xD + 1$, via the identity $(xD)^n = \sum_{k=0}^{n} \left\{{n \atop k}\right\} x^k D^k$. The right-hand side here is the normal order of the word $(xD)^n$. The Weyl algebra appears in quantum mechanics, and has received attention both from physicists and combinatorialists. A focus has been on understanding the coefficients of the normal order of an arbitrary word in the algebra. Much of the literature has tackled the problem algebraically; with Engbers and graduate student Justin Hilyard [28] we found a complete and simple combinatorial interpretation that directly extends the standard interpretation of $\left\{{n \atop k}\right\}$: for every word w there's a graph G_w such that the normal order coefficients of w are the graph Stirling numbers of G_w.

Many authors have considered Stirling numbers with restrictions on the allowed sizes of blocks. In particular r-restricted numbers — allowing only blocks of size $\leq r$ — have received much attention. These are understood algebraically, but a fascinating combinatorial question remains. The matrix $\left[\left\{{n \atop k}\right\}\right]_{n,k\geq 0}$ of Stirling numbers has an inverse with entries whose signs form a predictable checkerboard pattern, and there is a classical explanation for this, via combinatorics — the inverse entries count, up to predictable sign, particular sets. Choi et al. (2006) asked if inverses of matrices of r-restricted numbers can be similarly combinatorially interpreted. This is easy for $r = 2$, but for $r \geq 3$ they identified a significant obstacle.

With Engbers and Smyth, we took up the challenge, and have made significant progress. In [37] we showed that for every restriction, the inverse of the associated matrix can be combinatorially interpreted, as the difference in size between two explicitly defined sets. For many restrictions (including r-restricted numbers for all even r) we went further, explaining each inverse entry as, up to predictable sign, the size of a single set. The question of when we can get this stronger interpretation turns out to hinge on an unexpected problem from analysis: for which $R \subseteq \mathbb{N}$ with $1 \in R$, $0 \notin R$ does the "thinned" exponential series $\sum_{n\in R} x^n/n!$ have a compositional inverse with alternating power series? For all R for which we know that the inverse is alternating, we know this only through a combinatorial (bijective) argument, not an analytic one. Characterizing all such R is an ongoing project with Engbers and Smyth.

A recent project in the realm of generalized Stirling numbers involves total non-negativity of matrices — the property that all determinants of square submatrices are non-negative. This is an important concept connecting aspects of algebra, analysis and combinatorics. Many combinatorial matrices, such as those of binomial coefficients and Stirling numbers, are known to be totally non-negative. With graduate student Adrian Pacurar [40] we placed these examples in a common framework (in terms of polynomial change-of-basis matrices), and found a substantial new family of totally non-negative matrices, including matrices of rook numbers of Ferrers boards, and of graph Stirling numbers of chordal graphs. Our approach involved adding a novel twist to the standard machinery of weighted planar networks. We are actively seeking a complete characterization of totally non-negative change-of-basis matrices.

H-coloring graphs. An important notion linking problems in combinatorics, statistical physics and theoretical computer science is that of an H-coloring of a graph — an adjacency-preserving function between vertex sets of graphs. H-colorings directly generalize two central notions in graph theory: that of an independent set, and that of a proper coloring.

The extremal question — maximizing/minimizing given parameters over given families of graphs — is fundamental in graph theory. The extremal question for H-colorings — for graph H and family $\mathcal{G}$, which G in $\mathcal{G}$ admit the most/fewest H-colorings? — has been an active and fruitful one, dating back to work of Wilf and Linial (1980's).

Generalizing a result of Kahn (2001), in older work with Tetali [6] we settled the maximization question for all H and for $\mathcal{G}$ the family of regular bipartite graphs, and posed the question of what happens when the biparticity hypothesis is dropped. This has generated great interest — at least 16 papers with 18 distinct authors — and has been resolved in numerous special cases, though many questions remain. My recent contributions include [23] (giving, inter alia, asymptotic bounds on the maximum number of H-colorings of regular graphs, for every H), [29] (significantly improving these bounds in the special case of proper coloring, a problem spectacularly resolved by Sah et al. (2018)), and [27] (an expository paper explaining the powerful entropy method, and highlighting open questions in the area). My undergraduate senior thesis student Luke Sernau has also contributed in a sole-authored paper, settling some special cases and providing constructions to refute a conjecture of mine.

In older work [18] I introduced the question of maximizing the count of independent sets in graphs with given minimum degree, and this has since been the subject of at least 5 papers with 11 distinct authors, the latest of which, by Gah et al. (2015), gave a complete resolution. My contribution has included joint work with Engbers [26, 34], settling the natural conjecture for small values of the minimum degree, and broadening the problem to general H-colorings with connectivity restrictions. Focussing on $\mathcal{G}$ the family of trees we found the minimizing tree for a large family of H's, and gave a new proof of an old result of Siderenko fully solving the maximization problem for every H. We then extended our proof methods to make significant inroads on the problem of maximizing the count of H-colorings of 2-connected graphs. In this broader formulation of the problem, many interesting questions remain.

