BANACH SPACE THEORY AND LOCAL OPERATOR THEORY

KENNETH R. DAVIDSON AND STANISLAW J. SZAREK

Contents

Connections between Banach space theory and classical operator theory on Hilbert space are numerous. First, one generalizes to the Banach space context notions and results involving operators on a Hilbert space. Second, more often than not the study of the former area also involves *linear* operators, and so one uses methods developed in one of the fields to attack questions in the other. And thirdly, spaces of operators are typically Banach spaces, and so we may study them as such using the methods of geometry of Banach spaces. This is, of course, not a complete list of links. In this survey we shall focus on those aspects of the latter two which are related to the local theory of Banach spaces and the asymptotic properties of finite dimensional normed spaces, as the dimension approaches infinity.

More specifically, we shall address the following issues.

I. Problems in operator theory which are local in the sense that they reduce to questions about finite matrices, and moreover the questions that result are similar in flavor to ones studied in the context of Banach spaces. For example, they might involve estimates which are independent of the dimension of the space on which the matrices act, or require the determination of approximate isomorphic or almost isometric asymptotics for some quantity as a function of the dimension. At the same time, the estimates needed are usually not optimal, a feature that would be characteristic to a representative Banach algebra or C*-algebra problem.

II. The asymptotic behavior of large dimensional random matrices, with particular attention to their spectral properties. This is a topic which has been of major interest in Banach space geometry and in the theory of operator algebras. It is also connected to numerous other fields, and has lately become a very hot subject in view of apparent analogies between the distribution of eigenvalues of large random matrices and of zeros of the Riemann ζ function. We shall not go deeply into these connections, but shall mention some of them in passing.

More generally, we shall try to bring up examples where the methods of one of the fields appear to have relevance to the other. This relevance may be described by a theorem, or just hinted by a vague analogy. As these connections between Banach space theory and operator theory are known to relatively few people on both sides, we hope that popularizing them will give a boost to both areas.

Before proceeding any further we shall use a simple example to clarify what we mean by isomorphic, almost isometric and isometric estimates. Let $\mathcal{T}_n$ denote the triangular truncation map on the $n \times n$ matrices which sends a matrix $A = [a_{ij}]_{i,j=1}^n$ to $\mathcal{T}_n A = [b_{ij}]$ where

$$b_{ij} = \begin{cases} a_{ij} & \text{if} \quad i \le j \\ 0 & \text{if} \quad i > j \end{cases}$$

Suppose that we are interested in the behavior of $\|\mathcal{T}_n\|$ as n tends to infinity. Several answers of varying degrees of precision (respectively isomorphic, almost isometric and isometric) are possible:

(a) $\|\mathcal{T}_n\|$ is of order $\log n$, meaning that there are universal constants $c, C > 0$ such that for all n, $c \log n \le \|\mathcal{T}_n\| \le C \log n$.

(b) $\|\mathcal{T}_n\|$ is equivalent to $\frac{1}{\pi} \log n$ [**8**], meaning that

$$\lim_{n\to\infty} \frac{\|\mathcal{T}_n\|}{\log n} = \frac{1}{\pi}$$

(c) $\|\mathcal{T}_n\| = f(n)$, where f is an explicit function.

We shall also say that (a) and (b) describe respectively, the rough and precise asymptotic order of $\|\mathcal{T}_n\|$. As above and unless otherwise indicated, c, c_1, C, C' etc. will stand for universal positive constants whose numerical values may vary between occurrences.

We now describe those topics which we will address in this article. In Part I, we will treat questions of local operator theory focusing on the following areas:

• Almost commuting versus nearly commuting problems: given a pair of (e.g. selfadjoint) matrices whose commutant is small in norm, is it close to a pair of (selfadjoint) matrices which exactly commute?

• Spectral distance problems: given two matrices (normal, unitary, etc.) which are close in norm, are their spectra close in some appropriate sense?

• Approximation of large matrices by direct sums of smaller ones: for example, the Herrero–Szarek result on non-reducibility of matrices and the examples of Szarek and Voiculescu of exotic quasidiagonal operators.

• The Kadison-Singer problem (reformulated as the paving problem): does there exist a positive integer k such that, given a square matrix A with zero diagonal, it is always possible to find *diagonal* projections $P_1, P_2, \ldots, P_k$ such that

$$\sum_{i=1}^{k} P_i = I \qquad \text{and} \qquad \|\sum_{i=1}^{k} P_i A P_i\| \leq \frac{1}{2}\|A\|?$$

• Questions about hyper-reflexivity: in particular, do $T_n \otimes T_n$, the tensor products of two copies of the algebra of $n \times n$ upper triangular matrices, have a common distance estimate analogous to Arveson's distance formula for nests?

We shall not elaborate on the Sz.Nagy–Halmos similarity problem recently solved by Pisier, see **[122]** or **[58]**. Even though it conforms quite well to the spirit of this article, it is in its heart, a complete boundedness/operator space question. We do not embark on those areas as both them and the similarity problem itself are well covered in the literature, cf. **[118, 124, 123]** and (in this collection) **[125]** for the former and **[55]** for the latter. On the other hand, we shall mention other less known questions about finite matrices that fit into our framework.

In each case, we shall try to sketch the connections to infinite operator theory. In particular, these problems relate to the Brown–Douglas–Fillmore theory, similarity and perturbations of operator algebras, approximation and distance estimates from spectral data, quasidiagonality, and the uniqueness of extensions of pure states. Depending on whether a problem has been solved or not, we shall either hint at the ingredients of the proof or describe the current state of the art and the consequences for the original problem. We point out that, in addition to their significance for the infinite dimensional theory, many of the problems stated above are also clearly important from the numerical linear algebra point of view.

In Part II, we shall signal various ways in which random matrices interact with operator theory and the geometry of Banach spaces. We start with an overview of the subject by presenting a sampler of classical (and not so classical) results on asymptotic spectral properties of large random matrices. We shall show how some of these classical results can be rather routinely obtained using standard methods of Banach space theory, e.g., by exploiting the measure concentration phenomenon and various tools of probability in Banach spaces. Although some of these arguments are folklore in the Banach space circles, they are generally not known to the wider mathematical public. Our purpose is to just illustrate this approach to the topic of random matrices, and so we shall not aim at the strongest or most general results that could be so obtained.

We shall also sketch the links between the subject of random matrices and Voiculescu's free probability and state some other subtle and interesting questions (mostly of the almost isometric nature) about norms and similar parameters of random matrices. These questions are not necessarily directly relevant to Banach space theory, but conceivably they can be approached with the aid of standard tools of the area.

We feel that it would be a worthwhile project to review and unify the closely related random matrix results and rather dissimilar methods pertaining to those results coming from, among others, mathematical physics, probability and Banach space theory. In addition to clarifying the picture, the benefits could include stating results in a form immediately utilizable in other fields. This is often a problem. For example, large deviation and mathematical physics formulations frequently involve rescaling the quantities in question and then investigating their asymptotics as the dimension tends to ∞. On the other hand, applications to the geometry of Banach spaces, convexity and computational

complexity typically require estimates that are valid for a wide range of parameters in *any* given dimension.

Finally, in the last section, we mention some of the technologies used in operator theory to exhibit phenomena analogous to the ones obtained in the Banach space theory via random methods. These technologies typically involve representation theory, but there are also links to graph theory and arithmetic geometry, to name a few.

I. Local Operator Theory

We will discuss these issues through a few important problems which are in various states of solution, as mentioned in the introduction.

Ia. **Almost Commuting Hermitian Matrices.**

There has been dramatic progress made on the following matrix problem:

Problem I.1. *Given $\varepsilon > 0$, is there a $\delta > 0$ so that: whenever A and B are $n \times n$ Hermitian matrices of norm one, for any $n \geq 1$, such that $\|AB - BA\| < \delta$, then there are Hermitian matrices A_1 and B_1 which exactly commute such that $\|A - A_1\| < \varepsilon$ and $\|B - B_1\| < \varepsilon$?*

The condition $\|AB - BA\| < \delta$ colloquially says that the matrices A and B *almost commute*; while the conditions $\|A - A_1\| < \varepsilon$ and $\|B - B_1\| < \varepsilon$ for a commuting Hermitian pair A_1 and B_1 say that A and B *nearly commute*. So the question becomes:

Are almost commuting Hermitian matrices nearly commuting?

An interesting reformulation of the problem turns this into a question about normal matrices. The matrix $T = A + iB$ satisfies

$$[T, T^*] := TT^* - T^*T = -2i(AB - BA).$$

Thus the Hermitian pair $\{A, B\}$ is almost commuting when T is *almost normal*, and exactly commuting when T is normal. Hence the pair is near to a commuting pair precisely when T is close to a normal matrix. Conversely, if T is an almost normal or nearly normal matrix, then the real and imaginary parts $A = (T + T^*)/2$ and $B = (T - T^*)/2i$ are almost commuting or nearly commuting, respectively. Thus another equivalent question is: *Are almost normal matrices nearly normal?*

As with many of the problems we discuss in this article, the crucial point is the dimension-free character of the estimates. Indeed, the problem could be reformulated for *compact* Hermitian matrices. However that was not the appropriate way to go for the solution, as we shall see.

Dimension dependent results were obtained early on by Pearcy and Shields [**119**] who establish that $\delta = \varepsilon^2/n$ will suffice. This was sharpened by the second author [**138**] to $\delta = c\varepsilon^{13/2}/\sqrt{n}$. We shall see however that these estimates are not nearly of the correct order.

It is interesting to note that there is a negative answer to the analogue of Problem I.1 for *arbitrary* operators on Hilbert space. Consider a weighted shift on ℓ^2 with basis $\{e_k : k \geq 1\}$ given by

$$S_n e_k = \frac{\min\{k, n\}}{n} e_{k+1} \quad \text{for} \quad k \geq 1.$$

It is easy to see that $[S_n, S_n^*] = -\frac{1}{n} P_n$ where P_n is the orthogonal projection onto $\operatorname{span}\{e_1, \ldots, e_n\}$. Thus for large n, this operator has small self-commutator; and thus its real and imaginary parts are almost commuting. However, every S_n is a compact perturbation of the unilateral shift S_1. This operator is a proper isometry which is Fredholm of index -1 and has essential spectrum $\sigma_e(S_1) = \mathbb{T}$. Therefore each S_n is also Fredholm of index -1 with essential spectrum $\sigma_e(S_n) = \mathbb{T}$, as these properties are invariant under compact perturbations. It is also straightforward to show that any operator T with $\|S_1 - T\| < 1$ has Fredholm index -1. So again, this is also true for S_n. A normal operator N has index 0 because $\ker N = \ker N^*$. Thus each S_n is at least distance 1 from a normal operator, and in fact this is exact.

There are many variants of our problem, and most of them have turned out to have negative answers. Here is one example which has received a lot of attention. Fix a basis $\{e_j : 0 \leq j \leq n\}$ for $\mathbb{C}^{n+1}$, and consider the matrices

$$\begin{aligned} Ae_j &= (1 - \tfrac{2j}{n}) e_j && \text{for} \quad 0 \leq j \leq n \\ Be_j &= \tfrac{2}{n+1} \sqrt{(j+1)(n-j)}\, e_{j+1} && \text{for} \quad 0 \leq j \leq n-1 \\ Be_n &= 0. \end{aligned}$$

Easy computations show that A is Hermitian, $\|[B, B^*]\| < \frac{4}{n}$ and $\|[A, B]\| < \frac{2}{n}$. Voiculescu [**151**] showed that the triple $\{A, \operatorname{Re} B, \operatorname{Im} B\}$ are almost commuting but not near to a commuting triple of Hermitian matrices. A refinement of this by the first author [**49**] showed that $\{A, B\}$ was not close to a commuting pair $\{A', B'\}$ where A' is self-adjoint but B' is not required to be normal. Finally Choi [**45**] showed that this pair is far (about .5) from any commuting pair with no conditions on A' or B'. Choi's obstruction involved the determinant of a polynomial in A and B.

Curiously, the matrix B above *is* close to a normal matrix; and the pairs $\{A, \operatorname{Re} B\}$ and $\{A, \operatorname{Im} B\}$ are also close to commuting. This was

shown in [**49**] using a method introduced by Berg [**24**]. Berg developed a technique for perturbing direct sums of weighted shifts that are almost normal to obtain normal operators when no index obstruction prevents it, as in the example cited above. This was a special case of a deep problem in operator theory.

This leads to a connection with the Brown–Douglas–Fillmore theory [**41, 42**] of essentially normal operators. An operator T is *essentially normal* if $[T, T^*]$ is compact. The original operator theory question they analyzed was to find when there is a compact perturbation of T which is normal. It is evident that if $T - \lambda I$ is Fredholm, then the Fredholm index must be zero. The operator T has image $t = \pi(T)$ in the Calkin algebra $\mathcal{L}(\mathcal{H})/\mathcal{K}$. The Calkin algebra is a C*-algebra, and t is a normal element of it. Thus $\mathrm{C}^*(t)$ is isomorphic to $\mathrm{C}(X)$ where $X = \sigma(t) = \sigma_e(T)$. For each bounded component U_i of $\mathbb{C} \setminus \sigma(t)$, the index $n_i = \mathrm{ind}(t - \lambda_i)$ is independent of the choice of a point $\lambda_i \in U_i$. A normal operator will have $n_i = 0$ for every 'hole'. And it is possible to construct an essentially normal operator T with $\sigma_e(T) = X$ and any prescribed indices $n_i = \mathrm{ind}(T - \lambda_i I)$. The Brown–Douglas–Fillmore theorem establishes that these are the only obstructions.

Theorem I.2 (Brown–Douglas–Fillmore)**.** *Suppose that T is an essentially normal operator. Then there is a compact operator K such that $T - K$ is normal if and only if* $\mathrm{ind}(T - \lambda I) = 0$ *for every λ in $\mathbb{C} \setminus \sigma_e(T)$.*

Surprisingly, the proof of BDF did not follow usual operator theoretic lines. Instead they observed that for any operator T above, there is a short exact sequence

$$0 \longrightarrow \mathcal{K} \xrightarrow{i} \mathrm{C}^*(T) + \mathcal{K} \xrightarrow{\pi} \mathrm{C}(X) \longrightarrow 0.$$

Thus T determines an extension of the compact operators by $\mathrm{C}(X)$. Conversely any such extension determines, up to a compact perturbation, an essentially normal operator such that $\pi(T) = z$, where z is the identity function on X. A third viewpoint identifies this extension with the monomorphism that identifies $\mathrm{C}(X)$ with the subalgebra $\mathrm{C}^*(t)$ of the Calkin algebra. They put a natural equivalence on extensions which corresponds to identifying operators which are unitarily equivalent up to a compact perturbation. The set of equivalence classes $\mathrm{Ext}(X)$ becomes a semigroup under the operation of direct sum. Moreover this construction makes sense for any compact metric space X, not just subsets of the plane. They established that, in fact, $\mathrm{Ext}(X)$ is a group. Moreover it is a generalized homology theory paired in a natural way

with topological K-theory. This opened the door to a new era in C*-algebras in which K-theory and topological methods came to play a central role.

Even the existence of a zero element is non-trivial. In the case of a single operator or planar X, this reduces to the Weyl–von Neumann–Berg theorem [**23**] which shows that every normal operator is a small compact perturbation of a diagonalizable operator. This is easily generalized to arbitrary metric spaces, and it shows that any representation of $\mathrm{C}(X)$ on Hilbert space is close in an appropriate sense to a representation by diagonal operators. BDF show by elementary methods that any essentially normal operator T is unitarily equivalent to a small compact perturbation of $T \oplus D$, where D is a diagonal normal operator with $\sigma(D) = \sigma_e(D) = \sigma_e(T)$.

This leads us somewhat afield from the original problem. However the connection is made via *quasidiagonality*. A set of operators $\mathcal{T}$ is quasidiagonal if there is an increasing sequence P_k of finite rank projections converging strongly to I such that $\lim_{n\to\infty} \|[T, P]\| = 0$ for every $T \in \mathcal{T}$. In particular, if T is quasidiagonal and $\varepsilon > 0$, then there is a sequence P_k such that

$$T - \sum_{k\geq 1} (P_k - P_{k-1})T(P_k - P_{k-1}) = K$$

is compact and $\|K\| < \varepsilon$. Indeed, one just drops to an appropriate subsequence of the original one. Thus $T - K$ has the form $\sum_{k\geq 1}^{\oplus} T_k$ where T_k act on finite dimensional spaces. Such operators are called *block diagonal*.

Salinas [**133**] showed that the closure of the zero element in $\mathrm{Ext}(X)$ in the topology of pointwise–norm convergence, where we think of an extension as a monomorphism from $\mathrm{C}(X)$ into $\mathcal{L}(\mathcal{H})/\mathcal{K}$), is the set of all quasidiagonal extensions. In particular, unlike the planar case, the trivial element need not be closed. He used this to establish homotopy invariance of the Ext functor. In the case of a single operator, this shows that if T is essentially normal with zero index data, then it has a small compact perturbation which is block diagonal, say

$$T - K = T' = \sum_{k\geq 1}{}^{\oplus} T_k.$$

Moreover, it follows that $[T', T'^*] = \sum_{k\geq 1}^{\oplus} [T_k, T_k^*]$. Therefore

$$\lim_{k\to\infty} \|[T_k, T_k^*]\| = 0.$$

A positive solution to our original question would yield normal matrices N_k such that $\lim_{k\to\infty} \|T_k - N_k\| = 0$. Hence $N = \sum_{k\geq 1}^{\oplus} N_k$ is a normal operator such that $T - N$ is compact.

This was the motivation for the work of the first author [**49**]. If $T = \sum_{k\geq 1}^{\oplus} T_k$ is essentially normal with $\sigma_e(T) = X$, then after a small compact perturbation, T may be replaced with $T \oplus D$ where D is any normal operator with eigenvalues in X. Thus each summand T_k can be replaced by $T_k \oplus D_k$, where D_k is diagonal. An absorption theorem is proven:

Theorem I.3 (Davidson)**.** *If T is any $n \times n$ matrix, then there are an $n \times n$ normal matrix N and an $2n \times 2n$ matrix M such that*

$$\|M - T \oplus N\| \leq 75\|[T, T^*]\|^{1/2}.$$

One difficulty in applying this to the BDF problem is that any holes in the spectrum of T are obliterated by summing with a normal with eigenvalues dense in the unit disk.

In [**25**], Berg and the first author showed that similar operator-theoretic techniques can be used to show first that every essentially normal operator with zero index data is quasidiagonal. And then refinements of the absorption principle which control the spectrum were used to prove the planar version of BDF. This had the advantage of providing quantitative estimates for 'nice' spectra.

But the correct positive answer to our question came again from a more abstract approach. Huaxin Lin [**100**] considered $T = \sum_{k\geq 1}^{\oplus} T_k$ as an element of the von Neumann algebra $\mathfrak{M} = \prod_{k\geq 1} \mathfrak{M}_{n_k}$. A positive solution is equivalent to constructing a normal operator $N = \sum_{k\geq 1}^{\oplus} N_k$ asymptotic to T, that is, verifying $\lim_{k\to\infty} \|T_k - N_k\| = 0$. This is a perturbation by the ideal $\mathfrak{J} = \sum_{k\geq 1} \mathfrak{M}_{n_k}$ of elements $J = \sum_{k\geq 1}^{\oplus} J_k$ where $\lim_{k\to\infty} \|J_k\| = 0$. Thus T determines an element $t = q(T)$ in the quotient algebra $\mathfrak{M}/\mathfrak{J}$. The problem is reduced to *lifting* each normal element of $\mathfrak{M}/\mathfrak{J}$ to a normal element of $\mathfrak{M}$. This was accomplished by a long tortuous argument.

