4. Combinatorial Discrepancy

In this chapter, we are going to investigate the combinatorial discrepancy, an exciting and significant subject in its own right. From Section 1.3, we recall the basic definition: If X is a finite set and $\mathcal{S} \subseteq 2^X$ is a family of sets on X, a *coloring* is any mapping $\chi: X \to \{-1, +1\}$, and we have $\mathrm{disc}(\mathcal{S}) = \min_\chi \max_{S \in \mathcal{S}} |\chi(S)|$, where $\chi(S) = \sum_{x \in S} \chi(x)$.

In Section 4.1, we prove some general upper bounds for $\mathrm{disc}(\mathcal{S})$ expressed in terms of the number and size of the sets in $\mathcal{S}$, and also a bound in terms of the maximum degree of $\mathcal{S}$. Section 4.2 discusses a technique for bounding discrepancy from below, related to matrix eigenvalues. Section 4.3 reviews variations on the notion of discrepancy, such as the linear discrepancy and the hereditary discrepancy, and it gives another general lower bound, in terms of determinants. The subsequent section considers set systems with discrepancy 0 and those with hereditary discrepancy at most 1. (The material in Sections 4.2 through 4.4 will not be used in the rest of this book.)

In Section 4.5, we introduce one of the most powerful techniques for upper bounds in discrepancy theory: the partial coloring method. Section 4.6 deals with a refinement of the partial coloring method, called the entropy method. With this approach, bounds obtained by the partial coloring method can often be improved by logarithmic factors. For several important problems, this it is the only known technique leading to asymptotically tight bounds.

4.1 Basic Upper Bounds for General Set Systems

We begin with the following question. Let X be an n-point set and let $\mathcal{S}$ be a set system on X having m sets. What is the maximum possible value, over all choices of $\mathcal{S}$, of $\mathrm{disc}(\mathcal{S})$? We will be most interested in the case when $n \leq m$ (more sets than points). This is what we usually have in geometric situations, and it also turns out that the $m < n$ case can essentially be reduced to the $m = n$ case (see Theorem 4.9).

A quite good upper bound for the discrepancy is obtained by using a random coloring.

4.1 Lemma (Random coloring lemma). *Let $\mathcal{S}$ be a set system on an n-point set X. For a random coloring $\chi: X \to \{+1, -1\}$, the inequalities*

J. Matoušek, *Geometric Discrepancy*, Algorithms and Combinatorics 18,
DOI 10.1007/978-3-642-03942-3_4, © Springer-Verlag Berlin Heidelberg 2010

$$|\chi(S)| \le \sqrt{2|S|\ln(4|\mathcal{S}|)}$$

hold for all sets $S \in \mathcal{S}$ simultaneously with probability at least $\frac{1}{2}$.

Note that if we know that $\mathcal{S}$ has at most m sets and have no information about their sizes, we get the upper bound $\mathrm{disc}(\mathcal{S}) = O(\sqrt{n \log m})$. Moreover, the above formulation shows that a random coloring gives better discrepancy for smaller sets, and this may be useful in some applications.

Proof. This is similar to considerations made in the proof of Theorem 3.1 (upper bound for the discrepancy for discs), and actually simpler. For any fixed set $S \subseteq X$, the quantity $\chi(S) = \sum_{x \in S} \chi(x)$ is a sum of $s = |S|$ independent random ± 1 variables. Such a sum has a binomial distribution, with standard deviation $\sqrt{s}$, and the simplest form of the Chernoff tail estimate (see Alon and Spencer [AS00]) gives

$$\Pr\left[|\chi(S)| > \lambda\sqrt{s}\,\right] < 2e^{-\lambda^2/2}.$$

Hence, if we set $\lambda = \sqrt{2\ln(4|\mathcal{S}|)}$, the above bound becomes $1/(2|\mathcal{S}|)$, and, with probability at least $\frac{1}{2}$, a random coloring satisfies $|\chi(S)| \le \sqrt{2|S|\ln(4|\mathcal{S}|)}$ for all $S \in \mathcal{S}$. □

The following theorem is a small improvement over the lemma just proved, at least if the set sizes are not much smaller than n:

4.2 Theorem (Spencer's upper bound). *Let $\mathcal{S}$ be a set system on an n-point set X with $|\mathcal{S}| = m \ge n$. Then*

$$\mathrm{disc}(\mathcal{S}) = O\left(\sqrt{n \log(2m/n)}\right).$$

In particular, if $m = O(n)$ then $\mathrm{disc}(\mathcal{S}) = O(\sqrt{n})$.

We will prove this result in Section 4.6. A probabilistic construction shows that this bound is tight in the worst case (see Exercise 1 or Alon and Spencer [AS00]). For $m = n$, there is a simple constructive lower bound based on Hadamard matrices, which we present in Section 4.2.

Another important upper bound, which we will not use but which is definitely worth mentioning, is this:

4.3 Theorem (Beck–Fiala theorem). *Let $\mathcal{S}$ be a set system on an arbitrary finite set X such that $\deg_{\mathcal{S}}(x) \le t$ for all $x \in X$, where $\deg_{\mathcal{S}}(x) = |\{S \in \mathcal{S}\colon x \in S\}|$. Then $\mathrm{disc}(\mathcal{S}) \le 2t - 1$.*

Proof. Let $X = \{1, 2, \ldots, n\}$. To each $j \in X$, assign a real variable $x_j \in [-1, 1]$ which will change as the proof progresses. Initially, all the x_j are 0. In the end, all x_j will be $+1$ or -1 and they will define the required coloring.

At each step of the proof, some of the variables x_j are "fixed" and the others are "floating;" initially all variables are floating. The fixed variables have values $+1$ or -1 and their value will not change anymore. The floating variables have values in $(-1, 1)$. At each step, at least one floating variable becomes fixed. Here is how this happens.

Call a set $S \in \mathcal{S}$ *dangerous* if it contains more than t elements j with x_j currently floating, and call S *safe* otherwise. The following invariant is always maintained:

$$\sum_{j \in S} x_j = 0 \quad \text{for all dangerous } S \in \mathcal{S}. \tag{4.1}$$

Let F be the current set of indices of the floating variables, and let us regard (4.1) as a system of linear equations whose unknowns are the floating variables. This system certainly has a solution, namely the current values of the floating variables. Since we assume $-1 < x_j < 1$ for all floating variables, this solution is an interior point of the cube $[-1, 1]^F$. We want to show that there also exists a solution lying on the boundary of this cube, i.e. such that at least one unknown has value $+1$ or -1. The crucial observation is that the number of dangerous sets at any given moment is smaller than the number of floating variables (this follows by a simple double counting of incidences of the floating indices j with the dangerous sets). Hence our system of linear equations has fewer equations than unknowns, and therefore the solution space contains a line. This line intersects the boundary of the cube $[-1, 1]^F$ at some point z. The coordinates of this point specify the new value of the floating variables for the next step; however, the variables x_j for which $z_j = \pm 1$ become fixed.

This step is iterated until all the x_j become fixed. We claim that their values specify a coloring with discrepancy at most $2t - 1$. Indeed, consider a set $S \in \mathcal{S}$. At the moment when it became safe, it had discrepancy 0 by (4.1). At this moment it contained at most t indices of floating variables. The value of each of these floating variables might have changed by less than 2 in the remaining steps (it might have been -0.999 and become $+1$, say). This concludes the proof. □

Remark. Beck and Fiala conjectured that in fact $\mathrm{disc}(\mathcal{S}) = O(\sqrt{t})$ holds under the assumptions of their theorem but no proof is known. The Beck–Fiala theorem remains the best known bound in terms of the maximum degree alone (except for a tiny improvement of the bound $2t - 1$ to $2t - 3$).

Remark on Algorithms. For statements establishing upper bounds for discrepancy of a set system $(X, \mathcal{S})$, it is interesting to learn whether they provide a polynomial-time algorithm (polynomial in $|X|$ and $|\mathcal{S}|$) for computing a coloring with the guaranteed discrepancy. For the Random coloring lemma, a randomized algorithm is obvious, and it can be made deterministic (slower but still polynomial) by the method of conditional probabilities; see [AS00]. The proof of the Beck–Fiala theorem 4.3 also provides a polynomial algorithm, but the proof of Spencer's upper bound 4.2 does not—it is a big challenge to find one.

Bibliography and Remarks. Spencer's theorem 4.2 is from Spencer [Spe85]; alternative proofs were given by Gluskin [Glu89] via Minkowski's lattice point theorem and by Giannopoulos [Gia97] using the Gaussian measure. The Beck–Fiala theorem is from [BF81] (and the improvement from $2t-1$ to $2t-3$ is in [BH97]). Exercise 1 is in the spirit of a lower-bound proof presented in [AS00]. For more bounds related to the Beck–Fiala theorem see the remarks to Section 4.3 and the end of Section 5.5.

In theoretical computer science, an intriguing question is an efficient computation of a coloring with small discrepancy for a given set system. In cases where randomized algorithms are easy, such as for the Random coloring lemma, the task is to find an efficient deterministic counterpart (i.e. to *derandomize* the algorithm). A related question is to parallelize the algorithms efficiently. Some such issues are treated in Spencer [Spe87] already; a sample of newer references are Berger and Rompel [BR91] and Srinivasan [Sri97].

Exercises

1. (a)* Let $S = \sum_{i=1}^n S_i$ be a sum of n independent random variables, each attaining values $+1$ and -1 with equal probability. Let $P(n,\Delta) = \Pr[S > \Delta]$. Prove that for $\Delta \le n/C$,
$$P(n,\Delta) \ge \tfrac{1}{C}\exp\bigl(-\tfrac{\Delta^2}{Cn}\bigr),$$
where C is a suitable constant. That is, the well-known Chernoff bound $P(n,\Delta) \le \exp(-\Delta^2/2n)$ is close to the truth. (For very precise lower bounds, proved by quite different methods, see Feller [Fel43].)
(b)* Let $X = \{1,2,\ldots,n\}$ be a ground set, let $\chi\colon X \to \{+1,-1\}$ be any fixed coloring of X, and let R be a random subset of X (a random subset means one where each i is included with probability $\frac{1}{2}$, the choices being independent for distinct i). Prove that for any $\Delta \ge 0$, $\Pr[|\chi(R)| \ge \Delta] \ge P(n,2\Delta)$, where $P(.,.)$ is as in (a).
(c) Let $\mathcal{R}$ be a system of $m \ge n$ independently chosen random subsets of $\{1,2,\ldots,n\}$, and let $c_1 > 0$ be a sufficiently small constant. Use (a), (b) to show that $\operatorname{disc}(\mathcal{R}) > c_1\sqrt{n\log(2m/n)}$ holds with a positive probability, provided that $m \le 2^{c_1 n}$; that is, Theorem 4.2 is asymptotically tight.
2. (Discrepancy of the product of set systems) Let $\mathcal{S}$ and $\mathcal{T}$ be set systems (on finite sets). We let $\mathcal{S}\times\mathcal{T} = \{S\times T\colon S\in\mathcal{S}, T\in\mathcal{T}\}$.
(a) Show that $\operatorname{disc}(\mathcal{S}\times\mathcal{T}) \le \operatorname{disc}(\mathcal{S})\operatorname{disc}(\mathcal{T})$.
(b)* Find an example with $\operatorname{disc}(\mathcal{S}) > 0$ and $\operatorname{disc}(\mathcal{S}\times\mathcal{S}) = 0$.
These results are due to Doerr; see [DSW04].

4.2 Matrices, Lower Bounds, and Eigenvalues

Let $(X, \mathcal{S})$ be a set system on a finite set. Enumerate the elements of X as $x_1, x_2, \ldots, x_n$ and the sets of $\mathcal{S}$ as $S_1, S_2, \ldots, S_m$, in some arbitrary order. The *incidence matrix* of $(X, \mathcal{S})$ is the $m \times n$ matrix A, with columns corresponding to points of X and rows corresponding to sets of $\mathcal{S}$, whose element a_{ij} is given by

$$a_{ij} = \begin{cases} 1 & \text{if } x_j \in S_i \\ 0 & \text{otherwise.} \end{cases}$$

As we will see, it is useful to reformulate the definition of discrepancy of $\mathcal{S}$ in terms of the incidence matrix. Let us regard a coloring $\chi\colon X \to \{-1, +1\}$ as the column vector $(\chi(x_1), \chi(x_2), \ldots, \chi(x_n))^T \in \mathbf{R}^n$. Then the product $A\chi$ is the row vector $(\chi(S_1), \chi(S_2), \ldots, \chi(S_m)) \in \mathbf{R}^m$. Therefore, the definition of the discrepancy of $\mathcal{S}$ can be written

$$\operatorname{disc}(\mathcal{S}) = \min_{x \in \{-1,1\}^n} \|Ax\|_\infty,$$

where the norm $\|.\|_\infty$ of a vector $y = (y_1, y_2, \ldots, y_m)$ is defined by $\|y\|_\infty = \max_i |y_i|$. The right-hand side of the above equation can be used as a definition of discrepancy for an arbitrary real matrix A.

