Blueprint: Norm Discretization

Mario Ullrich

October 23, 2026

Contents

1 The problem and the result **3**

2 Matrices and operators in the Loewner order **8**
2.1 Comparison with multiples of the identity 8
2.2 Traces of products 9
2.3 The trace of a positive operator 10
2.4 From the trace of an inverse to the Loewner order 12
2.5 Rank-one updates 13

3 Averages of quadratic forms **15**
3.1 The Gram matrix 15
3.2 The Gram operator 16
3.3 A reproducing kernel Hilbert space 17
3.4 A point from two averages 18

4 The lower potential and its barrier **20**
4.1 The potential and the verifier 20
4.2 The verifier passes on average 22
4.3 From the potential to the frame bound 22

5 The upper potential and its barrier **24**
5.1 The upper potential and the verifier 24
5.2 The upper verifier passes on average 26
5.3 From the upper potential to the frame bound 27

6 The construction and the proof **28**
6.1 The barrier iteration 28
6.2 Proof of the discretization theorem 29

7 Discretization of the uniform norm and the L_p-norms **31**
7.1 The proof of the Kiefer–Wolfowitz theorem 35

References **39**

Chapter 1

The problem and the result

Let (Ω, μ) be a measure space and ι a finite nonempty index set with $m = \#\iota$ elements. Two objects are given: a finite family $a = (a_k)_{k\in\iota}$ of square-integrable functions on Ω, read as a map $a : \Omega \to \mathbb{C}^\iota$, and a square-integrable map $b : \Omega \to H$ into a separable complex Hilbert space H. We ask whether the integral of $|f|^2$ can be replaced by finitely many weighted point evaluations, so that the first family stays a frame from below and b stays a frame from above. A statement about weighted sums of $|f(x_i)|^2$ is a statement about the matrix $\sum_i w_i\, a(x_i)a(x_i)^*$ or the operator $\sum_i w_i\, b(x_i)b(x_i)^*$, compared in the Loewner order: $A \preceq B$ means that $B-A$ is positive semidefinite. That is the form the argument takes.

The *Gram operator* of b is

$$G(b) = \int_\Omega b(x)b(x)^*\, d\mu(x),$$

a Bochner integral of the rank-one operators $b(x)b(x)^* = \langle b(x), \cdot\rangle\, b(x)$ on H. For $H = \mathbb{C}^\iota$ it is a matrix, so the family a has the *Gram matrix* $G(a)$, in Lean `Discretization.gram`. For $\Lambda > 0$ with $G(b) \preceq \Lambda \cdot 1$, the number $M = \operatorname{Tr} G(b)/\Lambda$ is the *effective dimension* of $G(b)$ relative to Λ; it is at most $\dim H$, and much smaller as soon as the eigenvalues of $G(b)$ decay.

The following theorem is Theorem 3 of Chkifa, Dolbeault, Krieg and Ullrich [3], for a second family with values in a separable Hilbert space. Its first bound is their factor $(1 - \sqrt{(m-1)/n})^2 \lambda_{\min}(I)$ in Loewner form.

Theorem 1.1 (Discretization)**.** *Let (Ω, μ) be a measure space, let ι be a finite nonempty set, and let H be a nonzero separable complex Hilbert space. Let $a : \Omega \to \mathbb{C}^\iota$ be square-integrable in every coordinate with $I = G(a) = \int_\Omega a(x)a(x)^*\, d\mu(x)$ positive definite, let $b : \Omega \to H$ be square-integrable,*

and let $J = G(b) = \int_\Omega b(x)b(x)^*\,d\mu(x)$ *be its Gram operator; assume that* J *is injective. Let* $\Lambda > 0$ *satisfy* $J \preceq \Lambda \cdot 1$*, and put* $m = \#\iota$ *and* $M = \int_\Omega \|b(x)\|^2\,d\mu(x)/\Lambda$*. Then for every* $n \geq m$ *there are points* $x_1, \ldots, x_n \in \Omega$ *and weights* $w_1, \ldots, w_n > 0$ *with*

$$\left(1 - \sqrt{\tfrac{m-1}{n}}\right)^2 \cdot I \;\preceq\; \sum_{i=1}^{n} w_i\, a(x_i)a(x_i)^*$$

in the matrix order and

$$\sum_{i=1}^{n} w_i\, b(x_i)b(x_i)^* \;\preceq\; \left(1 + \sqrt{\tfrac{M-1}{n}}\right)^2 \Lambda \cdot 1$$

in the order of operators.

[Lean: `Discretization.Infinite.bss_generalized`]

The proof is given in Chapter 6, after the construction it rests on. No side condition on m or on M is needed; the potential argument there needs two, and three further cases of the proof remove them. Both hypotheses on J can be read on b (Theorem 3.2): $J \preceq \Lambda\cdot 1$ is the bound $\int_\Omega |\langle u, b\rangle|^2\,d\mu \leq \Lambda\|u\|^2$ for every $u \in H$, and J is injective exactly when $\langle u, b(\cdot)\rangle = 0$ almost everywhere only for $u = 0$.

Remark 1.2 (A countable family)**.** *A countable family* $b = (b_k)_{k\in\kappa}$ *of square-integrable functions with* $\sum_k \|b_k\|_{L_2}^2 < \infty$ *is the case* $H = \ell_2(\kappa)$ *and* $b(x) = (b_k(x))_k$*. The Gram operator is the Gram matrix* $(\int_\Omega b_k\overline{b_l}\,d\mu)_{k,l}$*,* $M = \sum_k \|b_k\|_{L_2}^2/\Lambda$*, and the upper bound reads* $\sum_i w_i|\sum_k c_k\, b_k(x_i)|^2 \leq (1 + \sqrt{(M-1)/n})^2\Lambda\|c\|^2$ *for every* $c \in \ell_2(\kappa)$*. Nothing pointwise is assumed: by Tonelli,* $\sum_k |b_k(x)|^2 < \infty$ *almost everywhere, and the points are chosen there. This is the form of Theorem 3 of [3]; in Lean it is* `Discretization.Countable.bss_generalized`*, and for finite* κ*, proved with matrices,* `Discretization.bss_generalized`*. Read through quadratic forms, the lower bound is* $(1 - \sqrt{(m-1)/n})^2 \int_\Omega |f|^2\,d\mu \leq \sum_i w_i|f(x_i)|^2$ *for* f *in the span of the first family; for finite* κ *both bounds in this form are* `Discretization.exists_discretization`*.*

For $b = a$, Theorem 1.1 gives the sparsification theorem of Batson, Spielman and Srivastava [1] in a slightly stronger form: $\sqrt{(m-1)/n}$ in place of their $\sqrt{m/n}$, which comes from [3]. Part (b) below is their statement for finitely many vectors, over $\mathbb{C}$.

Corollary 1.3 (Sparsification)**.** *Let ι be a finite nonempty set and put $m = \#\iota$.*

(a) Let (Ω, μ) be a measure space and let $a : \Omega \to \mathbb{C}^\iota$ be square-integrable in every coordinate with $\int_\Omega a(x)a(x)^*\, d\mu(x) = 1$. Then for every $n \geq m$ there are points $x_1, \dots, x_n \in \Omega$ and weights $w_1, \dots, w_n > 0$ with

$$\left(1 - \sqrt{\tfrac{m-1}{n}}\right)^2 \cdot 1 \;\preceq\; \sum_{i=1}^{n} w_i\, a(x_i)a(x_i)^* \;\preceq\; \left(1 + \sqrt{\tfrac{m-1}{n}}\right)^2 \cdot 1.$$

(b) Let Ω be a finite set and let $v_k \in \mathbb{C}^\iota$, for $k \in \Omega$, be vectors with $\sum_{k\in\Omega} v_k v_k^* = 1$. Then for every $n \geq m$ there are weights $s_k \geq 0$, at most n of them nonzero, with

$$\left(1 - \sqrt{\tfrac{m-1}{n}}\right)^2 \cdot 1 \;\preceq\; \sum_{k\in\Omega} s_k\, v_k v_k^* \;\preceq\; \left(1 + \sqrt{\tfrac{m-1}{n}}\right)^2 \cdot 1.$$

[Lean: `Discretization.bss Discretization.bss_of_sum_eq_one`]

Proof. (a) This is Theorem 1.1 for $b = a$, $H = \mathbb{C}^\iota$ and $I = 1$: the Gram operator of the second family is then the identity, $\Lambda = 1$ bounds it, and its effective dimension is $\operatorname{Tr} 1/1 = m$, so both frame constants are built from the same $r = \sqrt{(m-1)/n}$.