Theorem I.4 (Lin)**.** *Given $\varepsilon > 0$, is there a $\delta > 0$ so that whenever A and B are $n \times n$ Hermitian matrices of norm one, for any $n \geq 1$, and $\|AB - BA\| < \delta$, then there are Hermitian matrices A_1 and B_1 which exactly commute such that $\|A - A_1\| < \varepsilon$ and $\|B - B_1\| < \varepsilon$.*

A remarkable new proof of Lin's theorem is now available due to Friis and Rørdam [**71**]. They take the same approach, but base their argument only on two elementary facts about $\mathfrak{M}/\mathfrak{J}$. It has stable rank one,

meaning that the invertible elements are dense. And every Hermitian element can be approximated by one with finite spectra. Both of these results can be obtained by using well-known facts about matrices on finite dimensional spaces. From this, a short argument cleverly using no more than the basic functional calculus yields the desired lifting of normal operators. Thus the original question has a positive solution. See Loring's book [**102**] for a C*-algebraic view of this problem and its solution. In a second paper [**72**], they show how the planar case of the Brown–Douglas–Fillmore Theorem follows.

An important matrix question remains:

Problem I.5. *What is the dependence of δ on ε for almost commuting Hermitian matrices?*

Optimal estimates for δ should be $O(\varepsilon^2)$ as in the absorption results. However the Lin and Friis–Rørdam results yield only qualitative information, and do not provide information on the nature of the optimal δ function.

A related question deals with pairs of almost commuting pairs of unitary matrices. Voiculescu [**151**] constructed the pair $Ue_j = e_{j+1 \mod n}$ and $Ve_j = e^{2\pi ij/n}e_j$ for $1 \le j \le n$. It is evident that $\|[U, V]\| < 2\pi/n$. However he showed that this pair is not close to a commuting pair of unitaries. He speculated that both this example and the commuting triples example occur because of topological obstructions. Asymptotically, the spectrum approximates a torus (or a two-sphere for Hermitian triples), and thus there is a heuristic justification for this viewpoint. Loring [**101**] showed that this was indeed the case for this pair of unitaries by establishing a K-theoretic obstruction coming from the non-trivial homology of the two-sphere. Exel and Loring [**67**] obtained an elementary determinant obstruction that shows that this pair of unitaries is far from any commuting pair, unitary or not.

Ib. **Unitary Orbits of Normal Matrices.** Consider a normal matrix N in $\mathfrak{M}_k$. By the spectral theorem, there is an orthonormal basis which diagonalizes N. So N is unitarily equivalent to $\mathrm{diag}(\lambda_1, \ldots, \lambda_k)$ where the diagonal entries form an enumeration of the spectrum of N including multiplicity. Two normal matrices are unitarily equivalent precisely when they have the same spectrum and the same multiplicity of each eigenvalue. A natural question from numerical analysis and from the approximation theory of operators asks how the spectrum can change under small perturbations. The answer one gets depends very much on what kind of perturbations are allowed. In particular, if the

perturbed matrix is also normal, then one obtains a more satisfying answer than if the perturbation is arbitrary.

Consider first the case of a Hermitian matrix A. Then the eigenvalues may be enumerated so that $\lambda_1 \geq \cdots \geq \lambda_k$. If B is a small perturbation of A, say $\|A-B\| < \varepsilon$, then the Hermitian matrix $R = (B+B^*)/2$ also satisfies $\|A - R\| < \varepsilon$, and so we may assume that the perturbation B is Hermitian to begin with. In that case, write the eigenvalues as $\mu_1 \geq \cdots \geq \mu_k$. Then a 1912 argument due to Weyl **[161]** gives:

Theorem I.6 (Weyl)**.** *Let A and B be $k \times k$ Hermitian matrices with eigenvalues ordered as $\lambda_1 \geq \cdots \geq \lambda_k$ and $\mu_1 \geq \cdots \geq \mu_k$, respectively. Then*

$$\sup_{1 \leq j \leq k} |\mu_j - \lambda_j| \leq \|A - B\|.$$

Moreover, for each A, this value is attained by some such B.

Indeed, suppose for example that $\lambda_j \geq \mu_j$. Consider the spectral subspace for A corresponding to $\{\lambda_1, \ldots, \lambda_j\}$ and the spectral subspace of B for $\{\mu_j, \ldots, \mu_k\}$. Since the dimensions of these spaces add up to more than k, they contain a common unit vector x. Then $\langle Ax, x\rangle$ lies in the convex hull of $\{\lambda_1, \ldots, \lambda_j\}$, whence is $\geq \lambda_j$. Similarly $\langle Bx, x\rangle \leq \mu_j$. Hence

$$\|A - B\| \geq \langle (A-B)x, x\rangle = \langle Ax, x\rangle - \langle Bx, x\rangle \geq \lambda_j - \mu_j.$$

Weyl's theorem shows that a perturbation of Hermitian matrices of norm ε cannot move any eigenvalue more than ε. One can reformulate this to say that if A and B are Hermitian with eigenvalues as above, then $\|A-B\|$ is bounded below by the *spectral distance* $\sup_{1 \leq j \leq k} |\mu_j - \lambda_j|$. This spectral distance is achieved when A and B are simultaneously diagonalizable with eigenvalues matched in increasing order.

The *unitary orbit* $\mathcal{U}(A)$ of a matrix A is the set of all matrices which are unitarily equivalent to it. In finite dimensions, this is a closed set because the unitary group is compact. However, for operators this is no longer the case. The unitary orbit of a normal matrix consists of all normal matrices with the same spectral data. Thus one must be able to express the distance between these two unitary orbits solely in terms of the two spectra. Weyl's result shows that in the Hermitian case, this distance is precisely the spectral distance.

More recently, significant attention has been paid to the more difficult normal case. Here is it no longer trivial to define a spectral distance. However, if M and N are normal matrices which are simultaneously diagonalizable, then this diagonalization amounts to a pairing

of the two sets of eigenvalues. Among all possible permutations, there is one of least distance, and this is called the spectral distance:

$$\text{sp. dist}(\boldsymbol{\lambda}, \boldsymbol{\mu}) := \inf_{\pi \in \mathcal{S}_k} \max_{1 \le j \le k} |\lambda_j - \mu_{\pi(j)}|$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ represent the eigenvalues repeated according to their multiplicities. (This is really the quotient metric on ℓ_∞^k divided by the natural action of the symmetric group on $\mathbb{C}^n$.) Evidently,

$$\text{dist}(\mathcal{U}(M), \mathcal{U}(N)) \le \text{sp. dist}(\sigma(M), \sigma(N))$$

where $\sigma(M)$ denotes the spectrum of M including multiplicity.

Many results in the literature (cf. **[31, 137]**) show that under special hypotheses, the spectral distance is again the exact answer. See **[29]** for a full treatment of these ideas. However a recent result of Mueller (private communication) shows that even in three dimensions, it is possible that $\text{dist}(\mathcal{U}(M), \mathcal{U}(N)) < \text{sp. dist}(\sigma(M), \sigma(N))$. His example is explicit, but it is fair to say that we do not really understand why this strict inequality occurs.

So it is of great interest that a result of Bhatia, Davis and McIntosh **[30]** obtains bounds independent of dimension:

Theorem I.7 (Bhatia–Davis–McIntosh)**.** *There is a universal constant c $(> 1/\pi)$ such that for all normal matrices M and N,*

$$\text{dist}(\mathcal{U}(M), \mathcal{U}(N)) \ge c \, \text{sp. dist}(\sigma(M), \sigma(N)).$$

This result is obtained from a study of the Rosenblum operator on $\mathfrak{M}_k$ given by $\tau_{M,N}(X) = NX - XM$. It is well known that this map is injective if $\sigma(M)$ and $\sigma(N)$ are disjoint. In the case in which M and N are normal, there is a useful integral formula for the inverse which allows norm estimates. This integral formula requires finding a function f in $L^1(\mathbb{R}^2)$ with Fourier transform $\hat{f}$ satisfying $\hat{f}(x, y) = \frac{1}{x+iy}$ for $x^2 + y^2 \ge 1$. The constant depends on $\|f\|_1$, which may be chosen less than π.

Estimates of this type allow one to estimate the angle between spectral subspaces of nearby normal operators. For example, if $E = E_M(K)$ and $F = E_N(L)$ are the spectral projections of M and N respectively for sets K and L with $\text{dist}(K, L) = \delta$, then

$$\|EF\| \le \frac{1}{c\delta} \|A - B\|.$$

When this is small, the ranges of E and F are almost orthogonal.

These results can be extended to normal operators on infinite dimensional space. Here the unitary orbit of a normal operator is not closed. The unitary invariants are given by the spectral theorem in

terms of the spectral measure and a multiplicity function. However, the Weyl–von Neumann–Berg theorem [**23**] shows that the closure of the unitary orbit of every normal operator contains a diagonal operator. The multiplicity and spectral measure do not play a role in this closure except for *isolated* eigenvalues. Indeed, two normal operators have the same unitary orbit if and only if they have the same essential spectrum, and the same multiplicity at each isolated eigenvalue. (See [**54**, Theorem II.4.4].) It turns out that one can define an analogue of the spectral distance for spectra of normal operators. Once this is accomplished, the finite dimensional methods extend to show that the distance between unitary orbits is equal to the spectral distance for Hermitian operators [**15**] and greater than a universal constant times the spectral distance for normal operators [**50**].

Some of these results are valid for other norms on the set of matrices. Although we are usually interested in the operator norm because of its central role, other norms occur naturally. For example, the Schatten p-norms are defined as the ℓ^p norm of the singular values. In particular, $p = 2$ yields the Hilbert–Schmidt norm, sometimes also called the Frobenius norm, which is particularly tractable. More generally, there is a large class of norms which are *unitarily invariant* in the strong sense that $\|A\|_\tau = \|UAV\|_\tau$ for all unitaries U and V; for $\mathfrak{M}_n$, there is a one-to-one correspondence between such norms and *symmetric* norms on $\mathbb{C}^n$ (or $\mathbb{R}^n$) applied to the sequence of singular values. For example, the Ky Fan norms are given by the sum of the k largest singular values. The set of (properly normalized) unitarily invariant norms is a closed convex family and the Ky Fan norms are the extreme points.

One can define the spectral distance between the unitary orbits of two normal matrices with respect to any unitarily invariant norm by

$$\operatorname{sp.dist}_\tau(\boldsymbol{\lambda}, \boldsymbol{\mu}) := \inf_{\pi \in \mathcal{S}_k} \| \operatorname{diag}\left(\lambda_j - \mu_{\pi(j)}\right) \|_\tau$$

(Again, this is really a quotient metric induced by the symmetric norm on $\mathbb{C}^n$ realted to $\|\cdot\|_\tau$.) A classical result for the Hilbert–Schmidt norm is due to Hoffman and Wielandt [**87**]. Note that there is no constant needed.

Theorem I.8 (Hoffman–Wielandt)**.** *If M and N are normal matrices, then*

$$\operatorname{sp.dist}_2(\sigma(M), \sigma(N)) \leq \|M - N\|_2.$$

For general unitarily invariant norms, a result of Mirsky [**112**] extends Weyl's Theorem for Hermitian matrices. A more recent result of Bhatia [**28**] extends this to certain normal pairs.

Theorem I.9 (Mirsky–Bhatia)**.** *Suppose that M and N are normal matrices such that $M - N$ is also normal (e.g. M and N Hermitian). Then for every unitarily invariant norm,*

$$\text{sp. dist}_\tau(\sigma(M), \sigma(N)) \leq \|M - N\|_\tau.$$

For unitary matrices, Bhatia, Davis and McIntosh **[30]** show that

Theorem I.10. *Suppose that U and V are unitary matrices. Then for every unitarily invariant norm,*

$$\frac{2}{\pi} \text{sp. dist}_\tau(\sigma(U), \sigma(V)) \leq \|U - V\|_\tau.$$

More results along these lines may be found in Bhatia's books **[27, 29]**.

When matrices are not normal, there are generally no good estimates of distance in terms of their spectra. A simple example is to take an orthonormal basis $e_1, \ldots, e_n$ for $\mathbb{C}^n$, and set J_n to be the Jordan nilpotent $J_n e_i = e_{i+1}$ for $1 \leq i < n$ and $J_n e_n = 0$. Then consider $T = J_n + \varepsilon e_1 e_n^*$. We see that $\|T - J_n\| = \varepsilon$ while $\sigma(J_n) = \{0\}$ and $\sigma(T)$ consists of the nth roots of ε. So the spectral distance is $\varepsilon^{1/n}$. This is much more dramatic than the normal case. We refer the reader to the books **[27, 29]** for a lot more material on these types of questions.

Likewise when calculating the eigenvalues of a normal matrix numerically, we are dealing with its perturbation or approximation which is not necessarily normal. A small perturbation of a normal matrix may have spectrum which does not approximate the spectrum of the normal at all. Indeed, Herrero **[82]** shows that there is a normal matrix N in $\mathfrak{M}_k$ with $1 \in \sigma(N)$ and a nilpotent matrix Q such that $\|N - Q\| < 5k^{-1/2}$. This allowed him to establish an important infinite dimensional result:

Theorem I.11 (Herrero)**.** *Suppose that N is a normal operator on Hilbert space with connected spectrum containing 0. Then N is the norm limit of a sequence of nilpotent operators.*

One way to construct such examples is to take Q to be a weighted shift $Qe_i = a_i e_{i+1}$ for $1 \leq i < k$ and $Qe_k = 0$. This is evidently nilpotent. However, if the weights slowly increase from 0 to 1 and back down again, it will follow that $\|Q\| = 1$ and $\|Q^*Q - QQ^*\|$ is small on the order of $O(k^{-1})$. We saw in the previous section that the matrix Q must be close to a normal matrix. For weighted shifts, a method known as Berg's technique **[24]** (cf. **[82]**) allows explicit perturbations to a normal matrix with estimates on the order of $C\|Q^*Q - QQ^*\|^{1/2}$, which yields the desired example.

Self-adjoint examples can be constructed, but the estimates are strikingly different. Hadwin [**81**] solved this problem by relating the problem of approximating a Hermitian matrix by a nilpotent one to the norm of triangular truncation, mentioned in the introduction.

Theorem I.12 (Hadwin)**.** *For $k \geq 1$, let*

$$\delta_k = \inf\{\|A - Q\| : A, Q \in \mathfrak{M}_k,\ A = A^*,\ Q^k = 0\}.$$

Then $\lim_{k\to\infty} \delta_n \log n = \dfrac{\pi}{2}$.

However, for the numerical analysis problem of calculating eigenvalues of a matrix which is *a priori* known to be normal, this deterioration of estimates is inessential. Indeed, if A_0 is a given small but not necessarily normal perturbation of an unknown normal matrix A, and A_1 is *any* normal approximant of A_0, then by Theorem I.7, the spectra of A and A_1 are close. However this argument is not constructive, as it does not tell us how to find A_1. This leads to a refinement of Problems I.1 and I.5: *given a matrix A_0 which is known to be nearly normal, find a* specific *normal approximant with an* explicit *error estimate.*

In infinite dimensions, the description of those operators which are limits of nilpotents is a central result in the approximation theory of operators due to Apostol, Foiaş and Voiculescu[**10**], cf. [**82**, Ch. 5].

Theorem I.13. *An operator T is the limit of a sequence of nilpotent operators if and only if*

(i) *T has connected spectrum and essential spectrum containing* 0*, and*
(ii) *whenever $T - \lambda I$ is semi-Fredholm, the Fredholm index is* 0*.*

These conditions are easily seen to be necessary by elementary means. The special case of normal limits mentioned above follows from the finite dimensional approximations and an application of the continuous functional calculus. However, the general results require a good model theory for a dense set of Hilbert space operators. This requires interplay between infinite dimensional methods and analytic function theory. We return to some considerations of this result in the next section.

Ic. **Quasidiagonality.** The class of quasidiagonal operators and block diagonal operators were introduced in section Ia. A priori, one might expect that quasidiagonal operators behave more like matrices than arbitrary operators. However, in many ways, it has proven to be a difficult class of operators to deal with. However, they do occur in important ways in both operator theory and C*-algebras. Here we will limit our attention to questions related to questions about matrices.

The class of block diagonal operators of the form $T = \sum_{j\geq 1}^{\oplus} T_j$ which most closely mimics matrices are those in which the summands T_j are of uniformly bounded dimension, say m. Such operators, after a rearrangement of basis, can be written as a direct sum of $k \times k$ matrices with diagonal operator entries for $1 \leq k \leq m$. The Weyl–von Neumann–Berg Theorem shows that every normal operator is the norm limit of diagonalizable operators. More generally, the same methods show that any finite set of commuting normal operators can be simultaneously approximated by diagonal operators. Thus the closure of the set of operators which are $k \times k$ matrices with diagonal entries is the set of $k \times k$ matrices with commuting normal entries. These operators are called k-normal. From a more algebraic point of view, the C*-algebra generated by an m-normal operator has the property that there is a family of irreducible representations of dimension at most m which separate points in the algebra. Conversely this condition implies that the operator is the direct sum of k-normal operators for $k \leq m$.

In this section, we will concern ourselves with a variety of questions dealing with the approximation of operators in the quasidiagonal class. There are also many connections with C*-algebras which will have to be neglected.

Consider first the question of whether every quasidiagonal operator can be approximated by m-normal operators with m unlimited. The first attempt would be to try to approximate large finite matrices by a direct sum of smaller matrices. This works very well for weighted shifts using *Berg's technique* [**24**]. This method has been exploited by Herrero [**84**] in obtaining good estimates for the distance between unitary orbits of weighted shifts. With Berg, the first author [**25**] expanded this method to work for general block tri-diagonal forms in our proof of a quantitative Brown–Douglas–Fillmore Theorem [**41, 42**]. Further evidence comes from the infinite dimensional setting where Voiculescu's celebrated generalized Weyl–von Neumann Theorem [**150**] implies, among other things, that every operator is the norm limit of operators with many reducing subspaces. Indeed, every operator T is an arbitrarily small compact perturbation of an operator unitarily equivalent to $T' = T \oplus A \oplus A \oplus \ldots$, where A is the image of T under a $*$-representation of $\mathrm{C}^*(T)$ which annihilates $\mathcal{K}(\mathcal{H}) \cap \mathrm{C}^*(T)$. This is especially good since T and T' have the same closed unitary orbit (whence T' is also the limit of operators unitarily equivalent to T).

In spite of this positive evidence, the desired decomposition of large matrices into smaller ones is not possible.

Theorem I.14 (Herrero–Szarek)**.** *There is a computable constant $\delta > 0$ so that for every $n \geq 2$, there is an $n \times n$ matrix A_n of norm one which cannot be approximated within δ by any matrix which decomposes as an orthogonal direct sum of smaller matrices.*

The proof is probabilistic. It measures the set of reducible operators, and shows that small balls centered there cannot cover the whole ball of $\mathfrak{M}_n$. While the computed value of δ is quite small, 1.712×10^{-7}, Herrero conjectured that δ is probably $1/2$.

This negative evidence is not sufficient in itself to answer the question of approximation by m-normals. Nevertheless, the union of the set of m-normals over all finite m is not norm dense in the set of diagonal operators. This was established by two different methods. The second author [**139**] pushed the probabilistic argument harder and was able to show the existence of block diagonal operators which cannot be uniformly approximated by m-normals.

Voiculescu [**152, 153**] used more algebraic methods to find an obstruction to approximation. He showed that if T can be approximated by m-normals, then the injection of $\mathrm{C}^*(T)$ into $\mathcal{B}(\mathcal{H})$ is a nuclear map. This is equivalent to saying that $\mathrm{C}^*(T)$ can be imbedded as a subalgebra of a nuclear C*-algebra. Due to the work of Kirchberg, this class is known to be the set of exact C*-algebras. Voiculescu then shows by non-constructive methods that many block-diagonal operators do not have this property. In the second paper, property T groups are exploited to yield a more constructive approach. These papers point to the possibility of explicit examples by using explicit representations of explicit groups with property T such as $SL_3(\mathbb{Z})$. However they stop short of actually producing such an example.

On the other hand, Dadarlat [**47**] has positive results which show that this exactness condition could be precisely the obstruction. He shows that if T is quasidiagonal and $\mathrm{C}^*(T)$ contains no non-zero compact operators, then T is the limit of m-normal operators if and only if $\mathrm{C}^*(T)$ is exact.

