Discrepancy games

Noga Alon *
Schools of Mathematics and Computer Science
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University, Tel Aviv 69978, Israel
and IAS, Princeton, NJ 08540, USA
nogaa@tau.ac.il

Michael Krivelevich†
School of Mathematical Sciences
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University, Tel Aviv 69978, Israel
krivelev@post.tau.ac.il

Joel Spencer ‡
Department of Mathematics and Computer Science
Courant Institute, NYU, NY, USA
spencer@cs.nyu.edu

Tibor Szabó
Institut für Theoretische Informatik
ETH Zentrum, IFW B48.1, CH - 8092 Zürich, Switzerland
szabo@inf.ethz.ch

Submitted: Apr 8, 2005; Accepted: Aug 27, 2005; Published: to be inserted
Mathematics Subject Classifications: 90D42, 05C65

Abstract

We investigate a game played on a hypergraph $H = (V, E)$ by two players, Balancer and Unbalancer. They select one element of the vertex set V alternately until all vertices are selected. Balancer wins if at the end of the game all edges $e \in E$ are roughly equally distributed between the two players. We give a polynomial time

*Research supported in part by a USA-Israel BSF grant, by the Israel Science Foundation, by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University and by the Von Neumann Fund.

†Research supported in part by USA-Israel BSF Grant 2002-133 and by grant 64/01 from the Israel Science Foundation.

‡Research supported in part by a USA-Israel BSF Grant.

algorithm for Balancer to win provided the allowed deviation is large enough. In particular, it follows from our result that if H is n-uniform and has m edges, then Balancer can achieve having between $n/2 - \sqrt{\ln(2m)n/2}$ and $n/2 + \sqrt{\ln(2m)n/2}$ of his vertices on every edge e of H. We also discuss applications in positional game theory.

1 Introduction

In the classical theory of Maker/Breaker positional games a hypergraph $H = (V, E)$ is given and the players, Maker and Breaker, take turns in occupying a previously unoccupied element of the "board" V. The goal of Breaker is to *prevent* Maker from fully occupying an edge. The well-known criterion of Erdős and Selfridge [4] provides a strategy for Breaker to win. In a more general setting, the criterion of Beck [2] ensures that Breaker can select more than $\alpha|e|$ elements of each edge $e \in E$, for some $\alpha \geq 0$.

In these Maker/Breaker-type games Breaker does not care about fully occupying an edge himself. In the so-called Avoider/Forcer-type games this is the only thing player "Avoider" cares about *not doing*. More precisely Avoider wins the game against Forcer if at the end of the game he does not occupy any edge. More generally, Avoider wins if he occupies less than $(1-\alpha)|e|$ elements from any edge $e \in E$, for some $\alpha \geq 0$. Lu [7] obtained criteria similar to the ones of Erdős and Selfridge, and of Beck for this case.

In the present paper we investigate a game where one of the players, called Balancer, must achieve the goals of both Breaker and Avoider. Balancer's main difficulty is that while as Breaker he cannot hurt himself by selecting any particular vertex, as Avoider he can. Similarly, as Avoider he cannot hurt himself by *not* occupying any particular vertex while as Breaker he can. The classical Erdős-Selfridge-type criteria do not immediately generalize to this setting.

Now we give a precise formulation of our game. Given a hypergraph $H = (V, E)$, the game Discrepancy is played by two players, called Balancer and Unbalancer. They take turns in occupying previously unoccupied elements of V. The game ends when all elements are occupied by one of the players. Balancer's aim is to achieve a situation where each edge $e \in E(H)$ has about the same number of Balancer's and Unbalancer's vertices. In order to quantify this, assume that $E(H) = \{e_1, \ldots, e_m\}$, and in addition to H a target vector $\mathbf{b} = (b_1, \ldots, b_m\}$, $b_j \geq 0$, is given. Let B and U be the subsets occupied by Balancer and Unbalancer, respectively, at the end of the game. Then Balancer wins the $(H, \mathbf{b})$-game if for every edge $e_j \in E(H)$ one has: $||B \cap e_j| - |U \cap e_j|| \leq b_j$, otherwise the game is won by Unbalancer.