Expressing discrepancy via incidence matrices helps in obtaining lower bounds. For many lower bound techniques, it is preferable to consider the L_2-discrepancy instead of the worst-case discrepancy. In our case, this means replacing the max-norm $\|.\|_\infty$ by the usual Euclidean norm $\|.\|$, which is usually easier to deal with. Namely, we have

$$\operatorname{disc}(\mathcal{S}) \geq \operatorname{disc}_2(\mathcal{S}) = \min_\chi \left(\frac{1}{m} \sum_{i=1}^m \chi(S_i)^2 \right)^{1/2} = \frac{1}{\sqrt{m}} \cdot \min_{x \in \{-1,1\}^n} \|Ax\|.$$

Since $\|Ax\|^2 = (Ax)^T(Ax) = x^T(A^TA)x$, we can further rewrite

$$\operatorname{disc}_2(\mathcal{S}) = \left(\frac{1}{m} \min_{x \in \{-1,1\}^n} x^T(A^TA)x \right)^{1/2}. \tag{4.2}$$

Now we present the example of n sets on n points with discrepancy about $\sqrt{n}$ promised in Section 4.1. To this end, we first recall the notion of an *Hadamard matrix*. This is an $n \times n$ matrix H with entries $+1$ and -1 such that any two distinct columns are orthogonal; in other words, we have $H^TH = nI$, where I stands for the $n \times n$ identity matrix. Since changing all signs in a row or in a column preserves this property, one usually assumes that the first row and the first column of the considered Hadamard matrix consist of all 1's. We will also use this convention.

Hadamard matrices do not exist for every n. For example, it is clear that for $n \geq 2$, n has to be even if an $n \times n$ Hadamard matrix should exist. The existence problem for Hadamard matrices is not yet fully solved, but various constructions are known. We recall only one simple recursive construction, providing a $2^k \times 2^k$ Hadamard matrix for all natural numbers k. Begin with the 1×1 matrix $H_0 = (1)$, and, having defined a $2^{k-1} \times 2^{k-1}$ matrix H_{k-1}, construct H_k from four blocks as follows:

$$\begin{pmatrix} H_{k-1} & H_{k-1} \\ H_{k-1} & -H_{k-1} \end{pmatrix}.$$

Thus, we have

$$H_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

The orthogonality is easy to verify by induction.

4.4 Proposition (Hadamard set system). *Let H be an $n \times n$ Hadamard matrix, and let $\mathcal{S}$ be the set system with incidence matrix $A = \frac{1}{2}(H + J)$, where J denotes the $n \times n$ matrix of all 1's (in other words, A arises from H by changing the -1's to 0's). Then* $\operatorname{disc}(\mathcal{S}) \geq \operatorname{disc}_2(\mathcal{S}) \geq \frac{\sqrt{n-1}}{2}$.

Proof of Proposition 4.4. We have $A^T A = \frac{1}{4}(H+J)^T(H+J) = \frac{1}{4}(H^T H + H^T J + J^T H + J^T J) = \frac{1}{4}(nI + nR + nR^T + nJ)$, where $R = \frac{1}{n} H^T J$ is the matrix with 1's in the first row and 0's in the other rows. Therefore, for any $x \in \{-1,1\}^n$, we get

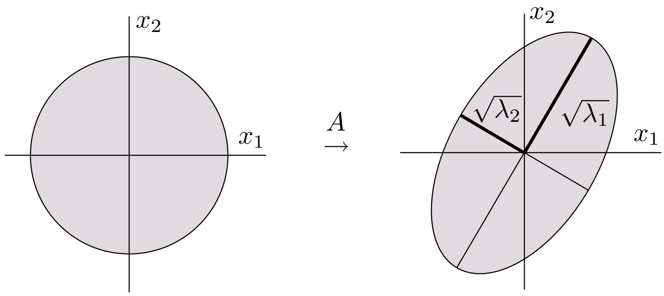
$$\begin{aligned} \frac{1}{n} \cdot x^T (A^T A) x &= \frac{1}{4} \left(\sum_{i=1}^n x_i^2 + 2x_1 \left(\sum_{i=1}^n x_i \right) + \left(\sum_{i=1}^n x_i \right)^2 \right) \\ &= \frac{1}{4} \left(\sum_{i=2}^n x_i^2 + (2x_1 + x_2 + \cdots + x_n)^2 \right) \\ &\geq \frac{1}{4} \left(\sum_{i=2}^n x_i^2 \right) = \frac{n-1}{4}, \end{aligned}$$

and so $\operatorname{disc}(\mathcal{S}) \geq \frac{\sqrt{n-1}}{2}$ follows from (4.2). □

A slightly different treatment of this result is outlined in Exercise 3. The proof just given used the estimate $\operatorname{disc}(\mathcal{S}) \geq \left(\frac{1}{m} \min_{x \in \{-1,1\}^n} x^T (A^T A) x \right)^{1/2}$. Often it is useful to take the minimum on the right-hand side over a larger set of vectors: instead of the discrete set $\{-1,1\}^n$, we minimize over all real vectors with Euclidean norm $\sqrt{n}$. (Combinatorially, this means that we allow "generalized colorings" assigning real numbers to the points of X, and we only require that the sum of squares of these numbers is the same as if we used ± 1's.) So we have

$$\begin{aligned}\mathrm{disc}_2(\mathcal{S}) &\geq \left(\frac{1}{m}\cdot\min_{\|x\|=\sqrt{n}} x^T(A^TA)x\right)^{1/2}\\ &= \left(\frac{n}{m}\cdot\min_{\|x\|=1} x^T(A^TA)x\right)^{1/2}. \end{aligned}\tag{4.3}$$

The right-hand side of this inequality can naturally be called the *eigenvalue bound* for the discrepancy of $\mathcal{S}$. This is because for any real $m\times n$ matrix A, the matrix $B = A^TA$ is positive semidefinite, it has n nonnegative eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and for the smallest eigenvalue λ_n we have $\lambda_n = \min_{\|x\|=1} x^TBx$. (The eigenvalues of A^TA are often called the *singular values* of the matrix A.) All this is well-known in linear algebra, and not too difficult to see, but perhaps it is useful to recall a geometric view of the situation. For simplicity, suppose that $m = n$. Then the mapping $x \mapsto Ax$ is a linear mapping of $\mathbf{R}^n$ into $\mathbf{R}^n$, and it maps the unit sphere onto the surface of an ellipsoid (possibly a flat one if A is singular). The quantity $\min_{\|x\|=1}\|Ax\|$ is the length of the shortest semiaxis of this ellipsoid. At the same time, the lengths of the semiaxes are $\sqrt{\lambda_1}$, $\sqrt{\lambda_2}$,..., $\sqrt{\lambda_n}$, where the λ_i are eigenvalues of A^TA as above. Here is an illustration for $n = 2$:



For $m > n$, the mapping $x \mapsto Ax$ maps $\mathbf{R}^n$ to an n-dimensional linear subspace of $\mathbf{R}^m$, and the image of the unit ball is an ellipsoid in this subspace.

The eigenvalue bound can be smaller than the L_2-discrepancy, but the eigenvalues of a matrix are efficiently computable and there are various techniques for estimating them. The following theorem summarizes our discussion:

4.5 Theorem (Eigenvalue bound for discrepancy). *Let $(\mathcal{S}, X)$ be a system of m sets on an n-point set, and let A denote its incidence matrix. Then we have*

$$\mathrm{disc}(\mathcal{S}) \geq \mathrm{disc}_2(\mathcal{S}) \geq \sqrt{\frac{n}{m}\cdot\lambda_{min}},$$

where λ_{min} denotes the smallest eigenvalue of the $n\times n$ matrix A^TA.

A very neat application of the eigenvalue bound concerns the discrepancy of a finite projective plane (Exercise 5.1.5). A more advanced example is the lower bound for the discrepancy of arithmetic progressions (a version due to Lovász; see Exercise 5). Moreover, numerous lower bounds in the Lebesgue-measure setting are in fact continuous analogues of the eigenvalue bound, in the sense that they hold for the L_2-discrepancy and for point sets with arbitrary weights, although they usually are not stated in this form (see Chapters 6 and 7). On the other hand, we should remark that for the Hadamard set system from Proposition 4.4, the eigenvalue bound fails miserably (Exercise 4). This can be fixed by deleting one set and one point (Exercise 3).

Bibliography and Remarks. An early result in combinatorial discrepancy was Roth's $\Omega(n^{1/4})$ lower bound on the discrepancy of arithmetic progressions [Rot64]. This beautiful proof uses harmonic analysis. Lovász suggested a version based on the eigenvalue bound, which is outlined in Exercise 5. Roth thought that the bound $n^{1/4}$ was too small and that the discrepancy was actually close to $n^{1/2}$. This was disproved by Sárközy (see [ES74]), who established an $O(n^{1/3+\varepsilon})$ upper bound. Beck [Bec81b] obtained the near-tight upper bound of $O(n^{1/4}\log^{5/2})$, inventing the powerful partial coloring method (see Section 4.5) for that purpose. The asymptotically tight upper bound $O(n^{1/4})$ was finally proved by Matoušek and Spencer [MS96]. (Proofs of slightly weaker upper bounds are indicated in Exercises 4.5.7 and 5.5.4.) Knieper [Kni98] generalized Roth's lower bound to the set system of *multidimensional arithmetic progressions*; these are the subsets of $\{1,2,\ldots,n\}^d$ of the form $A_1\times A_2\times\cdots\times A_d$, where each A_i is an arithmetic progression in $\{1,2,\ldots,n\}$ (note that the size of the ground set is n^d rather than n). The lower bound is $\Omega(n^{d/4})$ (d fixed), and this is easily seen to be asymptotically tight, using the $O(n^{1/4})$ bound for the case $d=1$ and the observation on product hypergraphs in Exercise 4.1.2 .

Proposition 4.4, with the approach as in Exercise 3 below, is due to Olson and Spencer [OS78]. The eigenvalue bound (Theorem 4.5) is attributed to Lovász and Sós in [BS95]. For another convenient formulation of the eigenvalue bound see Exercise 1.

Exercises

1. Let $(X,\mathcal{S})$ and A be as in Theorem 4.5. Show that if D is a diagonal matrix such that A^TA-D is positive semidefinite, then $\mathrm{disc}_2(\mathcal{S})\geq\sqrt{\frac{1}{m}\mathrm{trace}(D)}$, where $\mathrm{trace}(M)$ denotes the sum of the diagonal elements of a square matrix M.
2. Let the rows of a $2^k\times 2^k$ matrix H be indexed by the k-component $0/1$ vectors, and let the columns be indexed similarly. Let the entry of H corresponding to vectors x and y be $+1$ if the scalar product $\langle x,y\rangle$ is

even and -1 if $\langle x, y\rangle$ is odd. Check that H is a Hadamard matrix (in fact, the same one as constructed recursively in the text).

3. Let H be a $4n \times 4n$ Hadamard matrix, and let A be the $(4n-1)\times(4n-1)$ matrix arising from H by deleting the first row and first column and changing the -1's to 0's.
(a) Verify that A is the incidence matrix of a 2-$(4n-1, 2n-1, n-1)$-design, meaning that there are $4n-1$ points, all sets have size $2n-1$, and any pair of distinct points is contained in exactly $n-1$ sets.
(b) Show that the eigenvalue bound (Theorem 4.5) gives at least $\sqrt{n}$ for the matrix A.
4.* Let $A = \frac{1}{2}(H+J)$ be the incidence matrix of the set system as in Proposition 4.4. Show that the eigenvalue bound (Theorem 4.5) is quite weak for A, namely that the smallest eigenvalue of $A^T A$ is $O(1)$. (Note that in the proof of Proposition 4.4, we used the fact that all components of x are ± 1.)
5. (Discrepancy of arithmetic progressions) Let k be an integer; let $n = 6k^2$, and define the set system $\mathcal{A}_0$ on $\{0, 1, \ldots, n-1\}$ ("wrapped arithmetic progressions of length k with difference $\leq 6k$") as follows:
$$\begin{aligned}\mathcal{A}_0 = \{\{&i, (i+d) \bmod n, (i+2d) \bmod n, \ldots, \\ &(i+(k-1)d) \bmod n\}\colon d = 1, 2, \ldots, 6k,\ i = 0, 1, \ldots, n-1\}\end{aligned}$$
(a)** Use Theorem 4.5 to prove that $\operatorname{disc}(\mathcal{A}_0) = \Omega(n^{1/4})$.
(b) Deduce that the system of all arithmetic progressions (without wrapping) on $\{0, 1, \ldots, n-1\}$ has discrepancy $\Omega(n^{1/4})$.
6. Call a mapping $\chi\colon X \to \{+1, -1, 0\}$ *perfectly balanced* on a set system $(X, \mathcal{S})$ if $\chi(S) = 0$ for all $S \in \mathcal{S}$. Define $g(m)$ as the maximum possible size of X such that there exists a system $\mathcal{S}$ of m sets on X for which any perfectly balanced mapping χ is identically 0.
(a)* Show that $g(m)$ equals the maximum number n of columns of an $m \times n$ zero-one matrix A such that $\sum_{i\in I} a_i \neq \sum_{j\in J} a_j$ whenever I, J are distinct subsets of $\{1, 2, \ldots, n\}$, where a_i denotes the ith column of A.
(b) Deduce the bound $g(m) = O(m \log m)$ from (a).
(c) Prove that the discrepancy of an arbitrary system of m sets is always bounded by the maximum possible discrepancy of m sets on $g(m)$ points.

These results are from Olson and Spencer [OS78]. They also show that the bound in (b) is asymptotically tight.

4.3 Linear Discrepancy and More Lower Bounds

The discrepancy of a set system $\mathcal{S}$ can be thought of as a certain measure of "complexity" of $\mathcal{S}$. But from this point of view, it is not a very good measure, since $\operatorname{disc}(\mathcal{S})$ may happen to be small just "by chance." For example, let X be a disjoint union of two n-point sets Y and Z, and let $\mathcal{S}$ consist of all sets $S \subseteq X$ with $|S \cap Y| = |S \cap Z|$. Then $\operatorname{disc}(\mathcal{S}) = 0$ although we feel that $\mathcal{S}$ is nearly as complicated as the system of all subsets of X. A better concept from this point of view is the *hereditary discrepancy* of $\mathcal{S}$, defined as

$$\operatorname{herdisc}(\mathcal{S}) = \max_{Y \subseteq X} \operatorname{disc}(\mathcal{S}|_Y).$$

(Or, using our previously introduced notation, we can also write $\operatorname{herdisc}(\mathcal{S}) = \max_{0 \le m \le n} \operatorname{disc}(m, \mathcal{S})$.) As the just mentioned example shows, the hereditary discrepancy can be much larger than the discrepancy.

