Problems that I would like Somebody to Solve

Sam Spiro

December 14, 2020

Contents

1 What is this? 1

2 The Main Problems 2
   2.1 Small Quasikernels 2
   2.2 Ballot Permutations and Odd Order Permutations 2
   2.3 Polynomial Relations Between Matrices of Graphs 3
   2.4 The Turán problem in Random Hypergraphs 4

3 Some Smaller Problems 5
   3.1 Slow Tribonacci Walks 5
   3.2 An Adversarial Chernoff Bound 5
   3.3 Saturation Games 6
   3.4 Online Card Guessing Games 7
   3.5 Increasing Subsequences 7

4 Hall of Fame 8

1 What is this?

This is a (very informal) collection of problems that are of personal interest to me, and it is definitely not supposed to be an exhaustive list of the hardest/most interesting problems in math. Feel free to reach out to me if you have any questions regarding these questions and/or if you spot any glaring typos.

Organization. Section 2 contains the main problems that I’m interested in. Solving any of them will put you in the “Hall of Fame” of Section 4 as thanks for putting the problem to rest.
Section 3 consists of some other problems that I would still appreciate being resolved, but not to the level that I’ll sleep better at night knowing that they’ve been solved. In particular, many of these problems are relatively elementary to state, and I would not be surprised if several of them were easy to solve. I also emphasize that the order that the problems appear in the listing is based solely off of when I added them to the list, and in particular the ordering does not represent any particular favoritism from me.

2 The Main Problems

2.1 Small Quasikernels

Let $D$ be a digraph. Given a set $S$, we define $N^+(S) = \bigcup_{v \in S} N^+(v)$, where $N^+(v)$ is the out-neighborhood of $v$. We say that a set $K \subseteq V(D)$ is a kernel of $D$ if (1) $N^+(K) \cap K = \emptyset$ (that is, $K$ is an independent set of the underlying graph of $D$), and (2) $N^+(K) \cup S = V(D)$ (that is, every vertex is either in $K$ or can be reached by a vertex in $K$ in one step).

Not every digraph has a kernel (take any directed cycle of odd length), but it is not too hard to prove that every digraph has a quasikernel. This is a set $Q \subseteq V(D)$ such that (1) $N^+(Q) \cap Q = \emptyset$ and such that (2) $N^+(N^+(Q)) \cup N^+(Q) \cup Q = V(D)$. That is, it is an independent set such that every vertex can be reached from $Q$ in at most two steps.

Given that every digraph has a quasikernel, it is natural to ask how small of a quasikernel one can find. One quickly realizes that it can be quite large: any source of $D$ must belong to a quasikernel of $D$. Thus the most natural setting to consider is when $D$ has no sources, and in this case the following was conjectured by Erdős and Székely.

Conjecture 2.1. Every digraph $D$ with no sources has a quasikernel of size at most $|V(D)|/2$.

Some progress has been made, see [5], but overall the conjecture is far from being resolved.

2.2 Ballot Permutations and Odd Order Permutations

Let $S_n$ be the group of permutations acting on a set of $n$ elements. We say that a permutation $\pi$ is an odd order permutation if the order of $\pi$ is odd in $S_n$, which is equivalent to $\pi$ being the product of only odd cycles. For example, $\pi = (3, 1, 4)(2, 5, 6, 7, 9)$ is an odd order permutation since it has order 15 in $S_9$. We let $P_n$ denote the set of odd order permutations of size $n$.

We now consider elements of $S_n$ as words written in 1-line notation. Given a word $\pi$, we say that position $i$ is an ascent if $\pi_i < \pi_{i+1}$ and that it is a descent otherwise. We say that $\pi$ is a ballot permutation if the words $\pi_1 \cdots \pi_k$ have at least as many ascents as descents for all $k$. For example, $\pi = 31452$ is not ballot since 31 has one descent and no ascents, but $\sigma = 14352$ is ballot. We let $B_n$ denote the set of all ballot permutations of size $n$. The following was proven by Bernardi, Duplantier, and Nadeau [1].
Theorem 2.2 ([1]). For all \( n \),

\[
|B_n| = |P_n| = \begin{cases} 
(n-1)!!^2 & n \text{ even}, \\
(n!!(n-2))! & n \text{ odd}, 
\end{cases}
\]

where \((2m-1)!! := (2m-1) \cdot (2m-3) \cdots 3 \cdot 1\)

I conjectured that a stronger version of Theorem 2.2 was true, and this was proven by Wang and Zhao [14]. Based on their proof and further computations, I conjecture that an even stronger version of Theorem 2.2 is true.