A more convenient (yet completely equivalent) game description is as follows: Balancer labels each of his vertices by $+1$, Unbalancer's labels are -1. Let $f : V \to \{-1, +1\}$ be the labeling in the end of the game. Define $f(e_j) = \sum_{v \in e_j} f(v)$. Then the game is won by Balancer if and only if $|f(e_j)| \leq b_j$ for every $j = 1, \ldots, m$.

The study of games where both Breaker- and Avoider-type goals should be achieved by one of the players was first suggested in [6] and a method was developed to deal with such situations. In particular a player can successfully create a pseudo-random graph of density 1/2 and consequently, as it follows from properties of pseudo-random graphs [5], create, for example, $(1/4 - o(1))n$ pairwise edge-disjoint Hamiltonian cycles.

In this note we suggest a completely different approach, whose correctness is proved with basically no calculation. An added advantage of our strategy is that it supplies a polynomial time algorithm for Balancer to win even if Unbalancer has unbounded computational power. Our main result is the following.

Theorem 1 *Let $H = (V, E)$ be a hypergraph with edge set $E = \{e_1, \dots, e_m\}$. Let further $\mathbf{b} = (b_1, \dots, b_m)$ be a target vector. Assume that $|V|$ is even and Balancer moves first. Then he has a winning strategy for $(H, \mathbf{b})$ provided*

$$\sum_{i=1}^{m} e^{-\frac{b_i^2}{2|e_i|}} \leq \frac{1}{2}.$$

In particular Balancer has a winning strategy if $b_j \geq \sqrt{2\ln(2m)|e_j|}$ for $j = 1, \dots, m$.

In a sense the theorem gives a strategy for Balancer to create a pseudo-random coloring of the board while playing against an adversary. The deviation between the colors in each set is comparable to the one one would get from a random coloring.

A stronger result of a special case of Theorem 1 was proved in [6, Lemma 3.]. There it is shown that if $b_j = |e_j|$ (i.e. Balancer's goal is that no hyperedge is "monochromatic"), then a condition $\sum_{i=1}^{m} 2^{-|e_i|} < 1/4$ suffices for Balancer to win.

2 Playing Discrepancy

In this section we present two proofs of the main result. The first one involves essentially no computation, but gives a slightly weaker result obtained by replacing the 1/2 in the statement of the theorem by 1/8. The second one requires a limited amount of computation, but proves the slightly stronger assertion as stated in the theorem and has another algorithmic advantage explained before the proof.

First proof of Theorem 1 (weaker version). We will call the conditions $f(e_j) \leq b_j$ a *positive condition* and $f(e_j) \geq -b_j$ a *negative condition.* Balancer wins if at the end of the game all of these conditions are met. We define the weight W_j^+ (at the start of the game) of the j-th positive condition as follows: Let $X_1, \dots, X_{|e_j|}$ be independent uniform ± 1 random variables. Let W_j^+ be the probability that $X_1 + \ldots + X_i > b_j$ for *some* $0 \leq i \leq |e_j|$. For a negative condition the weight W_j^- is the probability that some $X_1 + \ldots + X_i < -b_j$. When $i = 0$ the sum is interpreted as zero.

We define the weight W to be the sum of all the weights W_j^+ and W_j^-. Observe that the condition of the theorem (with 1/2 replaced by 1/8) implies $W < 1/2$. Indeed, for W_j^+ we need to estimate the probability that the random walk of length $|e_j|$, starting at zero *ever* exceeds b_j. Notice that by the "mirror principle" this is less than twice the probability that the random walk of length $|e_j|$ ends up bigger than b_j, which has probability at most $e^{-\frac{b_j^2}{2|e_j|}}$ by the Chernoff bound. The estimation of W_j^- is analogous.

Hence the following claim implies Theorem 1.

Claim. If the initial weight W satisfies: $W < 1/2$, then Balancer has a winning strategy.