Another useful concept is the *linear discrepancy* of $\mathcal{S}$. It arises in the in the following "rounding" problem. Each point $x \in X$ is assigned a weight $w(x) \in [-1, 1]$. We want to color the points by $+1$'s and -1's in such a way that the sum of the colors in each set $S \in \mathcal{S}$ is close to the total weight of its points. The discrepancy of $\mathcal{S}$ with respect to the given weights is thus

$$\min_{\chi: X \to \{-1,1\}} \max_{S \in \mathcal{S}} |\chi(S) - w(S)|,$$

and the linear discrepancy of $\mathcal{S}$ is the supremum of this quantity over all choices of the weight function $w: X \to [-1, 1]$. (The usual discrepancy corresponds to the case $w(x) = 0$ for all x.) The linear discrepancy can again be defined for an arbitrary matrix A. Namely, we have

$$\operatorname{lindisc}(A) = \max_{w \in [-1,1]^n} \min_{x \in \{-1,1\}^n} \|A(x - w)\|_\infty.$$

Clearly $\operatorname{lindisc}(A) \ge \operatorname{disc}(A)$ for any matrix A, and one cannot expect to bound lindisc in terms of disc. But, perhaps surprisingly, the following bound in terms of the hereditary discrepancy holds:

4.6 Theorem (lindisc $\le$ 2·herdisc). *For any set system $\mathcal{S}$, we have*

$$\operatorname{lindisc}(\mathcal{S}) \le 2 \operatorname{herdisc}(\mathcal{S}).$$

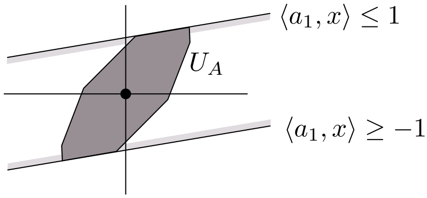
The same inequality holds between the linear and hereditary discrepancies for an arbitrary real matrix A.

This result can be proved in way somewhat similar to the proof of the transference lemma (Proposition 1.8). A proof in this spirit can be found in Spencer [Spe87], but here we give another proof using a geometric interpretation of the various notions of discrepancy. Unlike to the geometric discrepancy mostly treated in this book, the geometry here is introduced into the picture somewhat artificially, but once it is there it constitutes a powerful tool.

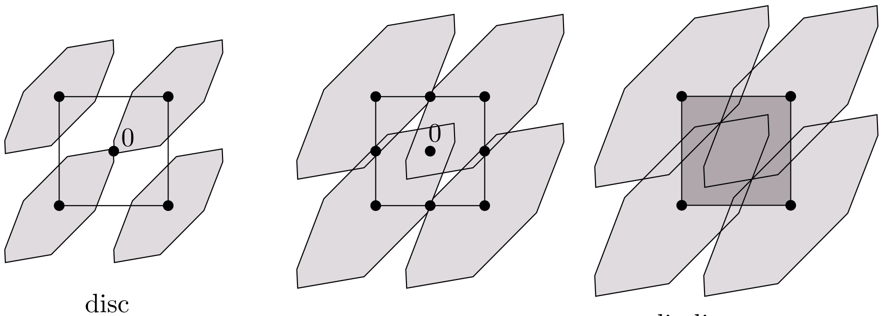
Let A be an $m \times n$ matrix, and let us define the set

$$U_A = \{x \in \mathbf{R}^n : \|Ax\|_\infty \le 1\}.$$

This U_A is a convex polyhedron symmetric about the origin, as is illustrated in the following picture for $n = 2$:



If a_i denotes the ith row of A, U_A is the intersection of the $2m$ halfspaces $\langle a_i, x\rangle \le 1$ and $\langle a_i, x\rangle \ge -1$, as is marked in the picture for $i = 1$. For any vector $x \in \mathbf{R}^n$, we have $\|Ax\|_\infty = \min\{t \ge 0\colon x \in tU_A\}$, and so disc$(A)$ is the smallest t such that the convex body tU_A contains a vertex of the cube $[-1,1]^n$. In other words, disc(A) is the smallest t such that for some vertex $a \in \{-1,1\}^n$, the translated body $tU_A + a$ contains the origin. This geometric interpretation of disc(A) allows perhaps the best comparison with the other notions of discrepancy introduced above. Their geometric interpretation is indicated in the following picture:



We can imagine that at time $t = 0$, we start growing a similar copy of U_A from each vertex of the cube $[-1,1]^n$, in such a way that at time t, we have a congruent copy of tU_A centered at each vertex. The reader is invited (and recommended) to check that

- disc(A) is the first moment when some of the growing bodies swallows the origin,
- herdisc(A) is the first moment such that for each face F of the cube (of each possible dimension), the center of F is covered by some of the bodies centered at the vertices of F, and
- lindisc(A) is the first moment such that the whole cube is covered.

Proof of Theorem 4.6 (lindisc $\le$ 2·herdisc). In view of the above geometric interpretation, it suffices to prove the following statement.

If U is a closed convex body such that $\bigcup_{a\in\{-1,1\}^n}(U+a)$ covers all the points of $\{-1,0,1\}^n$ then the set $C = \bigcup_{a\in\{-1,1\}^n}(2U+a)$ covers the whole cube $[-1,1]^n$.

Indeed, if $\operatorname{herdisc}(A) \le t$ then the body $U = tU_A$ satisfies even a stronger assumption, namely that each point $v \in \{-1,0,1\}^n$ is covered by the copy of U centered at one of the vertices closest to v.

Since U is closed, it is enough to prove that C covers all dyadic rational points in $[-1,1]^n$, i.e. all points $v = \frac{z}{2^k} \in [-1,1]^n$ for some integer vector $z \in \mathbf{Z}^n$. We proceed by induction on k, where the case $k=0$ follows immediately from the assumption. Consider some $v = \frac{z}{2^k} \in [-1,1]^n$. Since all components of $2v$ are in the interval $[-2,2]$, there is a vector $b \in \{-1,1\}^n$ such that $2v - b \in [-1,1]^n$. Since $2v - b = \frac{z+2^{k-1}b}{2^{k-1}}$, the inductive hypothesis for $k-1$ provides a vector $a \in \{-1,1\}^n$ such that $2v - b \in 2U + a$. Therefore, we obtain

$$v \in U + \frac{a+b}{2}.$$

The vector $\frac{a+b}{2}$ has all entries in $\{-1,0,1\}$, and so by the assumption on U, it is covered by some $U + c$ for $c \in \{-1,1\}^n$. Hence

$$v \in U + (U + c) = 2U + c,$$

where the last equality uses the convexity of U. This proves Theorem 4.6. □

A Lower Bound in Terms of Determinants. The hereditary discrepancy of a set system can be lower-bounded in terms of determinants of submatrices of the incidence matrix.

4.7 Theorem (Determinant lower bound). *For any set system $\mathcal{S}$, we have*

$$\operatorname{herdisc}(\mathcal{S}) \ge \frac{1}{2} \max_k \max_B |\det(B)|^{1/k},$$

where B ranges over all $k \times k$ submatrices of the incidence matrix of $\mathcal{S}$. An analogous bound also holds for the hereditary discrepancy of an arbitrary $m \times n$ real matrix A.

This is a consequence of the bound "lindisc ≤ 2·herdisc" (Theorem 4.6) and of the following lemma:

4.8 Lemma. *Let A be an $n \times n$ matrix. Then* $\operatorname{lindisc}(A) \ge |\det(A)|^{1/n}$.

Proof. Let $t = \operatorname{lindisc}(A)$ and set $U = tU_A$. By the above geometric interpretation of the linear discrepancy, the sets $U + a$ for $a \in \{-1,1\}^n$ cover the whole cube $[-1,1]^n$, and therefore the sets $U + a$ for $a \in 2\mathbf{Z}^n$ cover the whole space. Hence

$$\operatorname{vol}(U) = t^n \operatorname{vol}(U_A) \ge \operatorname{vol}([-1,1]^n) = 2^n.$$

On the other hand, the linear mapping $x \mapsto Ax$ changes the volume by the factor $|\det(A)|$ (since it maps the unit cube to a parallelepiped of volume $|\det(A)|$), and since U_A is the inverse image of the cube $[-1,1]^n$, we get

$\mathrm{vol}(U_A) = |\det(A)|^{-1}2^n$. Together with the previous inequality for $\mathrm{vol}(U_A)$, this gives $t \geq |\det(A)|^{1/n}$. □

It is instructive to compare the determinant lower bound and the eigenvalue lower bound (Theorem 4.5). For simplicity, let us consider the case of a square matrix A first, in which case the eigenvalue bound becomes $\sqrt{\lambda_{min}}$. We recall the geometric interpretation of the eigenvalue bound: $\sqrt{\lambda_{min}}$ is the length of the shortest semiaxis of the ellipsoid E that is the image of the unit ball $B(0,1)$ under the linear mapping $x \mapsto Ax$. The ratio $\mathrm{vol}(E)/\mathrm{vol}(B(0,1))$ equals, on the one hand, the product of the semiaxes of E, i.e. $\sqrt{\lambda_1\lambda_2\cdots\lambda_n}$, and on the other hand, it is equal to $|\det A|$. Therefore, since λ_{min} is the smallest among the n eigenvalues of A^TA, we get $\sqrt{\lambda_{min}} \leq |\det A|^{1/n}$. Thus, for a square matrix, the determinant lower bound for discrepancy is never weaker than the eigenvalue lower bound (and it can be much stronger if the ellipsoid E happens to be very flat). Also for non-square matrices A, the determinant lower bound is never smaller than the eigenvalue bound, possibly up to a small constant multiplicative factor; see Exercise 7. But one should not forget that the eigenvalue bound estimates discrepancy, while the determinant bound only applies to hereditary discrepancy.

Few Sets on Many Points. Set systems coming from geometric settings typically have more sets than points, so we are mainly interested in this case. For studying discrepancy of set systems with few sets and many points, the following result is important:

4.9 Theorem. *Let $(X,\mathcal{S})$ be a set system such that $\mathrm{disc}(\mathcal{S}|_Y) \leq K$ for all $Y \subseteq X$ with $|Y| \leq |\mathcal{S}|$. Then $\mathrm{disc}(\mathcal{S}) \leq 2K$.*

Proof. This is a nice application of the concept of linear discrepancy. We note that if w and w_0 are two weight functions on X such that $w(S) = w_0(S)$ for all $S \in \mathcal{S}$ then the discrepancy of any coloring χ for w is the same as that for w_0. We also have

4.10 Lemma. *Let $(X,\mathcal{S})$ be a set system, $|X| = n \geq |\mathcal{S}| = m$, and let $w\colon X \to [-1,1]$ be a weight function. Then there exist an n-point set $Y \subseteq X$ and a weight function $w_0\colon X \to [-1,1]$ such that $w_0(S) = w(S)$ for all $S \in \mathcal{S}$ and $w_0(x) = \pm 1$ for all $x \in X \setminus Y$.*

The proof of this lemma is quite similar to the proof of the Beck–Fiala theorem 4.3 and we leave it as Exercise 1. From the lemma and the observation above it, we get that $\mathrm{lindisc}(\mathcal{S}) \leq \sup_{Y\subseteq X,\,|Y|=n} \mathrm{lindisc}(\mathcal{S}|_Y)$. The left-hand side of this inequality is at least $\mathrm{disc}(\mathcal{S})$ while the right-hand side is at most $2\max_{Y\subseteq X,\,|Y|\leq n} \mathrm{disc}(\mathcal{S}|_Y)$ by the bound "lindisc $\leq 2\cdot$herdisc" (Theorem 4.6). This proves Theorem 4.9. □

The theorem just proved plus Spencer's upper bound (Theorem 4.2) imply that an arbitrary set system with m sets has discrepancy $O(\sqrt{m})$.

Bibliography and Remarks. Hereditary discrepancy and linear discrepancy were introduced by Lovász et al. [LSV86]. (In [BS95], linear discrepancy is called the *inhomogeneous discrepancy.*) Lovász et al. [LSV86] also established Theorem 4.6 and Theorem 4.7, and our presentation mostly follows their proofs.

We should remark that they use definitions of discrepancy giving exactly half of the quantities we consider. They work with 0/1 vectors instead of $-1/1$ vectors, which is probably somewhat more natural in the context of linear discrepancy.

