Research Statement

Wesley Pegden*

November 2, 2009

My research is in the broad area of Discrete Mathematics, and has included research in graph theory, convex geometry, and combinatorial game theory. Topics of successful research have included:

- A new kind of probabilistic approach to combinatorial games, using a 'Lefthanded' generalization of the Local Lemma to show the existence of winning strategies in sequence games [21]
- Optimum density constructions of k-chromatic-critical triangle-free graphs ($k \geq 6$), giving the first natural families of graphs for which we can determine the optimum density of critical members [17]
- (with J. Beck and S. Vijay) Exponential lower bounds on the Hales-Jewett number and the 'halving' Hales-Jewett number, which imply the existence of infinitely many Tic-Tac-Toe games which cannot be solved by Ramsey theory [7]
- A surprising characterization of Euclidean shapes resilient to a simple erosion operation, including certain convex sets and 'fractal-like' shapes [20]
- A finite goal set in the plane, which a first player cannot build *before* a perfectly-playing second player in a Euclidean game; this is in contrast to a surprising theorem of Beck which asserts that *any* finite goal set can be built by the first player (not necessarily before the second player) [19]
- The expansion of vertex-transitive graphs of infinite degree, which, unlike in the case of finite-degree, can be shown to be unimodal. [18]

I discuss my research in each of these areas in the sections below, outlining some directions for future research that I find particularly interesting.

Thue-type games and an ordered extension of the Local Lemma [21]
One of the biggest surprises of combinatorial game theory is the role of randomness. It seems that for many games, gameplay by perfect players has unexpected connections with random play. The theory of positional games pioneered by Beck [5] takes advantage of this connection by 'derandomizing' the probabilistic tools of extremal combinatorics to prove the existence of deterministic winning strategies in a surprising variety of settings. In spite of the many outstanding successes of this theory, however, significant gaps remain; the most prominent is the lack of a game-theoretic

*Department of Mathematics, Rutgers University (New Brunswick), 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019. Email: pegden@math.rutgers.edu

analog of the Lovász Local Lemma (this is the subject of Beck's Neighborhood Conjecture, the major open problem of this derandomized probabilistic theory of combinatorial games). In my research on nonrepetitive sequence games, however, I have shown that an appropriate generalization of the Local Lemma can sometimes be applied *directly* to games (the proofs remain probabilistic, there is no 'derandomization').

The study of nonrepetitive sequences was begun by Axel Thue, who in 1906 constructed an infinite ternary sequence 2102012101202... with no adjacent identical blocks [24]. Notice that there can be no such binary sequence: after 010 (or 101) both possible continuations introduce a repetition. Thue's result, discovered independently by many researchers (due in part to its many surprising applications), is the beginning of the study of nonrepetitive sequences, an area of research with many possible directions. One such direction is the subject of a theorem of Beck (1981) [4], which essentially asserts that if we neglect short blocks, then even in a binary sequence, we can ensure that long identical blocks are not only nonadjacent, but exponentially far apart.

Beck's result is proved probabilistically, using the asymmetric version of the Lovász Local Lemma to show that a random finite sequence has positive probability of having no too-close identical long blocks, and using compactness to extend this to an infinite sequence. The proof is thus quite nonconstructive, and no explicit examples of such sequences are known, even if we relax the requirement of exponential distance to anything superlinear.

My recent research concerns game-theoretic generalizations of Beck's theorem and other classical results on nonrepetitive sequences. Imagine that I want to build a sequence which is nonrepetitive in some sense, but only get to choose every other digit, while an adversary gets to choose those in between (our turns alternate, so that a player's choice of a particular digit can be made based on all the digits coming before it). I prove in [21] that, for example, it is possible for Player 1 to ensure exponential distance between long identical blocks in this situation, giving a game-theoretic analog to Beck's theorem. In fact, even if the underlying game is biased, in the sense that my adversary gets to make several moves between each of mine, I can still ensure exponential distance, the base of the exponent decreasing as the bias of the game is increased. In a certain sense, these results show that even though we know of no explicit sequences of the type guaranteed to exist by Beck's theorem, very robust construction strategies to produce them do exist.