Proof. Assume that in the beginning of the game all vertices $v \in V(H)$ are labeled $f(v) = 0$ by f. Then as the game proceeds Balancer changes the labels of his vertices from 0 to 1, while Unbalancer changes his from 0 to -1. We extend the definition of the weight of a condition for an arbitrary stage of the game. Let $f : V \to \{0, -1, +1\}$ denote a current labeling. For an edge $e_j \in E$ define

$$m_j \;=\; |\{v \in e_j : f(v) = 0\}| \;.$$

In case a condition was violated earlier in the game (irrespective of the current value of $f(e_j)$) let the weight of the condition be 1. Otherwise, we define the weight (for the positive condition) as the probability that $f(e_j)+X_1+\ldots+X_i > b_j$, for some $0 \le i \le m_j$ where the $X_i = \pm 1$ are independent and uniform random variables. The definition of the weight for a negative condition is similar. When $i = 0$ the sum is $f(e_j)$ so the weight is 1 if the condition is currently breached. The weight W is the sum of the weights over all the conditions j. Balancer's strategy is now simple to describe: He always moves so as to minimize the weight. At the end of the game the weight is simply the number of conditions that have been breached at some point during the game. Hence to infer the theorem it suffices to show that the weight does not increase in a single round.

Observe that for any unclaimed vertex x, the decrease in the value of W if x is taken by Balancer is equal to the increase of the value of W if x is taken by Unbalancer. Indeed, for each condition, its weight $\hat{W}$ is the average of the condition's weight after Balancer takes x and its weight after Unbalancer takes x. Hence it suffices to show that after an arbitrary move x by Balancer, Unbalancer, by moving $y \neq x$, *does not increase the weight by more* than he could have by moving y before the move of Balancer. Since the weight is the sum of the weights of the conditions it suffices to show this for a single condition j. All cases are trivial except when both x, y are in the set e_j. Say the original weight of the condition was $\hat{W}$ and let $\hat{W}^x$ be the new weight after Balancer takes x and $\hat{W}^y$ be the new weight after Unbalancer takes y (but without a move by Balancer!) and $\hat{W}^{xy}$ be the new weight after they both move in the set. So we would like to infer $\hat{W}^{xy} - \hat{W}^x \le \hat{W}^y - \hat{W}$ or equivalently $\hat{W}^{xy} \le \hat{W}^x + \hat{W}^y - \hat{W}$. Note that $\hat{W} = \frac{\hat{W}^x + \hat{W}^y}{2}$ since the weight is the average of the new weights with a random first move. Thus we would like to see $\hat{W}^{xy} \le \hat{W}$. Notice that looking at the definition of the weight this is tautological! The original $\hat{W}$ was the probability that a random walk starting at a point of a certain length (the number of still unchosen vertices) ever reaches a boundary while $\hat{W}^{xy}$ is the probability that the random walk starting at the same point which has length two less ever reaches the boundary.

(Notice a little subtlety here: since before the current round the total weight W was less than 1/2, for each condition j with $x \in e_j$ we had: $\hat{W} < 1/2$, implying that $f(e_j) < b_j - 1$. Hence Balancer cannot violate condition j in his turn, and thus indeed $\hat{W}^{xy} \le \hat{W}$.) □

We next present another proof of Theorem 1 (in its strong version), which contains some limited calculations compared to our first proof, but has the advantage that when actually playing, Balancer can calculate the value of each weight very easily (as opposed to the method of the first proof, where the weight of a condition is defined by a probability that is harder to compute).

Second proof of Theorem 1.

The proof is inspired in part by the method of conditional expectations with pessimistic estimators, see [8] or [1], Chapter 15.