Voiculescu's Weyl–von Neumann Theorem is an important tool is the approximation theory of Hilbert space operators. So when working with quasidiagonal operators, it is natural to want to use this result and stay within the set of quasidiagonal operators. Generally one wants to approximate T by operators unitarily equivalent to $T \oplus A$ or $T \oplus A^{(\infty)}$ where $A = \rho(\pi(T))$ and ρ is a faithful representation of $\mathrm{C}^*(T)/(\mathcal{K}(\mathcal{H}) \cap \mathrm{C}^*(T))$. We mention in passing that this somewhat resembles the "Pełczyński decomposition method" used in Banach space

theory. So the question arises whether ρ can be chosen to be quasidiagonal when T is quasidiagonal. Indeed a C*-algebra is called quasidiagonal if it has a faithful quasidiagonal representation. Clearly this property is preserved by subalgebras, but our question asks about specific *quotients*. It turns out to have a negative answer. Wassermann [**158, 159**] constructs several counterexamples. He uses the fact that the reduced C*-algebra of a non-amenable group cannot be quasidiagonal. His first example uses a connection of the free group to Anderson's example [**7**] of a C*-algebra $\mathfrak{A}$ for which $\mathrm{Ext}(\mathfrak{A})$ is not a group. His second paper exploits property T. On the other hand, in [**56**] it is shown that the quotient of a *nuclear* quasidiagonal C*-algebra by the compact operators remains quasidiagonal.

Theorem I.13 of Apostol, Foiaş and Voiculescu [**10**] characterized the closure of the nilpotent operators. An interesting corollary is that the closure of the algebraic operators, those satisfying a polynomial identity, is the set of operators satisfying only the second condition (ii) whenever $T - \lambda I$ is semi-Fredholm, the Fredholm index is 0. Another deep result of Apostol, Foiaş and Voiculescu [**9**] characterizes this class as the set of *biquasitriangular* operators.

A *triangular* operator T is one which has an orthonormal basis $\{e_n : n \geq 1\}$ in which the basis of T is upper triangular. Say that T is *quasitriangular* if it is the limit of triangular operators, normally with respect to different bases. An operator is biquasitriangular if both T and T^* are quasitriangular. Since every matrix can be put into triangular form, it is easy to see that every block diagonal operator is triangular. Consequently, every quasidiagonal operator if biquasitraingular. The converse is far from correct.

Within the class of quasidiagonal operators, one may ask for a description of the closure of all nilpotent or all algebraic quasidiagonal operators. Since quasidiagonal operators trivially satisfy condition (ii) of Theorem I.13, the natural conjecture might well be that property (i) should describe the closure of quasidiagonal nilpotents while the closure of quasidiagonal algebraic operators might be all quasidiagonal operators. Unfortunately, the answer to the nilpotent question is no; and the algebraic question remains open.

Herrero [**85**] considered the operator $\left[\begin{smallmatrix} I & I \\ 0 & D \end{smallmatrix}\right]$ where D is diagonal with eigenvalues dense in $[0, 1]$. This operator is the direct sum of 2×2 matrices, and so is 2-normal. The spectrum and essential spectrum is $[0, 1]$ which is connected and contains the origin. So this operator is the norm limit of nilpotent operators. However, there is a trace obstruction to being the limit of nilpotent quasidiagonal operators.

A natural finite dimensional question could play a central role in solving this problem.

Problem I.15. *If T is an $n \times n$ matrix, can the distance from T to the set* Nil_k *of nilpotent matrices of order k be estimated in terms of the quantity $\|T^k\|^{1/k}$?*

On the other hand, it is easy to exhibit Herrero's operator as the limit of quasidiagonal algebraic operators—just approximate D by diagonal operators with finite spectrum. So he asked whether every quasidiagonal operator is the limit of quasidiagonal algebraic operators. In [**56**], a special case of this conjecture is verified.

Theorem I.16 (Davidson–Herrero–Salinas)**.** *Suppose that T is quasidiagonal, its essential spectrum does not disconnect the plane, and $\mathrm{C}^*(T)/(\mathcal{K}(\mathcal{H})\cap\mathrm{C}^*(T))$ has a faithful quasidiagonal representation. Then T is the limit of quasidiagonal algebraic operators.*

We mention a problem from [**56**] in the spirit of this survey article. A positive solution would provide a positive answer to the algebraic approximation question.

Problem I.17. *Given $\varepsilon > 0$, is there a constant $C(\varepsilon)$ independent of n such that for every matrix T with $\|T\| \le 1$, there exists a diagonal operator D and an invertible operator W such that*

$$\|T - WDW^{-1}\| \qquad \textit{and} \qquad \|W\|\,\|W^{-1}\| \le C(\varepsilon).$$

A compactness argument shows that for each fixed dimension n, there is a constant $C(\varepsilon, n)$ which works.

Id. **Extensions of pure states and matrix paving.** In this section, we will discuss a well-known problem of Kadison and Singer [**89**], which asks whether every pure state on the algebra $\mathcal{D}$ of diagonal operators on ℓ_2 with respect to the standard orthonormal basis extends uniquely to a (necessarily pure) state on $\mathcal{L}(\ell_2)$. We note that $\mathcal{D}$ is a generic discrete maximal abelian subalgebra or masa. It is proved in [**89**] (see also [**4**]) that the answer to the uniqueness question is negative for non-discrete masas.

The interest in the Kadison-Singer problem lies, in particular, in the fact that a positive solution would shed new light on the structure of pure states on $\mathcal{L}(\ell_2)$. In fact it would go a long way towards a simple characterization of such states. Indeed, pure states on $\mathcal{D} \simeq \ell_\infty \simeq C(\beta\mathbb{N})$ are just point evaluations at elements of $\beta\mathbb{N}$, the Stone–Čech compactification of $\mathbb{N}$. The canonical extension of a state from $\mathcal{D}$ to $\mathcal{L}(\ell_2)$ is obtained by composing the state with the conditional

expectation E which takes each T in $\mathcal{L}(\ell_2)$ to its diagonal part in $\mathcal{D}$. In other words, if $\mathcal{U}$ is the ultrafilter associated to a pure state φ on $\mathcal{D}$, then the extension ψ to $\mathcal{L}(\ell_2)$ is given by

$$\psi(T) = \varphi(E(T)) = \lim_{n\in\mathcal{U}} \langle Te_n, e_n\rangle.$$

This suggests the open question raised in [**89**, §5] whether *every* pure state on $\mathcal{L}(\ell_2)$ is an extension of a pure state on some discrete masa, or indeed a given masa. However, it has been shown in [**6**], cf. [**4**, §7.3], that the two questions are in fact equivalent. See also [**5**] for a related positive result for the Calkin algebra. On the other hand, every pure state on $\mathcal{L}(\ell_2)$ is of the form $\varphi(T) = \lim_{n\in\mathcal{U}} \langle Tx_n, x_n\rangle$ for some sequence (x_n) of unit vector in ℓ_2 and some ultrafilter in $\mathbb{N}$ [**164**]. Moreover, if φ is not a vector state, the weak limit $w\text{-}\lim_{n\in\mathcal{U}} x_n$ is necessarily 0. Since weak-null *sequences* in ℓ_2 can be refined to be asymptotically orthogonal, it follows that the difficulty in settling both this and the uniqueness question lies in the difference between limits over ultrafilters and regular limits.

At the first sight, the Kadison-Singer problem seems to be a strictly infinite-dimensional question. However, it was shown by Anderson ([**4**]) that it is equivalent to a finite dimensional question known as the *paving problem*.

Problem I.18. *Does there exist a positive integer k such that, for any $n \geq 1$ and any matrix $A \in \mathfrak{M}_n$ with zero diagonal, one can find* diagonal *projections $P_1, P_2, \ldots, P_k \in \mathfrak{M}_n$ such that*

$$\text{(i)} \quad \sum_{i=1}^{k} P_i = I \qquad \textit{and} \qquad \text{(ii)} \quad \Big\| \sum_{i=1}^{k} P_i A P_i \Big\| \leq \frac{1}{2}\|A\|?$$

This is clearly the kind of question that fits into our framework. To clarify the connection we will sketch the argument. Observe first that Problem I.18 is formally equivalent to the following:

Given $\epsilon > 0$, does there exist a positive integer k such that, for any matrix T in $\mathcal{L}(\ell_2)$, one can find diagonal *projections $P_1, P_2, \ldots, P_k$ such that $\sum_{i=1}^{k} P_i = I$ and*

$$\text{(ii}') \quad \Big\| \sum_{i=1}^{k} P_i(T - E(T))P_i \Big\| \equiv \Big\| \sum_{i=1}^{k} P_i T P_i - E(T) \Big\| \leq \varepsilon\|T\| \ ?$$

Indeed, one gets ε in place of $1/2$ in the condition (ii) via iteration, with k depending on ε. Passing from finite matrices with a uniform estimate on k to infinite matrices considered as operators on ℓ_2, one obtains the partition of $\mathbb{N}$ corresponding to the decomposition of the

identity on ℓ_2 into a sum of projections from finite partitions via a diagonal argument.

Let φ be an extension to $\mathcal{L}(\ell_2)$ of a pure state on $\mathcal{D}$. We claim that

$$\varphi(D_1TD_2) = \varphi(D_1)\varphi(T)\varphi(D_2) \quad \text{for} \quad D_1, D_2 \in \mathcal{D},\ T \in \mathcal{L}(\ell_2) \tag{1}$$

We note that *a priori* φ is multiplicative only on $\mathcal{D}$ as a point evaluation on $C(\beta\mathbb{N}) \simeq \mathcal{D}$. We postpone the proof of (1) and show first how it allows to complete the argument showing that an affirmative answer to Problem I.18 implies that $\varphi(T) = \varphi(E(T))$ for $T \in \mathcal{L}(\ell_2)$.

Let $\epsilon > 0$. Fix $T \in \mathcal{L}(\ell_2)$ with $\|T\| \leq 1$, and let $P_1, P_2, \ldots, P_k$ be the projections given by the affirmative answer to infinite variant of Problem I.18. We now identify $\mathcal{D}$ with ℓ_∞. $P_1, P_2, \ldots, P_k$ correspond then to a partition of unity in ℓ_∞ into a sum of k indicator functions of subsets. In this identification, $\varphi|_{\mathcal{D}}$ is obtained as a limit with respect to certain ultrafilter $\mathcal{U}$. It follows that among the numbers $\varphi(P_i)$, $i = 1, 2, \ldots, k$, exactly one is equal to 1 and the others are all 0, depending on whether the corresponding subset belongs to $\mathcal{U}$ or not. Accordingly, by (1),

$$\varphi\big(\sum_{i=1}^{k} P_iTP_i\big) = \sum_{i=1}^{k} \varphi(P_iTP_i) = \sum_{i=1}^{k} \varphi(P_i)\varphi(T)\varphi(P_i) = \varphi(T).$$

Therefore

$$\begin{aligned} \varepsilon &\geq \|\sum_{i=1}^{k} P_iTP_i - E(T)\| \\ &\geq \Big|\varphi\big(\sum_{i=1}^{k} P_iTP_i\big) - \varphi\big(E(T)\big)\Big| = \big|\varphi(T) - \varphi(E(T))\big| \end{aligned}$$

Since $\epsilon > 0$ was arbitrary, it follows that $\varphi(T) = \varphi(E(T))$, as required.

It remains to prove (1). To this end, consider the GNS representation of $(\mathcal{L}(\ell_2), \varphi)$. There is a Hilbert space $\mathcal{H}$, a norm one vector $x \in \mathcal{H}$ and a $*$-representation π of $\mathcal{L}(\ell_2)$ on $\mathcal{L}(\mathcal{H})$ such that for all T in $\mathcal{L}(\ell_2)$,

$$\varphi(T) = \langle \pi(T)x, x\rangle \ .$$

An elementary argument shows then that x must be an eigenvector for each $\pi(D)$, $D \in \mathcal{D}$ with an eigenvalue $\varphi(D)$. Hence, for any T in $\mathcal{L}(\ell_2)$,

$$\begin{aligned} \varphi(D_1TD_2) &= \langle \pi(D_1TD_2)x, x\rangle = \langle \pi(D_1)\pi(T)\pi(D_2)x, x\rangle \\ &= \langle \pi(T)\pi(D_2)x, \pi(D_1)^*x\rangle = \langle \pi(T)(\varphi(D_2)x), \overline{\varphi(D_1)}x\rangle \\ &= \varphi(D_2)\varphi(D_1)\langle \pi(T)x, x\rangle = \varphi(D_2)\varphi(D_1)\varphi(T). \end{aligned}$$

The converse, uniqueness implies paving [**4**, (3.6)], can be proved in a similar spirit. One way is to show first that paving is implied by the *relative Diximier property*, which says that for $T \in \mathcal{L}(\ell_2)$,

$$\mathcal{D}(T) := \overline{\text{conv}\{U^*TU : U \in \mathcal{D},\ U \text{ unitary}\}} \cap \mathcal{D} \neq \emptyset.$$

This immediately implies that $E(T)$ belongs to $\mathcal{D}(T)$. This is done in very much the same way as the argument presented above. On the other hand, if $E(T)$ is not in $\mathcal{D}(T)$, then $E(T)$ can be separated from $\mathcal{D}(T)$ by a functional φ. Heuristically, in view of the balanced nature of $\mathcal{D}(T)$, φ may be assumed to be positive and with more work, an extension of a pure state on $\mathcal{D}$. By construction, φ is different form the canonical extension obtained by composing with the conditional expectation E.

There has been a lot of work on Problem I.18 in the intervening years, mostly yielding partial and related results. In particular, it was shown by Berman, Halpern, Kaftal and Weiss [**26**], cf. [**37, 88**], that the answer is positive for matrices with nonnegative entries. A major progress was achieved by Bourgain and Tzafriri. Their work was motivated principally by applications to local structure of L_p-spaces, see [**88**] in this collection for more details. First, in [**36**], they showed that for Problem I.18, there exists a diagonal projection P with $\|PAP\| \leq \frac{1}{2}\|A\|$ and $\text{rank}P \geq \delta n$, where δ is a universal positive constant. A closely related result was obtained earlier by Kashin [**92**]. It then clearly follows by iteration that there is a decomposition of identity $P_1, P_2, \ldots, P_k$ verifying the conditions (i) and (ii) of Problem I.18 such that $k = O(\log n)$.

Then in [**38**], they obtained by far the strongest results to date. Problem I.18 is solved in the affirmative when the absolute values of entries of the matrix A are relatively small, specifically $O(1/(\log n)^{1+\eta})$ for some $\eta > 0$, (Theorem 2.3). Their solution also applies to the cases of Hankel and Laurent matrices with certain regularity properties But the major accomplishment is a saturation result that follows. As earlier, we identify ℓ_∞ with the algebra $\mathcal{D}$. Similarly, the power set $\mathcal{P}(\mathbb{N}) \equiv \{0,1\}^{\mathbb{N}}$ is thought of as a subset of ℓ_∞. In particular, $\sigma \subset \mathbb{N}$ is associated with the sequence $(\sigma_j) \in \ell_\infty$, the indicator function of σ, and with the diagonal projection P_σ in $\mathcal{D}$.

Theorem I.19 (Bourgain–Tzafriri)**.** *Given $\epsilon > 0$ and $T \in \mathcal{L}(\ell_2)$, one can find a positive measure ν supported on the w^*-compact set*

$$K = K(T, \epsilon) := \left\{\sigma \subset \mathbb{N} : \|P_\sigma(T - E(T))P_\sigma\| \leq \epsilon\|T\|\right\} \subset \ell_\infty$$

of total mass $\|\nu\| \le C\epsilon^{-2}$ *for which*

$$\int_K \sigma_j d\nu(\sigma) \ge 1 \quad \text{for all} \quad j \ge 1,$$

where C *is a universal numerical constant.*

For clarity, we state also the finite dimensional version of Theorem I.19 from which the Theorem easily follows by a diagonal argument.

Proposition I.20. *There is a universal constant* $c > 0$ *so that given* $\epsilon > 0$, $n \in \mathbb{N}$ *and a matrix* $A \in \mathfrak{M}_n$ *with zero diagonal, there exist a finite sequence of nonnegative weights* (t_i) *with* $\sum_i t_i = 1$ *and diagonal projections* (P_i) *such that*

$$\|P_i A P_i\| \le \epsilon \|A\| \quad \text{for all} \quad i \ge 1 \qquad \text{and} \qquad \sum_i t_i P_i \ge c\epsilon^2 I.$$

Let us point out first that the last condition in this Proposition implies by trace evaluation that the rank of at least one of the P_i's is at least $c\epsilon^2 n$, thus recovering the result from **[36]** mentioned above with the best possible dependence on ϵ. Note also that except for that optimal dependence, it is enough to prove Proposition I.20 for some $\epsilon \in (0, 1)$.

We wish to emphasize that Theorem I.19 is indeed very close to implying the affirmative answer to the Kadison–Singer problem. Indeed, if we somewhat carelessly apply a (pure) state φ to both sides of the assertion of Theorem I.19, we "obtain"

$$\varphi\Big(\int_K \sigma d\nu\Big) = \int_K \varphi(\sigma) d\nu \ge 1. \tag{2}$$

Hence there is a $\sigma \in K$ such that $\varphi(\sigma) = \varphi(P_\sigma) > 0$. Thus $\varphi(P_\sigma) = 1$ because φ is multiplicative on $\mathcal{D}$ and P_σ is idempotent. We now proceed as in the derivation of the original Kadison–Singer problem from the paving problem. As $\sigma \in K$, one has $\|P_\sigma(T - E(T))P_\sigma\| \le \epsilon\|T\|$ and so

$$|\varphi(T - E(T))| = |\varphi(P_\sigma(T - E(T))P_\sigma)| \le \epsilon\|T\|$$

for all $\varepsilon > 0$. Therefore $\varphi(T) = \varphi(E(T))$, whence $\varphi(T)$ is determined by its diagonal part.

The weak point of this "argument" lies in the fact that the integral $\int_K \sigma d\nu(\sigma)$ makes only weak* sense, while in the equality (2) we implicitly used weak convergence. The argument would work if the measure ν was atomic, though. Still, Theorem I.19 shows that there is an abundance of diagonal projections P_σ verifying $\|P_\sigma(T - E(T))P_\sigma\| \le \varepsilon\|T\|$. That abundance just isn't formally strong enough to guarantee that

the collection of such σ's will intersect every ultrafilter. We show now a simple example to that effect.

Let $d \in \mathbb{N}$ and let I_d be the set of all words of length $2d$ in the alphabet $\{A, B\}$ consisting of d A's and d B's. So $n := \#I_d = (2d)!/(d!)^2)$. Next, for $s = 1, 2, \ldots, 2d$, let σ_s be the set of those words in I_d whose sth letter is A. Clearly, $\Sigma_d := \{\sigma_1\, \sigma_2, \ldots, \sigma_{2d}\}$ provides a cover of I_d. It is not a minimal cover, but $d + 1$ sets are required.

Let Σ'_d be the hereditary subset of the power set $\mathcal{P}(I_d)$ generated by Σ_d, namely

$$\Sigma'_d := \{\sigma : \sigma \subset \sigma_s \text{ for some } 1 \leq s \leq 2d\}.$$

Then every subcover, or partition, of I_d consisting of elements of Σ'_d must have at least $d + 1 = O(\log n)$ elements. On the other hand, it is easily seen that $\frac{1}{2d} \sum_{s=1}^{2d} \chi_{\sigma_s} = \frac{1}{2}$. Thus the set of projections $\{P_\sigma : \sigma \in \Sigma'_d\}$ verifies the condition in the assertion of Proposition I.20, but does not verify the condition of Problem I.18 with k independent of n. An infinite example verifying the condition of Theorem I.19, but not that of the infinite variant of Problem I.18 is routinely obtained by identifying $\mathbb{N}$ with $\bigcup_d I_d$ and setting

$$\Sigma := \{\sigma \subset \mathbb{N} : \sigma \cap I_d \in \Sigma'_d \quad \text{for all} \quad d \geq 1\}.$$

Of course, this is just a formal example. There is no *a priori* reason why Σ produced above would correspond to an actual operator $T \in \mathcal{L}(\ell_2)$.