Another potentially useful lower bound for the discrepancy of an $m \times n$ matrix A is

$$\mathrm{lindisc}(A) \geq \frac{2\,\mathrm{vol}(\mathrm{conv}(\pm A))^{1/n}}{c_n^{2/n}}. \tag{4.4}$$

Here $c_n = \pi^{n/2}/\Gamma(\frac{n}{2}+1)$ is the volume of the n-dimensional unit ball and $\mathrm{conv}(\pm A)$ denotes the convex hull of the $2m$ vectors a_1, $-a_1$, a_2, $-a_2$, ..., a_m, $-a_m$, where $a_i \in \mathbf{R}^n$ stands for the ith row of A. This lower bound is obtained from $\mathrm{lindisc}(A) \geq 2\,\mathrm{vol}(U_A)^{-1/n}$ (see the proof of Lemma 4.8) using so-called *Blaschke's inequality*, stating that $\mathrm{vol}(K)\,\mathrm{vol}(K^*) \leq c_n^2$ holds for any centrally symmetric convex body K in $\mathbf{R}^n$. Here

$$K^* = \{x \in \mathbf{R}^n\colon \langle x,y\rangle \leq 1 \text{ for all } y \in K\}$$

is the *polar body* of K. For $K = U_A$, it turns out that $U_A^* = \mathrm{conv}(\pm A)$. The inequality (4.4) is due to Lovász and Vesztergombi [LV89], who used it to estimate the maximum possible number $m(n,d)$ of distinct rows of an integral $m \times n$ matrix A with $\mathrm{herdisc}(A) \leq d$. They proved that this $m(n,d)$ is between $\binom{n+d}{n}$ and $\binom{n+2\pi d}{n}$. If one asks a similar question for a set system, i.e. what is the maximum possible number of distinct sets in a set system $\mathcal{S}$ on n points with $\mathrm{herdisc}(\mathcal{S}) \leq d$, then a precise answer can be given—see Exercise 5.2.5.

The next few remarks concern the relationship of the linear and hereditary discrepancies. In the inequality $\mathrm{lindisc}(\mathcal{S}) \leq 2\,\mathrm{herdisc}(\mathcal{S})$ (Theorem 4.6), the constant 2 cannot be improved in general; see Exercise 3. On the other hand, the inequality is always strict, and in fact, $\mathrm{lindisc}(\mathcal{S}) \leq 2(1-\frac{1}{2m})\,\mathrm{herdisc}(\mathcal{S})$ holds for any set system $\mathcal{S}$ with m sets (Doerr [Doe00]).

There is a simple example showing that the hereditary discrepancy of a matrix cannot be bounded in terms of the linear discrepancy [LSV86], namely the single-row matrix $(1,2,4,\ldots,2^{n-1})$ (Exercise 4). The question of whether the hereditary discrepancy of a set system can be estimated by a function of the linear discrepancy seems to be open. On the one hand, there is a set system such that any system containing it as an induced subsystem has linear discrepancy at least 2

[Mat00]. On the other hand, the hereditary discrepancy can be strictly bigger than the linear discrepancy (Exercise 6), and so the situation cannot be too simple.

The fact that the determinant lower bound for the hereditary discrepancy in Theorem 4.7 is never asymptotically smaller than the eigenvalue lower bound for disc_2 in Theorem 4.5 (Exercise 7) is a simple consequence of observations of Chazelle [Cha99] (I haven't seen it explicitly mentioned anywhere). Chazelle's result actually says that if the eigenvalue lower bound for some system $\mathcal{S}$ on n points equals some number Δ then there is a subsystem $\mathcal{S}_0 \subseteq \mathcal{S}$ of at most n sets with $\mathrm{herdisc}(\mathcal{S}_0) = \Omega(\Delta)$. So, in some sense, the eigenvalue bound is always "witnessed" by the hereditary discrepancy of at most n sets, no matter how many sets the original set system may have. Little seems to be known about possible strengthenings and analogues of this result. One related observation is that for a set system $\mathcal{S}$ on an n-point set with a large eigenvalue bound, all the systems $\mathcal{S}' \subseteq \mathcal{S}$ consisting of n sets may have the eigenvalue bound very small (Exercise 8).

A result somewhat weaker than Theorem 4.9, namely that the discrepancy of a system of m sets is always bounded by the maximum possible discrepancy of m sets on $O(m \log m)$ points, was first obtained by Olson and Spencer [OS78] (see Exercise 4.2.6). The fact that any m sets have discrepancy $O(\sqrt{m})$ was proved by Spencer [Spe85].

We have mentioned a natural generalization of the notion of discrepancy from incidence matrices of set systems to arbitrary real matrices. A different and interesting notion of matrix discrepancy arises from *vector sum* problems. Having n vectors $v_1, \ldots, v_n \in \mathbf{R}^m$ of norm at most 1, we ask for a choice of n signs $\varepsilon_1, \ldots, \varepsilon_n$ so that the vector $w = \sum_{i=1}^{n} \varepsilon_i v_i$ is as short as possible. We have a whole class of problems, since the vectors v_i can be measured by one norm in $\mathbf{R}^m$ (supremum norm, L_1-norm, Euclidean norm, etc.) and the vector w can be measured by another, possibly different, norm. Let us mention that for the case when both the norms are the supremum norm, an extension of Spencer's theorem shows that the norm of w can be made $O(\sqrt{m})$.

A famous conjecture of Komlós asserts that if all the v_i have Euclidean length at most 1 then the supremum norm of w can be bounded by an absolute constant (this is a generalization of the Beck–Fiala conjecture mentioned in Section 4.1; see Exercise 9). The current best result on Komlós' conjecture is due to Banaszczyk [Ban98]. He proves the following more general result: There is an absolute constant c such that if K is a convex body in $\mathbf{R}^m$ with $\gamma_m(K) \geq \frac{1}{2}$, then for any vectors $v_1, v_2, \ldots, v_n \in \mathbf{R}^m$ of Euclidean norm at most 1 there exist signs $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \in \{-1, +1\}$ such that $v_1\varepsilon_1 + v_2\varepsilon_2 + \cdots + v_n\varepsilon_n \in cK$. Here γ_m denotes the m-dimensional

Gaussian measure whose density at a point x is $(2\pi)^{-m/2}e^{-\|x\|^2/2}$; this is the density of the normalized normal distribution. This theorem improves an earlier result of Giannopoulos [Gia97], where the conclusion was that $v_1\varepsilon_1+v_2\varepsilon_2+\cdots+v_n\varepsilon_n \in c(\log n)K$ (this was already sufficient for proving Spencer's upper bound 4.2). Banaszczyk's result easily implies that in the situation of Komlós' conjecture, $\|w\|_\infty = O(\sqrt{\log n})$ can be achieved. Further, for the Beck–Fiala conjecture, this yields that the discrepancy of a set system of maximum degree t on n points is $O(\sqrt{t\log n})$, which is the best known bound in a wide range of the parameters n and t. (We will prove a weaker bound in Section 5.5.)

More about these and related subjects can be found, for instance, in Beck and Sós [BS95], Alon and Spencer [AS00], Bárány and Grinberg [BG81], and Spencer [Spe87].

Exercises

1. Prove Lemma 4.10.
2. Find a set system $(X,\mathcal{S})$ and a set $A \subseteq X$ such that $\mathrm{disc}(\mathcal{S}) = 0$ but $\mathrm{disc}(\mathcal{S}\cup\{A\})$ is arbitrarily large.

   Remark. It is not known whether an example exists with $\mathrm{herdisc}(\mathcal{S}) \le 1$ and with $\mathrm{disc}(\mathcal{S}\cup\{A\})$ large.
3. Show that the the set system $\{\{1\},\{2\},\ldots,\{n\},\{1,2,\ldots,n\}\}$ has hereditary discrepancy 1 and linear discrepancy at least $2-\frac{2}{n+1}$.
4. Show that the $1\times n$ matrix $(2^0,2^1,2^2,\ldots,2^{n-1})$ has hereditary discrepancy at least 2^{n-1} and linear discrepancy at most 2.
5. Let A be an $m\times n$ real matrix, and set

$$\Delta = \max_{w\in\{-1,0,1\}} \min_{x\in\{-1,1\}} \|A(x-w)\|_\infty$$

   ("linear discrepancy with weights $-1,0,1$"). Prove that $\mathrm{lindisc}(A) \le 2\Delta$.
6. Show that the set system $\{\{1,2\}, \{1,3\}, \{2,3,4\}\}$ has hereditary discrepancy 2 and linear discrepancy strictly smaller than 2.
7. (Relation of the determinant and eigenvalue bounds) Let $(X,\mathcal{S})$ be a system of m sets on an n-point set, $m\ge n$, and let A be the incidence matrix of $\mathcal{S}$.

   (a)* Put $\Delta = \left(\frac{n}{m}\det(A^TA)^{1/n}\right)^{1/2}$. Prove the existence of a subsystem $\mathcal{S}_0 \subseteq \mathcal{S}$ consisting of n sets with $\mathrm{herdisc}(\mathcal{S}_0) = \Omega(\Delta)$. Use the Binet–Cauchy theorem from linear algebra, asserting that for any $m\times n$ real matrix A, $m\ge n$, we have $\det(A^TA) = \sum_B \det(B)^2$, where B ranges over all $n\times n$ submatrices of A.

   (b)* Prove that if Δ_{eig} is the eigenvalue lower bound from Theorem 4.5 and Δ_{det} is the determinant lower bound from Theorem 4.7 then $\Delta_{eig} = O(\Delta_{det})$.

(c) Show that the lower bound in Theorem 4.7 for $\mathrm{herdisc}(\mathcal{S})$, and consequently also the eigenvalue bound in Theorem 4.5, is never bigger than $O(\sqrt{n})$.

8. Let A be an $(n+1) \times n$ zero-one matrix obtained from an $(n+2) \times (n+2)$ Hadamard matrix by deleting the first row and the first two columns and changing -1's to 0's. Show that the eigenvalue lower bound for A is $\Omega(\sqrt{n})$ (this is similar to Exercise 4.2.3), and that for any $n \times n$ submatrix B of A, the eigenvalue bound is only $O(1)$. Therefore, unlike the determinant lower bound, the eigenvalue lower bound for $n+1$ sets on n points need not be "witnessed" by n sets on these points.
9. Verify that Komlós' conjecture implies the Beck–Fiala conjecture. Komlós' conjecture says that there is a constant K such that for any vectors $v_1, v_2, \ldots, v_n$ of unit Euclidean norm, there exist signs $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ such that $\|\varepsilon_1 v_1 + \varepsilon_2 v_2 + \cdots + \varepsilon_n v_n\|_\infty \leq K$. The Beck–Fiala conjecture states that $\mathrm{disc}(\mathcal{S}) \leq C\sqrt{t}$ for any set system $\mathcal{S}$ of maximum degree t.
10. Let $A = \frac{1}{2}(H + J)$ be the $n \times n$ incidence matrix of a set system as in Proposition 4.4. Derive an $\Omega(\sqrt{n})$ lower bound for $\mathrm{herdisc}(A)$ using the determinant lower bound (Theorem 4.7); use the specific Hadamard matrices H_k of size $2^k \times 2^k$ whose construction was indicated above Proposition 4.4.

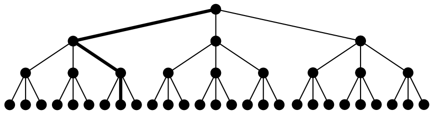
4.4 On Set Systems with Very Small Discrepancy

A very important class of set systems are those with hereditary discrepancy at most 1 (note that requiring hereditary discrepancy to be 0 leads to rather trivial set systems). Such set systems are called *totally unimodular*. They are of interest in polyhedral combinatorics, theory of integer programming, etc., and there is an extensive theory about them, which has been developing more or less independently of discrepancy theory. Here we only touch this subject very briefly, but it is useful to be aware of its existence.

An Example Destroying Several Conjectures. The following question has been open for some time: if $\mathcal{S}_1$ and $\mathcal{S}_2$ are two set systems on the same ground set, can $\mathrm{disc}(\mathcal{S}_1 \cup \mathcal{S}_2)$ be upper-bounded by some function of $\mathrm{disc}(\mathcal{S}_1)$ and $\mathrm{disc}(\mathcal{S}_2)$? The following important example shows that this is not the case. Even the union of two set systems with the best possible behavior in terms of discrepancy, namely with herdisc $= 1$, can have arbitrarily large discrepancy.

4.11 Proposition (Hoffman's example). *For an arbitrarily large number K, there exist set systems $\mathcal{S}_1$, $\mathcal{S}_2$ such that* $\mathrm{herdisc}(\mathcal{S}_1) \leq 1$, $\mathrm{herdisc}(\mathcal{S}_2) \leq 1$, *and* $\mathrm{disc}(\mathcal{S}_1 \cup \mathcal{S}_2) \geq K$.

Proof. The ground set of both set systems is the set of edges of the complete K-ary tree T of depth K (a picture shows the case $K = 3$).



The sets in $\mathcal{S}_1$ are the edge sets of all root-to-leaf paths (the picture shows one of them drawn thick). The sets of $\mathcal{S}_2$ are the "fans" in the tree: for each non-leaf vertex v, we put the set of the K edges connecting v to its successors into $\mathcal{S}_2$. The bound $\mathrm{herdisc}(\mathcal{S}_2) \leq 1$ is obvious, and $\mathrm{herdisc}(\mathcal{S}_1) \leq 1$ is simple and it is left as Exercise 1. Finally $\mathrm{disc}(\mathcal{S}_1 \cup \mathcal{S}_2) \geq K$ follows from a Ramsey-type observation: whenever the edges of T are each colored red or blue, there is either a red root-to-leaf path or a vertex with all successor edges blue. $\square$

Based on this example, one can refute several other plausible-looking conjectures about discrepancy (see Exercises 2 and 3).

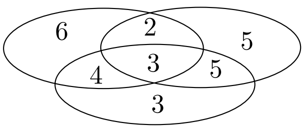
An Etude in Discrepancy Zero. We consider the following function: let $f(n)$ denote the smallest number of sets of size n each that constitute a system with nonzero discrepancy. For instance, we have $f(n) = 1$ for every odd n. The question is whether $f(n)$ can be bounded by some universal constant K for all n. The answer is negative:

4.12 Theorem. *We have* $\limsup_{n\to\infty} f(n) = \infty$.

This theorem is included mainly because of the beauty of the following proof.