Let α be a real parameter. Define a function

$$G(x) = cosh(\alpha x) = \frac{e^{\alpha x} + e^{-\alpha x}}{2} \ .$$

Here are two basic properties of $G(x)$. Both are straightforward to verify and can be found, for example, in Chapter 15 or Appendix A of [1].

$$\frac{G(x+1) + G(x-1)}{2} = G(x)G(1) \ ; \tag{1}$$

$$1 \leq G(1) \leq e^{\frac{\alpha^2}{2}} \ . \tag{2}$$

Recall the definition of f from the first proof and for an edge $e_j \in E$ define

$$\begin{aligned} f_j &= \sum_{v \in e_j} f(v)\,; \\ c_j &= |\{v \in e_j : f(v) \in \{-1, +1\}\}| \ . \end{aligned}$$

Let also

$$G_j(x) = cosh(\alpha_j x) = \frac{e^{\alpha_j x} + e^{-\alpha_j x}}{2} \ ,$$

where $\alpha_j = \frac{b_j}{|e_j|}$. Define now the *potential function* ϕ by:

$$\begin{aligned} \phi_j &= \frac{G_j(f_j)}{G_j(1)^{c_j}} \ , \quad j = 1, \ldots, m\,; \\ \phi &= \sum_{j=1}^{m} e^{-\frac{b_j^2}{2|e_j|}} \phi_j . \end{aligned}$$

The intuition behind this choice is given by the following observation: for every edge e_j, if ϕ_j is its current potential, and a label of a vertex $v \in e_j$ with $f(v) = 0$ is chosen randomly with $Pr[f(v) = 1] = Pr[f(v) = -1] = 0.5$, then by Property (1) of G_j the expected value of the new potential ϕ'_j is ϕ_j. Indeed,

$$E[\phi'_j] = \frac{G_j(f_j+1) + G_j(f_j-1)}{2G_j(1)^{c_j+1}} = \frac{G_j(f_j)}{G_j(1)^{c_j}} = \phi_j \ . \tag{3}$$

Hence, for any unchosen vertex $v \in V(H)$, the decrease in the value of ϕ, when v is taken by Balancer, is equal to the increase of the value of ϕ when v is taken by Unbalancer.

Here is a strategy for Balancer: he picks a previously unchosen vertex $u \in V(H)$ and labels it $f(u) = 1$ so as to minimize the new value of the potential function. We claim that if Balancer follows this strategy, then after each round the value of the potential function does not increase.

It suffices to show that after an arbitrary move by Balancer, Unbalancer cannot increase the value of the potential function by more than he could have before the move of Balancer. Indeed, by (3) and the sentence following it, Balancer, choosing an optimal vertex, can decrease the value of ϕ by at least the same amount Unbalancer increases with his move.

Since the potential function ϕ is a linear combination of potentials ϕ_j, it is enough to check this for each edge e_j separately. It is easy to see that the only non-trivial case is when both vertices (x of Balancer and y of Unbalancer) belong to e_j. We are thus to verify the inequality

$$\frac{G_j(f_j)}{G_j(1)^{c_j+2}} - \frac{G_j(f_j+1)}{G_j(1)^{c_j+1}} \leq \frac{G_j(f_j-1)}{G_j(1)^{c_j+1}} - \frac{G_j(f_j)}{G_j(1)^{c_j}}\ , \tag{4}$$

or equivalently,

$$G_j(f_j) - G_j(f_j+1)G_j(1) \leq G_j(f_j-1)G_j(1) - G_j(f_j)G_j^2(1)\,.$$

Since by (3) $G_j(f_j-1)G_j(1)+G_j(f_j+1)G_j(1) = 2G_j(f_j)G_j^2(1)$, the last inequality reduces to $G_j(f_j) \leq G_j(f_j)G_j^2(1)$, which is valid as $G_j(1) \geq 1$ by (2).