To this end, given a permutation \( \pi \) in 1-line notation, we define the peak triple set \( T(\pi) \) to consist of all the sets \( \{a,b,a'\} \) such that \( \pi_i \pi_{i+1} \pi_{i+2} = aba' \) for some \( i \) and such that \( b > a,a' \).

That is, the largest element of each triple in \( T(\pi) \) is a peak in \( \pi \) with the other two elements the letters directly to the left/right of the peak (but we don’t record whether \( a \) or \( a' \) appears first in \( \pi \)). For example, \( T(14352) = \{\{1,4,3\}, \{3,5,2\}\} \).

Given any set of triples \( X \), we define \( B_n(d,X) \) to be the set of ballot permutations of order \( n \) with exactly \( d \) descents and with \( T(\pi) = X \).

We wish to develop analogs of descents and peak triples for permutations written in cycle notation. Given a cycle \( c = (c_1, \ldots, c_k) \) of a permutation \( \pi \), we let \( \text{asc}'(c) \) denote the number of cyclic ascents of \( c \), i.e. the number of ascents in the word \( c_1 c_2 \cdots c_k c_1 \).

We similarly define \( \text{des}'(c) \) to be the number of cyclic descents of \( c \). We let \( M(c) = \min \{ \text{asc}'(c), \text{des}'(c) \} \). For example, if \( c = (4,2,8,5,6) \), we have \( \text{asc}'(c) = 2 \), \( \text{des}'(c) = 3 \), and hence \( M(c) = 2 \).

For a permutation \( \pi = c_1 c_2 \cdots c_k \) written in cycle notation, we define \( M(\pi) = \sum_k M(c_i) \). For example, if \( \pi = (1,3,9)(4,2,8,5,6)(7) \), then \( M(\pi) = 1+2+0 = 3 \).

Given a cycle \( c = (c_1, \ldots, c_k) \), we define \( T'(c) \) to consist of all sets \( \{a,b,a'\} \) such that the word \( w = c_k c_1 \cdots c_i c_1 \) has \( w_i w_{i+1} w_{i+2} = aba' \) for some \( i \) and such that \( b > a,a' \). If \( \pi \) is written in cycle notation, we let \( T'(\pi) = \bigcup T'(\pi) \) where the union ranges over all cycles of \( \pi \). For example, if \( \pi = (1,3,9)(4,2,8,5,6) \) we have \( T'(\pi) = \{\{3,9,1\}, \{2,8,5\}, \{5,6,4\}\} \).

With all this in mind, we let \( P_n(d,X) \) denote the set of odd order permutations \( \pi \) of order \( n \) with \( M(\pi) = d \) and \( T'(\pi) = X \).

Conjecture 2.3. For all \( n,d,X \), we have

\[
|B_n(d,X)| = |P_n(d,X)|.
\]

We note that if this conjecture were true, then fixing \( n \) and summing over all \( d,X \) gives Theorem 2.2. Other theorems and conjectures in the area follow from Conjecture 2.3, see [13, 14]. I have used a computer to verify this conjecture through \( n = 8 \). I know how to prove this whenever \( d = (n-1)/2 \) and \( n \) is odd, and I have been told that proving this for \( d = 1 \) is not too hard.

2.3 Polynomial Relations Between Matrices of Graphs

Let me first note that this is probably the most niche problem of this entire document, and in particular I may be the only person in the world who cares about it. Nevertheless, if you’re
looking for a (potentially) easy problem to solve to get into the “Hall of Fame”, this is the problem for you.

The general question is: given two matrices $X,Y$ associated to a graph $G$, does there exist an integer $r$ and a polynomial $f$ such that $X^r = f(Y)$? For example, let $A$ be the adjacency matrix of $G$, $D$ its diagonal matrix of degrees (i.e. $D_{uu} = d_u$), and $L = D - A$ its combinatorial Laplacian. If $G$ is $d$-regular, then we have $A = dI - L$, so we get a positive answer with $r = 1$ and $f(x) = d - x$. As another example, if $G$ is $(d_1,d_2)$-biregular, meaning it’s bipartite with every vertex of one part having degree $d_1$ and the other all having degree $d_2$, then one can check that $A^2 = (d_1I - L)(d_2I - L)$, so this gives another family of examples.