Proof. For contradiction, suppose that $f(n) \leq K$ for all n. This means that for every n there is a set system $\mathcal{S}^{(n)} = \{S_1^{(n)}, S_2^{(n)}, \ldots, S_K^{(n)}\}$ consisting of K n-element sets such that $\mathrm{disc}(\mathcal{S}^{(n)}) \geq 1$. Let us fix one such $\mathcal{S}^{(n)}$ for each n.

A system of 3 sets, say, can be described by giving the number of elements in each field of the Venn diagram, as an example illustrates:



Similarly, the system $\mathcal{S}^{(n)} = \{S_1^{(n)}, S_2^{(n)}, \ldots, S_K^{(n)}\}$ is determined by an integer vector indexed by nonempty subsets of $\{1, 2, \ldots, K\}$. Namely, for each nonempty index set $I \subseteq \{1, 2, \ldots, K\}$, we let $s_I^{(n)}$ be the number of elements that belong to all $S_i^{(n)}$ with $i \in I$ and to no $S_j^{(n)}$ with $j \notin I$. In this way, the set system $\mathcal{S}^{(n)}$ determines an integer vector $s^{(n)} \in \mathbf{R}^{2^K-1}$. The condition that all sets of $\mathcal{S}^{(n)}$ have size n implies that $\sum_{I:\, j \in I} s_I^{(n)} = n$ for all $j = 1, 2, \ldots, K$.

Similarly, a red-blue coloring χ of the ground set $\mathcal{S}^{(n)}$ can be described by an integer vector $c \in \mathbf{R}^{2^K-1}$, where this time the component c_I tells us how many elements colored red lie in all the sets $S_i^{(n)}$ with $i \in I$ and in none of the sets $S_j^{(n)}$ with $j \notin I$.

Let us put $\sigma^{(n)} = \frac{1}{n} s^{(n)}$, and let us consider the following system of linear equations and inequalities for an unknown vector $\gamma \in \mathbf{R}^{2^K-1}$:

$$\begin{array}{ll} 0 \le \gamma_I \le \sigma_I^{(n)} & \text{for all nonempty } I \subseteq \{1, 2, \ldots, K\} \\ \sum_{I:\, j \in I} \gamma_I = \frac{1}{2} & \text{for } j = 1, 2, \ldots, K. \end{array} \tag{4.5}$$

Let $\Gamma^{(n)} \subseteq \mathbf{R}^{2^K-1}$ denote the set of all real vectors γ satisfying the system (4.5) for a given n. If c were an integer vector encoding a coloring of the ground set of $\mathcal{S}^{(n)}$ with zero discrepancy, then we would conclude that $\frac{1}{n}c \in \Gamma^{(n)}$. But we assume that no such c exists, and so $\Gamma^{(n)}$ contains no vector γ with $n\gamma$ integral.

To arrive at a contradiction, we will look at the values of n of the form $q!$ for $q = 1, 2, \ldots$; the important thing about these values is that all the numbers up to q divide $q!$. So let us consider the vectors $\sigma^{(q!)}$, $q = 1, 2, 3, \ldots$. This is an infinite and bounded sequence of vectors in $\mathbf{R}^{2^K-1}$, and hence it has a cluster point; call it σ. Let $(n_1, n_2, \ldots)$ be a subsequence of the sequence $(1!, 2!, 3!, \ldots)$ such that the $\sigma^{(n_k)}$ converge to σ.

Let us choose a rational vector $\bar{\sigma}$ with $\frac{1}{2}\sigma \le \bar{\sigma} \le \frac{2}{3}\sigma$ (the inequalities should hold in each component), and let $\bar{\Gamma}$ denote the solution set of the system (4.5) with $\bar{\sigma}$ replacing $\sigma^{(n)}$. We have $\frac{1}{2}\sigma \in \bar{\Gamma}$ and so $\bar{\Gamma} \neq \emptyset$. At the same time, the inequalities and equations defining $\bar{\Gamma}$ have all coefficients rational, and hence $\bar{\Gamma}$ contains a rational vector $\bar{\gamma}$. This $\bar{\gamma}$ satisfies $\bar{\gamma}_I < \sigma_I$ (strict inequalities) for all I with $\sigma_I \neq 0$, and hence also $\bar{\gamma} \in \Gamma^{(n_k)}$ for all large enough k. But since the n_k were selected among the numbers $1!$, $2!$, $3!$,..., we get that for sufficiently large k, $n_k\bar{\gamma}$ is a vector of integers and hence it encodes a zero-discrepancy coloring of $\mathcal{S}^{(n_k)}$. This contradiction finishes the proof of Theorem 4.12. □

Bibliography and Remarks. For a basic overview and references concerning the theory of total unimodularity, the reader may consult Schrijver [Sch95]. A matrix A is called totally unimodular if the determinant of each square submatrix of A is 0, 1 or -1 (this implies, in particular, that the entries of A are 0's and ± 1's). A famous theorem of Ghouila-Houri [GH62] asserts, in our terminology, that a matrix consisting of 0's and ± 1's is totally unimodular if and only if its hereditary discrepancy is at most 1; see Exercise 4. On the other hand, the linear discrepancy of a totally unimodular matrix can be arbitrarily close to 2; see Exercise 4.3.3.

The question about bounding $\mathrm{disc}(\mathcal{S}_1 \cup \mathcal{S}_2)$ in terms of bounding $\mathrm{disc}(\mathcal{S}_1)$ and $\mathrm{disc}(\mathcal{S}_2)$ was raised by Sós. Hoffmann's example is cited as an oral communication from 1987 in Beck and Sós [BS95]. The conjectures in Exercises 2 and 3 were stated in Lovász et al. [LSV86].

Theorem 4.12 is due to Alon et al. [AKP+87], who also give fairly precise quantitative bounds for $f(n)$ in terms of the number-theoretic structure of n, more precisely in terms of the smallest number not dividing n.

Exercises

1. Verify the assertion $\mathrm{herdisc}(\mathcal{S}_1) \le 1$ in the proof of Proposition 4.11.
2. Let $(X, \mathcal{S})$ be a set system and let $(\mathcal{S}, \mathcal{S}^*)$ be the set system dual to $\mathcal{S}$; explicitly $\mathcal{S}^* = \{\{S \in \mathcal{S} \colon x \in S\} \colon x \in X\}$. Using Proposition 4.11, show that $\mathrm{herdisc}(\mathcal{S}^*)$ cannot in general be bounded by any function of $\mathrm{herdisc}(\mathcal{S})$.
3. Using Proposition 4.11, show that $\mathrm{herdisc}(\mathcal{S})$ cannot be bounded from above by any function of $\max_k \max_B |\det(B)|^{1/k}$, i.e. of the right-hand side of the inequality in Theorem 4.7, where B is a $k \times k$ submatrix of the incidence matrix of $\mathcal{S}$.
4. (On Ghouila-Houri's theorem)
   (a) Show that if A is a nonsingular $n \times n$ totally unimodular matrix (the definition was given above the exercises), then the mapping $x \mapsto Ax$ maps $\mathbf{Z}^n$ bijectively onto $\mathbf{Z}^n$.
   (b)* Show that if A is an $m \times n$ totally unimodular matrix and b is an m-dimensional integer vector such that the system $Ax = b$ has a real solution x, then it has an integral solution as well.
   (c)* (Kruskal–Hoffmann theorem—one implication) Let A be an $m \times n$ totally unimodular matrix and let $u, v \in \mathbf{Z}^n$ and $w, z \in \mathbf{Z}^m$ be integer vectors. Show that if the system of inequalities $u \le x \le v$, $w \le Ax \le z$ (the inequalities should hold in each component) has a real solution then it has an integer solution as well. Geometrically speaking, all the vertices of the polytope in $\mathbf{R}^n$ determined by the considered system are integral.
   (d)* Prove that the discrepancy of a totally unimodular set system with all sets of even size is 0.
   (e) Prove that the hereditary discrepancy of a totally unimodular set system is at most 1 (this is one of the implications in Ghouila-Houri's theorem for 0/1 matrices).

4.5 The Partial Coloring Method

Here we introduce one of the most powerful methods for producing low-discrepancy colorings.

Let X be a set. A *partial coloring* of X is any mapping $\chi \colon X \to \{-1, 0, +1\}$. For a point $x \in X$ with $\chi(x) = 1$ or $\chi(x) = -1$, we say that x is *colored by* χ, while for $\chi(x) = 0$ we say that x is *uncolored*.

4.13 Lemma (Partial coloring lemma). *Let $\mathcal{F}$ and $\mathcal{M}$ be set systems*[1] *on an n-point set X, $|\mathcal{M}| > 1$, such that $|M| \leq s$ for every $M \in \mathcal{M}$ and*

$$\prod_{F \in \mathcal{F}} (|F| + 1) \leq 2^{(n-1)/5}. \tag{4.6}$$

Then there exists a partial coloring $\chi: X \to \{-1, 0, +1\}$, such that at least $\frac{n}{10}$ elements of X are colored, $\chi(F) = 0$ for every $F \in \mathcal{F}$, and $|\chi(M)| \leq \sqrt{2s \ln(4|\mathcal{M}|)}$ for every $M \in \mathcal{M}$.

For brevity, let us call a partial coloring that colors at least 10% of the points a *no-nonsense partial coloring.*

Intuitively, the situation is as follows. We have the "few" sets of $\mathcal{F}$, for which we insist that the discrepancy of χ be 0. Each such $F \in \mathcal{F}$ thus puts one condition on χ. It seems plausible that if we do not put too many conditions then a coloring χ randomly selected among those satisfying the conditions will still be "random enough" to behave as a true random coloring on the sets of $\mathcal{M}$. In the lemma, we claim something weaker, however: instead of a "true" coloring $\chi: X \to \{+1, -1\}$ we obtain a no-nonsense partial coloring χ, which is only guaranteed to be nonzero at a constant fraction of points. (And, indeed, under the assumptions of the Partial coloring lemma, one cannot hope for a full coloring with the discrepancy stated. For example, although every system of $\frac{n}{10 \log n}$ sets on n points has a no-nonsense partial coloring with zero discrepancy, there are such systems with discrepancy about $\sqrt{n/\log n}$.)

Proof of Lemma 4.13. Let $\mathcal{C}_0$ be the set of all colorings $\chi: X \to \{-1, +1\}$, and let $\mathcal{C}_1$ be the subcollection of colorings χ with $|\chi(M)| \leq \sqrt{2s \ln(4|\mathcal{M}|)}$ for all $M \in \mathcal{M}$. We have $|\mathcal{C}_1| \geq \frac{1}{2}|\mathcal{C}_0| = 2^{n-1}$ by the Random coloring lemma 4.1.

Now let us define a mapping $b: \mathcal{C}_1 \to \mathbf{Z}^{|\mathcal{F}|}$, assigning to a coloring χ the $|\mathcal{F}|$-component integer vector $b(\chi) = (\chi(F): F \in \mathcal{F})$ (where the sets of $\mathcal{F}$ are taken in some arbitrary but fixed order). Since $|\chi(F)| \leq |F|$ and $\chi(F) - |F|$ is even for each F, the image of b contains at most

$$\prod_{F \in \mathcal{F}} (|F| + 1) \leq 2^{(n-1)/5}$$

distinct vectors. Hence there is a vector $b_0 = b(\chi_0)$ such that b maps at least $2^{4(n-1)/5}$ elements of $\mathcal{C}_1$ to b_0 (the pigeonhole principle!). Put $\mathcal{C}_2 = \{\chi \in \mathcal{C}_1: b(\chi) = b_0\}$. Let us fix an arbitrary $\chi_1 \in \mathcal{C}_2$ and for every $\chi_2 \in \mathcal{C}_2$, let us define a new mapping $\chi': X \to \{-1, 0, 1\}$ by $\chi'(x) = \frac{1}{2}(\chi_2(x) - \chi_1(x))$. Then $\chi'(F) = 0$ for all $F \in \mathcal{F}$, and also $\chi'(M) \leq \sqrt{2s \ln(4|\mathcal{M}|)}$ for all $M \in \mathcal{M}$. Let $\mathcal{C}_2'$ be the collection of the χ' for all $\chi_2 \in \mathcal{C}_2$.

To prove the lemma, it remains to show that there is a partial coloring $\chi' \in \mathcal{C}_2'$ that colors at least $\frac{n}{10}$ points of X. The number of mappings $X \to \{-1, 0, +1\}$ with fewer than $\frac{n}{10}$ nonzero elements is bounded by

[1] $\mathcal{F}$ for "few" sets, $\mathcal{M}$ for "minute" (or also "many") sets.

$$N = \sum_{0 \le q < n/10} \binom{n}{q} 2^q;$$

we will show that $N < |\mathcal{C}_2'|$. We may use the estimate

$$\sum_{0 \le i \le z} \binom{n}{i} a^i \le \left(\frac{ean}{z}\right)^z \tag{4.7}$$

(valid for any $n \ge z > 0$ and any real $a \ge 1$),[2] which in our case yields that

$$N < \left(\frac{2en}{n/10}\right)^{n/10} < 60^{n/10} < 2^{6n/10} < 2^{4(n-1)/5} \le |\mathcal{C}_2'|.$$

Hence there exists a partial coloring $\chi' \in \mathcal{C}_2'$ with at least $\frac{n}{10}$ points colored. □

Suppose that we want a low-discrepancy coloring of a set system $(X, \mathcal{S})$. How do we apply the Partial coloring lemma? Usually we look for an auxiliary set system $\mathcal{F}$ such that

- $\mathcal{F}$ has sufficiently few sets. More exactly, it satisfies the condition $\prod_{F \in \mathcal{F}}(|F|+1) \le 2^{(n-1)/5}$, where $n = |X|$.
- Each set $S \in \mathcal{S}$ can be written as a disjoint union of some sets from $\mathcal{F}$, plus some extra set M_S which is small (smaller than some parameter s, for all $S \in \mathcal{S}$).