The methods of **[36, 37, 38]** are quite sophisticated. Without going into details, we mention that the first step in finding large subset $\sigma \subset \{1, 2, \ldots, n\}$ for which $\|P_\sigma A P_\sigma\|$ is small involves a random procedure. The first approximation is $\sigma = \{j : \xi_j = 1\}$, where $\xi_1, \xi_2, \ldots, \xi_n$ is a sequence of independent Bernoulli selectors. These are random variables satisfying $\mathbf{P}(\xi_j = 1) = 1 - \mathbf{P}(\xi_j = 0) = \delta$ for $j = 1, 2, \ldots, n$ for some properly chosen $\delta \in (0, 1)$. The norm of $P_\sigma A P_\sigma$ is then the maximum of a random process, which is analyzed using decoupling inequalities (see, e.g., **[95]**), majorized by a more manageable maximum of a Gaussian process, and estimated via metric entropy and Dudley's majoration (see **[99**, (12.2)]). This works if the entries of the matrix A are rather small, such as the $O(1/(\log n)^{1+\eta})$ bound mentioned earlier. Such random subsets also yield a partition of $\{1, 2, \ldots, n\}$. In the general case, one only obtains an estimate on the norm of $P_\sigma A P_\sigma$ as a map from ℓ_2^n to $\ell_1^n\|$. In the case of Proposition I.20, one has instead a weighted ℓ_1^n. The final step uses the Little Grothendieck Theorem **[121]**. Alternatively, some of the steps may be done by using the measure concentration phenomena (cf. section IIb) already employed in

[**93**] or majorizing measures, cf. [**144**]. See [**88**] in this collection for more details on some of the above arguments and related issues.

We conclude this section by commenting on the type of examples that need to be analyzed in hope of further progress. To be a potential counterexample, a matrix $A = (a_{ij})_{i,j=1}^n$ (meaning a sequence of $n \times n$ matrices that together provide a counterexample) must have the following features:
(i) $\|A\|$ must be much smaller than $\|(|a_{ij}|)\|$, or otherwise we could apply the argument that works for nonnegative entries.
(ii) $|a_{ij}|$ do not admit a (uniform) $O(1/(\log n)^{1+\eta})$ bound.
(iii) On the other hand, the substantial entries of A must be abundant, or otherwise one could avoid them by the same combinatorial argument that works for nonnegative entries.
(iv) The combinatorial structure of the substantial part of A must be quite rigid to distinguish between the conditions from Proposition I.20 and Problem I.18.

One structure that comes to mind is related to incidence matrices of Ramanujan graphs (see e.g. [**103**]). Let $B = (b_{ij})$ be such a matrix corresponding to a d-regular graph on n vertices. Then $\|B\| = d$, and it is achieved on the eigenvector $(1, 1, \ldots, 1)$, while all the remaining eigenvalues are bounded by $2\sqrt{d-1}$. So the $n \times n$ matrix $(\frac{1}{2}(b_{ij} - k/n)/\sqrt{d-1})$ is of norm at most 1, and appears to enjoy the features (i)-(iv), some of which are admittedly vague. The question would then be to determine whether matrices obtained this way from various constructions of Ramanujan graphs can be paved. It seems at the first sight that new techniques are required for any kind of answer.

Ie. **Hyper-reflexivity.** If $\mathcal{A}$ is an operator algebra contained in $\mathcal{L}(\mathcal{H})$, then $\operatorname{Lat}\mathcal{A}$ denotes the lattice of all of its invariant subspaces. Dually, given a collection $\mathcal{L}$ of subspaces, $\operatorname{Alg}\mathcal{L}$ denotes the algebra of all operators leaving each element of $\mathcal{L}$ invariant. The algebra $\mathcal{A}$ is *reflexive* if $\operatorname{Alg}\operatorname{Lat}\mathcal{A} = \mathcal{A}$.

There is a quantitative version of reflexivity which has proven to be a powerful tool when it is available. Notice that if P_L is the projection onto an invariant subspace L of $\mathcal{A}$, then for any operators $T \in \mathcal{L}(\mathcal{H})$ and $A \in \mathcal{A}$,

$$\|P_L^\perp T P_L\| = \|P_L^\perp (T - A) P_L\| \le \|T - A\|.$$

Hence the inequality

$$\beta_{\mathcal{L}}(T) := \sup_{L \in \mathcal{L}} \|P_L^\perp T P_L\| \le \operatorname{dist}(T, \mathcal{A}).$$

The algebra $\mathcal{A}$ is *hyper-reflexive* if there is a constant C such that

$$\operatorname{dist}(T, \mathcal{A}) \leq C\beta_{\mathcal{L}}(T) \quad \text{for all} \quad T \in \mathcal{B}(\mathcal{H}).$$

We also say that an operator A is reflexive or hyper-reflexive if the unital WOT-closed algebra $W(A)$ generated by A is reflexive or hyper-reflexive respectively.

In finite dimensions, the question of which operators are reflexive was solved by Deddens and Fillmore [**63**] in terms of the Jordan form. Evidently, every reflexive subalgebra of $\mathfrak{M}_n$ is also hyper-reflexive; but there is no a priori estimate of the constant even in two dimensions. For example, it is very easy to show that the 2×2 diagonal matrices $\mathcal{D}_2$ are hyper-reflexive with constant 1. Similarity by an operator S will preserve hyper-reflexivity, but can change the constant by as much as the condition number $\|S\| \, \|S^{-1}\|$. Thus it may not be too surprising that the hyper-reflexivity constant of $S_t \mathcal{D}_2 S_t^{-1}$ increases to $+\infty$ with t when $S_t = \left[\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right]$.

There are two classical situations in which the existence of hyper-reflexivity has played a key role. The first is nest algebras. A nest $\mathcal{N}$ is a chain of subspaces containing $\{0\}$ and $\mathcal{H}$ which is closed under intersections and closed spans. The nest algebra is $\mathcal{T}(\mathcal{N}) = \operatorname{Alg}\mathcal{N}$ is the algebra of all operators with an *upper-triangular form* with respect to this chain. In particular, the algebra $\mathcal{T}_n$ of $n \times n$ upper triangular matrices with respect so some basis is the prototypical finite dimensional example. An early result of Ringrose [**129**] shows that $\mathcal{T}(\mathcal{N})$ is reflexive, as well as the fact that $\operatorname{Lat}\mathcal{T}(\mathcal{N}) = \mathcal{N}$. Arveson [**13**] showed that nest algebras are hyper-reflexive with constant 1, so that one obtains the distance formula:

Theorem I.21 (Arveson)**.** *Let $\mathcal{N}$ be a nest, and let A be an arbitrary operator in $\mathcal{B}(\mathcal{H})$. Then*

$$\operatorname{dist}(A, \mathcal{T}(\mathcal{N})) = \sup_{N \in \mathcal{N}} \|P_N^{\perp} A P_N\|.$$

An easy proof [**128**], cf. [**52**], can be based on the well-known matrix filling lemma of Parrott [**116**] and Davis–Kahane-Weinberger [**62**] valid for matrix or operator entries:

$$\inf_{X \in \mathcal{L}(\mathcal{H})} \left\| \begin{bmatrix} A & B \\ X & D \end{bmatrix} \right\| = \max \left\{ \left\| \begin{bmatrix} A & B \end{bmatrix} \right\|, \left\| \begin{bmatrix} B \\ D \end{bmatrix} \right\| \right\}.$$

In particular, Lance [**97**] used this distance formula to establish that close nests are similar. Likewise, this fact was a key ingredient in the first author's Similarity Theorem [**48**] stating that two nest algebras are similar if and only if there is a dimension preserving isomorphism

between the underlying nests. The search for generalizations of this will lead to an open question below.

The second context in which this concept occurs is the von Neumann algebras. Christensen [**46**] showed that many von Neumann algebras (specifically all whose commutant do not have certain type II_1 summands) are hyper-reflexive. This is closely related to derivations. Recall that a *derivation* is a map δ of an algebra $\mathcal{A}$ into an $\mathcal{A}$-bimodule $\mathcal{M}$ satisfying the product rule $\delta(ab) = a\delta(b) + \delta(a)b$. The derivation is *inner* if there is an element $x \in \mathcal{M}$ such that $\delta(a) = ax - xa$. Christensen showed:

Theorem I.22. *For a von Neumann algebra $\mathfrak{N}$ contained in $\mathcal{L}(\mathcal{H})$, the following are equivalent:*

(i) *Every derivation δ of $\mathfrak{N}'$ into $\mathcal{L}(\mathcal{H})$ is inner.*
(ii) *There is a constant C such that* $\operatorname{dist}(X, \mathfrak{N}) \leq C\|\delta_X|_{\mathfrak{N}'}\|$ *for all $X \in \mathcal{L}(\mathcal{H})$.*

The second condition actually asserts that $\mathfrak{N}$ is hyper-reflexive. The invariant subspaces of any C*-algebra correspond to projections in the commutant. Every von Neumann algebra is spanned by its projections. Thus von Neumann's famous double commutant theorem, which asserts that $(\mathfrak{N}')' = \mathfrak{N}$, shows that every von Neumann algebra is reflexive. Moreover, the convex hull of the symmetries $2P - I$, for P projections in $\mathcal{P}(\mathfrak{N})$, is the whole unit ball of the self-adjoint part of $\mathfrak{N}$. Now a simple calculation shows that

$$\begin{aligned}\|(2P - I)X - X(2P - I)\| &= 2\|PXP^\perp - P^\perp XP\| \\ &= 2\max\{\|PXP^\perp\|, \|P^\perp XP\|\}.\end{aligned}$$

Thus since every operator is the sum of its real and imaginary parts,

$$\|\delta_X|_{\mathfrak{N}'}\| \leq 4 \sup_{P \in \mathcal{P}(\mathfrak{N})} \|P^\perp XP\|.$$

So (ii) is equivalent to hyper-reflexivity.

For all injective von Neumann algebras, the constant is at most 4 [**46**]. And for abelian von Neumann algebras, the constant is no more than 2 [**130**]. Even for the 3×3 diagonal matrices, the constant is greater than 1. Indeed, it is exactly $\sqrt{3/2}$ [**57**]. So unlike the nest case, most examples involve non-trivial constants, not exact formulae.

Now we turn our consideration to the class of WOT-closed reflexive algebras which contain a masa, known as CSL algebras. The terminology is short for *commutative subspace lattice algebra* because any invariant subspace is, in particular, invariant for the masa, and therefore corresponds to a projection in the masa. So the orthogonal projections

onto all of these invariant subspaces commute with each other. These algebras were introduced in a seminal paper by Arveson [**14**].

The finite dimensional versions occur in many contexts, and are also called incidence algebras or digraph algebras in other parts of the literature. A masa in $\mathfrak{M}_n$ is unitarily equivalent to the diagonal algebra $\mathcal{D}_n$. Any algebra containing $\mathcal{D}_n$ is determined by the standard matrix units E_{ij} which belong to the algebra. Moreover, since this is an algebra, the set of matrix units is transitive in the sense that if E_{ij} and E_{jk} belong, then so does E_{ik}. Thus one may associate a directed graph to the algebra with n vertices representing the standard basis, and including a directed edge from node j to node i if E_{ij} lies in the algebra. In this way, we obtain the graph of a transitive relation.

Not all CSL algebras are hyper-reflexive. Davidson and Power [**61**] constructed finite dimensional examples with arbitrarily large distance constants. This is more subtle than the easy example mentioned earlier. However, Larson [**98**] showed that this construction was part of a general mechanism for producing examples of this type. On a positive note, Davidson and Pitts [**59**] showed that if there is a dimension preserving lattice isomorphism between two CSL algebras, then the two lattices are approximately unitarily equivalent. This was established for nests by Andersen [**3**], and was a key step in obtaining similarity invariants. Pitts [**126**] was able to obtain good perturbations results for the algebras without using hyper-reflexivity in spite of the fact that hyper-reflexivity was used in an important way in the nest case. But other possible extensions of results for nests to this more general context have been hampered by the lack of hyper-reflexivity.

For this reason, attention has been focussed on an important finite dimensional case (see [**52, 53**]). The algebra $\mathcal{A}_n = \mathcal{T}_n \otimes \mathcal{T}_n$ represents the subalgebra of $\mathfrak{M}_{n^2}$ consisting of $n \times n$ upper-triangular matrices with $n \times n$ upper-triangular matrix entries. The problem becomes:

Problem I.23. *Is there is a finite upper bound to the hyper-reflexivity constants of all of the algebras $\mathcal{A}_n$?*

This is a typical situation where one understands well the one dimensional case, but not the higher dimensional or multivariable case. Similar difficulty appears when studying, e.g., multi-indexed orthogonal expansions.

A few other examples of hyper-reflexivity are known and are worth mentioning. The unilateral shift generates the analytic Toeplitz algebra, and the first author [**51**] showed that it is hyper-reflexive. Then

with Pitts [**60**], he showed that the algebra generated by the left regular representation of the free semigroup on n letters is also hyper-reflexive. Popescu [**127**] generalized this to a wider family of semigroups. Recently, Bercovici [**22**] used predual techniques to establish hyper-reflexivity constants for the large class of algebras with property $\mathfrak{X}_{\theta,\gamma}$. This property, which we do not define precisely, allows approximation of any weak-$*$ continuous functional on the algebra by a sequence of rank one functionals tending to infinity in a very strong sense. These notions arose in using predual techniques to establish reflexivity for single operators beginning with the celebrated theorem of Scott Brown [**43**] that every subnormal operator has invariant subspaces. Olin and Thomson [**115**] showed that subnormal operators are reflexive. And we mentioned above that the prototypical subnormal operator is hyper-reflexive. Many subnormal operators fit into Bercovici's criterion, but the general question of hyper-reflexivity for subnormal operators remains open.

II. Random matrices

In this chapter, we shall present a selection of classical and not-so-classical results on asymptotic spectral properties of random matrices that are related one way or the other to the geometry of Banach spaces. As was the case with other topics, these connections come in several flavors. First, facts about random matrices are being applied to the geometry of Banach spaces. Second, the methods of the Banach spaces yield new results or offer an alternative perspective on random matrices. Finally, there are connections via vague analogies between the fields.

IIa. **The overview.** Random matrices appeared in Banach space theory in an explicit way some time in the '70s, for example in [**20**], [**21**] and [**91**], even though one can claim that their spirit was already present e.g. in probabilistic proofs of Dvoretzky theorem, cf. [**74**] in this collection. Let us quote here a result from [**20**]:

Theorem II.1. *Let $q \in [2,\infty)$ and for some $m,n \in \mathbb{N}$, let $A = (a_{ij})$ be an $m \times n$ random matrix whose entries $a_{ij} = a_{ij}(\omega)$ are independent, mean-zero real random variables with $|a_{ij}| \le 1$ for all i,j. Then*

$$\mathbf{E}\,\|A \,:\, \ell_2^n \to \ell_q^m\| \le K \max\{m^{1/q}, n^{1/2}\},$$

where $\mathbf{E}$ denotes the expectation and $K = K(q)$ is a numerical constant depending only on q.

This is a very typical statement as far as the asymptotic theory of finite dimensional Banach spaces is concerned. We have an estimate

giving the correct *rough* asymptotic order (see the introduction for terminology) and involving a universal numerical constant. Usually we do not know, and often do not really care about, the optimal value of that constant, i.e., the *precise* asymptotic order. This could be viewed as an unsatisfactory situation from the point of view of other related fields. We shall return to this issue later on. On the other hand, in spite of the "asymptotic" qualification, we do have above an inequality valid for any m, n, a crucial detail for applicability to fields like geometry of Banach spaces, computational complexity and approximation theory.

On the other hand, random matrices were of interest to statisticians at least since the '20s, and to theoretical physicists at least since the '50s, see **[108]**. Perhaps the most celebrated result coming from the latter area is the *Wigner's Semicircle Law* **[162, 163]** which says that, under some weak regularity assumptions, the spectra of large symmetric random matrices are, in a sense, virtually deterministic and their spectral densities are, after proper normalization, asymptotically semicircular. More precisely, Wigner proved a somewhat weaker statement:

Theorem II.2. *For each $n \in \mathbb{N}$, let $A = A^{(n)}(\omega)$ be an $n \times n$ random matrix whose entries $a_{ij} = a_{ij}^{(n)}(\omega) : \Omega \to \mathbb{R}$ are symmetric real random variables satisfying for each $n \geq 1$ and $1 \leq i, j \leq n$,*

(i) $\mathbf{E}|a_{ij}|^2 = 1/n$.
(ii) $a_{ij}^{(n)}$ *are independent for* $1 \leq i \leq j \leq n$.
(iii) $a_{ij}^{(n)} = a_{ji}^{(n)}$.
(iv) *For each $m \geq 1$, there is a $\beta_m < \infty$ independent of n so that* $\mathbf{E}|a_{ij}|^m \leq \beta_m$.

Then, for any $\nu \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathbf{E}\ \mathrm{Tr}(A^\nu) = \int_{\mathbb{R}} x^\nu \, d\mu(x), \tag{3}$$

where Tr *is the normalized trace in the respective dimension and μ is the measure on $\mathbb{R}$ whose density w is given by*

$$w(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2} & \text{if } x \in [-2,2] \\ 0 & \text{if } x \notin [-2,2] \end{cases}$$

The measure μ above is often referred to as the *standard semicircular distribution*. Theorem II.2 says in effect that, for large n, the random measure

$$\mu_n = \mu_n(\omega) := \frac{1}{n}\sum_{j=1}^{n} \delta_{\lambda_j(A)},$$

where δ_x is the Dirac measure at x and $\lambda_1(B) \geq \cdots \geq \lambda_n(B)$ are eigenvalues of $B \in \mathfrak{M}_n^{sa}$, is approximately standard semicircular when considered as a measure on $\Omega \times \mathbb{R}$. In fact, as was implicitly assumed by physicists and proved later (cf. **[12]**), a much stronger statement holds: for large n the random measures $\mu_n(\omega)$ are "close" to μ with probability close to 1. In particular, $\mu_n(\omega)$, the empirical measure associated to the spectrum of $A^{(n)}$, is nearly deterministic. We emphasize that this is far from being a formal consequence of Theorem II.2. It is possible in principle that the assertion of the theorem holds while, for any $\omega \in \Omega$, $A(\omega)$ is a multiple of identity: just consider a scalar random variable ξ whose law is standard semicircular and a random matrix $A := \xi I$. However, we do have

Theorem II.3. *In the notation and under the hypotheses of Theorem* II.2 *one has, for any interval* $\mathcal{I} \subset \mathbb{R}$,

$$\lim_{n\to\infty} \frac{\#\{\textit{eigenvalues of } A^{(n)} \textit{contained in } \mathcal{I}\}}{n} = \mu(\mathcal{I}) \quad \textit{almost surely} .$$

There are similar results for the asymptotic distribution of singular values of random non-selfadjoint matrices; e.g., when the hypothesis (ii) in Theorem II.2 is replaced by (ii) $a_{ij}^{(n)}$ *are independent for* $1 \leq i, j \leq n$ and the hypothesis (iii) is dropped. The limiting distribution is then a *quartercircle law* (note that singular values are necessarily nonnegative), supported in the interval $[0, 2]$ and given by the density $\frac{1}{\pi}\sqrt{4 - x^2}$. Similarly, one may consider (see **[105]**, **[157]**) large rectangular matrices such that the ratio of the sides is roughly fixed, cf. Theorem II.13, and many other *ensembles*.

The example preceding Theorem II.3 notwithstanding, the implication Theorem II.2 $\Rightarrow$ Theorem II.3 is in fact a formal consequence of rather standard local theory techniques, even though this was not the way how it was historically shown. We shall sketch the argument later in this chapter, after Theorem II.7; see also Corollary II.12 and Theorem II.17. As a bonus, one obtains fairly strong estimates on the probabilities implicit in the almost sure part of Theorem II.3. That proof, and many other arguments presented here, while being a folklore among Banach space theorists, are not exactly a common knowledge in the random matrix circles. As opposed to the techniques emphasized here, standard tools used in that area involved the *moment method* (roughly, working directly – via a heavy duty combinatorics – with the moments involved in the limit in (3), cf. the remarks following Problem II.18), the more precise *Stieltjes transform method* introduced

in **[105]** and later adopted in and developed by the free probability approach, or, in the case of classical random matrices, analyzing the explicit *formulae for joint densities* of eigenvalues or singular values, see (8) below.