The most surprising aspect of these game-theoretic results is that they are proved probabilistically, using a Local-Lemma based approach: let Player 1 play randomly against any fixed strategy of Player 2, and show that the probability that Player 1 prevails in a finite version of the game is positive. Attempts to apply the classical Local Lemma in this setting fail, however, for the same reason that the Local Lemma has historically not been applied to any games: the unknown strategy of Player 2 introduces a mess of dependencies which removes any possibility of demonstrating the kind of "only local dependence" criterion needed to apply the Local Lemma. In a nutshell, how can one argue that events in a game are independent (when Player 1 is playing randomly), when Player 2's moves that influence each event are allowed to depend on previous moves in the game according to some arbitrary fixed strategy? For example, given any fixed pair of intervals, it seems unavoidable that the corresponding bad event (that the intervals get colored identically) will depend on *all* events corresponding to pairs coming later in the sequence, since these are affected by Player 2's strategy, which is allowed to depend on the previous moves made in the game. The most problematic dependencies are all 'one directional', however; dependencies on earlier intervals can be handled in a conventional way by making use of the 'lopsided' version of the Local Lemma due to Erdős and Spencer [11]. The proofs in [21] succeed by generalizing the Local Lemma to an ordered or 'Lefthanded' version, which allows one to ignore all dependencies in one direction.

To make this clearer, recall that the Local Lemma can be seen as generalizing the basic fact that if $\mathcal{A}$ is a family of independent events of probability < 1, then $\mathrm{P}(\bigcap_{A\in\mathcal{A}} \bar{A}) > 0$. The 'lopsided' version of the Local Lemma due to Erdős and Spencer [11] can be seen as generalizing the fact that we can drop the requirement of independence for this conclusion so long as we have that $\mathrm{P}(\bar{A}|\bigcap_{\mathcal{C}} \bar{C}) > 0$ for all A and all families $\mathcal{C} \subset \mathcal{A}$. In this context, the 'Lefthanded' extension of the Local Lemma can be seen as generalizing the fact that if $\mathcal{A} = \{A_1, \ldots, A_m\}$, then it is sufficient to have that $\mathrm{P}\left(\bar{A}_i|\bigcap_{\mathcal{C}} \bar{C}\right) > 0$ for all A_i and all $\mathcal{C} \subset \{A_j | j < i\}$). (For the ordered Local Lemma I prove in [21], the ordering $A_1, A_2, \ldots$ is required to be consistent in certain ways with the dependency graph being used.) In the context of sequence games, the Lefthanded Local Lemma allows us to ignore the problematic dependencies on future events.

Applying the Lefthanded extension of the Local Lemma also proves game-based analogs of several other classical results on nonrepetitive sequences, including Thue's original theorem, and results on pattern avoidance. (The Lefthanded Local Lemma of [21] has also now been used by Grytczuk, Przybyło, and Zhu in a game-free situation, to get a near-optimal bound on the 'Thue choice number' [14].)

The results in [21] represent the first successful applications of a Local Lemma to combinatorial game theory. There are many natural games where there is a strong Local-Lemma based probabilistic intuition suggesting the existence (or not) of a winning strategy for a player. Previous successes, however, have came from 'derandomizing' the intuition to give constructive proofs; and in the cases where this failed, the results suggested by the Local Lemma remained unproven. (See [6] for a discussion.) In fact, the central open problem of Beck's 'derandomized' theory of positional games [5] (the 'Neighborhood Conjecture') essentially asks for a general game-theoretic analog of the Local Lemma for positional games. The success in generalizing the Local Lemma to allow it to apply directly to the Thue-type games suggests that, in at least some cases, a Local-Lemma based intuition can be converted into a Local Lemma based proof by developing a suitable generalization of the Local Lemma which allows the discarding of the most problematic dependencies. There is hope that this technique could work in general settings; for example, a first target is proving game-theoretic analogs of the results on nonrepetitive colorings of graphs due to Alon, Grytczuk, and others (*e.g.*,, [1, 9, 12]).