(b) The counting measure on Ω has Gram matrix $\sum_k v_k v_k^* = 1$, so (a) applies to $k \mapsto v_k$. A vector chosen several times receives the sum of its weights, $s_k = \sum_{i:x_i=k} w_i$, which leaves both sums unchanged, and a nonzero s_k belongs to one of the n chosen points. □

Read through quadratic forms, with the kernel sections of a reproducing kernel Hilbert space as second family, the theorem becomes a statement about norms, Corollary 4 of [3]: the L_2-norm on the span of the first family is discretized from below, and the norm of the space from above.

Corollary 1.4 (Discretization of the L_2-norm)**.** *Let (Ω, μ) be a measure space, let ι be a finite nonempty set, and let $a : \Omega \to \mathbb{C}^\iota$ be square-integrable in every coordinate with positive definite Gram matrix $I = \int_\Omega a(x)a(x)^*\, d\mu(x)$. Let H be a nonzero separable complex Hilbert space of measurable functions on Ω on which every point evaluation is continuous, let K be its reproducing kernel, and assume $\int_\Omega K(x,x)\, d\mu(x) < \infty$. Let $\Lambda > 0$ satisfy $\int_\Omega |g|^2\, d\mu \leq \Lambda \|g\|_H^2$ for every $g \in H$, and assume that*

the only $g \in H$ vanishing almost everywhere is $g = 0$. Put $m = \#\iota$ and $M = \int_\Omega K(x,x)\,d\mu(x)/\Lambda$. Then for every $n \geq m$ there are points $x_1, \ldots, x_n \in \Omega$ and weights $w_1, \ldots, w_n > 0$ with

$$\left(1 - \sqrt{\tfrac{m-1}{n}}\right)^2 \int_\Omega |f|^2\,d\mu \;\leq\; \sum_{i=1}^{n} w_i\,|f(x_i)|^2$$

for every $f \in \operatorname{span}\{a_k\}$, and with

$$\sum_{i=1}^{n} w_i\,|g(x_i)|^2 \;\leq\; \left(1 + \sqrt{\tfrac{M-1}{n}}\right)^2 \Lambda\,\|g\|_H^2$$

for every $g \in H$.

[Lean: `Discretization.RKHS.exists_discretization`]

Proof. Take for $b(x)$ the kernel section k_x of Theorem 3.4: b is square-integrable, $|\langle g, k_x\rangle| = |g(x)|$ and $\|k_x\|^2 = K(x,x)$. By Theorem 3.2(b) and (c) the two assumptions on H say that the Gram operator of b satisfies $J \preceq \Lambda \cdot 1$ and is injective, and $\int_\Omega \|b\|^2\,d\mu = \int_\Omega K(x,x)\,d\mu$, so Theorem 1.1 applies. Its two bounds are inequalities between quadratic forms. For $f = \sum_k \overline{c_k}\,a_k$ one has $\sum_i w_i |f(x_i)|^2 = c^* A c$ with $A = \sum_i w_i\,a(x_i)a(x_i)^*$, and $\int_\Omega |f|^2\,d\mu = c^* I c$ by Theorem 3.1(a); this is the first bound. For $g \in H$, $\langle g, \sum_i w_i\,k_{x_i}k_{x_i}^*\,g\rangle = \sum_i w_i |\langle g, k_{x_i}\rangle|^2 = \sum_i w_i |g(x_i)|^2$; this is the second. □

The first bound involves only the functions of the span, so the case $I = 1$ already gives it: $I^{-1/2}a$ spans the same space and has Gram matrix 1.

The constant Λ is the squared norm of the embedding $H \hookrightarrow L_2(\mu)$, and the assumption on vanishing functions says that this embedding is injective.

Every square-integrable map $b : \Omega \to H$ with injective Gram operator gives such a space: the functions $\langle u, b(\cdot)\rangle$, with norm $\|u\|$ and kernel $K(x,y) = \langle b(x), b(y)\rangle$, for which $\int_\Omega K(x,x)\,d\mu = \int_\Omega \|b\|^2\,d\mu$. So the corollary covers every such b; in Lean this form is `Discretization.Infinite.exists_discretization`.

On a probability space the lower bound of Theorem 1.1 holds with equal weights, a one-sided discretization of the L_2-norm. With equal weights and of the order of m points, Limonova and Temlyakov [6, Theorem 1.1] prove such a discretization for spaces satisfying a Nikol'skii-type inequality, and Bartel, Schäfer and Ullrich [2] for arbitrary spaces, up to constants. The

form below is the case $p = 2$ of Proposition 8 of [3], whose argument is followed here: for $b \equiv 1$ in $H = \mathbb{C}$ one has $J = 1$ and $\Lambda = M = 1$, and the upper bound $\sum_i w_i \le 1$ of the theorem comes, in the case $M \le 1 + 1/n$ of the proof, from a construction that bounds every single weight by $1/n$.

Corollary 1.5 (One-sided discretization)**.** *Let (Ω, μ) be a probability space, let ι be a finite nonempty set with $m = \#\iota$ elements, and let $a : \Omega \to \mathbb{C}^\iota$ be square-integrable in every coordinate with $I = \int_\Omega a(x)a(x)^*\, d\mu(x)$ positive definite. Then for every $n \ge m$ there are points $x_1, \dots, x_n \in \Omega$, not necessarily distinct, with*

$$\Big(1 - \sqrt{\tfrac{m-1}{n}}\Big)^2 \int_\Omega |f|^2\, d\mu \;\le\; \frac{1}{n}\sum_{i=1}^{n} |f(x_i)|^2 \qquad \text{for every } f \in \operatorname{span}\{a_k\}.$$

[Lean: `Discretization.exists_one_sided_discretization`]

Proof. Put $r = \sqrt{(m-1)/n}$. It suffices to find weights $0 < w_i \le 1/n$ with $(1-r)^2 \cdot I \preceq \sum_i w_i\, a(x_i)a(x_i)^*$: the quadratic form of this inequality at the coefficient vector of f is $(1-r)^2 \int_\Omega |f|^2\, d\mu \le \sum_i w_i |f(x_i)|^2$ by Theorem 3.1(a), and raising each w_i to $1/n$ only enlarges the right-hand side. The first family is normalized as in the proof of Theorem 1.1, which leaves the weights unchanged, so let $I = 1$.

For $m \ge 2$, run the construction of the case $M \le 1 + 1/n$ in the proof of Theorem 1.1 with the constant upper verifier $U \equiv n$, whose average is n because μ is a probability measure. It starts from $A_0 = (\delta m/r) \cdot 1$ with $\delta = (1-r)/n$, tracks only the lower matrix with Theorem 4.3, finds each point with Theorem 4.4(b), and gives the lower bound by Theorem 4.5. Each point x_i satisfies $U(x_i) < L(x_i)$ for the lower verifier L and receives the weight $w_i = 1/L(x_i)$, so $w_i\, n \le 1$.

For $m = 1$ the lower constant is 1. The lower verifier $L(x) = n|a(x)|^2$ and $U \equiv n$ have the same average n, so Theorem 3.5(b) gives a point y with $n \le n|a(y)|^2$. Used n times with the weight $1/(n|a(y)|^2) \le 1/n$, it makes $\sum_i w_i |a(x_i)|^2 = 1$. □

In the Loewner order the corollary reads $(1 - \sqrt{(m-1)/n})^2 \cdot I \preceq \frac{1}{n}\sum_i a(x_i)a(x_i)^*$ (`Discretization.bss_lower_equal_weights`).