In other directions, it has been determined that much weaker assumptions on regularity of the entries suffice: one needs just a little bit more than the existence of second moments, with uniform estimates (more precisely, a Lindeberg type condition), and the symmetry hypothesis may be replaced by $\mathbf{E}a_{ij} = 0$ (see, e.g, **[117, 75]**).

We point out that assertions of Theorems II.1 and II.3 are not comparable. On one hand, Theorem II.3 doesn't say anything about the norms of matrices. It is consistent with its assertion that $o(n)$ largest (in absolute value) eigenvalues are far outside the interval $[-2, 2]$, the support of w. On the other hand, as indicated earlier, Theorem II.1, being "isomorphic" in nature, doesn't give the precise order of the norm (even when $q = 2$ and $m = n$), and it doesn't address the question of the asymptotic distribution of the eigenvalues (or singular values in the non-selfadjoint case). Again, common strengthenings of results of the two kinds have been obtained (see, e.g., **[40, 73, 16]**), but we do not know of any *really* satisfactory argument that encompasses simultaneously the two aspects of the picture.

In the opposite direction, precise estimates were obtained on the (very small) probability that a specific eigenvalue (or singular value) of a random Gaussian matrix is far away from its theoretical value predicted by the corresponding Semicircle Law result. A sample such large deviation result, motivated by questions in geometry of Banach spaces and numerical analysis, is Theorem II.4 below.

Before stating the theorem, a few words about the level of generality of our discussion. More often than not, we shall concentrate on the central Gaussian case, and most of the arguments will be specific to that case. In particular, throughout this chapter $G = G^{(n)}$ will stand for an $n \times n$ random matrix whose entries g_{ij}, $1 \le i \ne j \le n$, are independent identically distributed Gaussian random variables following the $N(0, 1/n)$ law. Similarly, $A = A^{(n)}$ will usually stand for a Gaussian selfadjoint matrix verifying the hypotheses of Theorem II.2. However, both the results and the methods employed are representative of much more general setting. It would be a useful project to work out in detail the consequences of general concentration results for product measures (see **[145, 146]** and their references) and related "probability in Banach spaces" tools to the setting of random matrices. As indicated earlier, very few papers produced in the random matrix circles

employ those methods and, to our knowledge, no vigorous research in this direction was attempted. It is clear that some consequences would be just a formal repetition of the arguments sketched in this section. Chances are that, to get best results, one may need to apply those tools rather creatively. The payoff can be substantial since important part of research in random matrices turns around "universality" of the limit laws, i.e., their independence of the particular probabilistic model involved, the "symmetries" of that model being the only meaningful parameter.

Theorem II.4. ([**141, 142**]) *Let $G = G^{(n)}$ be as above and $s_1 \geq s_2 \geq \cdots \geq s_n$ be its singular values. Then*

$$\mathcal{P}\left(s_{n-d+1} \geq \beta\frac{d}{n}\right) \leq \exp\left(-c\beta^2 d^2\right), \quad \mathcal{P}\left(s_{n-d+1} \leq \alpha\frac{d}{n}\right) \leq (C\alpha)^{d^2}$$

for $1 \leq d \leq n$, $\beta \geq \beta_0$ and $\alpha \geq 0$. Above, c, C and β_0 are universal positive constants. Apart from the precise values of the constants, the estimates are optimal except possibly for the for the first one when $n - d = o(n)$.

Analogous results can be obtained in the complex case and, for eigenvalues, in the selfadjoint case. The optimality of the Theorem means here that there are similar lower estimates with c, C replaced by other positive universal constants. However, being an isomorphic result, the Theorem doesn't detect smaller deviations from the values predicted by the Quartercircle or Semicircle Law. Much more precise results of the same nature were obtained very recently in the case when d is asymptotically a fixed proportion of n, [**19**]. One of the concepts employed there is the *free* (noncommutative) *entropy* [**156**], arrived at by following on the ideas sketched in section IId below. It is quite likely that the methods generalize to the case $d = o(n)$, but probably not to the edge of the of the spectrum, i.e., the case when d is nearly equal to n; cf. Theorem II.8 below and the comments following it. It would be potentially useful to clarify the relationship between Theorem II.4 and these results.

To complete the overview, we shall mention, besides the large deviation results mentioned above and the results in the spirit of Wigner's Semicircle Law, that there is a large body of research dealing with microlocal analysis of the spectrum $\{\lambda_1(A(\omega)), \ldots, \lambda_n(A(\omega))\}$, and particularly the gaps $\lambda_j - \lambda_{j+1}$ in that spectrum. These are natural objects to consider in view of the original physical motivation, the eigenvalues being interpreted there as energy levels. Further attention to this direction was brought by the largely experimental results [**114, 70**] relating

the properly defined distribution of such gaps and the corresponding gaps between zeros of the Riemann ζ function; cf. [**113, 132**]). In what follows, we shall state a result in the spirit of [**109**], the fundamental work in this direction. It deals with a complex analogue of $A^{(n)}$ called the Gaussian unitary ensemble or GUE (defined more precisely in item (iii) the next section), for which Theorems II.2 and II.3 hold with the same limiting semicircle distribution.

We need to introduce first some notation: given $B \in \mathfrak{M}_n^{sa}$ with eigenvalues $\lambda_1(B) \geq \cdots \geq \lambda_n(B)$ and $x \in \mathbb{R}$, we define $\delta_B(x)$ to be the length of the interval $(\lambda_{j-1}(B), \lambda_j(B)]$ containing x, with the convention that $\delta_B(x) = \infty$ if $x > \lambda_1(B)$ or $x \leq \lambda_n(B)$ and $\delta_B(x) = 0$ if x is a multiple eigenvalue.

Theorem II.5. *There exists a probability density σ supported on $[0, \infty)$ such that whenever $x \in (-2, 2)$ and $w(x) = \frac{1}{2\pi}\sqrt{4 - x^2}$, the density of the standard semicircle distribution, then, for any $s \geq 0$,*

$$\lim_{n\to\infty} \mathcal{P}\left(\delta_{GUE}(x) \leq \frac{s}{n\, w(x)}\right) = \int_0^s \sigma(u)\, du.$$

We point out that the scaling of $\delta_{GUE}(x)$ by the quantity $nw(x)$ in Theorem II.5 was to be expected: by Theorem II.3, for large n and a short interval $\mathcal{I}$ containing x, the number of eigenvalues of the random matrix that belong to $\mathcal{I}$ will be, with probability close to 1, approximately $n \int_{\mathcal{I}} w(u)\, du \approx n\, w(x)|\mathcal{I}|$, and so the gaps should average about $\frac{1}{n\, w(x)}$.

IIb. **Concentration of measure and its consequences.** In this section we shall sketch some applications of the *measure concentration* phenomenon, well known and widely applied in local theory of Banach spaces (see the article [**134**] in this collection), to the subject of random matrices. Most of these applications have been a folklore among some of the experts in the former field, but to the best of our knowledge, they haven't been presented anywhere in a systematic fashion. One of possible starting points is the Gaussian isoperimetric inequality [**33**], which we shall present here in the functional form (see [**120**] for the last assertion). Recall that a function F defined on a metric space (X, d) is called *Lipschitz with constant* L if $|F(x) - F(x')| \leq L\, d(x, x')$ for all $x, x' \in X$. Let $\gamma = \gamma_n$ be the standard Gaussian measure on $\mathbb{R}^n$ with density $(2\pi)^{-n/2}\, e^{-|x|^2/2}$, where $|\cdot|$ is the usual Euclidean norm. As usual, $\Phi(t) := \gamma_1((-\infty, t])$ is the cumulative distribution function of the $N(0, 1)$ Gaussian random variable.

Theorem II.6. *Let F be a function on $\mathbb{R}^n$ which is Lipschitz with constant L with respect to the Euclidean metric and let $M = M_F$ be the median value of F with respect to γ_n. Then, for any $t > 0$,*

$$\mathcal{P}\left(F \geq M + t\right) \leq 1 - \Phi(t) < \exp\left(-t^2/2L^2\right). \tag{4}$$

One has the same upper estimate $\exp\left(-t^2/2L^2\right)$ *if the median M is replaced by the expected value $\int_{\mathbb{R}^n} F \, d\gamma_n$.*

For future reference we point out here that, for a convex function, its median with respect to a Gaussian measure does not exceed the expected value; see **[64]**, **[96]** or Cor. 1.7.3 in **[69]**.

The relevance of Theorem II.6 to random (Gaussian) matrices is based on the elementary and well-known fact, described in section Ib of this article, that spectral parameters like singular values (resp. eigenvalues in the selfadjoint or unitary case) are Lipschitz functions with respect to the matrix elements; see **[30, 29]** for related more general results. In particular, for each $k \in \{1, \ldots, n\}$, the k-th largest singular value $s_k(X)$ (resp. the eigenvalue $\lambda_k(X)$) is Lipschitz with constant 1 if $X = (x_{jk})_{j,k=1}^n$ is considered as an element of the Euclidean space $\mathbb{R}^{n^2}$ (resp. the submanifold of $\mathbb{R}^{n^2}$ corresponding to the matrices). If one insists on thinking of X as a matrix, this corresponds to considering the underlying Hilbert-Schmidt metric. Accordingly, in the context of applying Theorem II.6, the Lipschitz constant of $s_k(G^{(n)})$ is $1/\sqrt{n}$ (because of the variances of the entries being $1/n$, which corresponds to the identification $G = G^{(n)} = 1/\sqrt{n}\, X$). Respectively, for a Gaussian selfadjoint matrix $A = A^{(n)}$ verifying the hypotheses of Theorem II.2, the Lipschitz constant of $\lambda_k(A^{(n)})$ is $\sqrt{2/n}$. The additional 2 is a consequence of the same variable appearing twice, in the jk-th and kj-th position. The above comments, and hence the results below, carry essentially word for word to the following often considered variants, all Gaussian unless explicitly stated otherwise.

(i) A variant of $A = A^{(n)}$ in which variances of the diagonal entries are assumed to be $2/n$ rather than $1/n$, called often the Gaussian orthogonal ensemble or GOE. This is in fact the same as $\sqrt{2}$ times the real part of $G^{(n)}$. Note for future reference that GOE can be represented as $Y + A$, where Y is a (diagonal) Gaussian random matrix independent of A, and so many results for GOE transfer formally to A.

(ii) The complex non-selfadjoint case, all the entries being independent and of the form $g + ig'$, where g, g' are independent *real* $N(0, 1/2n)$ Gaussian random variables.

(iii) The complex selfadjoint case: formally the same conditions as in Theorem II.2 (except for the obvious modification in the symmetry condition), but the above-diagonal entries are as in (ii) while the diagonal entries are real $N(0, 1/n)$'s. Again, this is the real part of the matrix in (ii) times $\sqrt{2}$, and is frequently referred to as the Gaussian unitary ensemble or GUE.

(iv) Rectangular, real anti-symmetric or complex anti-selfadjoint matrices.

(v) Orthogonal or unitary matrices distributed uniformly on $SO(n)$, resp. $U(n)$.

It is also easily seen that in the first three cases above the Lipschitz constants are respectively $\sqrt{2/n}$, $1/\sqrt{2n}$ and $1/\sqrt{n}$. The anti-symmetric/anti-selfadjoint/rectangular cases are treated the same way, and there are equally useful results for orthogonal/unitary matrices cf. **[111, 79]**. There appears to be no easily available exposition of the unitary case, even though all ingredients are available, cf. **[154]**. Still, a word of caution is needed. As noted in section Ia, eigenvalues are not very regular functions of general (non-normal) matrices.

Combining the above remarks and Theorem II.6 we get

Theorem II.7. *Given $n \in \mathbb{N}$, there exist positive scalars $s_1, s_2, \ldots, s_n$ such that the singular values of the $n \times n$ real Gaussian random matrix $G = G^{(n)}$ satisfy*

$$\mathcal{P}\left(|s_k(G) - s_k| \geq t\right) < 2\exp\left(-nt^2/2\right) \tag{5}$$

for all $t \geq 0$ and $k = 1, 2, \ldots, n$. This holds, in particular, if s_k's are the medians or the expected values of $s_k(G)$.

Similar results hold for the eigenvalues of the real symmetric Gaussian random matrix $A = A^{(n)}$ with $-nt^2/4$ in the exponent on the right side of (5)*. Moreover, the corresponding deterministic sequence $\lambda_1, \lambda_2, \ldots, \lambda_n$ can be assumed to be symmetric, i.e., $\lambda_k = -\lambda_{n-k+1}$ for $k = 1, 2, \ldots, n$. Likewise, related results hold for other Gaussian random matrices, in particular those described in* (i)-(iii) *above.*

A shortcoming of the above result is that it doesn't see the possible relationships between the deterministic sequences $s_1, s_2, \ldots, s_n$ (or $\lambda_1, \lambda_2, \ldots, \lambda_n$) for different n's. Still, it allows to formally deduce statements in the spirit of Theorem II.3 from the corresponding Theorem II.2 like results. Consider, as an illustration, the ensemble $A = A^{(n)}$. Once we know that there is a rough estimate on, say, $\mathbf{E}\|A\|$ (e.g., of the type of Theorem II.1), the analogue of (5) for A implies via an

elementary calculation

$$\left|\mathbf{E}\ \mathrm{Tr}(A^\nu) - \int_{\mathbb{R}} x^\nu \, d\mu^{(n)}(x)\right| \leq \frac{C_\nu}{\sqrt{n}} \tag{6}$$

for $\nu \in \mathbb{N}$, where $\mu^{(n)} := \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j}$ is the deterministic measure involving the expected values or medians of the eigenvalues of A and C_ν is a constant depending only on ν. Combining this inequality with Theorem II.2 we deduce that $\mu^{(n)} \to \mu$, the semicircular distribution from Theorem II.2, weakly as $n \to \infty$. The estimate on the right hand side of (5) and the Borel-Cantelli lemma imply then the assertion of Theorem II.3 for our ensemble $A^{(n)}$ or for any ensemble for which a result of the Theorem II.7 variety holds. We emphasize that this is independent from how the matrices $A^{(n)}$ are stochastically related for different n's. We shall present even stronger results (Corollary II.12, Theorem II.17) in the next two sections.

Another consequence of Theorem II.7 is that, in the normalization we use, fluctuations of singular values or eigenvalues of $n \times n$ (Gaussian) random matrices are at most $O(n^{-1/2})$: if $t \gg n^{-1/2}$, the exponent in (5) becomes "large negative". We need to point out that, in all likelihood, this is not an optimal result, and certainly not for the full scale of parameters. For example, it is conjectured that fluctuations of eigenvalues in the bulk of the spectrum (i.e., neither $k = o(n)$ nor $n - k = o(n)$) are of the order $O(n^{-1})$, and the conjecture is supported by the large deviation results quoted in the preceding section (see Theorem II.4 and the paragraph following it). There are several results in that direction of varying degrees of generality. In the Gaussian case, an improvement to the $n^{-1/2}$ result "nearly" follows from the approach presented and the results stated in this article. For example, if we knew that the differences between the *deterministic* λ_j's from Theorem II.7 were of order $1/n$, we could argue that, for a $\lambda_k(A)$ to be θ or more away from its central value λ_k, approximately the same would have to be true for $c\theta n$ neighboring eigenvalues and so, by Theorem I.8, the square of the Hilbert-Schmidt deviation from the most likely spectral picture would be of order $t^2 = \theta^3 n$; substituting this value, and $L^2 = 2/n$, into an estimate of type (4) we would obtain a meaningful estimate whenever $\theta/n^{-2/3}$ was large. Another promising approach would be in exploiting the representation of $G^{(n)}$ (which works also for $A^{(n)}$) found in [**135**]. A related fact which is worth pointing out is that there are two sources contributing to the quantity in (6): the deviation of the $\lambda_j(A)$'s from their central values λ_j's, and the difference between the

deterministic measures $\mu^{(n)}$ and the limit semicircular distribution μ, and both of them are of interest.

For the very edge of the spectrum, i.e., the largest or smallest eigenvalues and perhaps a few adjacent ones and in particular for the norm of the matrix, the fluctuations are, in some cases, known to be of order $n^{-2/3}$. A sample result (see **[148]**) is

Theorem II.8. *There exists an increasing function φ on $\mathbb{R}$ such that the largest eigenvalue $\lambda_1(GOE)$ of the $n \times n$ Gaussian orthogonal ensemble satisfies*

$$\lim_{n\to\infty} \mathcal{P}\left(\lambda_1(GOE) \le 2 + \tau n^{-2/3}\right) = \varphi(\tau) \tag{7}$$

for $\tau \in \mathbb{R}$, the convergence being uniform on compact subsets of $\mathbb{R}$. Similar result (with a different φ) holds for the Gaussian unitary ensemble.

The reader will notice that exactly this size of fluctuations is predicted by the semicircle law itself: if we choose $s > 0$ so that $1/2\pi \int_{2-s}^{2} \sqrt{4-x^2}dx = 1/n$, then $s \approx (\frac{3\pi}{2n})^{2/3}$. On the other hand, it is easy to see that (7) does not accurately predict the order of the probabilities *without* passing to the limit. For example, if $\tau \ge n^{1/6}$, which corresponds to $t \ge 1$ in Theorem II.7, (5) yields the correct asymptotic order of $\log \mathcal{P}\left(\lambda_1(GOE) \le 2 + t\right)$, which, because of a difference in scaling, is inconsistent with the behavior that would have been suggested by (7). In view of possible applications in local theory, it would be potentially useful to recover the asymptotics on the probabilities involved, given by the arguments of **[148]**, for the *full* range of the parameters, a similar project to the one suggested in the paragraph following Theorem II.4 in the context of large deviation results.

The proof of Theorem II.8, and similarly, the proofs of Theorem II.5, II.4 and the other large deviation results, uses the explicit formulae for joint densities of eigenvalues (resp. singular values), which are of the form

$$c(\beta, n) \prod_{1\le j<k\le n} |\lambda_j - \lambda_k|^\beta \exp\big(-\beta \sum_{1\le k\le n} \lambda_k^2/2\big), \tag{8}$$

where $\beta = 1$ or 2 depending on whether the context is real or complex and $c(\beta, n)$ is the normalizing numerical coefficient. Accordingly, one can not expect that it generalizes much beyond the Gaussian and some other classical random matrices. However, there is a strong circumstantial evidence that at least Theorem II.8 is much more universal. It has been shown in **[136]** that the moments $\mathbf{E}\ \mathrm{Tr}(A^\nu)$, where $A = A^{(n)}$ are real symmetric matrices satisfying any of our variance

and independence assumptions and such that the laws of the appropriately normalized matrix elements are uniformly sub-Gaussian, exhibit, as $n, \nu \to \infty$, an asymptotic behavior which is consistent with the assertion of the Theorem. A concentration result for the norm in similar degree of generality but going in a somewhat different direction can be found in [**39**]. As suggested earlier, it would be an interesting project to figure out relevant consequences of measure concentration phenomenon so successfully exploited in local theory.

The phenomenon signalized in Theorem II.8 hasn't been exploited in local theory of Banach spaces, and so it is not known whether, conversely, any of the methods of that area are relevant here. It would be extremely interesting to find a general concentration principle which would imply to small fluctuations of the norm and/or eigenvalues in the bulk of the spectrum.

IIc. **Norm of a random matrix and the Slepian-Gordon lemma.** In this section we shall present a simple argument giving exact asymptotics for extreme singular values (basically norms) and eigenvalues of some symmetric Gaussian random matrices (in particular all *real* Gaussian matrices considered here). The argument is based on the well-known *Slepian's lemma* from probability and its generalization due to Gordon. For greater transparence we shall state the relevant special cases of both variants.