We then define $\mathcal{M} = \{M_S \colon S \in \mathcal{S}\}$. The Partial coloring lemma yields a partial coloring χ which has zero discrepancy on all sets of $\mathcal{F}$, and so $|\chi(S)| = |\chi(M_S)| = O(\sqrt{s \log |\mathcal{S}|})$. In this way, some 10% of points of X are colored. We then look at the set of yet uncolored points, restrict the system $\mathcal{S}$ on these points, and repeat the construction of a partial coloring. In $O(\log n)$ stages, everything will be colored. This scheme has many variations, of course.

Here we describe an application in an upper bound for the so-called *Tusnády's problem*: What is the combinatorial discrepancy for axis-parallel rectangles in the plane, i.e. $\operatorname{disc}(n, \mathcal{R}_2)$? By the transference lemma (Proposition 1.8), this discrepancy is asymptotically at least as large (for infinitely many n) as the Lebesgue-measure discrepancy $D(n, \mathcal{R}_2)$, and the latter quantity is known to be of the order $\log n$ (see Proposition 2.2 and Schmidt's theorem 6.2). But obtaining tight bounds for Tusnády's problem seems quite hard, and the subsequent theorem gives nearly the best known upper bound.

---

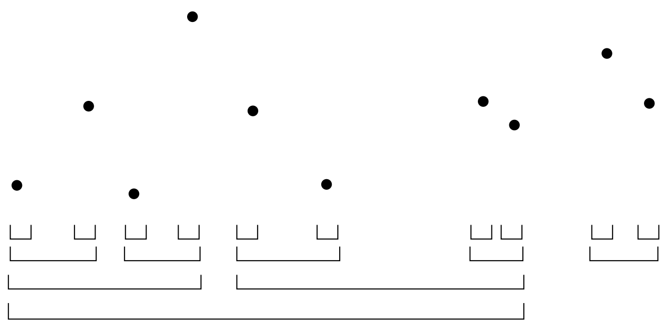
[2] Here is a few-line proof of (4.7): By the binomial theorem, we have $(1+ax)^n \ge \sum_{0 \le i \le z} \binom{n}{i} a^i x^i$, so for $0 < x \le 1$ we get $\sum_{0 \le i \le z} \binom{n}{i} a^i \le \sum_{0 \le i \le z} \binom{n}{i} a^i x^{i-z} \le (1+ax)^n / x^z \le e^{axn}/x^z$ (since $1 + y \le e^y$ for all real y). The estimate follows by substituting $x = z/an$.

4.14 Theorem. *The combinatorial discrepancy for axis-parallel rectangles satisfies*

$$\mathrm{disc}(n, \mathcal{R}_2) = O(\log^{5/2} n \sqrt{\log\log n}\,).$$

The $\sqrt{\log\log n}$ factor can be removed from the bound by a more sophisticated method (Exercise 5.5.2) but currently it is not known how to improve the exponent of $\log n$, let alone what the correct bound is.

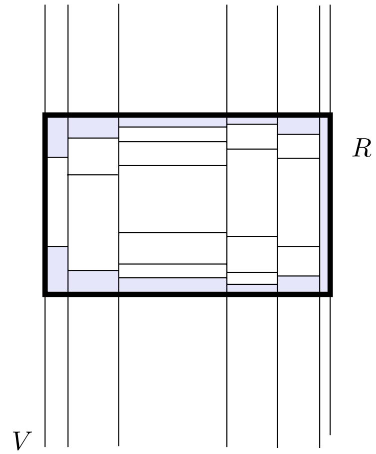
Proof. First we construct a partial coloring. Let $P \subset \mathbf{R}^2$ be an n-point set, and let $p_1, p_2, \ldots, p_n$ be its points listed in the order of increasing x-coordinates (without loss of generality, we may assume that all the x-coordinates and all the y-coordinates of the points of P are pairwise distinct). Define a *canonical interval of P in the x-direction* as a subset of P of the form $\{p_{k2^q+1}, p_{k2^q+2}, \ldots, p_{(k+1)2^q}\}$. Here is a schematic illustration:



Let $\mathcal{C}$ be the collection of all canonical intervals of P in the x-direction, of all possible lengths 2^q with $1 \le 2^q \le n$. By considerations analogous to the ones in the proof of Proposition 2.2 (Claim II), we see that any interval in P, of the form $\{p_i, p_{i+1}, \ldots, p_{i+j}\}$, can be expressed as a disjoint union of at most $2\lfloor \log_2 n + 1 \rfloor \le 2\log 2n$ sets of $\mathcal{C}$.

For each canonical interval $C \in \mathcal{C}$, consider the collection of all canonical intervals of C in the y-direction (defined analogously to the canonical intervals in x-direction). Discard those of size smaller than t, where t is a threshold parameter (to be determined later). Call the collection of the remaining canonical intervals in the y-direction $\mathcal{F}_C$, and put $\mathcal{F} = \bigcup_{C \in \mathcal{C}} \mathcal{F}_C$.

Let $\mathcal{R}_2|_P$ be the set system defined on P by axis-parallel rectangles. We claim that for any rectangle $R \in \mathcal{R}_2$, the set $P \cap R$ can be written as a disjoint union of some sets from $\mathcal{F}$ plus a set of at most $s = 4t \log_2 2n$ extra points. To see this, we extend the rectangle R to an infinite vertical strip V. The set $P \cap V$ can be decomposed into at most $2\log_2 2n$ disjoint sets from $\mathcal{C}$. For any C in this decomposition, $C \cap R$ is in fact an intersection of C with an infinite horizontal strip, and can thus be decomposed into disjoint sets from $\mathcal{F}_C$ plus an extra set consisting of at most $4t$ points. Such a decomposition is schematically depicted below:



From this, the claim follows.

For the considered rectangle R, let M_R denote the set of the at most $s = O(t \log n)$ extra points, i.e. the points of $R \cap P$ that are not covered by the sets from $\mathcal{F}$ in the decomposition (these are the points of P in the gray region in the above schematic picture). Define $\mathcal{M} = \{M_R \colon R \in \mathcal{R}_2\}$. We have $|\mathcal{M}| \leq \big|\mathcal{R}_2|_P\big| = O(n^4)$. We plan to apply the Partial coloring lemma for the set systems $\mathcal{F}$ and $\mathcal{M}$, so we need to choose the parameter t in such a way that $\prod_{\mathcal{F}}(|F|+1) \leq 2^{(n-1)/5}$. If $C \in \mathcal{C}$ has 2^q points, then $\mathcal{F}_C$ contains 2^{q-i} sets of size 2^i, $\lceil \log_2 t \rceil \leq i \leq q$. Thus,

$$\log_2\left(\prod_{F \in \mathcal{F}_C} (|F|+1)\right) \leq \sum_{i=\lceil \log_2 t \rceil}^{q} 2^{q-i} \log_2(2^i+1) = O\left(\frac{2^q \log t}{t}\right).$$

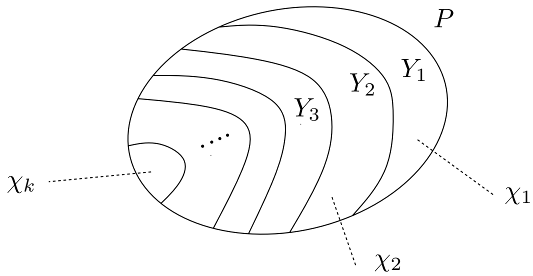
The system $\mathcal{C}$ contains $\lfloor n/2^q \rfloor$ sets C of size 2^q, so we have

$$\log_2\left(\prod_{F \in \mathcal{F}} (|F|+1)\right) \leq \sum_{q=\lceil \log_2 t \rceil}^{\lfloor \log_2 n \rfloor} \frac{n}{2^q}\, O\left(\frac{2^q \log t}{t}\right) = O\left(\frac{n \log n \log t}{t}\right).$$

We see that in order to satisfy the assumption $\prod_{\mathcal{F}}(|F|+1) \leq 2^{(n-1)/5}$ of the Partial coloring lemma, t should be chosen as $K \log n \log\log n$, for a sufficiently large constant K. Then the size of sets in $\mathcal{M}$ is bounded by $s = O(t \log n) = O(\log^2 n \log\log n)$.

From the Partial coloring lemma, we obtain a no-nonsense partial coloring χ satisfying $\operatorname{disc}(\chi, \mathcal{M}) = O(\sqrt{s \log n}) = O(\log^{3/2} n \sqrt{\log\log n}\,)$ and $\operatorname{disc}(\chi, \mathcal{F}) = 0$. For any rectangle $R \in \mathcal{R}_2$, we thus have $|\chi(P \cap R)| = |\chi(M_R)| = O(\log^{3/2} n \sqrt{\log\log n}\,)$.

To prove Theorem 4.14, we apply the construction described above iteratively, as the following drawing indicates:



We set $P_1 = P$, and we construct a partial coloring χ_1 as above. Let Y_1 be the set of points colored by χ_1 and let $P_2 = P \setminus P_1$ be the uncolored points. We produce a partial coloring χ_2 of P_2 by applying the above construction to the set system $\mathcal{R}_2|_{P_2}$, and so on. We repeat this construction until the size of the set P_k becomes trivially small, say smaller than a suitable constant—this means $k = O(\log n)$. Then we define $Y_k = P_k$ and we let χ_k be the constant mapping with value 1 on Y_k. Finally we put $\chi(p) = \chi_i(p)$ for $p \in Y_i$.

Let $R \in \mathcal{R}_2$ be a rectangle. We have

$$\begin{aligned} |\chi(P \cap R)| &\leq \sum_{i=1}^{k} |\chi_i(Y_i \cap R)| \leq \sum_{i=1}^{O(\log n)} O(\log^{3/2} n \sqrt{\log\log n}\,) \\ &= O(\log^{5/2} n \sqrt{\log\log n}\,). \end{aligned}$$

□

Remark on Algorithms. The method of partial colorings is not algorithmic; the problem stems from the use of the pigeonhole principle in the proof of the Partial coloring lemma. (In fact, the problem mentioned earlier, with making Spencer's upper bound 4.2 effective, comes from the same source.) In some of the applications, the use of partial colorings can be replaced by the Beck–Fiala theorem 4.3. While one usually loses a few logarithmic factors, one obtains a polynomial-time algorithm—see Exercises 4 and 6.

Bibliography and Remarks. The partial coloring method was invented by Beck [Bec81b]; this is the paper with the first near-optimal upper bound for the discrepancy of arithmetic progressions (see the remarks to Section 4.2). The method was further elaborated by Beck in [Bec88a]. For other refinements see Section 4.6.

In 1980, Tusnády raised the question whether, in our terminology, the combinatorial discrepancy for axis-parallel rectangles is bounded by a constant (the question originated in an attempt to generalize results of Komlós et al. [KMT75] to higher dimensions). This was answered negatively by Beck [Bec81a], who also proved the upper bound of $O(\log^4 n)$. (The order of magnitude of the discrepancy is not interesting from the point of view of Tusnády's application but it is an intriguing problem in its own right.) This was improved to

$O((\log n)^{3.5+\varepsilon})$ by Beck [Bec89a] and to $O(\log^3 n)$ by Bohus [Boh90] via a bound for the "k-permutation problem" (Exercise 5). The possibility of a further slight improvement, to $O(\log^{5/2} n\sqrt{\log\log n}\,)$, was noted by the author of this book in a draft of this chapter. Independently, an $O(\log^{5/2} n)$ bound was recently proved by Srinivasan [Sri97] by a related but different method (see also Exercise 5.5.2). However, a generalization of the proof method shown above for Tusnády's problem gives an $O(\log^{d+1/2} n\sqrt{\log\log n}\,)$ bound in dimension d (see Exercise 1 or [Mat99]), while the method of Srinivasan and the one indicated in Exercise 5.5.2 lead to worse bounds for $d > 2$. Another challenging problem is to determine the combinatorial L_2-discrepancy for axis-parallel boxes: while in the continuous setting, the L_2-discrepancy bounds are considerably better than the bounds for the worst-case discrepancy, no such improvement is known in the combinatorial setting.

Beck [Bec88a] investigated, as a part of more general questions, the Lebesgue-measure discrepancy for the family of translated and scaled copies of a fixed convex polygon P_0 in the plane and he proved an $O(\log^{4+\varepsilon} n)$ upper bound, with the constant of proportionality depending on ε and on P_0 (also see Beck and Chen [BC89]). His result in fact applies to a somewhat larger family. For a finite set H of hyperplanes in $\mathbf{R}^d$, let $\mathrm{POL}(H)$ denote the set of all polytopes $\bigcap_{i=1}^{\ell}\gamma_i$, where each γ_i is a halfspace with boundary parallel to some $h \in H$ (obviously, for each $h \in H$, it suffices to consider at most two γ_i parallel to h in the intersection). Beck's upper bound is valid for any family $\mathrm{POL}(H)$ with H a finite set of lines in the plane. Károlyi [Kár95a] studied a d-dimensional analogue of the problem and proved the upper bound $D(n, \mathrm{POL}(H)) = O((\log n)^{\max(3d/2+1+\varepsilon, 2d-1)})$ for any fixed H and an arbitrarily small $\varepsilon > 0$, with the constant of proportionality depending on H and on ε. He uses the partial coloring method plus a sophisticated way of decomposing the polytopes in $\mathrm{POL}(H)$ into "canonical" ones. Exercises 2 and 3 below indicate proofs of similar but quantitatively somewhat better bounds for the combinatorial discrepancy of $\mathrm{POL}(H)$ (at least for dimensions 2 and 3). A detailed discussion of these bounds is in [Mat99]. The best estimates for the Lebesgue-measure discrepancy for $\mathrm{POL}(H)$ have recently been obtained by Skriganov [Skr98], whose results imply an $O(\log^{d-1} n(\log\log n)^{1+\varepsilon})$ upper bound, for any fixed finite H in $\mathbf{R}^d$ (this paper is discussed in the remarks to Section 2.5).