Chapter 2

Matrices and operators in the Loewner order

The construction keeps a positive definite matrix above a multiple of the identity and a strictly positive operator, positive and invertible, above a multiple of the Gram operator, and controls each through one real number. This chapter collects what is needed about the Loewner order, the trace and rank-one updates, for matrices and for operators on a complex Hilbert space H; for $H = \mathbb{C}^\iota$ the operators are the matrices. Mathlib has no trace-class theory, so the trace of a positive operator is built here from a Hilbert basis.

2.1 Comparison with multiples of the identity

Lemma 2.1 (Comparison with multiples of the identity)**.** *Let ι be a finite set; all matrices below are in $\mathbb{C}^{\iota\times\iota}$.*

(a) Let A be Hermitian and c a real number with $\lambda_i(A) \le c$ for every eigenvalue. Then $A \preceq c \cdot 1$.

(b) Every positive semidefinite A satisfies $A \preceq (\operatorname{Tr} A) \cdot 1$.

(c) If A and B are positive definite with $A \preceq B$, then $B^{-1} \preceq A^{-1}$.

(d) If A is positive definite and $A^{-1} \preceq c\cdot 1$ for some $c > 0$, then $c^{-1}\cdot 1 \preceq A$.

[Lean: `Matrix.IsHermitian.le_smul_one Matrix.PosSemidef.le_trace_smul_one Matrix.PosDef.inv_le_inv_of_le Matrix.PosDef.smul_one_le_of_inv_le`]

Proof. (a) By the spectral theorem $c \cdot 1 - A$ is unitarily conjugate to the diagonal matrix with entries $c - \lambda_i(A) \ge 0$, and conjugation by a unitary preserves positive semidefiniteness.

(b) The trace is the sum of the eigenvalues of A, all of which are nonnegative, so every single eigenvalue is at most $\operatorname{Tr} A$, and (a) applies with $c = \operatorname{Tr} A$.

(c), (d) Both are read off Mathlib's antitonicity of the inverse in a C*-algebra with a star ordering, applied to the units of $\mathbb{C}^{\iota\times\iota}$ with the ℓ_2-operator norm; the inverse of $c \cdot 1$ is $c^{-1} \cdot 1$. □

Lemma 2.2 (Conjugation by the square root)**.** *Let A be a C*-algebra with a star ordering, let $B \in A$ be positive and invertible, let $X \in A$ and let $t > 0$. If $B^{-1/2} X B^{-1/2} \preceq t \cdot 1$, then $t^{-1} \cdot X \preceq B$.*

[Lean: `CStarAlgebra.inv_smul_le_of_conj_inv_sqrt_le`]

Proof. Conjugating the hypothesis by $B^{1/2}$ preserves the order. On the left the two inverse square roots cancel, on the right $t \cdot 1$ becomes $t \cdot B$; dividing by t gives the claim. □

2.2 Traces of products

Lemma 2.3 (Traces of products)**.** *Let ι be a finite set; all matrices below are in $\mathbb{C}^{\iota\times\iota}$.*

(a) For positive semidefinite P and Q we have $\operatorname{Tr}(PQ) \geq 0$. If P is positive definite, Q is positive semidefinite and $Q \neq 0$, then $\operatorname{Tr}(PQ) > 0$.

(b) Let Y be positive semidefinite and Z any matrix. Then

$$\left|\operatorname{Tr}(ZY)\right|^2 \;\leq\; \operatorname{Tr} Y \cdot \operatorname{Tr}(ZYZ^*).$$

[Lean: `Matrix.PosSemidef.trace_mul_nonneg Matrix.PosDef.re_trace_mul_pos`
`Matrix.PosSemidef.norm_trace_mul_sq_le`
`Matrix.PosSemidef.re_trace_mul_sq_le`]

Proof. (a) Write Q as a sum of rank-one matrices vv^*. Then $\operatorname{Tr}(Pvv^*) = v^*Pv \geq 0$, with strict inequality for $v \neq 0$ when P is positive definite, and at least one of the v is nonzero when $Q \neq 0$.

(b) A positive semidefinite Y makes $\langle P, Q\rangle = \operatorname{Tr}(QYP^*)$ a semi-inner product on matrices. Cauchy–Schwarz for the pair 1, Z is the inequality. No commutation of Y and Z is used. □

2.3 The trace of a positive operator

Lemma 2.4 (Parseval and Hilbert–Schmidt sums under adjoints)**.** *Let H be a complex Hilbert space with a Hilbert basis $e = (e_k)_{k\in\kappa}$.*

(a) For every $x \in H$ the family $\big(|\langle e_k, x\rangle|^2\big)_k$ has sum $\|x\|^2$.

(b) For every bounded operator T on H,

$$\sum_k \|Te_k\|^2 = \sum_k \|T^*e_k\|^2,$$

and the family on the left is summable if and only if the one on the right is.

(c) For every bounded operator T on H and every second Hilbert basis $e' = (e'_j)_{j\in\kappa'}$ of H,

$$\sum_k \|Te_k\|^2 = \sum_j \|T^*e'_j\|^2 \qquad \text{in } [0,\infty].$$

[Lean: `HilbertBasis.hasSum_norm_sq_inner`
`ContinuousLinearMap.tsum_norm_sq_adjoint`
`ContinuousLinearMap.summable_norm_sq_adjoint_iff`
`ContinuousLinearMap.tsum_ofReal_norm_sq_adjoint_of_basis`]

Proof. (a) Take real parts in Mathlib's Parseval identity $\sum_k \langle x, e_k\rangle\langle e_k, x\rangle = \langle x, x\rangle$.

(c) Expanding $\|Te_k\|^2$ by (a) along e' and $\|T^*e'_j\|^2$ along e gives the same double series $\sum_{j,k} |\langle e'_j, Te_k\rangle|^2$ with the two summations exchanged. The rearrangement is carried out in $[0,\infty]$, where every sum is defined and the interchange needs no hypothesis. For a self-adjoint T it says that the Hilbert–Schmidt sum does not depend on the basis.

(b) The case $e' = e$ of (c); summability is read off the value being finite. □

Part (b) is the hinge of the whole trace layer: it is what supplies the summability of $P^{1/2}Q$ in Theorem 2.6(b), whose adjoint $QP^{1/2}$ is visibly Hilbert–Schmidt.

Definition 2.5 (Trace of a positive operator)**.** Let H be a complex Hilbert space, let T be a positive operator on H, and let $e = (e_k)_{k\in\kappa}$ be a Hilbert basis of H. The *trace* of T is

$$\operatorname{Tr} T = \sum_{k\in\kappa} \operatorname{Re}\langle e_k, Te_k\rangle \in [0,\infty],$$

and T is *of finite trace* if $\operatorname{Tr} T < \infty$.

[Lean: `ContinuousLinearMap.traceAlong ContinuousLinearMap.trace`]

The terms are nonnegative, because $\operatorname{Re}\langle x, Tx\rangle = \|T^{1/2}x\|^2$, so $\operatorname{Tr} T = \sum_k \|T^{1/2}e_k\|^2$ is the squared Hilbert–Schmidt norm of $T^{1/2}$. Theorem 2.4(c) for the self-adjoint operator $T^{1/2}$ gives the same value in $[0,\infty]$ along every Hilbert basis, so the trace does not depend on e. For a positive operator P of finite trace and a positive operator Q write $\operatorname{Tr}(PQ) = \sum_k \operatorname{Re}\langle e_k, PQe_k\rangle$; by Theorem 2.6(b) this is the trace of the positive operator $P^{1/2}QP^{1/2}$, so it does not depend on e either.

In Lean the trace is a real number, `ContinuousLinearMap.trace`, computed along a Hilbert basis chosen once and for all (`HilbertBasis.chosen`). The sum is taken to be 0 when the family is not summable; every statement assumes finite trace, so this convention never enters. The same sum along a given basis e is `traceAlong e`, and for a positive operator the two agree (`ContinuousLinearMap.trace_eq_traceAlong`). The chosen basis is countable when H is separable and nonempty when H is nonzero. `IsFiniteTracePos J` bundles three properties of J: positive, of finite trace, and injective.