Somewhat surprisingly, the only graphs which have $A^r = f(L)$ for some $r,f$ are the regular and biregular graphs \cite{9}. In fact, I was able to prove that a similar phenomenon holds for any $X,Y \in \{A,L,Q,L\}$ where $Q = D + A$ is the signless Laplacian and $L = D^{-1/2}LD^{-1/2}$ is the normalized Laplacian, except for one irksome case.

**Conjecture 2.4** \cite{9}. If $G$ is a graph and $A^r = f(L)$ for some integer $r \geq 1$ and polynomial $f$, then $G$ is either regular or biregular.

I wrote \cite{9} after knowing only some very basic spectral graph theory, so every once in a while I think “this problem probably wasn’t really that hard. I should be able to solve it without much effort.” As you might have guessed, this has not turned out to be the case!

The most promising approach to solving Conjecture 2.4 that I came up with was the following.

**Conjecture 2.5** \cite{9}. Let $1$ denote the all 1’s vector. If $G$ is connected and $D^{1/2}1$ is an eigenvector of $A^2$, then $G$ is regular or biregular.

In \cite{9} I prove that this conjecture implies Conjecture 2.4, and computational data strongly suggests that this result is true.

### 2.4 The Turán problem in Random Hypergraphs

Given two hypergraphs $H, F$, we define $\text{ex}(H, F)$ to be the maximum number of edges in an $F$-free subgraph of $H$. We’re particularly interested in the case when $H = H^r_{n,p}$, which is the random $r$-uniform hypergraph obtained by keeping each edge of $K^r_n$ independently and with probability $p$. There are many, many open problems related to determining $\text{ex}(H^r_{n,p}, F)$, see for example \cite{7, 11, 12}; but here we’ll focus on just one case.

Let $C^3_4$ be the 3-uniform loose 4-cycle, which can be defined by having edges

$$\{1, 2, 3\}, \{3, 4, 5\}, \{5, 6, 7\}, \{7, 8, 1\}.$$

That is, it’s obtained from the graph $C_4$ by inserting an extra vertex into each edge. A standard deletion argument shows that

$$\mathbb{E}[\text{ex}(H^3_{n,n^{-2/3}}, C^3_4)] = \Omega(n^{4/3}),$$

and work of Mubayi and Yepremyan \cite{6} shows

$$\mathbb{E}[\text{ex}(H^3_{n,n^{-2/3}}, C^3_4)] \leq n^{13/9+o(1)}.$$
Problem 2.6. Improve upon either of these bounds for $\mathbb{E}[\text{ex}(H_{n,n^{2/3}}^3, C_4^3)]$.

Mubayi and Yepremyan [6] conjecture that the lower bound from the deletion argument is essentially best possible, and they provide some evidence that points in this direction. On the other hand, Nie, Verstraëte, and myself [7] proved that the analogous problem for loose triangles has a stronger lower bound than what you get from a deletion argument, which suggests that improvements might be possible for all loose cycles. Overall it’s very unclear what the correct answer should be here.

3 Some Smaller Problems

3.1 Slow Tribonacci Walks

Given a triple of positive integers $(w_1, w_2, w_3)$, recursively define $w_k = w_{k-1} + w_{k-2} + w_{k-3}$ for $k \geq 4$. We say that $(w_1, w_2, w_3)$ is an $n$-tribonacci walk if $w_s = n$ for some $s$. There are infinitely many $n$-tribonacic walks, e.g. those of the form $(42, n, x)$ for any $x$. To make things more interesting, we say that $(w_1, w_2, w_3)$ is an $n$-slow tribonacci walk if $w_s = n$ with $s$ as large as possible. For example, $(1, 1, 1)$ and $(42, 3, 42)$ are both 3-tribonacci walks, but only the first one is slow. Let $p(n)$ denote the number of $n$-slow tribonacci walks. For example, it’s easy to check that $p(3) = 1$, and $p(1) = p(2) = \infty$.

Question 3.1. Does there exist some absolute constant $c$ such that either $p(n) = \infty$ or $p(n) \leq c$ for all $n$?