Observe that before the game starts $\phi = \sum_{j=1}^{m} e^{-\frac{b_j}{2|e_j|}} \leq 1/2$. Hence if Balancer follows the suggested strategy, he can achieve $\phi \leq 1/2$ in the end. This implies:

$$\phi_j = e^{-\frac{b_j^2}{2|e_j|}} \frac{G_j(f_j)}{G_j(1)^{|e_j|}} \leq 1/2$$

for each edge e_j of H. Therefore by (2)

$$\frac{e^{\alpha_j|f_j|}}{2} \leq \frac{1}{2} e^{\frac{b_j^2}{2|e_j|}} G_j(1)^{|e_j|} \leq \frac{1}{2} e^{\frac{b_j^2}{2|e_j|}+\frac{\alpha_j^2|e_j|}{2}}\ ,$$

or

$$|f_j| \leq \frac{b_j^2}{2\alpha_j|e_j|} + \frac{\alpha_j|e_j|}{2} = \frac{b_j}{2} + \frac{b_j}{2} = b_j\ .$$

Therefore, Balancer wins, as claimed. □

3 Applications and remarks

1. The first proof of the theorem works under slightly more general circumstances, i.e. when the conditions are not necessarily required to be symmetric or possibly only one of the positive and negative conditions are required. More formally, let (a_j) and (b_j) be a sequence of negative and positive integers, respectively, not necessarily defined for every $j = 1, \ldots, m$. Then Balancer can achieve that at the end of the game $a_j \leq f(e_j) \leq b_j$ for every $j = 1, \ldots, m$, provided

$$\sum_{i=1}^{m} e^{-\frac{a_i^2}{2|e_i|}} + \sum_{i=1}^{m} e^{-\frac{b_i^2}{2|e_i|}} \leq \frac{1}{2}.$$

Note that the second proof trivially implies a slightly stronger assertion: Balancer can ensure that $|f(e_j)| \leq min(|a_j|, b_j)$ for all j.

2. If Unbalancer moves first and/or the board is of odd size, than the obtained target vector (b_j) might be an additive constant 1 or 2 away from the one provided by the formula of Theorem 1.

3. Theorem 1 is obviously tight up to a multiplicative factor when $H = \binom{[n]}{n/2}$. A less trivial setting is the so-called "maximum degree game". There V consists of the edges of the complete graph on $[n]$ and E is the $(n-1)$-uniform set system consisting of the n stars. Theorem 1 implies that Balancer can achieve having all of his degrees between $\frac{n}{2} - \sqrt{n \ln(2n)/2}$ and $\frac{n}{2} + \sqrt{n \ln(2n)/2}$. This improves the error term obtained in [6] for this problem by a constant factor. Beck [3] proved that Unbalancer can get one of his degrees to be at least $n/2 + c\sqrt{n}$, i.e., Theorem 1 is tight up to a logarithmic factor in this case.

4. The statement of Theorem 1 supersedes the method of [6] in several aspects. Namely it applies to the non-uniform case, it is more straightforward to apply, while giving comparable and often better results. To demonstrate this we derive a variant of one of the main results of [6] with an improved constant factor as a corollary of Theorem 1 of our paper. As it was shown in [6] the multiplicative constant we obtain in the following corollary in the bound for ϵ is at most $\sqrt[3]{2}$ away from being best possible.

Corollary 1 *(compare to [6, Theorem 1]) Let $\epsilon \geq (\log n/3n)^{1/3}$. Playing against Breaker on the edges of K_n, Maker can build a graph G with the following properties.*

- **P1:** *All degrees of G deviate from $n/2$ by at most $\sqrt{n \log(4n)/2}$.*
- **P2:** *Any pair S, T of disjoint subsets of $[n]$ with $|S|, |T| \geq \epsilon n$ is* ϵ-unbiased, *i.e. they satisfy*

$$\left| \frac{e_G(S,T)}{|S||T|} - \frac{1}{2} \right| \leq \epsilon,$$

where $e_G(S,T)$ is the number of $S - T$ edges in G.