Lemma 2.6 (Traces of positive operators)**.** *Let H be a complex Hilbert space.*

(a) Every positive operator T on H of finite trace satisfies $T \preceq \operatorname{Tr} T \cdot 1$.

(b) Let P be positive and of finite trace, let Q be positive, and let $e = (e_k)_k$ be a Hilbert basis of H. Then

$$\sum_k \operatorname{Re}\langle e_k, PQe_k\rangle = \sum_k \operatorname{Re}\big\langle P^{1/2}e_k,\, Q\,P^{1/2}e_k\big\rangle \;\geq\; 0,$$

and the value is positive as soon as Q is invertible and $P \neq 0$.

(c) Let $P \preceq P'$ be positive and of finite trace, and let Q be positive. Then $\operatorname{Tr}(PQ) \leq \operatorname{Tr}(P'Q)$.

[Lean: `ContinuousLinearMap.le_traceAlong_smul_one`
`ContinuousLinearMap.traceAlong_mul ContinuousLinearMap.traceAlong_mul_pos`
`Discretization.Infinite.traceAlong_mul_le_of_le`]

Proof. (a) For $x \in H$, writing x in a Hilbert basis and applying the Cauchy–Schwarz inequality to each term gives $\operatorname{Re}\langle x, Tx\rangle \leq \operatorname{Tr} T \cdot \|x\|^2$. This is the operator analogue of Theorem 2.1(b), and unlike the matrix proof it uses no diagonalisation.

(b) The identity is the cyclicity $\mathrm{Tr}(PQ) = \mathrm{Tr}(P^{1/2}QP^{1/2})$, read as a symmetry of the pairing $\sum_k \langle Se_k, Te_k\rangle$ under passing to adjoints; the required summability is Theorem 2.4(b). Every term on the right is nonnegative. For the strict inequality, invertibility of Q makes $Q^{1/2}$ injective, so a vanishing trace would force $P^{1/2}$ to annihilate the whole basis.

(c) The difference $\mathrm{Tr}(P'Q) - \mathrm{Tr}(PQ)$ is the sum of (b) for the positive operator $P' - P$ of finite trace, hence nonnegative. □

Invertibility of Q here is what replaces positive definiteness. A positive operator of finite trace on an infinite-dimensional space is compact, hence never bounded away from zero, so positive definiteness is unavailable and the hypotheses on the two operators have to be chosen differently: injectivity for J, and invertibility for the operator the construction moves.

2.4 From the trace of an inverse to the Loewner order

An operator is *strictly positive* if it is positive and invertible; this is Mathlib's `IsStrictlyPositive`, and for a matrix it means positive definite. The inverse is `Ring.inverse`, defined on all operators and equal to the inverse on the invertible ones.

The next lemma is the reason the whole argument can be run on two real numbers: a bound on the trace of an inverse is a bound in the Loewner order. The traces $\mathrm{Tr}\, A^{-1}$ and $\mathrm{Tr}(JB^{-1})$ are the lower and the upper potential of Chapter 4 and Chapter 5.

Lemma 2.7 (From the trace of an inverse to the Loewner order)**.** *(a) Let ι be a finite set and let $A \in \mathbb{C}^{\iota\times\iota}$ be positive definite. Then $\left(\mathrm{Tr}\, A^{-1}\right)^{-1} \cdot 1 \preceq A$.*

(b) If moreover $0 < \delta < (\mathrm{Tr}\, A^{-1})^{-1}$, then $A - \delta \cdot 1$ is positive definite.

(c) Let H be a nonzero complex Hilbert space, let J be a positive injective operator on H of finite trace, and let B be a strictly positive operator on H. Then $\mathrm{Tr}(JB^{-1}) > 0$ and

$$\left(\mathrm{Tr}(JB^{-1})\right)^{-1} \cdot J \;\preceq\; B.$$

[Lean: `Matrix.PosDef.inv_re_trace_smul_one_le` `Matrix.PosDef.sub_smul_one` `Discretization.Infinite.upperPotential_pos` `Discretization.upperPotential_pos` `Discretization.Infinite.inv_upperPotential_smul_le` `Matrix.PosDef.inv_re_trace_mul_smul_le`]

Proof. (a) Theorem 2.1(b) applied to A^{-1} gives $A^{-1} \preceq (\operatorname{Tr} A^{-1}) \cdot 1$, and Theorem 2.1(d) turns that around.

(b) Write $A - \delta \cdot 1 = \big(A - (\operatorname{Tr} A^{-1})^{-1} \cdot 1\big) + \big((\operatorname{Tr} A^{-1})^{-1} - \delta\big) \cdot 1$, a positive semidefinite matrix by (a) plus a positive definite one.

(c) The operator $B^{-1/2} J B^{-1/2}$ is positive and of finite trace, and its trace is again $\operatorname{Tr}(JB^{-1})$ by Theorem 2.6(b), which is positive since $J \neq 0$ on the nonzero space H. So Theorem 2.6(a) bounds it by $\operatorname{Tr}(JB^{-1}) \cdot 1$, and Theorem 2.2 carries that back to $J \preceq \operatorname{Tr}(JB^{-1}) \cdot B$. Lean proves (c) a second time for positive definite matrices, which is what the finite case uses. □

2.5 Rank-one updates

Lemma 2.8 (Rank-one updates)**.** *Let H be a complex Hilbert space and $u \in H$, and write uu^* for the rank-one operator $z \mapsto \langle u, z\rangle u$; for $H = \mathbb{C}^\iota$ it is the matrix uu^*, and $\langle u, Au\rangle = u^* A u$.*

(a) Let A be a self-adjoint invertible operator on H and t a scalar with $1 + t\langle u, A^{-1}u\rangle \neq 0$. Then $A + t\,uu^*$ is invertible and

$$\big(A + t\,uu^*\big)^{-1} = A^{-1} - \frac{t}{1 + t\langle u, A^{-1}u\rangle}\,\big(A^{-1}u\big)\big(A^{-1}u\big)^*.$$

For $H = \mathbb{C}^\iota$ consequently

$$\operatorname{Tr}\big(A + t\,uu^*\big)^{-1} = \operatorname{Tr} A^{-1} - \frac{t}{1 + t\,u^* A^{-1} u}\,u^* A^{-2} u.$$

(b) Let A be strictly positive and $w \geq 0$. Then $A + w\,uu^*$ is strictly positive. If moreover $w \operatorname{Re}\langle u, A^{-1}u\rangle < 1$, then $A - w\,uu^*$ is strictly positive.

(c) $uu^* \preceq \|u\|^2 \cdot 1$.

[Lean: `ContinuousLinearMap.inverse_add_smul_rankOne`
`Matrix.inv_add_smul_vecMulVec` `Matrix.trace_inv_add_smul_vecMulVec`
`ContinuousLinearMap.isStrictlyPositive_add_smul_rankOne`
`ContinuousLinearMap.isStrictlyPositive_sub_smul_rankOne`
`Matrix.PosDef.add_smul_vecMulVec` `Matrix.PosDef.sub_smul_vecMulVec`
`ContinuousLinearMap.rankOne_le_norm_sq_smul_one`
`Matrix.vecMulVec_le_norm_sq_smul_one`]

Proof. (a) The candidate is verified to be a two-sided inverse: multiplying it by $A + t\,uu^*$ collapses, by self-adjointness of A^{-1}, to 1 plus a multiple of a rank-one operator whose scalar factor vanishes by the choice of the coefficient. For matrices one product suffices; for operators both products are computed, and together they exhibit $A + t\,uu^*$ as a unit. Taking traces and using $\operatorname{Tr}\big((A^{-1}u)(A^{-1}u)^*\big) = u^*A^{-2}u$ gives the trace identity.