The 3-permutation problem discussed in Exercise 5 and Exercise 5.5.3 remains one of the most tantalizing questions in combinatorial discrepancy.

Exercises

1.* Consider a d-dimensional version of Tusnády's problem; generalize the method shown for the planar case to prove the upper bound $\mathrm{disc}(n, \mathcal{R}_d) = O(\log^{d+1/2} n \sqrt{\log\log n}\,)$ for any fixed d.
2. (Discrepancy for translates I)
   (a)* Let T_0 be a triangle in the plane, and let $\mathcal{T}$ denote the family of all translated and scaled copies of $\mathcal{T}$ (no rotation allowed). Show that there is a plane $\rho \subset \mathbf{R}^3$ such that if $\mathbf{R}^2$ is identified with ρ then any triangle $T \in \mathcal{T}$ can be written as $T = \rho \cap R$ for some axis-parallel box $R \in \mathcal{R}_3$. By the result of Exercise 1, this implies that $\mathrm{disc}(n, \mathcal{T}) = O(\log^{3.5} n \sqrt{\log\log n})$.
   (b) More generally, let H be a finite set of hyperplanes in $\mathbf{R}^d$, and define $\mathrm{POL}(H)$ as in the remarks above, i.e. as the set of all polytopes $\bigcap_{i=1}^{\ell} \gamma_i$, where each γ_i is a halfspace with boundary parallel to some $h \in H$. Using a suitable embedding of $\mathbf{R}^d$ into $\mathbf{R}^{|H|}$ and Exercise 1, derive that $\mathrm{disc}(n, \mathrm{POL}(H)) = O((\log n)^{|H|+1/2} \sqrt{\log\log n})$.
3. (Discrepancy for translates II)
   (a) Modify the proof of Theorem 4.14 to show that if $H_1, H_2, \ldots, H_k$ are families consisting of two lines each, where k is considered as a constant, then
   $$\mathrm{disc}(n, \mathrm{POL}(H_1) \cup \mathrm{POL}(H_2) \cup \cdots \cup \mathrm{POL}(H_k)) = O(\log^{5/2} n \sqrt{\log\log n})$$
   (the same bound as for axis-parallel rectangles), with the constant of proportionality depending on k. The notation $\mathrm{POL}(H)$ is as in Exercise 2(a).
   (b)* Using (a), improve the result of Exercise 2(a) to $\mathrm{disc}(n, \mathcal{T}) = O(\log^{5/2} n \sqrt{\log\log n})$.
   (c)* More generally, if H is a set of k lines in the plane, with k a constant, prove $\mathrm{disc}(n, \mathrm{POL}(H)) = O(\log^{5/2} n \sqrt{\log\log n})$, with the constant of proportionality depending on k.
   (d)** Generalize part (c) to dimension 3 (or even higher). That is, if H is a family of k planes in $\mathbf{R}^3$ with k fixed, then $\mathrm{disc}(n, \mathrm{POL}(H)) = O(\log^{3.5} n \sqrt{\log\log n})$. (Details of this can be found in [Mat99].)
4.* Prove an upper bound $\mathrm{disc}(n, \mathcal{R}_2) = O(\log^4 n)$ by using the Beck–Fiala theorem 4.3 instead of the Partial coloring lemma.
5. (The k-permutation problem) Let $X = \{1, 2, \ldots, n\}$, and let $\pi_1, \ldots, \pi_k$ be arbitrary permutations of X (bijective mappings $X \to X$). Define a set system $\mathcal{P}_k = \mathcal{P}(\pi_1) \cup \mathcal{P}(\pi_2) \cup \ldots \cup \mathcal{P}(\pi_k)$, where $\mathcal{P}(\pi)$ denotes the family of all initial segments along π; that is, $\mathcal{P}(\pi) = \{\{\pi(1), \pi(2), \pi(3), \ldots, \pi(q)\}\colon 1 \le q \le n\}$.
   (a)* Show that for $k = 2$, $\mathrm{disc}(\mathcal{P}_2) \le 1$ (for all choices of π_1, π_2).
   (b)* Use the Partial coloring lemma to prove $\mathrm{disc}(\mathcal{P}_k) = O(\log n)$ for any fixed k. What is the dependence of the constant of proportionality on k in the resulting bound? (Also see Exercise 5.5.3.)

(c) Prove that $\operatorname{disc}(\mathcal{P}_k)$ is not bounded by a constant independent of k. Let us remark that the question whether $\operatorname{disc}(\mathcal{P}_3)$ is bounded by some constant is well-known and probably difficult (the *three-permutation problem*).

6.* Prove an upper bound of $O(\log^2 n)$ for the discrepancy of a set system defined by 3 permutations as in Exercise 5 using the Beck–Fiala theorem 4.3 instead of the Partial coloring lemma.

7. Consider the set system $\mathcal{A}_n$ consisting of all arithmetic progressions in $\{1, 2, \ldots, n\}$, that is,

$$\mathcal{A}_n = \{\{a_0, a_0 + d, a_0 + 2d, \ldots\} \cap \{1, 2, \ldots, n\}\colon a_0, d \in \mathbf{N}\}.$$

(a)* Prove that $\mathcal{A}_n$ has a no-nonsense partial coloring with discrepancy $O(n^{1/4} \log^{3/4} n)$ (if you can't get this try to get at least a bigger power of $\log n$).
(b) Explain why (a) cannot be used iteratively in a straightforward manner to conclude that $\operatorname{disc}(\mathcal{A}_n) = O(n^{1/4} \log^{7/4})$.
(c)* Let $X \subseteq \{1, 2, \ldots, n\}$ be an m-element set. Show that the restriction of $\mathcal{A}_n$ on X also has a no-nonsense partial coloring with discrepancy $O(n^{1/4} \log^{3/4} n)$. This already implies $\operatorname{disc}(\mathcal{A}_n) = O(n^{1/4} \log^{7/4} n)$.
(d) Why doesn't the Beck–Fiala theorem seem to be directly applicable for getting a bound close to $n^{1/4}$ in this problem?

Remark. A slightly better upper bound will be proved in Exercise 5.5.4.

4.6 The Entropy Method

We are going to discuss a refinement of the Partial coloring lemma 4.13 which can often save logarithmic factors in discrepancy bounds. For instance, suppose that we have a set system $\mathcal{S}$ and two auxiliary set systems $\mathcal{F}$ and $\mathcal{M}$ as in the Partial coloring lemma, such that any set of $\mathcal{S}$ can be expressed as a disjoint union of a set from $\mathcal{F}$ and a set from $\mathcal{M}$. If the assumptions of the Partial coloring lemma are met (in particular, this means that $\mathcal{F}$ has somewhat fewer than n sets) then the lemma gives us a partial coloring where the sets of $\mathcal{F}$ have zero discrepancy, while the sets of $\mathcal{M}$ have discrepancy roughly as if colored randomly. The exactly zero discrepancy of the sets of $\mathcal{F}$ is somewhat wasteful, however, since it would be quite sufficient to make their discrepancy of the same order as the discrepancy of the sets of $\mathcal{M}$. With this idea in mind, let us look at the proof of the Partial coloring lemma again.

In that proof, we have exhibited two (full) colorings χ_1 and χ_2 differning on many elements and satisfying $\chi_1(S) = \chi_2(S)$ for all sets $S \in \mathcal{F}$. We now want to relax the latter condition, and only require that $|\chi_1(S) - \chi_2(S)| < 2\Delta_S$, where Δ_S is the required bound for the discrepancy of S. If this condition is satisfied, then the "difference coloring" $\chi = \frac{1}{2}(\chi_1 - \chi_2)$ has $|\chi(S)| < \Delta_S$.

If we fix some discrepancy bound Δ_S for each of the considered sets S, we need not distinguish between sets of two types anymore as we did in the Partial coloring lemma. The sets $F \in \mathcal{F}$ in that lemma would simply have $\Delta_F = 1$, while the sets $M \in \mathcal{M}$ would have $\Delta_M = \sqrt{2s\ln(4|\mathcal{M}|)}$. So we work with a single set system $\mathcal{S}$, but we will typically need some knowledge about the distribution of the sizes of sets.

For an application of the pigeonhole principle as in the proof of the Partial coloring lemma, we need to replace the inequality $|\chi_1(S) - \chi_2(S)| < 2\Delta_S$ by an equality condition (so that we can assign a pigeonhole to every coloring). A suitable replacement for this inequality is

$$\operatorname{round}\left(\frac{\chi_1(S)}{2\Delta_S}\right) = \operatorname{round}\left(\frac{\chi_2(S)}{2\Delta_S}\right),$$

where $\operatorname{round}(x) = \lfloor x + \frac{1}{2} \rfloor$ denotes rounding to the nearest integer.

For a coloring $\chi\colon X \to \{-1,+1\}$ and a set $S \in \mathcal{S}$, let us put

$$b_S = b_S(\chi) = \operatorname{round}\left(\frac{\chi(S)}{2\Delta_S}\right),$$

and let $b = b(\chi)$ be the vector $(b_S\colon S \in \mathcal{S})$ (the sets of $\mathcal{S}$ are taken in some order fixed once and for all). The value of $b(\chi)$ is the pigeonhole where the pigeon χ is supposed to live.

The set $\mathcal{C}$ of all possible colorings $\chi\colon X \to \{-1,+1\}$ is partitioned into classes according to the value of the vector $b(\chi)$. We want to show that there exists a big class, since in a big class we have two colorings χ_1, χ_2 differing in sufficiently many points. Their difference coloring $\frac{1}{2}(\chi_1 - \chi_2)$ will be the partial coloring giving discrepancy below Δ_S to each set $S \in \mathcal{S}$. Up to the definition of the classes, this is the same argument as in the proof of the Partial coloring lemma, and we have already calculated how big a big class should be: a class containing at least $2^{4n/5}$ colorings has some two colorings χ_1, χ_2 differing in at least $\frac{n}{10}$ components, and hence provides a no-nonsense partial coloring, i.e. one that colors at least $\frac{n}{10}$ points.

Thus, it remains to show the existence of a big class. To this end, it is very convenient to use entropy.

Entropy. Let Z be a discrete random variable attaining values in a finite set V. (Our main example of such a variable will be the $b(\chi)$ defined above, which is a function of the random coloring χ.) For $v \in V$, let p_v denote the probability of the event "$Z = v$." The *entropy* of Z, denoted by $H(Z)$, is defined by

$$H(Z) = -\sum_{v \in V} p_v \log_2 p_v.$$

One can think of entropy as the average number of bits of information we gain by learning the value of Z. For instance, if there are m equally likely values, then by learning which one was actually attained we gain $\log_2 m$ bits of information. If there are only two possible values, rain tomorrow and no rain tomorrow, and we're in the middle of a rainy season, then rain tomorrow brings almost no information, and a sunny day is a big surprise but extremely unlikely, so the total entropy is small (here entropy measures the "expected surprise," so to speak).

We need three basic properties of entropy.

1. (Good chance) *If* $H(Z) \leq K$ *then some value* $v \in V$ *is attained by* Z *with probability at least* 2^{-K}.
2. (Uniformity optimal) *If* Z *attains at most* k *distinct values, then* $H(Z) \leq \log_2 k$ *(with equality when* Z *is uniformly distributed on* k *values).*
3. (Subadditivity) *Let* $Z_1, Z_2, \ldots, Z_m$ *be arbitrary discrete random variables, and let* $Z = (Z_1, Z_2, \ldots, Z_m)$ *be the random vector with components* $Z_1, Z_2, \ldots, Z_m$. *Then we have* $H(Z) \leq H(Z_1) + H(Z_2) + \cdots + H(Z_m)$.

The first property is immediate from the definition of entropy. The other two need some work to prove but are not difficult either. The subadditivity is intuitively obvious from the "average information" interpretation of entropy. The inequality may be strict, for instance if the vector consists of several copies of the same random variable.

Partial Coloring from Entropy. Let a set system $\mathcal{S}$ and numbers Δ_S be given, and let the vector $b = b(\chi)$ be defined as above. If χ is a random coloring, then $b(\chi)$ is a random variable. If we could prove that its entropy $H(b)$ is at most $\frac{n}{5}$, then by the first property of entropy (good chance), some value $\bar{b}$ is attained by $b(\chi)$ with probability at least $2^{-n/5}$. This means that the class of colorings with $b(\chi) = \bar{b}$ has at least $2^{4n/5}$ members. Together with the previous considerations, we obtain

4.15 Lemma. *If* $H(b) \leq \frac{n}{5}$ *then there exists a no-nonsense partial coloring* χ *such that* $|\chi(S)| < \Delta_S$ *holds for all* $S \in \mathcal{S}$. □

To apply the method in specific examples, we need to estimate $H(b)$. As a first step, we use subadditivity: $H(b) \leq \sum_{S \in \mathcal{S}} H(b_S)$. It remains to estimate the entropies $H(b_S)$ for the individual sets and sum up. A large part of this calculation can be done once and for all. The distribution of b_S, and thus also its entropy, only depends on the size of S and on Δ_S. Our aim is thus to estimate the function $h(s, \Delta)$ defined as the entropy of b_S with $|S| = s$ and $\Delta_S = \Delta$.