When an exact solution to an enumeration problem is not feasible, it is natural to consider the broad "shape" of the space of solutions, as the underlying parameter(s) of the problem vary. There is a vast literature on the important role in combinatorics played by "global" properties of sequences, such as unimodality, log-concavity and asymptotic normality. For example, Alavi et al. (1987) asked whether the independent set sequence of any tree (the sequence whose kth term is the number of independent sets of size k) is unimodal — rises monotonely to a single peak and then falls. This fundamental question has attracted much attention but remains stubbornly open. With graduate student Justin Hilyard [35], we found many families of recursively defined trees for which the answer is yes; in follow-up work in preparation we are settling the question for some non-recursively defined families.

Alavi et al. showed that, unlike for trees, the independent set sequence of general graphs can be far from unimodal, exhibiting an arbitrary pattern of rises and falls. They posed an extremal question: how large are the graphs needed to exhibit every pattern? With undergraduates Katie Hyry and Kyle Weingartner, and graduate student Taylor Ball [42], we settled this question completely; we also answered a question of Alavi et al. on restrictions on the pattern of rises and falls that can be exhibited by the matching sequence of a graph.

Statistical physics. H-colorings are intimately related to hard-constraint spin models, an important class of statistical physics models describing random structures determined by local rules. The vertices of a graph G are sites occupied by particles, with edges representing pairs of bonded sites. Vertices of H are spins that a particle may have, and edges of H encode restrictions: a pair of spins can span a bond only if they are adjacent in H. A spin configuration on G is exactly an H-coloring. Important examples include the hard-core model (random independent sets) and the Potts model (random proper colorings).

For the Potts model, Kotecký (1985) conjectured that for all $q \geq 3$ the model on $\mathbb{Z}^d$ exhibits phase coexistence. Combinatorially this says that the set of 3-colorings of the lattice breaks into 6 clusters characterized by one (of 3) colors being dominant on one (of 2) partition classes of the lattice. Using sophisticated probabilistic and combinatorial tools, and building on earlier work with Kahn and Randall [5, 10], with Kahn, Sorkin and Randall [30] we recently settled this conjecture for $q = 3$, the first significant progress on Kotecký's conjecture (which has since been fully resolved by Peled and Spinka, building partly on our work). Some of these tools, due to Sapozhenko, are described in my expository note [36].

The existence of phase transition for the hard-core model on $\mathbb{Z}^d$ for sufficiently large values of its density parameter — combinatorially saying that a typical dense independent set from the lattice lies mostly in one of the two partition classes — was shown by Dobrushin (1968), but there was no good information on the (conjectural) phase transition point (the critical density at which this dichotomy first occurs) until my thesis work with Kahn [5]. This only gave useful information in high dimensions. In $\mathbb{Z}^2$ Sinclair et al. (2016), building on extensive previous work, gave a lower bound of around 2.5 on the critical density (the conjectured value is around 3.9). With Blanca, Tetali and Randall [25] we made a serious attack on the problem from the other direction, and obtained the first comparable upper bound on the phase transition point. With undergraduate Ethan Chen [39] we used tools from the theory of self avoiding walks to obtain an improved upper bound, of around 5.3. Developing combinatorial and discrete probabilistic tools to pin down the critical density in both low and high dimensions, and more fundamentally to establish its existence, is a major research goal.

David Galvin's bibliography (numbering same as on CV)

[1] D. Galvin, Two problems involving the notion of phase transition, *ProQuest Dissertations and Theses* (2002) (Ph.D. thesis).

[2] D. Galvin, On homomorphisms from the Hamming cube to $\mathbb{Z}$, *Israel J. Math.* **138** (2003), 189–213.

[3] D. Galvin, Entropy and graph homomorphisms, *Oberwolfach Rep.* **1** (2004), 30–32.

[4] D. Galvin and P. Tetali, Slow mixing of Glauber dynamics for the hard-core model on the hypercube, *Proc. ACM–SIAM SODA* (2004), 459–460.

[5] D. Galvin and J. Kahn, On phase transition in the hard-core model on $\mathbb{Z}^d$, *Combin. Probab. Comput.* **13** (2004), 137–164.

[6] D. Galvin and P. Tetali, On weighted graph homomorphisms, *DIMACS Ser. Discrete Math. and Theoret. Comput. Sci.* **63** (2004), *Graphs, Morphisms and Statistical Physics*, 97–104.

[7] R. D'Souza, D. Galvin, C. Moore and D. Randall, Global connectivity from local geometric constraints for sensor networks with various wireless footprints, *Proc. ACM–IEEE IPSN* (2006), 19–26.

[8] D. Galvin and P. Tetali, Slow mixing of Glauber dynamics for the hard-core model on regular bipartite graphs, *Random Structures Algorithms* **28** (2006), 427–443.

[9] D. Galvin, Bounding the partition function of spin systems, *Electron. J. Combin.* **13** (2006), #R72.

[10] D. Galvin and D. Randall, Torpid Mixing of Local Markov Chains on 3-Colorings of the Discrete Torus, *Proc. ACM–SIAM SODA* (2007), 376–384.

[11] D. Galvin, Sampling 3-colourings of regular bipartite graphs, *Electron. J. Probab.* **12** (2007), 481–497.