(b) The first is a strictly positive operator plus a positive one; it is positive, and invertible because it dominates an invertible positive operator. For the second, the hypothesis makes the Sherman–Morrison denominator positive, so (a) shows that $A - w\,uu^*$ is invertible and exhibits its inverse as a strictly positive operator plus a positive one; an invertible operator whose inverse is positive is itself positive.

(c) A self-adjoint element a of a C*-algebra satisfies $a \preceq \|a\| \cdot 1$, and uu^* is self-adjoint with $\|uu^*\| = \|u\|^2$. For matrices the formalisation derives $uu^* \preceq \|u\|^2 \cdot 1$ from Theorem 2.1(b), since the trace of uu^* is $\|u\|^2$. □

Lemma 2.9 (Determinant of a rank-one update)**.** *Let ι be a finite set with $m = \#\iota$ elements, let $v \in \mathbb{C}^\iota$, let $A \in \mathbb{C}^{\iota\times\iota}$ be invertible, $\beta \neq 0$ a scalar and α a scalar. Then*

$$\det(\beta A + \alpha\, vv^*) = \beta^{\,m-1} \det A \,\big(\beta + \alpha\, v^* A^{-1} v\big).$$

[Lean: `Matrix.det_smul_add_smul_vecMulVec`]

Proof. Factor βA out and apply the matrix determinant lemma $\det(A + vw^*) = \det A\,(1 + w^*A^{-1}v)$, which holds because the added matrix has rank one. The dependence on α is affine; this identity is what replaces the derivative of the determinant. □

Chapter 3

Averages of quadratic forms

The bridge between the analytic side of the problem and the algebra of matrices and operators is one identity: the average of a quadratic form along a family is a trace against its Gram matrix, or against its Gram operator for a family with values in a Hilbert space. Read from right to left it says that a trace is an average, which is how the construction finds a point.

3.1 The Gram matrix

Lemma 3.1 (Two identities for the Gram matrix)**.** *Let (Ω, μ) be a measure space, let ι be a finite set, let $a : \Omega \to \mathbb{C}^{\iota}$ be square-integrable in every coordinate, and write $G(a) = \int_{\Omega} a(x)a(x)^{*}\, d\mu(x)$.*

(a) For every matrix $Q \in \mathbb{C}^{\iota \times \iota}$,

$$\int_{\Omega} a(x)^{*} Q\, a(x)\, d\mu(x) = \operatorname{Tr}\big(Q\, G(a)\big).$$

(b) For every matrix $S \in \mathbb{C}^{\iota \times \iota}$,

$$G(Sa) = S\, G(a)\, S^{*}.$$

[Lean: `Discretization.gram Discretization.integral_quadForm Discretization.integral_re_quadForm Discretization.gram_mulVec`]

Proof. (a) Expand both sides as a double sum over ι and exchange the finite sum with the integral. The integrand is integrable by Cauchy–Schwarz for L_2-functions.

(b) Expand both sides entrywise and exchange the finite sums with the integral. □

3.2 The Gram operator

For a square-integrable map $b : \Omega \to H$ into a Hilbert space the Gram matrix becomes the Gram operator $J = \int_\Omega b(x)b(x)^*\, d\mu(x)$. It is positive, since its quadratic form is nonnegative, and on a separable space its trace is $\int_\Omega \|b\|^2\, d\mu$, by Parseval and monotone convergence (Theorem 3.3(a)). The only genuine assumption on J is therefore that it is injective, which is what positive definiteness of the Gram matrix says for a finite family, and Theorem 1.1 assumes nothing else.

Lemma 3.2 (The Gram operator)**.** *Let (Ω, μ) be a measure space, let H be a complex Hilbert space, let $b : \Omega \to H$ be square-integrable, and let $J = \int_\Omega b(x)b(x)^*\, d\mu(x)$.*

(a) The operators $b(x)b(x)^*$ are Bochner integrable, and J is positive with $\langle u, Ju\rangle = \int_\Omega |\langle u, b(x)\rangle|^2\, d\mu(x)$ for every $u \in H$.

(b) Let $\Lambda \in \mathbb{R}$. Then $J \preceq \Lambda \cdot 1$ if and only if $\int_\Omega |\langle u, b(x)\rangle|^2\, d\mu(x) \le \Lambda\|u\|^2$ for every $u \in H$.

(c) J is injective if and only if every $u \in H$ with $\langle u, b(x)\rangle = 0$ for almost every x is zero.

[Lean: `ContinuousLinearMap.integrable_rankOne_self`
`ContinuousLinearMap.inner_integral_rankOne_self`
`ContinuousLinearMap.nonneg_of_inner_eq_integral`
`ContinuousLinearMap.integral_rankOne_self_le_iff`
`ContinuousLinearMap.integral_rankOne_self_injective_iff`]

Proof. (a) The operator $b(x)b(x)^*$ has norm $\|b(x)\|^2$, which is integrable. Evaluation at u and the inner product with u are continuous linear maps, so both pass under the integral, and $\langle u, b(x)b(x)^*u\rangle = |\langle u, b(x)\rangle|^2$. This quadratic form is real and nonnegative, so J is positive.

(b) By (a) both sides say that the quadratic form of the self-adjoint operator $\Lambda \cdot 1 - J$ is nonnegative.

(c) The quadratic form vanishes at u exactly when $\langle u, b(\cdot)\rangle$ vanishes almost everywhere. If $Ju = 0$ it vanishes. Conversely, J is positive, and $\langle u, Ju\rangle = \|J^{1/2}u\|^2 = 0$ gives $J^{1/2}u = 0$ and hence $Ju = 0$. □

Lemma 3.3 (The trace of the Gram operator)**.** *Let (Ω, μ) be a measure space, let H be a separable complex Hilbert space, let $b : \Omega \to H$ be square-integrable, and let $J = \int_\Omega b(x)b(x)^*\, d\mu(x)$.*

(a) J is of finite trace, and $\operatorname{Tr} J = \int_\Omega \|b(x)\|^2\, d\mu(x)$.

(b) For every positive bounded operator Q on H,

$$\int_\Omega \operatorname{Re}\big\langle b(x), Q\, b(x)\big\rangle \, d\mu(x) = \operatorname{Tr}(JQ).$$

[Lean: `Discretization.Infinite.isFiniteTracePos_of_integral_rankOne`
`Discretization.Infinite.integral_norm_sq_eq_trace`
`ContinuousLinearMap.integral_re_inner_apply`]

Proof. (a) Choose a Hilbert basis (e_k) of H. Its index set is countable, because the open balls of radius $1/2$ around the basis vectors are pairwise disjoint and a separable space holds only countably many pairwise disjoint nonempty open sets. By Tonelli and Theorem 2.4(a), $\sum_k \langle e_k, Je_k\rangle = \sum_k \int_\Omega |\langle e_k, b\rangle|^2 \, d\mu = \int_\Omega \|b\|^2 \, d\mu < \infty$, which is the finite trace and its value.

(b) Choose a countable Hilbert basis (e_k) as in (a). Pointwise $\operatorname{Re}\langle b, Qb\rangle = \|Q^{1/2}b\|^2$, which Theorem 2.4(a) turns into $\sum_k |\langle Q^{1/2}e_k, b\rangle|^2$. Sum and integral may be exchanged because all terms are nonnegative and their integrals sum to a finite value, and the quadratic form of J turns the result into $\sum_k \operatorname{Re}\langle Q^{1/2}e_k, J\, Q^{1/2}e_k\rangle$. That equals $\operatorname{Tr}(JQ)$ because both are the squared Hilbert–Schmidt norm of $Q^{1/2}J^{1/2}$, once read through the operator and once through its adjoint, by Theorem 2.4(b) and Theorem 2.6(b). Countability of the index set of the basis enters here, in the interchange. □

3.3 A reproducing kernel Hilbert space

The second family of Theorem 1.4 is the map $x \mapsto k_x$ to the kernel sections of a reproducing kernel Hilbert space. It is square-integrable with values in a separable Hilbert space, and by Theorem 3.2(a) the quadratic form of its Gram operator is $\langle g, Jg\rangle = \int_\Omega |g|^2 \, d\mu$.