As we know, for large s and for χ random, $\chi(S)$ behaves roughly as a normal random variable with standard deviation $\sqrt{s}$. By passing from $\chi(S)$ to b_S, we "shrink" all values of $\chi(S)$ from an interval $[(2i-1)\Delta, (2i+1)\Delta)$ to the single value i, thus forgetting some information about $\chi(S)$ (and thereby lowering the entropy).

Let us put $\lambda = \Delta/\sqrt{s}$. For estimating $h(s, \Delta)$, we distinguish two cases. If $\lambda \geq 2$, then $\chi(S)$ almost always lies in the interval $[-\Delta, \Delta)$ where b_S is 0,

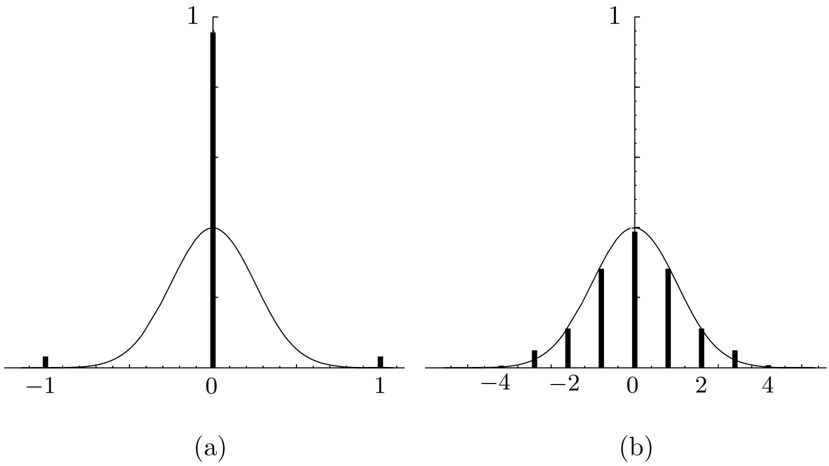


Fig. 4.1. Probability distributions of $\chi(S)/2\Delta$ and of b_S for $\lambda = 2$ (a), and for $\lambda = 0.4$ (b).

so the entropy will be small. Fig. 4.1(a) shows the probability distribution of $\chi(S)/2\Delta_S$ and the distribution of b_S for $\lambda = 2$ (with s large). On the other hand, for $\lambda < 2$, b_S has reasonable chance of attaining nonzero values, so the entropy will be larger (Fig. 4.1(b)). In other words, we do not violate the natural order of things so much by insisting that $|\chi(S)| < 10\sqrt{s}$ holds for the partial coloring, say, since a random coloring typically has this property anyway, and so we do not pay much entropy. On the other hand, requiring that $|\chi(S)| < \frac{\sqrt{s}}{10}$ already imposes quite a strong condition, so we need more entropy to compensate.

Having indicated what to expect, we do the actual calculation now.

The Case $\lambda \geq 2$. Let p_i denote the probability of $b_S = i$. We want to show that p_0 is very close to 1 and that the other p_i are small (Fig. 4.1(a)). For $i \geq 1$, we have

$$p_i \leq \Pr\left[\chi(S) \geq (2i-1)\Delta_S\right] = \Pr\left[\chi(S) \geq (2i-1)\lambda\sqrt{s}\,\right] \leq e^{-(2i-1)^2\lambda^2/2}$$

by Chernoff's inequality. By elementary calculus, the function $x \mapsto -x \log_2 x$ is nondecreasing on $(0, \frac{1}{e})$, and hence

$$-\sum_{i=1}^{\infty} p_i \log_2 p_i \leq \sum_{i=1}^{\infty} \frac{(2i-1)^2\lambda^2}{2\ln 2}\, e^{-(2i-1)^2\lambda^2/2}.$$

It is easy to check that the ratio of two successive terms in this series is smaller than $\frac{1}{4}$, and so by comparing the series with a geometric series we get that the sum is no larger than $\lambda^2 e^{-\lambda^2/2}$. By symmetry, the same bound applies for the contribution of the terms with $i \leq -1$.

For p_0 we derive

$$p_0 \geq 1 - \Pr\left[|\chi(S)| \geq \Delta_S\right] \geq 1 - 2e^{-\lambda^2/2}.$$

To estimate $\log_2 p_0$, we calculate that $2e^{-\lambda^2/2} < \frac{1}{3}$ for $\lambda \geq 2$, and we check that $\log_2(1-x) \geq -2x$ holds for $0 < x < \frac{1}{3}$ (more calculus). Therefore we have

$$-p_0 \log_2 p_0 \leq -\log_2 p_0 \leq -\log_2\left(1 - 2e^{-\lambda^2/2}\right) \leq 4e^{-\lambda^2/2}.$$

Altogether we obtain the estimate $h(s,\Delta) = H(b_S) \leq 6\lambda^2 e^{-\lambda^2/2}$. As one can check (by a computer algebra system, say), $6\lambda^2 e^{-\lambda^2/2}$ is bounded above by the simpler function $10e^{-\lambda^2/4}$ for all $\lambda \geq 2$.

The Case $\lambda < 2$. Here we can make use of the calculation done in the previous case, by the following trick. Let us decompose b_S into two parts $b_S = b'_S + b''_S$. The first addend b'_S is b_S rounded to the nearest integer multiple of $L = \lceil \frac{2}{\lambda} \rceil$; that is,

$$b'_S = L \operatorname{round}\left(\frac{b_S}{L}\right) = L \operatorname{round}\left(\frac{\chi(S)}{2L\Delta}\right).$$

Hence by the result of the $\lambda \geq 2$ case, we have $H(b'_S) = h(s, L\Delta) \leq 10e^{-(L\lambda)^2/4} \leq 4$, as $L\lambda \geq 2$. The second component, $b''_S = b_S - b'_S$, can only attain L different values, and thus its entropy is at most $\log_2 L$. Finally we obtain, by subadditivity,

$$\begin{aligned} h(s,\Delta) &= H(b_S) \leq H(b'_S) + H(b''_S) \leq 4 + \log_2 L \\ &\leq 4 + \log_2\left(\frac{2}{\lambda} + 1\right) \leq \log_2\left(16 + \frac{32}{\lambda}\right). \end{aligned}$$

The estimates in both cases can be combined into a single formula, as the reader is invited to check: $h(s,\Delta) \leq Ke^{-\lambda^2/4}\log_2(2 + 1/\lambda)$ for an absolute constant K. (This formula is a bit artificial, but it saves us from having to distinguish between two cases explicitly; it is a matter of taste whether it is better to write out the cases or not.) Plugging this into Lemma 4.15, we arrive at the following convenient device for applying the entropy method:

4.16 Proposition (Entropy method—quantitative version). *Let $\mathcal{S}$ be a set system on an n-point set X, and let a number $\Delta_S > 0$ be given for each $S \in \mathcal{S}$. Suppose that $\sum_{S\in\mathcal{S}} h(|S|, \Delta_S) \leq \frac{n}{5}$ holds, where the function $h(s,\Delta)$ can be estimated by*

$$h(s,\Delta) \leq Ke^{-\Delta^2/4s}\log_2\left(2 + \frac{\sqrt{s}}{\Delta}\right)$$

with an absolute constant K. Then there exists a no-nonsense partial coloring $\chi\colon X \to \{+1,-1,0\}$ such that $|\chi(S)| < \Delta_S$ for all $S \in \mathcal{S}$.

Often one only has upper bounds on the sizes of the sets in $\mathcal{S}$. In such a case, it is useful to know that the entropy contribution does not increase by decreasing the set size (while keeping Δ fixed). It suffices to check by elementary calculus that the function $s \mapsto e^{-\Delta^2/4s} \log_2(2+\sqrt{s}/\Delta)$ is nondecreasing in s.

Proof of Spencer's Upper Bound 4.2. We have a set system $\mathcal{S}$ on n points with m sets, $m \geq n$. We want to prove $\operatorname{disc}(\mathcal{S}) = O(\sqrt{n \ln(2m/n)})$. The desired coloring is obtained by iterating a partial coloring step based on Proposition 4.16.

To get the first partial coloring, we set $\Delta_S = \Delta = C\sqrt{n \ln(2m/n)}$ for all $S \in \mathcal{S}$, with a suitable (yet undetermined but sufficiently large) constant C. For the entropy estimate, we are in the region $\lambda \geq 2$, and so we have

$$\sum_{S \in \mathcal{S}} h(|S|, \Delta) \leq m \cdot h(n, \Delta) \leq m \cdot 10 e^{-\Delta^2/4n} = m \cdot 10 \left(\frac{n}{2m}\right)^{C^2} < \frac{n}{5}$$

for a sufficiently large C. Therefore, an arbitrary set system on n points with $m \geq n$ sets has a no-nonsense partial coloring with discrepancy at most $C\sqrt{n \ln(2m/n)}$.

Having obtained the first partial coloring χ_1 for the given set system $(\mathcal{S}, X_1)$, we consider the set system $\mathcal{S}_2$ induced on the set X_2 of points uncolored by χ_1, we get a partial coloring χ_2, and so on. The number of sets in $\mathcal{S}_i$ is at most m and the size of X_i is at most $\left(\frac{9}{10}\right)^i n$. We can stop the iteration at some step k when the number of remaining points drops below a suitable constant. The total discrepancy of the combined coloring is bounded by

$$\sum_{i=1}^{k} C \sqrt{\left(\tfrac{9}{10}\right)^i n \ln\left(\left(\tfrac{10}{9}\right)^i 2m/n\right)}.$$

After the first few terms, the series decreases geometrically, and thus the total discrepancy is $O\left(\sqrt{n \ln(2n/m)}\right)$ as claimed. □

Bibliography and Remarks. A refinement of Beck's partial coloring method similar to the one shown in this section was developed by Spencer [Spe85] for proving that the discrepancy of n sets on n points is $O(\sqrt{n})$. His method uses direct calculations of probability; the application of entropy, as suggested by Boppana for the same problem (see [AS00]), is a considerable technical simplification. As was remarked in Section 4.1, alternative geometric approaches to Spencer's result have been developed by Gluskin [Glu89] and by Giannopoulos [Gia97]; the latter paper can be particularly recommended for reading.

The possibility of taking set sizes into account and thus unifying the method, in a sense, with Beck's sophisticated applications of the partial coloring methods was noted in [Mat95]. Matoušek and Spencer

[MS96] use the method is a similar way, with an additional trick needed for a successful iteration of the partial coloring step, for proving a tight upper bound on the discrepancy of arithmetic progressions (see the remarks to Section 4.2). More applications of the entropy method can be found in [Sri97], [Mat96b], and [Mat98a].

Spencer's founding paper [Spe85] has the title "Six standard deviations suffice," indicating that the constant of proportionality can be made quite small: for instance, the discrepancy for n sets on n points is below $6\sqrt{n}$. The calculation shown above yields a considerably worse result. We were really wasteful only in the proof of the Partial coloring lemma, at the moment when we had a class of at least $2^{4n/5}$ colorings and concluded that it must contain two colorings at least $\frac{n}{10}$ apart. In reality, for instance, a class of this size contains colorings at least $0.48n$ apart. This follows from an isoperimetric inequality for the Hamming cube due to Kleitman [Kle66], which gives a tight bound on the number of points in a subset of the Hamming cube of a given diameter. Namely, any $\mathcal{C} \subseteq \{-1,+1\}^n$ of size bigger than $\sum_{j=0}^{\ell} \binom{n}{j}$ with $\ell \leq \frac{n}{2}$ contains two points differing in at least 2ℓ coordinates. This sum of binomial coefficients is conveniently bounded above by $2^{H(\ell/n)n}$, where $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ (here the customary notation $H(x)$ stands for a real function, not for the entropy of a random variable!). Using this result, it is sometimes possible to produce partial colorings with almost all points colored, say with at most $n^{0.99}$ uncolored points. Then much fewer than $\log n$ iterations of the partial coloring step are needed. An application, noted by Spencer, is given in Exercise 4.

Exercises

1. Prove the subadditivity property of entropy.
2. Prove that if a random variable Z attains at most k distinct values then $H(Z) \leq \log_2 k$.
3. (a) Let $\mathcal{S}$ be a system of n sets on an n-point set, and suppose that each set of $\mathcal{S}$ has size at most s. Check that the entropy method provides a no-nonsense partial coloring where each set of $\mathcal{S}$ has discrepancy at most $O(\sqrt{s})$.
   (b) Why can't we in general conclude that $\mathrm{disc}(\mathcal{S}) = O(\sqrt{s})$ for an $\mathcal{S}$ as in (a)? Show that this estimate is false in general.
4. (The discrepancy for m sets of size s) The goal is to show that for any m sets of size at most s, where $s \leq m$, the discrepancy is $O(\sqrt{s\log(2m/s)}\,)$. The important case, dealt with in (c), is an unpublished result of Spencer (private communication from September 1998).
   (a) Show that this bound, if valid, is asymptotically the best possible (at least for m bounded by a polynomial function of s, say).

(b) Why can we assume that n, the size of the ground set, equals m, and that $Cs \leq m \leq s^{1+\varepsilon}$ for arbitrary constants C and $\varepsilon > 0$?
(c)* With the assumptions as in (b), use Kleitman's isoperimetric inequality mentioned at the end of the remarks to this section to show that there is a partial coloring with at most s uncolored points for which each set has discrepancy at most $O(\sqrt{s \log(2m/s)}\,)$.
(d) Using (a)–(c), prove the bound claimed at the beginning of this exercise.