[12] D. Galvin, Sampling independent sets on the discrete torus, *Random Structures Algorithms* **33** (2008), 356–376.

[13] T. Carroll, D. Galvin and P. Tetali, Matchings and Independent Sets of a Fixed Size in Regular Graphs, *J. Combin. Theory Ser. A* **116** (2009), 1219–1227.

[14] D. Galvin, An upper bound for the number of independent sets in regular graphs, *Discrete Math.* **309** (2009), 6635–6640.

[15] D. Galvin, A threshold phenomenon for random independent sets in the discrete hypercube, *Combin. Probab. Comput.* **20** (2011), 27–51.

[16] D. Galvin and Y. Zhao, The number of independent sets in a graph with small maximum degree, *Graphs Combin.* **27** (2011), 177–186.

[17] D. Galvin, F. Martinelli, K. Ramanan and P. Tetali, The multi-state hard core model on a regular tree, *SIAM J. Discrete Math.* **25** (2011), 894–916.

[18] D. Galvin, Two problems on independent sets in graphs, *Discrete Math.* **311** (2011), 2105–2112.

[19] P. Cholak, D. Galvin and R. Solomon, Reverse Mathematics and infinite traceable graphs, *Math. Log. Q.* **58** (2012), 18–28.

[20] J. Engbers and D. Galvin, H-colouring bipartite graphs, *J. Combin. Theory Ser. B* **102** (2012), 726–742.

[21] D. Galvin, The independent set sequence of regular bipartite graphs, *Discrete Math.* **312** (2012), 2881–2892.

[22] J. Engbers and D. Galvin, H-coloring tori, *J. Combin. Theory Ser. B* **102** (2012), 1110–1133.

[23] D. Galvin, Maximizing H-colorings of a regular graph, *J. Graph Theory* **73** (2013), 66–84.

[24] D. Galvin and D. T. Thanh, Stirling numbers of forests and cycles, *Electron. J. Combin.* **20** (2013), #P73.

[25] A. Blanca, D. Galvin, D. Randall and P. Tetali, Phase Coexistence and Slow Mixing for the Hard-Core Model on $\mathbb{Z}^2$, *Lecture Notes in Comput. Sci.* **8096** (*Proc. APPROX/RANDOM 2013*) (2013), 379–394.

[26] J. Engbers and D. Galvin, Counting independent sets of a fixed size in graphs with a given minimum degree, *J. Graph Theory* **76** (2014), 149–168.

[27] D. Galvin, Three tutorial lectures on entropy and counting (expository article), arXiv:1406.7872.

[28] J. Engbers, D. Galvin and J. Hilyard, Combinatorially interpreting generalized Stirling numbers, *European J. Combin.* **43** (2015), 32–54.

[29] D. Galvin, Counting colorings of a regular graph, *Graphs Combin.* **31** (2015), 629–638.

[30] D. Galvin, J. Kahn, D. Randall and G. Sorkin, Phase coexistence and torpid mixing in the 3-coloring model on $\mathbb{Z}^d$, *SIAM J. Discrete Math.* **29** (2015), 1223–1244.

[31] D. Galvin, Asymptotic normality of some graph sequences, *Graphs Combin.* **32** (2016), 639–647.

[32] J. Carraher, D. Galvin, S. Hartke, J. Radcliffe and D. Stolee, On the independence ratio of distance graphs, *Discrete Math.* **339** (2016), 3058–3072.

[33] D. Chakrabarty and D. Galvin, Guest editors' foreword, *Theory Comput.* **12** (2016), *Special issue: APPROX-RANDOM 2014*, Article 8.

[34] J. Engbers and D. Galvin, Extremal H-colorings of trees and 2-connected graphs, *J. Combin. Theory Ser. B* **122** (2017), 800–814.

[35] D. Galvin and J. Hilyard, The independent set sequence of some families of trees, *Australas. J. Combin.* **70** (2018), 236–252.

[36] D. Galvin, Independent sets in the discrete hypercube (expository article), arXiv:1901.01991.

[37] J. Engbers, D. Galvin and C. Smyth, Restricted Stirling and Lah number matrices and their inverses, *J. Combin. Theory Ser. A* **161** (2019), 271–298.

[38] T. Carroll and D. Galvin, The game of plates and olives, *Electron. J. Combin.* **26** (2019), P1.18.

[39] A. Blanca, Y. Chen, D. Galvin, D. Randall and P. Tetali, Phase Coexistence for the Hard-Core Model on $\mathbb{Z}^2$, *Combin. Probab. Comput.* **28** (2019), 1–22.

[40] D. Galvin and A. Pacurar, Total non-negativity of some combinatorial matrices, submitted to journal Sept. 2018, revision submitted June 2019, arXiv:1807.08658.

[41] A. Chernikov, D. Galvin and S. Starchenko, Cutting lemma and Zarankiewicz's problem in distal structures, submitted to journal Aug. 2018, arXiv:1612.00908.

[42] T. Ball, D. Galvin, K. Hyry and K. Weingartner, Independent set and matching permutations, submitted to journal Feb. 2019, arXiv:1901.06579.