Lemma 3.4 (Kernel sections)**.** *Let H be a complex Hilbert space of functions on a set Ω on which every point evaluation $f \mapsto f(x)$ is continuous.*

(a) For every $x \in \Omega$ there is $k_x \in H$ with $\langle k_x, f\rangle = f(x)$ for every $f \in H$. With $K(x,y) = \langle k_y, k_x\rangle$ one has $K(x,y) = k_x(y)$, so $k_x = K(x,\cdot)$ and $K(x,x) = \|k_x\|^2$, and $|\langle f, k_x\rangle| = |f(x)|$ for every $f \in H$.

(b) Let (Ω, μ) be a measure space, let H be separable, and let every $f \in H$ be measurable. Then $x \mapsto k_x$ is strongly measurable, and it is square-integrable if $\int_\Omega K(x,x)\, d\mu(x) < \infty$.

[Lean: `Discretization.RKHS.evalCLM` `Discretization.RKHS.kernelSection`
`Discretization.RKHS.inner_kernelSection`

[Lean: `Discretization.RKHS.norm_inner_kernelSection Discretization.RKHS.kernel Discretization.RKHS.kernel_eq Discretization.RKHS.kernel_self Discretization.RKHS.aestronglyMeasurable_kernelSection Discretization.RKHS.memLp_kernelSection`]

Proof. (a) Evaluation at x is a continuous linear functional, so the Riesz representation theorem gives k_x. Then $K(x,y) = \langle k_y, k_x\rangle = k_x(y)$, $K(x,x) = \|k_x\|^2$, and $\langle f, k_x\rangle = \overline{\langle k_x, f\rangle} = \overline{f(x)}$.

(b) Take a Hilbert basis (e_j) of H, which is countable because H is separable. Then $k_x = \sum_j \langle e_j, k_x\rangle\, e_j$ with $\langle e_j, k_x\rangle = \overline{e_j(x)}$, so $x \mapsto k_x$ is the pointwise limit of finite sums of measurable maps. Its squared norm is $K(x,x)$. □

In Lean a Hilbert space of functions is a Hilbert space H together with a linear map $\varphi : H \to \mathbb{C}^\Omega$, $\varphi(f)(x)$ being the value of f at x. Injectivity of φ is not required; in Theorem 1.4 it follows from the assumption on vanishing functions.

3.4 A point from two averages

Lemma 3.5 (A good point from two averages)**.** *Let (Ω, μ) be a measure space and let $f, g : \Omega \to \mathbb{R}$ be integrable.*

(a) If $\int_\Omega g\, d\mu < \int_\Omega f\, d\mu$, then $g(x) < f(x)$ for at least one $x \in \Omega$.

(b) If f and g are nonnegative with $\int_\Omega g\, d\mu \le \int_\Omega f\, d\mu$ and $\int_\Omega f\, d\mu > 0$, then there is an $x \in \Omega$ with $g(x) \le f(x)$ and $f(x) > 0$.

[Lean: `Discretization.exists_lt_of_integral_lt Discretization.exists_le_of_integral_le`]

Proof. (a) If $f \le g$ everywhere then $\int f \le \int g$.

(b) If $f \le g$ everywhere then the two averages agree, so the nonnegative function $g - f$ has average zero and vanishes almost everywhere. At every point where f and g agree, the assumption forces f to vanish, so f vanishes almost everywhere, contradicting $\int f > 0$. □

The paper states step (a) as: the set of admissible points has positive measure. One point is all the construction uses. When the two averages are equal, which happens in the doubly degenerate case in the proof of

Theorem 1.1, a strict inequality is no longer available and (b) takes over. Both the nonnegativity of g and the conclusion $f(x) > 0$ in (b) are needed: for $f = (1, 0)$ and $g = (2, -1)$ on two points with the counting measure both averages are 1, and the only point with $g \leq f$ has $f = 0$.

Chapter 4

The lower potential and its barrier

The first family is carried through the construction by a positive definite matrix A, which must stay above a multiple of the identity. It is controlled by a single real number, and whether a candidate point may be used is decided by a function of that point. The second family is treated in the same way by an operator, in Chapter 5.

4.1 The potential and the verifier

Definition 4.1 (The lower potential and verifier)**.** Let ι be a finite set, let $A \in \mathbb{C}^{\iota\times\iota}$ be positive definite, and let $\delta > 0$ with $A - \delta \cdot 1$ positive definite. The *lower potential* of A is $\Phi(A) = \operatorname{Tr} A^{-1}$. With $Z = (A - \delta \cdot 1)^{-1}$, the *lower verifier* at $v \in \mathbb{C}^{\iota}$ is

$$L_A^{\delta}(v) = \frac{v^* Z^2 v}{\Phi(A - \delta \cdot 1) - \Phi(A)} - v^* Z v.$$

[Lean: `Discretization.lowerPotential Discretization.lowerVerifier`]

The lower potential is large when A has a small eigenvalue. Before each step A is shrunk by $\delta \cdot 1$, which raises the potential by an amount that is computed exactly, and that amount is the budget one new sampling point is allowed to consume. The verifier decides pointwise, without recomputing any inverse, whether a candidate point stays within that budget. A bound on the potential is a bound on the matrix, $\Phi(A)^{-1} \cdot 1 \preceq A$ (Theorem 2.7(a)).

The upper potential $\Psi_J(B) = \operatorname{Tr}(JB^{-1})$ and its verifier play the same role for the second family (Theorem 5.1).

Lemma 4.2 (How the lower potential changes)**.** *Let ι be a finite set.*

(a) Let $A \in \mathbb{C}^{\iota\times\iota}$ be positive definite and let $\delta > 0$ with $A - \delta\cdot 1$ positive definite. Then

$$\Phi(A-\delta\cdot 1) - \Phi(A) = \delta\,\operatorname{Tr}\big((A-\delta\cdot 1)^{-1}A^{-1}\big) > 0.$$

(b) Let $N \in \mathbb{C}^{\iota\times\iota}$ be positive definite, $v \in \mathbb{C}^{\iota}$ and $w \geq 0$. Then

$$\Phi\big(N + w\,vv^*\big) = \Phi(N) - \frac{w}{1 + w\,v^*N^{-1}v}\,v^*N^{-2}v.$$

[Lean: `Discretization.lowerPotential_sub_eq`
`Discretization.lowerPotential_add_smul_vecMulVec`]

Proof. (a) The identity is the resolvent identity $X^{-1} - Y^{-1} = X^{-1}(Y - X)Y^{-1}$ for the matrices $X = A - \delta\cdot 1$ and $Y = A$, whose difference is $\delta\cdot 1$. The sign is Theorem 2.3(a) applied to the product of two positive definite matrices.

(b) Insert the Sherman–Morrison formula of Theorem 2.8(a) and take traces. □

Lemma 4.3 (Barrier lemma for the lower potential)**.** *Let ι be a finite set, let $A \in \mathbb{C}^{\iota\times\iota}$ be positive definite, let $0 < \delta < \Phi(A)^{-1}$ and $v \in \mathbb{C}^{\iota}$, and let $w > 0$ satisfy $1/w \leq L_A^{\delta}(v)$. Then $A - \delta\cdot 1 + w\,vv^*$ is positive definite and*

$$\Phi\big(A - \delta\cdot 1 + w\,vv^*\big) \leq \Phi(A).$$

[Lean: `Discretization.lowerPotential_update_le`]

Proof. The matrix $A-\delta\cdot 1$ is positive definite by Theorem 2.7(b), and adding $w\,vv^*$ keeps it so by Theorem 2.8(b). The closed form of Theorem 4.2(b) for $N = A - \delta\cdot 1$, together with the positive gain of Theorem 4.2(a), reduces the inequality to one between real numbers, which is the hypothesis $1/w \leq L_A^{\delta}(v)$ multiplied out. □

4.2 The verifier passes on average

A point may be used if the upper verifier stays below the lower one. Such a point exists because the lower verifier is large on average.