Proof. Let $\mathcal{H}_1$ be the hypergraph whose members are the n stars of size $n-1$ in K_n. The members of the hypergraph $\mathcal{H}_2$ are the sets $[S,T]$ of edges for an arbitrary pair S, T of disjoint subsets of the vertex set with $|S| = |T| = \epsilon n$, where $[S, T]$ represents the set of all edges connecting a vertex of S with a vertex of T. An easy averaging argument shows that it is enough to guarantee property P2 for subsets of size exactly ϵn. Now let Maker act as Balancer in a game of Discrepancy on the hypergraph $\mathcal{H}_1 \cup \mathcal{H}_2$ with a target vector $b_j = \sqrt{2n \log(4n)}$ for $e_j \in \mathcal{H}_1$ and $b_j = 2\epsilon|S||T| = 2\epsilon^3 n^2$ for $e_j \in \mathcal{H}_2$. A straightforward substitution into Theorem 1 gives the corollary. □

5. In [9] Székely studies the function $\Delta_2(\mathcal{F})$, which is the largest bias Red can achieve on a member of $\mathcal{F}$ playing against Blue. More precisely, let $(X, \mathcal{F})$ be a hypergraph. Red and Blue select one element of the board X alternately. Then $\Delta_2(\mathcal{F})$ is the largest integer k such that Red has a strategy to occupy k elements *more* of some member A of $\mathcal{F}$ than Blue at some point during the game.

In part (i) of his Corollary he shows that if $\mathcal{F}$ is an n-uniform hypergraph, then $\Delta_2(\mathcal{F}) \leq cn^{1/2} \log^{1/2} |\mathcal{F}|$. Theorem 1 provides a similar result, but gives not only a one- but a two-sided discrepancy bound.

In part (ii) of his Corollary Székely shows that if $n = \max_{A \in \mathcal{F}} |A|$, then $\Delta_2(\mathcal{F}) \leq cn^{2/3} \log^{1/3} |\mathcal{F}|$. He later uses this result to derive an upper bound when the board X_n is the n-by-n integer lattice and $\mathcal{F}_n$ is the family of all maximal sets of lattice points contained in a line segment. Theorem 1 now easily implies an improvement of the Corollary and thus of the Theorem of Székely on line segments. Namely, $\Delta_2(\mathcal{F}) \leq cn^{1/2} \log^{1/2} |\mathcal{F}|$ in both cases. Actually an even stronger result is true, i.e. Blue can achieve this bound not only for one-, but for two-sided discrepancy against Red. Moreover the error term is not absolute, but depends on the size of the actual line segment.

4 An open question

We do not know how to generalize our approach to the biased game on H. In the $(p : q)$-game Balancer selects p and then Unbalancer q elements of the board. Balancer's goal is to have around $p/(p + q)$ vertices in every edge. It would be very interesting to obtain a criteria for Balancer's win. Such a result would have many applications as it can be used to create a pseudo-random graph of density $p/(p+q)$ which in turn guarantees Balancer's win in many other games (see [6]).

Acknowledgement. Thanks are due to Valentyn Vengerovskyi for spotting an inaccurate point in the first version of the paper.

References

[1] N. Alon and J. Spencer, The probabilistic method, Second Edition, John Wiley and Sons, 2000.

[2] J. Beck, Van der Waerden and Ramsey type games, *Combinatorica* **1**(2) (1981), 103–116.

[3] J. Beck, Deterministic graph games and a probabilistic intuition, *Combinatorics, Probability and Computing* **3** (1994), 13–26.

[4] P. Erdős and J. Selfridge, On a combinatorial game, *Journal of Combinatorial Theory* **14** (1973), 298–301.

[5] A. M. Frieze and M. Krivelevich, On packing Hamilton cycles in ϵ-regular graphs, *Journal of Combinatorial Theory Series B* **94** (2005), 157–172.

[6] A. M. Frieze, M. Krivelevich, O. Pikhurko and T. Szabó, The game of JumbleG, *Combinatorics, Probability and Computing*, to appear.

[7] X. Lu, A matching game, *Discrete Mathematics*, **94** (1991), 199-207.

[8] P. Raghavan, Probabilistic construction of deterministic algorithms: approximating packing integer programs, *Journal of Computer and System Sciences*, **37** (1988) 130–143.

[9] L. A. Székely, On two concepts of discrepancy in a class of combinatorial games, Finite and Infinite Sets, *Colloq. Math. Soc. János. Bolyai*, Vol. **37**, North-Holland, 1984, 679–683.