The end goal, however, is Beck's Neighborhood Conjecture. If proved, this would provide a derandomization of the Local Lemma for games in a very general setting, asserting the existence of drawing strategies in positional games played on hypergraphs whose maximum degree is not too large. This is analogous to the Erdős-Lovász 2-coloring theorem for which the Local Lemma was originally developed, which states that an n-uniform hypergraph with maximum degree $< 2^{n-3}/n$ has a proper 2-coloring.

An optimum density construction of k-critical triangle-free graphs [17]

A theme in many graph-theoretic questions concerns the degree of similarity between the families of bipartite and triangle-free graphs. Since the former consists of the graphs without any odd cycles at all, one might wonder to what extent the properties of bipartite graphs are forced just by their avoidance of triangles, which are, after all, the shortest kind of odd-cycle. One simple way in which these families are very close is in their membership: of course every bipartite graph is triangle-free, but it turns out that almost every triangle-free graph is bipartite as well.

On the other hand, these families differ greatly in the chromatic number of their members. Bipartite graphs have chromatic number 2 of course, while triangle-free graphs can have arbitrarily large chromatic number, as shown first by a recursive construction of Zykov [28]. Typical constructions of triangle-free graphs of large chromatic number contain relatively few edges, however. Zykov's construction has essentially a linear number of edges; another recursive construction due to Mycielski has $O(n^{\log_2 3})$ edges. This suggests that perhaps there is some kind of barrier to being triangle-free and having both lots of edges and a large chromatic number. Of course, given any k-chromatic triangle-free graph, we can take a disjoint union with a large complete bipartite graph to get a trivial example of a k-chromatic triangle-free graph which is dense (*i.e.*, has a quadratic number of edges). An old question of Erdős and Simonovits [10]—now essentially solved [8, 23]—avoided this kind of triviality by asking how large the minimum degree could be in a highly chromatic triangle-free graph.

Another way to avoid trivialities when asking about triangle-free graphs with lots of edges is to require the graphs in question to be *chromatic-critical*, say, with respect to removing edges. This means removing any edge from G decreases the chromatic number. A k-chromatic chromatic-critical graph is said to be k-critical. In my research I considered the question: *is there a family of triangle-free critical graphs of arbitrarily large chromatic number, and each with* $> cn^2$ *edges for some fixed* $c > 0$*?*

It turns out such a family does exist. Via a recursive construction, I give in [17] for each $k \geq 4$ an infinite family of k-critical triangle-free graphs with $\geq (c_k - o(1))n^2$ edges. (Observe that for $k = 3$ critical graphs are just the odd-cycles and so have a linear number of edges.) Here $c_4 = \frac{1}{16}$ (the construction in this case is identical to the Toft graph), $c_5 = \frac{5}{31}$, and, strikingly, for all $k \geq 6$ I can show $c_k = \frac{1}{4}$; this is best possible for $k \geq 6$ since (by Turán's Theorem) a graph with more than $\frac{n^2}{4}$ edges must contain a triangle. With some tricks I extend this to a similar result while simultaneously avoiding pentagons (the constants for the cases $k = 4, 5$ are worse). While this research was originally motivated by questions about the density of triangle-free graphs, it is also closely connected with the study of the density of critical graphs. The families of triangle-free graphs and pentagon-and-triangle-free graphs appear to be the only natural families of graphs where the correct asymptotic density of k-critical members is now known ($k \geq 6$).

In the case where we want to avoid larger odd cycles, I can only prove quadratic density for the case $k = 4$. A very natural research direction is to consider the density of k-critical graphs without ℓ cycles for $k \geq 5$ and odd $\ell \geq 7$. It is striking that we cannot prove for any such pair (k, ℓ) that there are infinite families of such graphs with $> (c - \varepsilon)n^2$ edges for *any* positive c, when for $\ell = 3, 5$ and for all $k \geq 6$, we can prove the best-possible constant. Part of the problem for the case of avoiding higher order odd-cycles stems from the fact that constructive techniques appear to be unavoidable when we want to show the existence of critical graphs with lots of edges.