Lemma 4.4 (The lower verifier passes on average)**.** *Let (Ω, μ) be a measure space, let ι be a finite nonempty set, and let $a : \Omega \to \mathbb{C}^\iota$ be square-integrable in every coordinate with $\int_\Omega a(x)a(x)^* \, d\mu(x) = 1$. Let $A \in \mathbb{C}^{\iota\times\iota}$ be positive definite and $0 < \delta < \Phi(A)^{-1}$.*

(a) The lower verifier satisfies

$$\int_\Omega L_A^\delta\big(a(x)\big) d\mu(x) > \frac{1}{\delta} - \Phi(A).$$

(b) Let $U : \Omega \to \mathbb{R}$ be integrable with $\int_\Omega U \, d\mu \le 1/\delta - \Phi(A)$. Then there is an $x \in \Omega$ with $U(x) < L_A^\delta\big(a(x)\big)$.

[Lean: `Discretization.integral_lowerVerifier_gt Discretization.UpperBarrier.exists_admissible_point`]

Proof. (a) Write $Y = A^{-1}$ and $Z = (A - \delta \cdot 1)^{-1}$. Theorem 3.1(a) turns the average into $\operatorname{Tr}(Z^2)/(\delta \operatorname{Tr}(YZ)) - \operatorname{Tr} Z$, the denominator being the gain of Theorem 4.2(a). Writing $t = \operatorname{Tr}(YZ)$ and $u = \operatorname{Tr}(YZ^2)$, the claim reduces to $\delta t^2 < u$, which is what Theorem 2.3(b) gives together with $\delta\Phi(A) < 1$.

(b) The lower verifier is a quadratic form in $a(x)$ and hence integrable. By (a) its average exceeds $1/\delta - \Phi(A) \ge \int_\Omega U \, d\mu$, and Theorem 3.5(a) gives the point. □

In the construction U is the upper verifier along the second family, whose average is below $1/\zeta + \Psi_J(B)$ by Theorem 5.4; the gap condition $1/\zeta + \Psi_J(B) \le 1/\delta - \Phi(A)$ then makes (b) applicable. Any weight w between the two reciprocals is admissible for Theorem 4.3 and for Theorem 5.3, and the construction takes the largest one, $w = 1/L_A^\delta$. Part (b) compares two averages of real functions and nothing else; in Lean it is proved once for every upper state, `Discretization.UpperBarrier.exists_admissible_point`.

4.3 From the potential to the frame bound

Lemma 4.5 (From the potential to the lower frame bound)**.** *Let ι be a finite nonempty set, let $n \ge 1$, and let $v_1, \dots, v_n \in \mathbb{C}^\iota$ and $w_1, \dots, w_n \in \mathbb{R}$.*

Let $r > 0$ with $r \neq 1$, put $\delta = (1-r)/n$ and $c_0 = \delta(nr^2+1)/r$, and let $A_n = c_0 \cdot 1 - n\delta \cdot 1 + \sum_{i=1}^{n} w_i\, v_i v_i^*$. If A_n is positive definite and $\Phi(A_n) \le r/\delta$, then

$$(1-r)^2 \cdot 1 \;\preceq\; \sum_{i=1}^{n} w_i\, v_i v_i^*.$$

[Lean: `Discretization.lower_bound_of_state Discretization.lower_frame_bound`]

Proof. By Theorem 2.7(a), $\Phi(A_n) \le r/\delta$ gives $(\delta/r) \cdot 1 \preceq A_n$; inserting the definition of A_n gives $(\delta/r + n\delta - c_0) \cdot 1 \preceq \sum_i w_i\, v_i v_i^*$, and $\delta/r + n\delta - c_0 = n\delta(1-r) = (1-r)^2$. The arithmetic is in `Discretization/Parameters.lean`. □

For $r = \sqrt{(m-1)/n}$ one has $nr^2 + 1 = m$; the lemma is stated in r because its proof never unfolds the square root.

Chapter 5

The upper potential and its barrier

The second family is carried through the construction by a strictly positive operator B on a separable Hilbert space H, which must stay above a multiple of the Gram operator $J = \int_\Omega b(x)b(x)^*\, d\mu(x)$ of Theorem 3.2, as the matrix A of Chapter 4 stays above a multiple of the identity. It is controlled by a single real number, and whether a candidate point may be used is decided by a function of that point. Nothing in the argument ties the number of points to the size of the second family: it is governed by the effective dimension alone. A finite second family is the case $H = \mathbb{C}^\kappa$, where every operator below is a matrix and every trace a matrix trace.

5.1 The upper potential and the verifier

Definition 5.1 (Upper potential and verifier of an operator)**.** Let H be a complex Hilbert space, let J be a positive injective operator on H of finite trace, let B be a strictly positive operator on H, that is positive and invertible, and let $\zeta > 0$. The *upper potential* of B is $\Psi_J(B) = \mathrm{Tr}\big(JB^{-1}\big)$, and with $X = (B + \zeta \cdot J)^{-1}$ the *upper verifier* at $u \in H$ is

$$U_B^\zeta(u) = \frac{\mathrm{Re}\langle u, XJXu\rangle}{\Psi_J(B) - \Psi_J(B + \zeta \cdot J)} + \mathrm{Re}\langle u, Xu\rangle.$$

[Lean: `Discretization.Infinite.upperPotential Discretization.Infinite.upperVerifier Discretization.upperPotential Discretization.upperVerifier`]

The upper potential is large when B is small in the directions where J is large. Before each step B is grown by $\zeta \cdot J$, which lowers the potential by an amount that is computed exactly, and the verifier decides whether a candidate point stays within that budget, as Φ and L_A^δ do for the first family (Theorem 4.1). A bound on the potential is a bound on the operator, $\Psi_J(B)^{-1} \cdot J \preceq B$ (Theorem 2.7(c)). In Lean both notions are also defined directly for matrices (`Discretization.upperPotential`, `Discretization.upperVerifier`). Lean proves the next three lemmas a second time for matrices, where the trace is the matrix trace, and the read-off Theorem 5.5 once for every upper state; the matrix proofs serve the finite case and keep the sparsification theorem elementary.

Lemma 5.2 (How the upper potential changes)**.** *Let H be a nonzero complex Hilbert space, let J be a positive injective operator on H of finite trace, and let B be a strictly positive operator on H.*

(a) For every $\zeta \geq 0$ the operator $B + \zeta \cdot J$ is strictly positive and

$$\Psi_J(B) - \Psi_J(B + \zeta \cdot J) = \zeta \operatorname{Tr}\big((JB^{-1}J)(B + \zeta \cdot J)^{-1}\big),$$

and for $\zeta > 0$ the right-hand side is positive.

(b) Let C be a strictly positive operator on H, let $u \in H$, and let $w \geq 0$ with $w \operatorname{Re}\langle u, C^{-1}u\rangle < 1$. Then

$$\Psi_J\big(C - w\,uu^*\big) = \Psi_J(C) + \frac{w}{1 - w \operatorname{Re}\langle u, C^{-1}u\rangle} \operatorname{Re}\langle u, C^{-1}JC^{-1}u\rangle.$$

[Lean: `Discretization.Infinite.upperPotential_sub_eq`
`Discretization.Infinite.upperPotential_add_smul_lt`
`ContinuousLinearMap.isStrictlyPositive_add_smul`
`Discretization.Infinite.upperPotential_sub_smul_rankOne`
`Discretization.upperPotential_sub_eq`
`Discretization.upperPotential_add_smul_lt Matrix.PosDef.add_smul`
`Discretization.upperPotential_sub_smul_vecMulVec`]

Proof. (a) The operator $B+\zeta\cdot J$ dominates B, so it is positive, and invertible because it dominates an invertible positive operator. The identity is the resolvent identity, which holds in any ring. The right-hand side is the trace of a product of two positive operators, the first of them $JB^{-1}J$, which is again of finite trace and positive; Theorem 2.6(b) gives the sign, and injectivity of J makes it strict.