If we instead look at Fibonacci walks (which are defined using the Fibonacci recurrence $w_k = w_{k-1} + w_{k-2}$), then one can show that $p(n) \leq 2$ for all $n \geq 2$ [3]. More generally if one looks at walks following a two-term recurrence of the form $w_k = \alpha w_{k-1} + \beta w_{k-2}$ with $\alpha, \beta \geq 1$ relatively prime, then $p(n) \leq \alpha^2 + 2\beta - 1$ for all but finitely many $n$ [10].

Computational data made it easy to conjecture the correct answer for two-term recurrences, but the situation is less clear for slow tribonacci walks. For example, $p(61) = 9$, which is fairly large given how small 61 is.

I’ll note that there are many other interesting problems related to the behavior of slow recurrences that were left unanswered in [3, 10]. However, as is often the case in number theory, these relatively easy to state questions are likely very difficult to solve. This being said, I do think that this tribonacci problem is tractable.

3.2 An Adversarial Chernoff Bound

Persi Diaconis, Ron Graham, Xiaoyu He, and myself [4] proved the following.

Theorem 3.2 ([4]). Let $X_i$ be independent Bernoulli random variables with $\text{Pr}[X_i = 1] = p$ and $\text{Pr}[X_i = 0] = 1 - p$. Let $S_t = \sum_{i \leq t} X_i$. There exist absolute constants $c_0, c_1$ such that for
all $\lambda > 0$ and integers $k_1 \geq k_0 \geq 2\lambda^{-1}$,
\[
\Pr[\exists t \in [k_0, k_1] : |S_t - pt| \geq \lambda pt] \leq \frac{c_0 k_1}{k_0} e^{-c_1 \lambda^3 pk_0}.
\]
That is, with high probability, for every $t$ in the interval $[k_0, k_1]$, every partial sum $S_t$ is close to its expectation. This is immediate for any given value of $t$ by the Chernoff bound (since each $S_t$ is a binomial random variable), but it does not follow from the Chernoff bound and a naive application of the union bound (this gives a bound like $k_1 e^{-\lambda^2 pk_0}$, which is much weaker if $k_0$ is very large).

**Question 3.3.** Does the bound of Theorem 3.2 hold with $\lambda^2$ instead of $\lambda^3$?

Note that $\lambda^2$ would be best possible because this is what one gets if $k_0 = k_1$. While the statement of Theorem 3.2 is fairly technical, the proof itself only required a slightly clever application of the union bound together with the Chernoff bound, so my hope is that more sophisticated probabilistic tools can be used to solve Question 3.3 without too much difficulty.

Secretly I’m interested in this because it would improve upon the error term for our main result in [4], but also I just think it’s of independent interest to determine how much concentration one can get for an “adversarial” binomial distribution.

### 3.3 Saturation Games

For a family of graphs $\mathcal{F}$, we say that a graph $G$ is $\mathcal{F}$-saturated if $G$ contains no graph of $\mathcal{F}$ as a subgraph, but adding any edge to $G$ would create a subgraph of $\mathcal{F}$. The $\mathcal{F}$-saturation game consists of two players, Max and Mini, who alternate turns adding edges to an initially empty graph $G$ on $n$ vertices (say with Max starting), with the only restriction being that $G$ is never allowed to contain a subgraph that lies in $\mathcal{F}$. The game ends when $G$ is $\mathcal{F}$-saturated. The payoff for Max is the number of edges in $G$ when the game ends, and Mini’s payoff is the opposite of this. We let $\text{sat}_g(n, \mathcal{F})$ denote the number of edges that the graph in the $\mathcal{F}$-saturation game ends with when both players play optimally, and we call this quantity the game $\mathcal{F}$-saturation number.

Bounding $\text{sat}_g(\mathcal{F}; n)$ seems to be pretty hard in general, and even the original problem of determining $\text{sat}_g(n, C_3)$ is still wide open. See [8] for further history and known bounds. In [8] I proved $\text{sat}_g(n, \mathcal{C}_\infty \setminus \{C_3\}) = O(n)$, where $\mathcal{C}_\infty$ is the set of all odd cycles. I also proved (somewhat indirectly) that $\text{sat}_g(n, \mathcal{C}_\infty \setminus \{C_{2k+1}\}) = \Omega(n^2)$ for $k \geq 3$. Given this, it is natural to ask the following.