The Hales-Jewett number and Tic-Tac-Toe [7]

The Hales-Jewett number $\mathrm{HJ}(n)$ is the minimum dimension d for which any 2-coloring of the $[n]^d$ hypercube induces a monochromatic geometric line; by *line* here we mean a winning set in the generalized hypercube version of Tic-Tac-Toe. The Hales-Jewett theorem [15], a cornerstone of Ramsey Theory, asserts that there are such sufficiently large dimensions d for any n; thus $\mathrm{HJ}(n)$ is a well-defined finite number. The upper bound that comes out of Hales' and Jewett's original proof involves the ridiculously huge Ackermann function; this was improved by Shelah [22] to a bound based on the *super-tower* function, a still quite ridiculously large function. For example, the best known bound for $n = 4$ is $\mathrm{HJ}(4) \leq 2^{2^{\cdot^{\cdot^{2}}}}$ where the tower has height 24. (An easy case study shows $\mathrm{HJ}(3) = 3$, on the other hand). On the other side of things, the original Hales-Jewett lower bound was just $\mathrm{HJ}(n) \geq n$.

In joint research with J. Beck and S. Vijay, we showed a new exponential lower bound $\mathrm{HJ}(n) > 2^{n/4}/3n^4$. What are the implications for Tic-Tac-Toe? We would like to be able to conclude that for sub-exponential dimensions d, there are terminal drawing positions in the $[n]^d$ Tic-Tac-Toe game (there are colorings with no monochromatic lines). However, observe that the only colorings resulting from a Tic-Tac-Toe game on the hypercube are *halving* colorings: the sizes of the color classes differ by at most 1. Define $\mathrm{HJ}^*_{\frac{1}{2}}(n)$ as the smallest d for which any halving 2-coloring of $[n]^d$ induces a monochromatic Tic-Tac-Toe line. Note that $\mathrm{HJ}^*_{\frac{1}{2}}(n) \leq \mathrm{HJ}(n)$. We showed that $\mathrm{HJ}^*_{\frac{1}{2}}(n) \geq \mathrm{HJ}(n-2)$ for large n, thus implying that the halving Hales-Jewett number has an exponential lower bound as well. This exponential lower bound on $\mathrm{HJ}^*_{\frac{1}{2}}(n)$ has an interesting game theoretic consequence: it implies that there are infinitely many pairs (n, d) for which there is a possible terminal drawing position in the $[n]^d$ Tic-Tac Toe game, and for which a first player can always force the appearance of a monochromatic line (though not necessarily before Player 2 does so). This implies that the win/draw status of infinitely many Tic-Tac-Toe games are delicate, in the sense that, if Player 1 has a winning strategy, it is in spite of the fact that a terminal drawing strategy exists, and if Player 2 has a drawing strategy, it is in spite of the fact that he cannot prevent Player 1 from achieving a monochromatic line—he can only prevent him from being the first to achieve one. Note that ordinary 3x3 Tic-Tac-Toe is an 'delicate draw' for Player 2 in this sense. We know from Patashnik's huge computer assisted work [16] that the 4^3 Tic-Tac-Toe game is a 'delicate win' for Player 1. In spite of the fact that that we know that infinitely many Tic-Tac-Toe games fall into one of these 'delicate' classes as a result of our exponential lower bounds, these are the only two examples where the correct class is known, and it is quite plausible that determining the class for any larger examples is a totally hopeless problem.

Euclidean shapes 'resilient to erosion' [20]

Given a set X in $\mathbb{R}^n$, define the *erosion* $e_r(X)$ of X by the radius r as the set of points of X at distance $\geq r$ from the complement X^C. This operation has been studied from a practical standpoint, as a model of pebble erosion; for example, [26] considers the the 'typical' limit shapes under this and related operations. In my research I considered a more theoretical question, namely, when is $e_r(X)$ equivalent to X under a Euclidean similarity transformation? When this is the case we say X is *resilient to erosion* by the radius r. For example, a disk, the body of a square, and the body of any triangle are all resilient to erosion by some positive radii.