Lemma II.9. *Let $(X_t)_{t\in T}$ and $(Y_t)_{t\in T}$ be two finite families of jointly Gaussian mean zero random variables such that*
(a) $\|X_t - X_{t'}\|_2 \leq \|Y_t - Y_{t'}\|_2$ *for* $t, t' \in T$.
Then

$$\mathbf{E} \max_{t\in T} X_t \leq \mathbf{E} \max_{t\in T} Y_t \ . \tag{9}$$

Similarly, if $T = \cup_{s\in S} T_s$ and
(b) $\|X_t - X_{t'}\|_2 \leq \|Y_t - Y_{t'}\|_2$ *if* $t \in T_s, t' \in T_{s'}$ *with* $s \neq s'$
(c) $\|X_t - X_{t'}\|_2 \geq \|Y_t - Y_{t'}\|_2$ *if* $t, t' \in T_s$ *for some* s
one has

$$\mathbf{E} \max_{s\in S} \min_{t\in T_s} X_{s,t} \leq \mathbf{E} \max_{s\in S} \min_{t\in T_s} Y_{s,t} \ .$$

Remark II.10. Lemma II.9 is usually stated with an additional hypothesis $\|X_t\|_2 = \|Y_t\|_2$ for $t \in T$, and yields the stronger assertion *for any* $\lambda \in \mathbb{R}$, $\mathcal{P}\left(\max_{t\in T} X_t > \lambda\right) \leq \mathcal{P}\left(\max_{t\in T} Y_t > \lambda\right)$; (9) is then an immediate consequence. Analogous comment applies to the second part of the Lemma. Let us also point out that when all T_s are singletons, the second part of the lemma (the Gordon version) reduces to the first

(the Slepian version), which in our formulation coincides with a result of Fernique [**68**].

A typical context in which Lemma II.9 shall be applied is the probability space $(\mathbb{R}^N, \gamma_N)$ with linear functionals as random variables. For $u \in \mathbb{R}^N$, set $Z_u := \langle \cdot, u\rangle$; for future reference we point out that the map $\mathbb{R}^N \ni u \to Z_u \in L_2(\mathbb{R}^N, \gamma_N)$ is an isometry. If $K \subset \mathbb{R}^N$, one has the gauge, or the Minkowski functional of the polar of K given by $\max_{u\in K} Z_u = \|\cdot\|_{K^\circ}$. This identity provides a link between Gaussian processes and convexity or geometry of Banach spaces. It is also clear that in this context T could be any bounded set, as long as we replace max and min by sup and inf. In particular, if $\Psi : \tilde{T} \to T$ is a surjective contraction (possibly nonlinear) between subsets of two Euclidean spaces, then

$$\mathbf{E}\|\cdot\|_{T^\circ} \le \mathbf{E}\|\cdot\|_{\tilde{T}^\circ}.$$

In our setting, the norm $\|\cdot\|_{T^\circ}$ is the operator norm on $n \times n$ real matrices identified with $\mathbb{R}^{n^2}$. For such a matrix X and $u, v \in \mathbb{R}^n$ we have

$$\langle Xu, v\rangle = \mathrm{tr}(X(v \otimes u)) = \langle X, u \otimes v\rangle_{tr} = Z_{u\otimes v}(X),$$

where, as usual, $v \otimes u$ stands for the rank one matrix $(u_j v_k)_{j,k=1}^n$, that is, the matrix of the map $x \to \langle x, v\rangle u$ and $\langle X, Y\rangle_{tr} := \mathrm{Tr}(XY^T)$ is the trace duality, (sometimes referred to as the Hilbert-Schmidt scalar product), which can also be thought of as the usual scalar product on $\mathbb{R}^{n^2}$; note that – as opposed to the remainder of the paper – we use here the standard, i.e., not normalized trace. Accordingly

$$\|X\| = \max_{u,v\in S^{n-1}} \langle Xu, v\rangle = \max_{u,v\in S^{n-1}} Z_{u\otimes v}(X).$$

The Gaussian process $X_{u,v} := Z_{u\otimes v}$, $u, v \in S^{n-1}$ is now going to be compared with $Y_{u,v} := Z_{(u,v)}$, where (u, v) is thought of as an element of $\mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$. In view of prior remarks, to show that the Slepian's lemma applies, one only needs to verify that, for $u, v, u', v' \in S^{n-1}$,

$$|u \otimes v - u' \otimes v'|^2 \le |(u, v) - (u', v')|^2 = |u - u'|^2 + |v - v'|^2, \tag{10}$$

an elementary exercise. On the other hand, for $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$Z_{(u,v)}(x, y) = \langle x, u\rangle + \langle y, v\rangle$$

whence

$$\max_{u,v\in S^{n-1}} Z_{(u,v)}(x, y) = |x| + |y|$$

(this is just saying that $\|\cdot\|_{(U\times V)^\circ} = \|\cdot\|_{U^\circ} + \|\cdot\|_{V^\circ}$) and so the assertion of Lemma II.9 translates to

$$n^{1/2}\mathbf{E}\|G^{(n)}\| \leq 2\int_{\mathbb{R}^n} |x|\, d\gamma_n(x).$$

By comparing with the second moment of $|\cdot|$, the last integral is easily seen to be $< n^{1/2}$. In fact, it equals $\sqrt{2}\Gamma(\frac{n+1}{2})/\Gamma(\frac{n}{2})$. Note also that $|\cdot|^2$ is distributed according to the familiar $\chi^2(n)$ law). The same argument, applied just to symmetric tensors $u\otimes u$, allows us to analyze $\lambda_1(GOE)$, the largest eigenvalue of the Gaussian orthogonal ensemble. The case of $\lambda_1(A^{(n)})$ can then be deduced formally (see the comment after the definition of GOE). In combination with Theorem II.6 and the remark following it, the above shows:

Theorem II.11. *Given $n \in \mathbb{N}$ consider the ensembles of $n\times n$ matrices G, A and GOE. If the random variable F equals either $\|G\|$, $\lambda_1(A)$ or $\lambda_1(GOE)$, then*

$$M_F < \mathbf{E}F < 2,$$

where M_F stands for the median of F. Consequently, for any $t > 0$,

$$\mathcal{P}\left(F \geq 2 + \sigma t\right) < 1 - \Phi(t) < \exp\left(-nt^2/2\right), \tag{11}$$

where $\sigma = 1$ in the case of $\|G\|$ and $\sqrt{2}$ for $\lambda_1(A)$ or $\lambda_1(GOE)$.

The beauty of the above result lies, in particular, in the inequalities being valid for all n rather than asymptotically. However, Theorem II.8 shows that *asymptotically* (11) is not optimal. We need to mention that the gist of Theorem II.11 and the argument given above can be extracted from the work of Gordon [**76, 77, 78**]; the same applies to Theorem II.13 that follows. The relevance of the approach to the "standard" results in random matrices was noticed and publicized by the second author.

Similarly as in the preceding section, it is possible to derive from (11) (via the Borel-Cantelli lemma) results on convergence in probability, for example

$$\lim_{n\to\infty} \lambda_1(A^{(n)}) = 2 \quad \text{almost surely.}$$

Indeed, (11) implies that $\limsup_{n\to\infty} \lambda_1(A^{(n)}) \leq 2$ almost surely. The reverse inequality for $\liminf$ is even easier: by Theorem II.3, for any $\epsilon > 0$, $\lim_{n\to\infty} \mathcal{P}(\exists k\ \lambda_k(A^{(n)}) \in [2-\epsilon, 2]) = 1$ and so, for n large enough, the median λ_1 of $\lambda_1(A^{(n)})$ is $\geq 2 - \epsilon$. We now get the conclusion by appealing to (5). It is now rather routine to deduce the following more precise version of Theorem II.7.

Corollary II.12. *Given $n \in \mathbb{N}$, set $\tilde{\lambda}_k := F^{-1}(\frac{2k-1}{2n})$, $k = 1, \ldots, n$, where F is the cumulative distribution function of the Semicircle Law described in Theorem* II.2 *(i.e. the measure $\mu^{(n)} := \frac{1}{n}\sum_{k=1}^{n} \delta_{\tilde{\lambda}_k}$ is the 'best' approximant of the semicircle distribution among measures with n atoms). Then*

$$\lim_{n\to\infty} \max_{1\le k\le n} |\lambda_k(A^{(n)}) - \tilde{\lambda}_k| = 0$$

almost surely, with analogous results for other ensembles.

An argument just slightly more involved than the proof of Theorem II.11 allows an analysis of the extreme singular numbers of rectangular Gaussian matrices with independent entries. It has been known for quite a while, cf. **[105, 157]**, that as the size of such matrices (appropriately normalized, with entries not necessarily Gaussian) increases with the ratio of the sides approaching $\beta \in (0, 1)$, then, as in Theorem II.3, the empirical measures corresponding to the singular values converge in distribution, or almost surely, to a deterministic measure supported on the interval $[1-\sqrt{\beta}, 1+\sqrt{\beta}]$ and often referred to as the *Marchenko–Pastur distribution* (or, more recently, in the free probability context, the *free Poisson distribution*, cf. the next section). Somewhat later it was determined **[73, 135, 165, 18]** that, under appropriate assumptions on the distribution of the entries, the extreme singular values do converge almost surely to the endpoints of the interval above.

Here we present the following special but elegant fact.

Theorem II.13. *Given $m, n \in \mathbb{N}$ with $m \le n$, put $\beta = m/n$ and consider the $n \times m$ random matrix Γ whose entries are real, independent Gaussian random variables following the $N(0, 1/n)$ law. Let the singular values be $s_1(\Gamma) \ge \cdots \ge s_m(\Gamma)$. Then*

$$1 - \sqrt{\beta} < \mathbf{E}s_m(\Gamma) < M_{s_1(\Gamma)} < \mathbf{E}s_1(\Gamma) < 1 + \sqrt{\beta} \tag{12}$$

and consequently, for any $t > 0$,

$$\max\left\{\mathcal{P}(s_1(\Gamma) \ge 1 + \sqrt{\beta} + t), \mathcal{P}(s_m(\Gamma) \le 1 - \sqrt{\beta} - t)\right\} < 1 - \Phi(t) < \exp\left(-nt^2/2\right).$$

The proof of the upper estimate in (12) is essentially the same as that of the first assertion of Theorem II.11. For the lower estimate, we consider the same families: $X_{u,v} := Z_{u\otimes v}$ and $Y_{u,v} := Z_{(u,v)}$ with $u \in S^{m-1}, v \in S^{n-1}$, but then we set $T_u = \{(u, v) : v \in S^{n-1}\}$. The hypothesis (c) is satisfied since (10) is an equality if $u = u'$, and so we

may use the second part of Lemma II.9 to obtain

$$n^{1/2}\mathbf{E}\max_{u\in S^{m-1}}\min_{v\in S^{n-1}}\langle \Gamma u, v\rangle \le \int_{\mathbb{R}^m}|x|\,d\gamma_m(x)-\int_{\mathbb{R}^n}|x|\,d\gamma_n(x).$$

The quantity on the right does not exceed $\sqrt{m}-\sqrt{n}$. This actually requires some work. What is clear is that the difference between the two expression $\to 0$ as $m,n\to\infty$. Since for any Γ,

$$\max_{u\in S^{m-1}}\min_{v\in S^{n-1}}\langle \Gamma u, v\rangle \equiv -s_m(\Gamma),$$

the first inequality in (12) follows. It remains to appeal again to Theorem II.6.

The arguments of this section can be modified to treat other classes of real matrices, even if the outcomes may be less elegant. It would be nice and potentially very fruitful to find an approach to the complex case(s) that is based on similar ideas. As complex (or symplectic) matrices can be thought of as real matrices with a special structure, the problem at hand is equivalent to considering simple expressions in real Gaussian matrices with matrix coefficients. For example, the complex analogue of $G^{(n)}$ ((ii) on our list) can be identified with a real matrix

$$\frac{1}{\sqrt{2}}\begin{bmatrix} G & -G' \\ G' & G\end{bmatrix} = \frac{1}{\sqrt{2}}\left(\begin{bmatrix}1&0\\0&1\end{bmatrix}\otimes G + \begin{bmatrix}0&-1\\1&0\end{bmatrix}\otimes G'\right) \tag{13}$$

where G and G' are independent copies of $G^{(n)}$. Accordingly, the proper context for the question appears to be that of *operator spaces*; more about related issues in the next section.

In another direction, it is quite clear that variations of the methods of this sections *can* be applied to a wide range questions including random factorizations, and particularly, estimating virtually any reasonable norm applied to Gaussian matrices. For example, one could consider for the operator norm with respect to underlying norms on $\mathbb{R}^n$ different than the Euclidean norm. One just needs to replace S^{n-1} by the spheres corresponding to the norms in question. This has been noticed early on, see [**44, 21, 77**]. However, more often than not, one only gets this way an approximate asymptotic order up to a constant like in Theorem II.1, rather than the precise asymptotics obtained in this section. We mention here an important question which perhaps did not receive enough publicity in the Banach space circles.

Problem II.14. *Show the existence and find*

$$\lim_{n\to\infty}\frac{\mathbf{E}\max_{\epsilon\in\{-1,1\}^n}\langle A^{(n)}\epsilon,\epsilon\rangle}{n}\ .$$

The quantity in the numerator, which is close to and of the same order as the norm of $A^{(n)}$ considered as an operator from ℓ_∞^n to ℓ_1^n, is related to *spin glass theory* [**147, 110**]. Even the existence of the limit is not clear. However, it is generally believed that it does exist and equals approximately 1.527. Mimicking the proof of Theorem II.11 yields $\sqrt{8/\pi} \approx 1.596$ as an upper estimate for every n, and hence for the lim sup, just slightly worse than the best known rigorous upper estimate of $2 - \sqrt{e}/4 \approx 1.588$ for the latter, due to B. Derrida. The best rigorous lower estimate is $4/\pi \approx 1.273$ [**1**]; we suspect that this one is most susceptible to attack due to improvements in technology in the intervening years.

A related, possibly simpler, problem would be calculating the *exact* asymptotic order of, say, the norm of $G^{(n)}$ considered as an operator from ℓ_p^n to ℓ_q^n. To the best of our knowledge, this is known only for very special values of p and q. Or, perhaps more appropriately, one could consider the analogous question for the matrices Γ from Theorem II.13.

IId. **Random matrices and free probability.** A significant recent development was the emergence, due largely to efforts of D. Voiculescu, of the area of *free probability*, and the realization of its connections to random matrices. Even though, *a priori*, free probability seems relevant only to the macroscopic features of the asymptotic spectral pictures of random matrices, it is hard to overestimate its effect on clarifying the subject. Several books on the topic appeared in recent years or are in the works and, accordingly, we shall only sketch here the basic ideas and mention a few results and problems that are relevant to the remainder of our discussion.

A noncommutative probability space is a pair $(\mathcal{A}, \varphi)$, where $\mathcal{A}$ is an algebra, usually over $\mathbb{C}$, with unit I and a state φ. A state is a linear functional verifying $\varphi(I) = 1$. In the classical case, we have an algebra of (measurable) functions on a probability space and the expectation. There are also Banach algebra C*-algebra probability spaces where the algebra in question is endowed with the appropriate additional structure, and the state φ respects that structure. Elements of $\mathcal{A}$ are thought of as random variables, and the distribution of a random variable x is the linear functional μ_x on the algebra of $\mathbb{C}[X]$ of complex polynomials in variable X defined by $\mu_x(p) := \varphi(p(x))$. In the C*-algebra context, the distribution of a normal element is actually represented by a measure supported on the spectrum of x (by the Gelfand-Neumark theorem). In the classical context, this measure is necessarily the law of x. One defines similarly *joint distributions* as

functionals on the algebra of noncommuting polynomials in the appropriate number of variables. The convergence in distribution is the weak-$*$ convergence: $\mu_{x_n}(p) \to \mu_x(p)$ for all polynomials p. This can be defined even if the x_n's belong to different probability spaces.

Three important examples of probability spaces are
(i) the C*-algebra $\mathfrak{M}_n$ of $n \times n$ matrices with the normalized trace Tr; the distribution of $A \in \mathfrak{M}_n$ is represented by the measure $\mu_A = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j(A)}$.
(ii) the *-algebra $\mathcal{A}^{(n)}$ of random $n \times n$ matrices $A(\omega)$ on some sufficiently rich classical probability space $(\Omega, \Sigma, \mathbf{P})$ verifying $\mathbf{E}\|A(\omega)\|^p < \infty$ for all $p < \infty$, with $\varphi = \mathbf{E}\,\mathrm{Tr}$.
(iii) An operator algebra $\mathcal{A} \subset L(\mathcal{H})$ with a vector state $\varphi_\xi(X) := \langle X\xi, \xi\rangle$, where $\xi \in \mathcal{H}$ is a norm one vector.

The fundamental concept of the theory is that of *freeness*. It is modelled after that of *independence*, which may be restated as follows: commuting subalgebras $\mathcal{A}_1, \mathcal{A}_2$ are independent if $\varphi(a_1a_2) = 0$ whenever $a_k \in \mathcal{A}_j$ with $\varphi(a_j) = 0$ for $j = 1, 2$. By analogy, one says that a family $(\mathcal{A}_j)_{j\in J}$ are *free* if for any *reduced* product $a = a_1a_2\ldots a_m$ where neighboring elements come from different $\mathcal{A}_j$'s, one has $\varphi(a) = 0$ whenever $\varphi(a_k) = 0$ for $k = 1, \ldots, m$. A family of random variables (or sets of such) is free (resp, *-free) if the algebras (resp, *-algebras) generated by them are free.

An early result of free probability is a free central limit theorem (CLT): *if $(x_j)_{j\in\mathbb{N}}$ is a sequence of free random variables normalized by $\varphi(x_j) = 0, \varphi(x_j^2) = 1$, $j \in \mathbb{N}$ and satisfying some mild technical conditions such as $\sup_{j\in\mathbb{N}} |\varphi(x_j^k)| < \infty$ for all $k \in \mathbb{N}$, then as $N \to \infty$, $\frac{x_1+\cdots+x_N}{\sqrt{N}}$ converges in distribution to the standard semicircular distribution.* The mysterious appearance of that distribution in this context was elucidated only after Theorem II.16 below was proved.

It is not clear from the above discussion that nontrivial free random variables actually do exist. A prototype of free algebras is as follows. Let G_1, G_2 be discrete groups and $G = G_1 \star G_2$ their free product. The group algebras $\mathcal{A}_j = \mathbb{C}(G_j)$, i.e., the formal linear combinations of elements of the corresponding group, considered as subalgebras of $\mathbb{C}(G)$ are free with respect to the state which assigns to an element of $\mathbb{C}(G)$ the coefficient of the unit e. Equivalently, one may consider $\mathbb{C}(G)$ as a canonical subalgebra of $L(\ell_2(G))$ with the vector state corresponding to $\xi = \delta_e$, the unit vector supported at e. However, a more useful model is provided by the creation/annihilation operators on the full Fock space. Given Hilbert space $\mathcal{H}$, the corresponding Fock space $\mathcal{F}(\mathcal{H})$ is defined as the orthogonal direct sum $\sum_{k\geq 0}^{\oplus} \mathcal{H}^{\otimes k}$ of all tensor powers of $\mathcal{H}$. The

0th power is identified with $\mathbb{C}\Omega$, where Ω is a fixed unit vector, which is thought of as an empty tensor product and is called the *vacuum vector*. The corresponding vector state φ_Ω is referred to as the vacuum state. Given $h \in \mathcal{H}$, one defines a shift operator $\ell(h)$ by $\ell(h)\eta = h \otimes \eta$ for elementary tensors η. We have

Fact II.15. *In a probability space $(L(\mathcal{F}(\mathcal{H})), \varphi_\Omega)$ the following hold*

(i) *If $(\mathcal{H}_j)_{j\in J}$ is a family of orthogonal subspaces of $\mathcal{H}$, then the corresponding family $(\{\ell(h) : h \in \mathcal{H}_j\})_{j\in J}$ of subsets of $L(\mathcal{F}(\mathcal{H}))$ is $*$-free.*

(ii) *If $h \in \mathcal{H}$ is a norm one vector, then $s(h) := \ell(h) + \ell(h)^*$ has the standard semicircular distribution.*

The crucial link between random matrices and free probability is provided by the following result of Voiculescu ([**154**], see also [**155**]):

Theorem II.16. *For each $n \in \mathbb{N}$, let $(A_j^{(n)})_{j\in J}$ be independent copies of the random $n \times n$ matrix $A^{(n)}$, considered as elements of the noncommutative probability space $(\mathcal{A}^{(n)}, \mathbf{E}\,\mathrm{Tr})$ defined above. Then, as $n \to \infty$, $(A_j^{(n)})_{j\in J}$ converges in distribution to $(s_j)_{j\in J}$, a family of free random variables, each of which has the standard semicircular distribution. The same is true for GOE, GUE and any ensembles of random matrices verifying the hypotheses of Theorem* II.2.