**Problem 3.4.** Prove non-trivial bounds on $\text{sat}_g(n, \mathcal{C}_\infty \setminus \{C_5\})$, where $\mathcal{C}_\infty = \{C_3, C_5, C_7, \ldots\}$.

I’d be happy to have even an $\omega(n)$ lower bound or any non-trivial asymptotic upper bound for this problem. Possibly a more tractable problem is the following.

**Problem 3.5.** Prove non-trivial bounds on $\text{sat}_g(n, C_k)$ for $k > 3$. 

For odd $k$, I proved an asymptotic upper bound of $\text{sat}(n, C_k) \leq \frac{4}{27} n^2 + o(n^2)$. In [2] the authors proved a non-trivial lower bound for $C_4$ if you play a “bipartite” version of the game, but I’d still like to see bounds proved in the original setting. Lastly, I’d like to know the following.

**Question 3.6.** Does there exist a finite set of (odd) cycles $C$ such that $\text{sat}_g(n, C) = O(n)$?

### 3.4 Online Card Guessing Games

Consider the following card game. You have a deck of $mn$ cards where there are $m$ copies of $n$ different card types (e.g. a usual deck of playing cards has $m = 4$ and $n = 13$). You want to try and shuffle the deck so that the cards appear in an “unpredictable way”. More precisely, each round you draw one of the remaining cards in the deck, and an opponent tries to guess what your card is. After they guess, you reveal and discard the card. You repeat this process until the deck is depleted. If your goal is to have your opponent correctly guess the fewest number of cards possible, what is your optimal strategy?

Consider the following strategy which I call the **greedy strategy**: if there are $r$ card types left in the deck, uniformly at random choose one of these card types. For example, if the deck has 100 copies of card type $a$ and 1 of card type $b$, then you draw each of $a$ and $b$ with equal probability. It’s not difficult to show that this strategy is the best you can do in any given round, but it’s far from clear that it’s optimal overall. For instance, if card type $b$ is drawin the previous example, then the opponent knows that the remaining cards in the deck are all $a$, so they can guarantee 100 points. Nevertheless, I conjecture the following.

**Conjecture 3.7.** The greedy strategy is optimal in this game.

I can prove that the greedy strategy asymptotically gives the score of an optimal strategy if $m$ is fixed and $n$ tends to infinity. At one point I thought I proved it was optimal, but my notes were unorganized and I couldn’t decipher them when I reread them.

This problem is really an online version of a more classical card problem, see [4]. There are some other problems in this area (both online and not online), but this is the main one that’s bugging me.

### 3.5 Increasing Subsequences

Let $\mathcal{S}_{m,n}$ denote the set of words which consist of $m$ copies of each of the symbols 1 through $n$. In particular, $\mathcal{S}_{1,n}$ consists of the set of permutations of size $n$.

Given $\pi \in \mathcal{S}_{m,n}$, define $L(\pi)$ to be the largest integer $p$ such that there exist $i_1 < \cdots < i_p$ with $\pi_{i_j} = j$ for all $1 \leq j \leq p$. That is, $\pi$ contains the sub-word $12 \cdots p$ for all $p \leq L(\pi)$. Define $\mathcal{L}_{m,n} = \mathbb{E}[L(\pi)]$, where in the expectation $\pi$ is chosen uniformly at random.

**Conjecture 3.8.** If $n$ is sufficiently large in terms of $m$, then

$$\mathcal{L}_{m,n} = \Theta(m).$$
In [4], it was proven (in a very indirect way) that \( \mathcal{L}_{m,n} \leq m + O(m^{3/4} \log m) \) provided \( n \) is sufficiently large in terms of \( m \). A simple random argument gives \( \mathcal{L}_{m,n} = \Omega(m/\log m) \) in this regime, and that’s the best lower bound we were able to prove. In fact, computational and heuristic evidence suggests that
\[
\mathcal{L}_{m,n} = m + \Theta(m^{1/2}),
\]
but for the moment I’d be satisfied with proving the correct order of magnitude.

4 Hall of Fame

A list of people who have successfully solved any of my main problems.

- Wand and Zhao [14] for solving my original conjecture on ballot permutations.

References