It turns out that the family of resilient sets is surprisingly rich, and includes many intricate examples. Nevertheless, in [20] I prove a complete characterization of resilient subsets $X \subset \mathbb{R}^n$. One part of the characterization deals with the natural case where the corresponding similarity transformation is distance-nonincreasing; in this case the resilient set must be convex and satisfy certain necessary and sufficient conditions on the positions of its regular supporting hyperplanes. The other case, dealing with sets that 'get bigger' upon erosion (the corresponding similarity transformation is distance-increasing), gives a correspondence between this kind of resilient set and 'scale-invariant' sets (*e.g.*, unbounded exact fractals). Thus, for example, there are unbounded resilient sets corresponding to the familiar fractals such as Sierpinski's triangle and the Koch snowflake.

A nonwinning goal set in a game played in the plane [19]

Consider a game played in the plane between 2 players; on each turn a player selects any previously unchosen point. The goal is to build a congruent copy of some finite goal set agreed upon ahead of time before the other player does so. If the goal set is the 4 corners of the unit square, for example, a simple case study shows that Player 1 has a winning strategy. Even easier to see is that he has a winning strategy whenever the goal set has size ≤ 3.

Combinatorial game theory often presents seemingly hopeless problems almost immediately. For example, it is unknown whether Player 1 can win if the goal set is the five vertices of a regular pentagon (and, in general, if it is the n vertices of a regular n-gon, $n \geq 5$). Nevertheless, a surprising theorem of J. Beck [5] asserts that if Player 1 is content just to build a congruent copy of the goal set at all, not necessarily before Player 2, then he can always assure this kind of *weak win*, for *any* finite goal set. Left open was the question of whether it might even be true that he can build any finite goal set *before* Player 2 does. To show this is not the case, I gave in [19] a 5-point goal set for which Player 2 has a drawing strategy.

Expansion of symmetric digraphs of infinite degree [18] [undergraduate work]
A vertex-transitive graph is one in which all vertices are equivalent under the automorphism group of the graph. Thus, for example, the vertices of the cube connected along its edges gives a vertex-transitive graph with 8 vertices and 12 edges. The *distance sequence* $\{f(k)\}$ of a vertex-transitive graph is the sequence whose kth term $f(k)$ is the number of vertices at distance exactly k from a fixed vertex v—observe that vertex-transitivity implies that the choice of v is irrelevant. For the cube, for example, we have that the distance sequence $\{f(k)\}$ is $1, 3, 3, 1$ (starting with $k = 0$). In general, for the d-dimensional hypercube, the sequence $\{f(k)\}$ is given by the binomial coefficients $\binom{d}{k}$; thus, at least in this particular case, the sequences are very nice, rising and then falling without any 'valleys'.

A conjecture of Babai [2] asserted that the distance sequences of all vertex-transitive graphs of finite degree are unimodal (they have no valleys $f(i) > f(j) < f(k)$, $i < j < k$). Watkins and Shearer [27] found counterexamples, but their examples have relatively shallow valleys, and there has been research on constraints that can be put on the sequences in general (*e.g.*, [3]). An unresolved question of Babai was whether the distance sequences could be pathological even when the degree of the vertex-transitive graph is infinite; for example, could we have a distance sequence $1, \aleph_0, 17, \aleph_0, \ldots$, or maybe $1, \aleph_1, \aleph_0, \aleph_1, \ldots$. It turns out that no such valleys can occur in the infinite case. I proved in [18] that, after the first term, the sequence is constant except possibly for the last term (which can be at most the degree), and gave constructions to show that all sequences allowed by this result actually occur. For the directed case, I showed that the infinite terms of the 'out-distance-sequence' of a vertex-transitive digraph of infinite degree are nonincreasing, but valleys *do* occur sometimes in the directed case, as the sequence can have a tail of finite increasing terms.

References

[1] N. Alon, J. Grytczuk. Nonrepetitive colorings of graphs, in *Electronic Notes in Discrete Mathematics* **22** (2005) 353–355.

[2] L. Babai, Problem 29. In: *Unsolved Problems, Summer Research Workshop in Algebraic Combinatorics*, K. Heinrich Ed., Mathematics Department, Simon Fraser University, Burnaby, B.C., Canada V5A 1S6 (1979), p. 8.

[3] László Babai and Mario Szegedy. Local Expansion of Symmetrical Graphs. *Combinatorics, Probability and Computing* **1** (1992) 1-11.