The fact that, for any $j \in J$, $A_j^{(n)}$ converges in distribution to s_j is of course equivalent to Theorem II.2. The new ingredient is the asymptotic freeness of large random matrices. It is worthwhile emphasizing that this asymptotic freeness in combination with the free CLT and the infinite divisibility of the Gaussian distribution do imply that the only possible limit distribution is semicircular.

As we indicated, free probability in general and Theorem II.16 in particular open new vistas on the subject of random matrices. For example, if P is a polynomial in m noncommuting variables, it follows that, for any $\nu \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathbf{E}\,\mathrm{Tr}(P(A_1^{(n)}, \ldots, A_m^{(n)})^\nu) = \varphi_\Omega(P(s_1, \ldots, s_m)^\nu), \tag{14}$$

where $A_j^{(n)}$ and s_j are as in Theorem II.16. In particular, if the polynomial P is formally selfadjoint (the variables themselves *are* assumed to be selfadjoint here), this leads to a measure concentrated on the spectrum of the bounded selfadjoint operator $P(s_1, \ldots, s_m)$ as the asymptotic spectral distribution of the random matrix $P(A_1^{(n)}, \ldots, A_m^{(n)})$. Asymptotic means *a priori* in the sense of Theorem II.2, but proceeding as in section IIb, one deduces:

Theorem II.17. *In the notation of Theorem* II.16*, let P be a selfadjoint polynomial in m noncommuting variables and let $n \in \mathbb{N}$. Then there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that, for $k = 1, \ldots, n$ and $t \geq 0$,*

$$\mathcal{P}\Big(\big|\lambda_k(P(A_1^{(n)},\ldots,A_m^{(n)}))-\lambda_k\big| \geq t\Big) < C\exp\left(-c_P n \min\{t^2, t^{2/d}\}\right) \tag{15}$$

where d is the degree of P, $c_P > 0$ depends only on P and C is a universal constant. One necessarily has $\max_{1\leq k\leq n}|\lambda_k| \leq C_P$, where C_P depends only on P, and λ_k's can be chosen to be the medians (or expected values) of $\lambda_k(P(A_1^{(n)}, \ldots, A_m^{(n)}))$. Moreover, as $n \to \infty$, the deterministic measures $\mu^{(n)} := \frac{1}{n}\sum_{k=1}^n \delta_{\lambda_k}$ converge weakly to μ_P, the law of the random variable $P(s_1, \ldots, s_m)$. Consequently, as $n \to \infty$, the random measures $\mu_n(\omega) := \frac{1}{n}\sum_{k=1}^n \delta_{\lambda_k(P(A_1^{(n)},\ldots,A_m^{(n)}))}$ converge (weakly) to μ_P almost surely.

The above generalizes to other ensembles of real or complex selfadjoint random matrices. The variant for singular values follows formally: they are square roots of eigenvalues for the polynomial P^*P. There is also a variant for orthogonal or unitary matrices, or for expressions involving additionally *constant* matrices with a limiting spectral distribution (see **[105]** for an early work in that direction), and one can similarly analyze at least some operator space type expressions cf. (13).

The proof of Theorem II.17 is almost the same as that of prior results. The only fine point is that, this time, the functions $\lambda_j(P(X_1, \ldots, X_m))$ are not "nicely" Lipschitz if P is not linear. However, polynomials *are* Lipschitz if the size of the arguments is under control, and Theorem II.11 implies that, with large probability, $P(A_1^{(n)}, \ldots, A_m^{(n)})$ is "not too large". One then gets the result by appealing to the original isoperimetric inequality underlying Theorem II.6. The factor $\min\{t^2, t^{2/d}\}$ reflects the variation of P: Lipschitz for small values of the arguments and exhibiting power growth of order d far away from 0.

Part of the appeal of Theorem II.17 lies in the definitive identification of the limit measure μ_P. Moreover, that measure often can be found explicitly using analytic methods of free probability (close to the earlier Stieltjes transform method introduced in **[105]**), see **[156]**. Also, this is the language in which it is easiest to explain the otherwise surprising coincidence between asymptotic spectral distributions of $|A^{(n)}|$ and $|G^{(n)}|$ (see Theorem II.3 and the remarks following it).

What is missing in the nonlinear context are analogues of Theorem II.11 and the resulting Corollary II.12. We do not know the exact asymptotics of the extreme eigenvalues of $P(A_1^{(n)}, \ldots, A_m^{(n)})$ or, equivalently, the norms $\|P(A_1^{(n)}, \ldots, A_m^{(n)})\|$.

Problem II.18. *In the notation of Theorem* II.17*, does* λ_1 *converge to the maximum of the spectrum of* μ_P*?*

An affirmative answer would imply the almost sure convergence of the eigenvalues $\lambda_k(P(A_1^{(n)}, \dots, A_m^{(n)}))$ to the same quantity and, more importantly, the fact that, asymptotically, $(A_1^{(n)}, \dots, A_m^{(n)})$ are equivalent to $(s_1, \dots, s_m)$ not only in distribution as in (14) or Theorem II.17, but also in the C*-algebra sense. Problem II.18 is of similar nature as Problem II.14: in both cases we do know, for somewhat similar reasons, the order of growth of the quantity in question, but we do not know exact asymptotics. However, the former seems to be much more accessible than the latter because of direct applicability of spectral methods. This is because, in the Euclidean case, $\|B\| \approx (\mathrm{Tr}(B^\nu))^{1/\nu}$ if B is an $n \times n$ matrix and ν is an even number such that $\nu/\log n$ is large, and so the problem reduces to properly estimating $\mathbf{E}\,\mathrm{Tr}(P(A_1^{(n)}, \dots, A_m^{(n)})^\nu)$ for such ν, n, in principle a straightforward combinatorial question. This approach, already present in the original Wigner's paper [**162**], was exploited by a number of authors, perhaps in the most sophisticated way in [**80**].

Another possibility would be to somehow use the ideas behind Theorem II.7 or II.11 to show by an *a priori* argument that λ_k's of Theorem II.17 vary slowly with k. This could be interesting for other reasons as the distribution of quantities like $s_1(G^{(n)}) - s_2(G^{(n)})$ are of interest in the theory of computational complexity, cf. [**142**]. For the quantity mentioned here, the needed information can probably be extracted from the methods of [**149**] But to our knowledge, no careful examination of the relevant arguments was made.

A nontrivial but possibly accessible test cases for any approach would be to find the exact asymptotic behavior of $\|(G^{(n)})^\nu\|$ for $\nu \geq 2$ or $\|U + V + U^* + V^*\|$, where U, V are independent and uniformly distributed on $U(n)$ or $SO(n)$. The predictions given by the free probability are $\sqrt{(\nu+1)^{\nu+1}/\nu^\nu} \approx \sqrt{e\nu}$ for large ν, and $\sqrt{3}/2$. Both were actually determined before the era of free probability, see [**160, 94**]. The latter reference applies in the random walk on the free group reformulation, cf. [**156**]. A known (asymptotic) estimate on $\|(G^{(n)})^\nu\|$ is $\nu + 1$ ([**17**]).

IIe. **Random vs. explicit constructions.** It is a quite frequent occurrence that existence of mathematical objects possessing a certain property is shown via nonconstructive methods. Roughly speaking, one produces a random variable whose values are those objects and then proves that the property in question is satisfied with nonzero (cf. [**2**]), or close to 1, probability. Two fields where this principle has

been successfully applied are combinatorics and analysis, particularly harmonic analysis and local theory of Banach spaces. Many developments in the latter area are described elsewhere in this collection [**104, 74, 35**], and some spillovers to local operator theory were mentioned in this survey. Here we shall introduce only the details needed to address some philosophical aspect of the issue.

We start by recalling a remarkable result of Kashin [**91**], cf. [**140, 143**], motivated by questions in approximation theory, which roughly asserts that the space ℓ_1^{2n} is an orthogonal sum (in the ℓ_2^{2n} sense) of two *nearly Euclidean* subspaces. More precisely:

Theorem II.19. *Given $n = 2m \in \mathbb{N}$, there exist two orthogonal m-dimensional subspaces $E_1, E_2 \subset \mathbb{R}^n$ such that*

$$\frac{1}{8}\|x\|_2 \le \frac{1}{\sqrt{n}}\|x\|_1 \le \|x\|_2 \quad \text{for all} \quad x_i \in E_i,\ i = 1, 2. \tag{16}$$

Moreover, for large n, this holds for nearly all decompositions $E_1 \oplus E_2$ with respect to the Haar measure on the Grassmann manifold $G_{n,m}$.

The existence of such a decomposition was surprising because, when considered on the entire space $\mathbb{R}^n$, the ratio between the ℓ_1 and ℓ_2 norms varies between 1 and $\sqrt{n}$.

Because of the wealth of examples of various objects (e.g. random Banach spaces) arising from or related to the one above, it would be of interest to have an *explicit* example of a *Kashin decomposition*. Having explicit E_k's would likely lead to more natural examples than the ones obtained by the probabilistic method of finite dimensional Banach spaces with various extremal properties. Indeed, an explicit example must have an explicit reason, and this should presumably be reflected by a presence of some additional structure. It may also conceivably lead to some useful algorithms.

It may be the right place to comment here on what exactly we mean by an explicit construction. We shall not give a general definition, but, in the case at hand, admissible descriptions of $E_1 \oplus E_2$ would include an algorithm yielding, for a given n, a basis of E_1 or the matrix of the orthogonal projection onto E_1. That algorithm would need to have a reasonable *worst case* performance, preferably polynomial in n. The worst case requirement is needed to exclude a strategy that would involve choosing $E_1 \oplus E_2$ at random and then somehow efficiently verifying whether it satisfies (16). We note that, at least at the first sight, even checking (16) for a given E_k appears to be an exponentially hard problem.

To the best of our knowledge, the largest explicit subspace E of ℓ_1^m, for which an assertion of type (16) holds is of dimension approximately $\sqrt{m}$ as opposed to $m/2$ in Theorem II.19—a long way to go. This follows, e.g., from the construction in [**131**], using finite fields and difference sets, and giving, for an even integer $p \geq 4$, an exact Λ_p-set. This leads to another question: finding explicit exact Λ_p-sets for other values of p; for probabilistic results see [**34, 144**] and cf. [**35**].

An example of a probabilistic construction followed by an explicit one is the work on approximation of quasidiagonal operators mentioned in section Ic [**139, 153, 159**]. The explicit approach used representations of groups with property T. It is quite likely that the argument could be refined to yield explicit matrices poorly approximable by reducible matrices cf. [**86**], also mentioned in section Ic. This is presumably not a very important problem, but could be a good starting point.

A somewhat similar example comes from another field, involving random graphs. Following the seminal work of Erdös [**66**], this became a powerful technique to show existence of graphs with various extremal properties. As some questions about graphs are very practical optimization problems (like design of a network), it was important to have explicit solutions. This is particularly relevant because verifying that a given large graph has some required property is sometimes computationally not feasible. For some important questions, explicit solution were found, see [**107, 103**] and much earlier work [**106**]. The constructions were based on properties of some arithmetic groups, sophisticated tools from arithmetic geometry and, again, the property T. Some of the graphs in question were already mentioned at the end of section Id.

The final topic we propose to analyse comes from the preceding section on free probability. Theorem II.16 can be interpreted as an assertion that two large random matrices are *nearly free*, in the sense of the pattern of the moments of monomials in the two matrices with respect to the normalized trace being approximately the same as it would have been the case if the two matrices were free. In fact, the same is true for, roughly speaking, one fixed and one random matrix [**155**]. A natural problem is to come up with explicit matrices having the same property. (It can not happen that two finite matrices are exactly free, except for the trivial cases.) The article [**32**], involving representations of symmetric groups, may be relevant here. This and the preceding two paragraphs do broadly suggest some of the areas of mathematics that might be pertinent to other explicit constructions.

References

[1] Aizenman, M., Lebowitz, J. L. and Ruelle, D., *Some rigorous results on the Sherrington-Kirkpatrick spin glass model*, Comm. Math. Phys. **112** (1987), 3–20.

[2] Alon, N., *Probabilistic proofs of existence of rare events*, Geometric aspects of functional analysis (1987–88), 186–201, Lecture Notes in Math., **1376**, Springer, Berlin-New York, 1989.

[3] Andersen, N.T., *Similarity of continuous nests*, Bull. London Math. Soc. **15** (1983), 131–132.

[4] Anderson, J., *Extensions, restrictions, and representations of states on C^*-algebras*, Trans. Amer. Math. Soc. **249** (1979), 303–329.

[5] Anderson, J., *Pathology in the Calkin algebra*, J. Oper. Theory, **2** (1979), 159–167.

[6] Anderson, J., *Extreme points of positive linear maps on $B(H)$*, J. Func. Anal. **31** (1979), 195–217.

[7] Anderson, J., *A C^*-algebra $\mathcal{A}$ for which* $\mathrm{Ext}(\mathcal{A})$ *is not a group*, Ann. Math. **107** (1978), 455–458.

[8] Angelos, J.R., Cowen, C.C. and Narayan, S.K., *Triangular truncation and finding the norm of a Hadamard multiplier*, Lin. Alg. Appl. **170** (1992), 117–135.

[9] Apostol, C., Foiaş, C. and Voiculescu, D., *Some results on nonquasitriangular operators. VI*, Rev. Roum. Math. Pures et Appl. **18** (1973), 1473–1494.

[10] Apostol, C., Foiaş, C. and Voiculescu, D., *On the norm-closure of nilpotents. II*, Rev. Roum. Math. Pures et Appl. **19** (1974), 549–577.

[11] Apostol, C. and Salinas, N., *Nilpotent approximations and quasinilpotent operators*, Pacific J. Math. **61** (1975), 327–337.

[12] Arnold, L., *On the asymptotic distribution of the eigenvalues of random matrices*, J. Math. Anal. Appl. **20** (1967), 262–268.

[13] Arveson, W.B., *Interpolation problems in nest algebras*, J. Func. Anal. **3** (1975), 208–233.

[14] Arveson, W.B., *Operator algebras and invariant subspaces*, Ann. Math. **100** (1974), 433–532.

[15] Azoff, E. and Davis, C., *On distances between unitary orbits of self-adjoint operators*, Acta Sci. Math. (Szeged) **47** (1984), 419–439.

[16] Bai, Z.D. and Yin, Y.Q., *Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix*, Ann. Probab. **16** (1988), 1729–1741.

[17] Bai, Z.D., Yin, Y.Q., *Limiting behavior of the norm of products of random matrices and two problems of Geman-Hwang*, Probab. Theory Related Fields **73** (1986), 555–569.

[18] Bai, Z.D. and Yin, Y.Q., *Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix*, Ann. Probab. **21** (1993), 1275–1294.

[19] Ben Arous, G., Guionnet, A., *Large deviations for Wigner's law and Voiculescu's non-commutative entropy*, Probab. Theory Related Fields **108** (1997), 517–542.

[20] Bennett, G., Goodman, V. and Newman, C.M., *Norms of random matrices*, Pacific J. Math. **59** (1975), 359–365

[21] Benyamini, Y. and Gordon, Y., *Random factorization of operators between Banach spaces*, J. Analyse Math. **39** (1981), 45–74

[22] Bercovici, H., *Hyper-reflexivity and the factorization of linear functionals*, J. Func. Anal. **158** (1998), 242–252.

[23] Berg, I.D., *An extension of the Weyl-von Neumann theorem to normal operators*, Trans. Amer. Math. Soc. **160** (1971), 365–371.

[24] Berg, I.D., *On approximation of normal operators by weighted shifts*, Mich. Math. J. **21** (1974), 377–383.

[25] Berg, I.D. and Davidson, K.R., *A quantitative version of the Brown-Douglas-Fillmore Theorem*, Acta. Math. **166** (1991), 121–161.

[26] Berman, K., Halpern, H., Kaftal, V. and Weiss, G., *Matrix norm inequalities and the relative Diximier property*, Int. Eq. Oper. Theory **11** (1988), 28–48.

[27] Bhatia, R., *Perturbation bounds for matrix eigenvalues*, Pitman Research Notes in Mathematics Series **162**, Longman Scientific and Technical Pub. Co., London, New York, 1987.

[28] Bhatia, R., *Analysis of spectral variation and some inequalities*, Trans. Amer. Math. Soc. **272** (1982), 323–332.

[29] Bhatia, R., *Matrix analysis*, Graduate Texts in Mathematics **169**, Springer-Verlag, New York, 1997.

[30] Bhatia, R., Davis, C. and McIntosh, A., *Perturbation of spectral subspaces and solution of linear operator equations*, Lin. Alg. Appl. **52/53** (1983), 45–67.

[31] Bhatia, R. and Holbrook, J., *Short normal paths and spectral variation*, Lin. Alg. Appl. **94** (1985), 377–382.

[32] Biane, P., *Representations of symmetric groups and free probability*, Adv. Math. **138** (1998), 126–181.

[33] Borell, C., *The Brunn-Minkowski inequality in Gauss space*, Invent. Math. **30** (1975), 207–216.

[34] Bourgain, J., *Bounded orthogonal systems and the $\Lambda(p)$-set problem*, Acta Math. **162** (1989), 227–245.

[35] Bourgain, J., *Lambda-p sets in analysis: results, problems and related aspects*, this volume.

[36] Bourgain, J. and Tzafriri, L., *Invertibility of "large" submatrices with applications to the geometry of Banach spaces and harmonic analysis*, Israel J. Math. **57** (1987), 137–224.

[37] Bourgain, J. and Tzafriri, L., *Restricted invertibility of matrices and applications*, Analysis at Urbana II (Berkson, E. R., Peck, N. T. and Uhl, J. J., eds) London Math. Soc. Lect. Notes **138**, Cambridge Univ. Press (1989), 61–107.

[38] Bourgain, J. and Tzafriri, L., *On a problem of Kadison and Singer*, J. Reine Angew. Math. **420** (1991), 1–43.

[39] Boutet de Monvel, A. and Shcherbina, M., *On the Norm of Random Matrices*, Math. Notes **57** (1995), 475–484.

[40] Bronk, B.V., *Accuracy of the semicircle approximation for the density of eigenvalues of random matrices*, J. Math. Phys. **5** (1964), 215–220.

[41] Brown, L.G., Douglas, R.G. and Fillmore, P.A. , *Unitary equivalence modulo the compact operators and extensions of C^*-algebras*, Proceedings of a conference on operator theory, Halifax, NS 1973, Lect. Notes in Math. **345**, Springer-Verlag, Berlin-Heidelberg-New York, 1973, 58–128.