[4] J. Beck. An application of Lovász Local Lemma: there exists an infinite 01-sequence containing no near identical intervals, in *Finite and Infinite Sets*, Eger 1981.

[5] J. Beck. *Combinatorial Games: Tic-Tac-Toe Theory*, Encyclopedia of Mathematics and its Applications **114**, Cambridge University Press, 2008.

[6] J. Beck. Surplus of Graphs and the Lovász Local Lemma, in *Building Bridges: Between Mathematics and Computer Science*, Bolyai Society Mathematical Studies, Vol. 19., M. Grötschel, G. Katona (Eds.) (2008).

[7] J. Beck, W. Pegden, and S. Vijay. The Hales-Jewett number is exponential: game-theoretic consequences, to appear in *Special volume dedicated to Klaus Roth on the occasion of his 80th birthday*, W. Chen, T. Gowers, H. Halberstam, W. Schmidt and B. Vaughan (Eds.).

[8] S. Brandt and S. Thomassé. Dense triangle-free graphs are 4-colorable: A solution to the Erdős-Simonovits problem. `http://www.lirmm.fr/~thomasse/liste/vega11.pdf`.

[9] S. Czerwiński, J. Grytczuk. Nonrepetitive colorings of graphs, in *Electronic Notes in Discrete Mathematics* **28** (2007) 453–459.

[10] P. Erdős and M. Simonovits. On a valence problem in extremal graph theory. *Discrete Mathematics*, 5:323–334, 1973.

[11] P. Erdős and J. Spencer, Lopsided Lovász Local Lemma and Latin transversals, *Discrete Applied Math.* **30** (1991) 151-154.

[12] J. Grytczuk. Pattern avoidance on graphs, *Discrete Mathematics* **307** (2007) 1341–1346.

[13] J. Grytczuk. Thue type problems for graphs, points and numbers, *Discrete Mathematics* **308** (2008) 4419–4429.

[14] J. Grytczuk, J. Przybyło, and X. Zhu. Thue choice number of paths. Manuscript (2009).

[15] A.W. Hales and R.I. Jewett. On regularity and positional games. *Trans. Amer. Math. Soc.*, **106** (1963), 222–229.

[16] O. Patashnik. Qubic: $4 \times 4 \times 4$ Tic-Tac-Toe. *Math. Mag.***53** (1980), 202–216.

[17] W. Pegden. Critical graphs without triangles: an optimum density construction, *(submitted)*.

[18] W. Pegden. Distance sequences in locally infinite vertex-transitive digraphs, *Combinatorica* **26** (2006) 577-585

[19] W. Pegden. A finite goal set in the plane which is not a winner, *Discrete Mathematics* **308** (2008) 6546-6551

[20] W. Pegden. Sets resilient to erosion, to appear in *Advances in Geometry*.

[21] W. Pegden. Thue-type games and highly nonrepetitive sequences., *(in preparation)*.

[22] S. Shelah. Primitive recursive bounds for van der Waerden numbers. *J. Amer. Math. Soc.*, **1** (1988), 683–697.

[23] C. Thomassen. On the chromatic number of triangle-free graphs of large minimum degree. *Combinatorica*, 22:591–596, 2002.

[24] A. Thue. Über unendliche Zeichenreichen, *Norske Vid Selsk. Skr. I. Mat. Nat. Kl. Christiana 7* (1906), pp. 1-22.

[25] A. Thue. Über die gegenseitigen Lage gleicher Teile gewisser Zeichenreihen, *Norske Vid Selsk. Skr. I. Mat. Nat. Kl. Christiana 1* (1912), pp. 1-67

[26] P.L. Várkonyi, G. Domokos, A.A. Sipos, Pebble shapes and abrasion processes, *Proc. of the First Workshop on Mathematical Geosciences*, June 11-12, 2008, Freiberg, Germany.

[27] Mark E. Watkins and James B. Shearer. Counterexamples to Two Conjectures about Distance Sequences. *Discrete Mathematics* **66** (1987) pp 289-298.

[28] A. Zykov. On some properties of linear complexes (in Russian). *Matem. Sbornik*, 24:163–187, 1949.