[42] Brown, L.G., Douglas, R.G. and Fillmore, P.A., *Extensions of C^*-algebras and K-homology*, Ann. Math. **105** (1977), 265–324.
[43] Brown, S.W., *Some invariant subspaces for subnormal operators*, Int. Eq. Oper. Theory **1** (1978), 310–333.
[44] Chevet, S., *Séries de variables aléatoires gaussiennes à valeurs dans $E\hat{\otimes}_\varepsilon F$. Application aux produits d'espaces de Wiener abstraits*, Séminaire sur la Géométrie des Espaces de Banach **(1977–1978)**, Exp. No. 19, 15 pp., École Polytech., Palaiseau, 1978.
[45] Choi, M.D., *Almost commuting matrices need not be nearly commuting*, Proc. Amer. Math. Soc. **102** (1988), 529–533.
[46] Christensen, E., *Perturbations of operator algebras II*, Indiana Univ. Math. J. **26** (1977), 891–904.
[47] Dadarlat, M., *On the approximation of quasidiagonal C*-algebras*, J. Func. Anal. **167** (1999), 69–78.
[48] Davidson, K.R., *Similarity and compact perturbations of nest algebras*, J. Reine Angew. Math. **348** (1984), 72–87.
[49] Davidson, K.R., *Almost commuting Hermitian matrices*, Math. Scand. **56** (1985), 222–240.
[50] Davidson, K.R., *The distance between unitary orbits of normal operators*, Acta. Sci. Math. (Szeged) **50** (1986), 213–223.
[51] Davidson, K.R., *The distance to the analytic Toeplitz operators*, Illinois J. Math. **31** (1987), 265–273.
[52] Davidson, K.R., *Nest Algebras*, Pitman Research Notes in Mathematics Series **191**, Longman Scientific and Technical Pub. Co., London, New York, 1988.
[53] Davidson, K.R., *Finite dimensional problems in operator theory*, The Gohberg Anniversary Collection, Operator Theory: Advances and Applications **40**, 187–201. Birkhaüser Verlag. Basel, 1989.
[54] Davidson, K.R., *C*-Algebras by Example*, Fields Institute Monograph Series **6**, Amer. Math. Soc., Providence, RI, 1996.
[55] Davidson, K.R., *Polynomially bounded operators*, NATO ASI Proceedings, Samos, Greece, August 1996, *Operator algebras and applications*, pp. 145–162, Kluger Acad. Pub., Dordrecht, 1997.
[56] Davidson, K.R., Herrero, D.A. and Salinas, N., *Quasidiagonal operators, approximation, and C*-algebras*, Indiana Univ. J. Math. **38** (1989), 973–998.
[57] Davidson, K.R. and Ordower, M., *Some Exact Distance Constants*, Lin. Alg. Appl. **208/209** (1994), 37–55.
[58] Davidson,K.R. and Paulsen, V.I., *Polynomially bounded operators*, J. Reine Angew. Math. **487** (1997), 153–170.
[59] Davidson, K.R. and Pitts, D.R., *Approximate unitary equivalence of completely distributive commutative subspace lattices*, Int. Eq. Oper. Theory **22** (1995), 196–211.
[60] Davidson, K.R. and Pitts, D.R., *Invariant subspaces and hyper-reflexivity for free semigroup algebras*, Proc. London Math. Soc. **78** (1999), 401–430.
[61] Davidson, K.R. and Power, S.C., *Failure of the distance formula*, J. London Math. Soc. **32** (1984), 157–165.
[62] Davis, C., Kahan, W.M. and Weinberger, W.F., *Norm preserving dilations and their applications to optimal error bounds*, SIAM J. Numer. Anal. **19** (1982), 445–469.

[63] Deddens, J.A. and Fillmore, P.A., *Reflexive linear transformations*, Lin. Alg. Appl. **10** (1975), 89–93.

[64] Dmitrovskiĭ, V.A. , *On the integrability of the maximum and the local properties of Gaussian fields*, Probability theory and mathematical statistics **1** (Vilnius, 1989), 271–284, "Mokslas", Vilnius, 1990.

[65] Dykema, K., *On certain free product factors via an extended matrix model*, J. Func. Anal. **112** (1993), 31–60.

[66] Erdös, P., *Graph theory and probability*, Canad. J. Math. **11** (1959), 34–38. *Graph theory and probability II*, Canad. J. Math. **13** (1961), 346–352.

[67] Exel, R. and Loring, T.A., *Determinants and almost commuting matrices*, Proc. Amer. Math. Soc. **106** (1989), 913–915.

[68] Fernique, X., *Des résultats nouveaux sur les processus gaussiens*, C. R. Acad. Sci. Paris Sér. A **278** (1974), 363–365.

[69] Fernique, X., *Fonctions Aléatoires Gaussiennes Vecteurs Aléatoires Gaussiennes*, Publications Centre de Recherches Mathématiques, 1997

[70] Forrester, P.J. and Odlyzko, A.M., *Gaussian unitary ensemble eigenvalues and Riemann ζ function zeros: a nonlinear equation for a new statistic*, Phys. Rev. E (3) **54** (1996), R4493–R4495

[71] Friis, P. and Rørdam, M., *Almost commuting self-adjoint matrices – a short proof of Huaxin Lin's theorem*, J. Reine Angew. Math. **479** (1996), 121–131.

[72] Friis, P. and Rørdam, M., *Approximation with normal operators with finite spectrum, and an elementary proof of the Brown–Douglas–Fillmore theorem*, preprint, 1998.

[73] Geman, S., *A limit theorem for the norm of random matrices*, Ann. Probab. **8** (1980), 252–261.

[74] Giannopoulos, A.A. and Milman, V.D., *Euclidean Structure in Finite Dimensional Normed Spaces*, this volume.

[75] Girko, V.L., *Theory of random determinants*, Translated from the Russian. Mathematics and its Applications (Soviet Series), **45**. Kluwer Academic Publishers Group, Dordrecht, 1990.

[76] Gordon, Y., *On Dvoretzky's theorem and extensions of Slepian's lemma*, Israel seminar on geometrical aspects of functional analysis (1983/84), **II**, 25 pp., Tel Aviv Univ., Tel Aviv, 1984.

[77] Gordon, Y., *Some inequalities for Gaussian processes and applications*, Israel J. Math. **50** (1985), 265–289.

[78] Gordon, Y., *Majorization of Gaussian processes and geometric applications*, Probab. Theory Related Fields **91** (1992), 251–267.

[79] Gromov, M. and Milman, V.D., *A topological application of the isoperimetric inequality*, Amer. J. Math. **105** (1983), 843–854.

[80] Haagerup, U. and Thorbjørnsen, S., *Random matrices and K-theory for exact C^*-algebras*, Doc. Math. **4** (1999), 341–450.

[81] Hadwin, D., *Triangular truncation and normal limits of nilpotent operators*, Proc. Amer. Math. Soc. **123** (1995), 1741–1745.

[82] Herrero, D.A., *Approximation of Hilbert space operators. I*, Research Notes in Math. **72**, Pitman Books Ltd., London, Boston, Melbourne, 1982).

[83] Herrero, D.A., *Normal limits of nilpotent operators*, Indiana Univ. Math. J. **23** (1974), 1097–1108.

[84] Herrero, D.A., *Unitary orbits of power partial isometries and approximation by block-diagonal nilpotents*, Topics in modern operator theory, Fifth Int. Conf. on Oper. Theory, Timişoara and Herculane (Romania 1980), OT: Advances and Applications **2**, Birkhaüser-Verlag, Basel, 1981), 171–210.

[85] Herrero, D.A., *A trace obstruction to approximation by block diagonal operators*, Amer. J. Math. **108** (1986), 451–484.

[86] Herrero, D.A. and Szarek, S.J., *How well can an $n \times n$ matrix be approximated by reducible ones?*, Duke J. Math. **53** (1986), 233–248.

[87] Hoffman, A.J. and Wielandt, H.W., *The variation of the spectrum of a normal matrix*, Duke J. Math. **20** (1953), 37–39.

[88] Johnson, W.B. and Schechtman, G., *Finite dimensional subspaces of L_p*, this volume.

[89] Kadison, R.V. and Singer, I.M., *Extensions of pure states*, Amer. J. Math. **81** (1959), 383–400.

[90] Kahane, J.-P., *Une inégalité du type de Slepian et Gordon sur les processus gaussiens*, Israel J. Math. **55** (1986), no. 1, 109–110.

[91] Kashin, B.S., *The widths of certain finite-dimensional sets and classes of smooth functions*, (Russian), Izv. Akad. Nauk SSSR Ser. Mat. **41** (1977), 334–351.

[92] Kashin, B.S., *A property of bilinear forms*, (Russian), Soobshch. Akad. Nauk Gruzin. SSR **97** (1980), 29–32.

[93] Kashin, B.S., *Some properties of matrices of bounded operators from the space l_2^n into l_2^m*, (Russian), Izv. Akad. Nauk Armyan. SSR Ser. Mat. **15** (1980), 379–394 (English translation: Soviet J. Contemporary Math. Anal. **15** (1980), 44–57, (1982)).

[94] Kesten, H., *Symmetric random walks on groups*, Trans. Amer. Math. Soc. **92** (1959), 336–354.

[95] Kwapień, S., *Decoupling inequalities for polynomial chaos*, Ann. Probab. **15** (1987), 1062–1071.

[96] Kwapień, S., *A remark on the median and the expectation of convex functions of Gaussian vectors*, Probability in Banach spaces **9** (Sandjberg, 1993), 271–272, Progr. Probab. **35**, Birkhaüser Boston, Boston, MA, 1994.

[97] Lance, E.C., *Cohomology and perturbations of nest algebras*, Proc. London Math. Soc. **423** (1981), 334–356.

[98] Larson, D.R., *Hyperreflexivity and a dual product construction*, Trans. Amer. Math. Soc. **294** (1986), 79–88.

[99] Ledoux, M. and Talagrand, M., *Probability in Banach spaces. Isoperimetry and processes.* Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **23**. Springer-Verlag, Berlin, 1991.

[100] Lin, H., *Almost commuting selfadjoint matrices and applications*, Operator algebras and their applications (Waterloo, ON, 1994/1995), 193–233, Fields Inst. Commun. **13**, Amer. Math. Soc. Providence, RI, 1997.

[101] Loring, T.A., *K-theory and asymptotically commuting matrices*, Canad. J. Math. **40** (1988), 197–216.

[102] Loring, T.A., *Lifting solutions to perturbing problems in C*-algebras*, Fields Institute Monograph Series **8**, Amer. Math. Soc., Providence, RI, 1997.

[103] Lubotzky, A., Phillips, R. and Sarnak, P., *Ramanujan graphs*, Combinatorica **8** (1988), 261–277.

[104] Mankiewicz, P. and Tomczak-Jaegermann, N., *Quotients of finite-dimensional Banach spaces; random phenomena*, this volume.

[105] Marčenko, V.A.; Pastur, L.A., *Distribution of eigenvalues in certain sets of random matrices*, (Russian), Mat. Sb. (N.S.) **72** (1967), 507–536.

[106] Margulis, G.A., *Explicit constructions of expanders*, (Russian) Problemy Peredači Informacii **9** (1973), 71–80. English translation: Problems Inform. Transmission **9** (1973), 325–332.

[107] Margulis, G.A., *Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators*, (Russian), Problemy Peredachi Informatsii **24** (1988), 51–60; English translation: Problems Inform. Transmission **24** (1988), 39–46.

[108] Mehta, M.L., *Random matrices. Second edition*, Academic Press, Inc., Boston, MA, 1991.

[109] Mehta, M.L. and Gaudin, M., *On the density of eigenvalues of a random matrix*, Nuclear Phys. **18** (1960), 420–427.

[110] Mézard, M., Parisi, G. and Virasoro, M.A., *Spin glass theory and beyond*, World Scientific Lecture Notes in Physics, **9**. World Scientific Publishing Co., Inc., Teaneck, NJ, 1987.

[111] Milman, V.D. and Schechtman, G. *Asymptotic theory of finite dimensional normed spaces. With an appendix by M. Gromov*, Lecture Notes Math. **1200**, Springer Verlag, Berlin-New York, 1986.

[112] Mirsky, L., *Symmetric gauge functions and unitarily invariant norms*, Quart. J. Math. (2) **11** (1960), 50–59.

[113] Montgomery, H.L., *The pair correlation of zeros of the zeta function*, Analytic number theory (Proc. Sympos. Pure Math., **XXIV**, St. Louis Univ., St. Louis, Mo., 1972), pp. 181–193. Amer. Math. Soc., Providence, R.I., 1973

[114] Odlyzko, A.M., *On the distribution of spacings between zeros of the zeta function*, Math. Comp. **48** (1987), 273–308.

[115] Olin, R.F. and Thomson, J.E., *Algebras of subnormal operators*, J. Func. Anal. **37** (1980), 271–301.

[116] Parrott, S.K., *On a quotient norm and the Sz.-Nagy Foiaş lifting theorem*, J. Func. Anal. **30** (1978), 311–328.

[117] Pastur, L.A., *Spectra of random selfadjoint operators*, (Russian) Uspehi Mat. Nauk **28** (1973), no. 1(169), 3–64.

[118] Paulsen, V.I., *Completely bounded maps and dilations*, Pitman Research Notes in Mathematics Series **146**, Longman Scientific and Technical Pub. Co., London, New York, 1986.

[119] Pearcy, C.M. and Shields, A.L., *Almost commuting matrices*, J. Func. Anal. **30** (1978), 311–328.

[120] Pisier, G., *Probabilistic methods in the geometry of Banach spaces*, *C.I.M.E. Varenna 1985.* Lecture Notes in Math. **1206**, 167–241, Springer Verlag, Berlin-New York, 1986.

[121] Pisier, G., *Factorization of linear operators and geometry of Banach spaces*, CBMS Regional Conference Series in Mathematics, **60**. Amer. Math. Soc., Providence, R.I., 1986.

[122] Pisier, G., *A polynomially bounded operator on Hilbert space which is not similar to a contraction*, J. Amer. Math. Soc. **10** (1997), 351–369.

[123] Pisier, G., *Similarity problems and completely bounded maps*, Lect. Notes in Math. **1618**, Springer-Verlag, Berlin, 1996.
[124] Pisier, G., *Operator spaces and similarity problems*, Proc. Inter. Congress Math., Vol. I (Berlin, 1998). Doc. Math. (1998), Extra Vol. I, 429–452.
[125] Pisier, G., *Operator spaces*, this volume.
[126] Pitts, D.R., *Close CSL algebras are similar*, Math. Ann. **300** (1994), 149–156.
[127] Popescu, G., *A generalization of Beurling's theorem and a class of reflexive algebras*, J. Operator Theory **41** (1999), 391–420.
[128] Power, S.C., *The distance to upper triangular operators*, Math. Proc. Camb. Phil. Soc. **88** (1980), 327–329.
[129] Ringrose, J.R., *On some algebras of operators*, Proc. London Math. Soc. **15** (1965), 61–83.
[130] Rosenoer, S., *Distance estimates for von Neumann algebras*, Proc. Amer. Math. Soc. **86** (1982), 248–252.
[131] Rudin, W., *Trigonometric series with gaps*, J. Math. Mech. **9** (1960), 203–227.
[132] Rudnick, Z. and Sarnak, P., *Zeros of principal L-functions and random matrix theory*, A celebration of John F. Nash, Jr., Duke Math. J. **81** (1996), 269–322.
[133] Salinas, N., *Homotopy invariance of* $\mathrm{Ext}(\mathcal{A})$, Duke Math. J. **44** (1977), 777–794.
[134] Schechtman, G., *Concentration, results and applications*, this volume.
[135] Silverstein, J.W., *The smallest eigenvalue of a large-dimensional Wishart matrix*, Ann. Probab. **13** (1985), 1364–1368.
[136] Sinaĭ, Ya.G. and Soshnikov, A.B., *A refinement of Wigner's semicircle law in a neighborhood of the spectrum edge for random symmetric matrices*, (Russian). Funk. Anal. i Prilozhen. **32** (1998), 56–79; English translation: Funct. Anal. Appl. **32** (1998), 114–131.
[137] Sunder, V.S., *Distance between normal operators*, Proc. Amer. Math. Soc. **84** (1982), 483–484.
[138] Szarek, S.J., *On almost commuting Hermitian operators*, Proc. Seventh GPOTS (Lawrence, KS, 1987), Rocky Mountain J. Math. **20** (1990), 581–589.
[139] Szarek, S.J., *An exotic quasidiagonal operator*, J.. Func. Anal. **89** (1990), 274–290.
[140] Szarek, S.J., *On Kashin's almost Euclidean orthogonal decomposition of* ℓ_1^n, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **26** (1978), 691–694.
[141] Szarek, S.J., *Spaces with large distance to* ℓ_∞^n *and random matrices*, Amer. J. Math. **112** (1990), 899–942.
[142] Szarek, S.J., *Condition numbers of random matrices*, J. Complexity **7** (1991), 131–149.
[143] Szarek, S.J. and Tomczak-Jaegermann, N., *On nearly Euclidean decomposition for some classes of Banach spaces*, Comp. Math. **40** (1980), 367–385.
[144] Talagrand, M., *Sections of smooth convex bodies via majorizing measures*, Acta Math. **175** (1995), 273–300.
[145] Talagrand, M., *New concentration inequalities in product spaces*, Invent. Math. **126** (1996), 505–563.
[146] Talagrand, M., *Concentration of measure and isoperimetric inequalities in product spaces*, Inst. Hautes Études Sci. Publ. Math. **81** (1995), 73–205.

[147] Talagrand, M., *Huge random structures and mean field models for spin glasses*, Proc. Inter. Congress Math. Vol. I (Berlin, 1998). Doc. Math. (1998), Extra Vol. I, 507–536.

[148] Tracy, C.A. and Widom, H., *Level-spacing distributions and the Airy kernel*, Comm. Math. Phys. **159** (1994), 151–174.

[149] Tracy, C.A. and Widom, H., *On orthogonal and symplectic matrix ensembles*, Comm. Math. Phys. **177** (1996), 727–754.

[150] Voiculescu, D., *A non-commutative Weyl-von Neumann theorem*, Rev. Roum. Math. Pures et Appl. **21** (1976), 97–113.

[151] Voiculescu, D., *Asymptotically commuting finite rank unitary operators without commuting approximants*, Acta Sci. Math. (Szeged) **45** (1983), 429–431.

[152] Voiculescu, D., *A note on quasidiagonal operators*, Topics in operator theory, Constantin Apostol Memorial Issue, O.T.: Advances and Applications **32**, Birkhaüser-Verlag, Basel, 1988, 265–274.

[153] Voiculescu, D., *Property T and approximation of operators*, Bull. London Math. Soc. **22** (1990), 25–30.

[154] Voiculescu, D., *Limit laws for random matrices and free products*, Invent. Math. **104** (1991), 201–220.

[155] Voiculescu, D., *A strengthened asymptotic freeness result for random matrices with applications to free entropy*, Inter. Math. Res. Notices **1** (1998), 41–63.

[156] Voiculescu, D., Dykema, K.J. and Nica, A. *Free random variables. A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups*, CRM Monograph Series **1**. Amer. Math. Soc., Providence, RI, 1992.

[157] Wachter, K.W., *The strong limits of random matrix spectra for sample matrices of independent elements*, Ann. Probab. **6** (1978), 1–18.

[158] Wassermann, S., *A separable quasidiagonal C^*-algebra with a nonquasidiagonal quotient by the compact operators*, Math. Proc. Cambridge Phil. Soc. **110** (1991), 143–145.

[159] Wassermann, S., *C*-algebras associated with groups with Kazhdan's property T*, Ann. Math. **134** (1991), 423–431.

[160] Wegmann, R., *The asymptotic eigenvalue-distribution for a certain class of random matrices*, J. Math. Anal. Appl. **56** (1976), 113–132.

[161] Weyl, H., *Der asymptotische Verteilungs gesetz der Eigenwerte linearer partieller Differentialgleichungen*, Math. Ann. **71** (1912), 441–479.

[162] Wigner, E.P., *Characteristic vectors of bordered matrices with infinite dimensions*, Ann. Math. **62** (1955), 548–564.

[163] Wigner, E.P., *On the distribution of the roots of certain symmetric matrices*, Ann. Math. **67** (1958), 325–327.

[164] Wils, I.M., *Stone-Cech compactifications and representations of operator algebras*, Ph. D. thesis, Catholic Univ. of Nijemgin, 1968.

[165] Yin, Y.Q., Bai, Z.D. and Krishnaiah, P.R., *On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix*, Probab. Theory Related Fields **78** (1988), 509–521.