(b) Insert the Sherman–Morrison formula of Theorem 2.8(a) with $t = -w$ and take traces, as for the lower potential in Theorem 4.2(b). □

Lemma 5.3 (Barrier lemma for the upper potential)**.** *Let H be a nonzero complex Hilbert space, let J be a positive injective operator on H of finite trace, let B be a strictly positive operator on H, let $\zeta > 0$, $u \in H$, and let $w > 0$ satisfy $U_B^\zeta(u) \le 1/w$. Then $B + \zeta \cdot J - w\,uu^*$ is strictly positive and*

$$\Psi_J\big(B + \zeta \cdot J - w\,uu^*\big) \le \Psi_J(B).$$

[Lean: `Discretization.Infinite.upperPotential_update_le`
`Discretization.upperPotential_update_le`]

Proof. Write $X = (B + \zeta \cdot J)^{-1}$. The first summand of $U_B^\zeta(u)$ is positive for $u \neq 0$, because $\mathrm{Re}\langle u, XJXu\rangle = \|J^{1/2}Xu\|^2$ and J is injective, so the hypothesis gives $w\,\mathrm{Re}\langle u, Xu\rangle < 1$: the Sherman–Morrison denominator is positive, and Theorem 2.8(b), applied to the strictly positive operator $B + \zeta \cdot J$ of Theorem 5.2(a), makes $B + \zeta \cdot J - w\,uu^*$ strictly positive. Finally Theorem 5.2(b) for $C = B + \zeta \cdot J$ and (a) reduce the inequality to the hypothesis $U_B^\zeta(u) \le 1/w$ multiplied out, as in the proof of Theorem 4.3. □

5.2 The upper verifier passes on average

For a square-integrable b with values in a separable space, the Gram operator $J = \int_\Omega b(x)b(x)^*\,d\mu(x)$ is positive and of finite trace (Theorem 3.3(a)), so its upper potential is defined.

Lemma 5.4 (The upper verifier passes on average)**.** *Let (Ω, μ) be a measure space, let H be a nonzero separable complex Hilbert space, let $b : \Omega \to H$ be square-integrable with injective Gram operator $J = \int_\Omega b(x)b(x)^*\,d\mu(x)$, let B be a strictly positive operator on H, and let $\zeta > 0$. Then $x \mapsto U_B^\zeta(b(x))$ is integrable and*

$$\int_\Omega U_B^\zeta\big(b(x)\big)\,d\mu(x) < \frac{1}{\zeta} + \Psi_J(B).$$

[Lean: `Discretization.Infinite.integrable_upperVerifier`
`Discretization.Infinite.integral_upperVerifier_lt`
`Discretization.integrable_upperVerifier`
`Discretization.integral_upperVerifier_lt`]

Proof. Both summands of the verifier are quadratic forms in $b(x)$, bounded by a multiple of $\|b(x)\|^2$, so the verifier is integrable. Write $W = B^{-1}$ and $X = (B + \zeta \cdot J)^{-1}$, and let $E = \Psi_J(B) - \Psi_J(B + \zeta \cdot J)$ be the gain of Theorem 5.2(a). By Theorem 3.3(b) for $Q = XJX$ and $Q = X$ the average equals $\mathrm{Tr}(JXJX)/E + \Psi_J(B+\zeta\cdot J)$. Since $X \preceq W$ by the antitonicity of the inverse, as in Theorem 2.1(c), we have $JXJ \preceq JWJ$, and Theorem 2.6(c) bounds the numerator by $\mathrm{Tr}(JWJX) = E/\zeta$. So the first summand is at most $1/\zeta$, and the second is strictly smaller than $\Psi_J(B)$ by Theorem 5.2(a). □

5.3 From the upper potential to the frame bound

Lemma 5.5 (From the potential to the upper frame bound)**.** *Let H be a nonzero complex Hilbert space, let J be a positive injective operator on H of finite trace, let $\Lambda > 0$ with $J \preceq \Lambda \cdot 1$, let $n \geq 1$, let $s \geq 1/n$ with $\mathrm{Tr}\, J = \Lambda(ns^2 + 1)$, put $\zeta = (1 + s)/n$ and $d_0 = \zeta\, \mathrm{Tr}\, J/s$, let $u_1, \ldots, u_n \in H$ and $w_1, \ldots, w_n \in \mathbb{R}$, and let $B_n = d_0 \cdot 1 + n\zeta \cdot J - \sum_{i=1}^n w_i\, u_i u_i^*$. If B_n is strictly positive and $\Psi_J(B_n) \leq s/\zeta$, then*

$$\sum_{i=1}^{n} w_i\, u_i u_i^* \;\preceq\; (1 + s)^2 \Lambda \cdot 1.$$

[Lean: `Discretization.UpperBarrier.upper_bound_of_state` `Discretization.UpperBarrier.upper_frame_bound`]

Proof. By Theorem 2.7(c), $\Psi_J(B_n) \leq s/\zeta$ gives $(\zeta/s) \cdot J \preceq B_n$, hence $\sum_i w_i\, u_i u_i^* \;\preceq\; d_0 \cdot 1 + (n\zeta - \zeta/s) \cdot J$. The coefficient $n\zeta - \zeta/s = \zeta(n - 1/s)$ is nonnegative exactly because $s \geq 1/n$, so $J \preceq \Lambda \cdot 1$ may be applied to it, and $d_0 + (n\zeta - \zeta/s)\Lambda = \zeta\Lambda(ns + n) = (1 + s)^2\Lambda$. In Lean this read-off is proved once for every upper state, `Discretization.UpperBarrier.upper_bound_of_state`; the arithmetic is in `Discretization/Parameters.lean`. □

For $s = \sqrt{(M-1)/n}$ one has $ns^2 + 1 = M$, and $s \geq 1/n$ is the condition $M \geq 1 + 1/n$; the lemma is stated in s because its proof never unfolds the square root.

Chapter 6

The construction and the proof

The construction carries a matrix over ι for the first family and an operator on H for the second, each controlled by its potential. This chapter runs it and proves Theorem 1.1.

6.1 The barrier iteration

Proposition 6.1 (The barrier iteration)**.** *Let (Ω, μ) be a measure space, let ι be a finite nonempty set, and let H be a nonzero separable complex Hilbert space. Let $a : \Omega \to \mathbb{C}^\iota$ be square-integrable in every coordinate with $\int_\Omega a(x)a(x)^*\, d\mu(x) = 1$, and let $b : \Omega \to H$ be square-integrable with injective Gram operator $J = \int_\Omega b(x)b(x)^*\, d\mu(x)$. Let $A_0 \in \mathbb{C}^{\iota\times\iota}$ be positive definite, let B_0 be a strictly positive operator on H, and let $\delta, \zeta > 0$ satisfy the initial gap condition*

$$\frac{1}{\zeta} + \Psi_J(B_0) \;\le\; \frac{1}{\delta} - \Phi(A_0).$$

For points $x_1, \dots, x_k \in \Omega$ and weights $w_1, \dots, w_k$ write

$$A_k = A_0 - k\delta\cdot 1 + \sum_{i=1}^{k} w_i\, a(x_i)a(x_i)^*, \qquad B_k = B_0 + k\zeta\cdot J - \sum_{i=1}^{k} w_i\, b(x_i)b(x_i)^*.$$

Then for every $k \ge 0$ there are points $x_1, \dots, x_k \in \Omega$ and weights $w_1, \dots, w_k > 0$ such that A_k is positive definite, B_k is strictly positive, $\Phi(A_k) \le \Phi(A_0)$ and $\Psi_J(B_k) \le \Psi_J(B_0)$.

[Lean: `Discretization.UpperBarrier.exists_points_weights`]

Proof. Induction on k; for $k = 0$ there is nothing to show. The invariant keeps the gap open, $1/\zeta + \Psi_J(B_k) \le 1/\zeta + \Psi_J(B_0) \le 1/\delta - \Phi(A_0) \le 1/\delta - \Phi(A_k)$, and in particular $\delta < \Phi(A_k)^{-1}$. By Theorem 5.4 the function $U(x) = U^{\zeta}_{B_k}(b(x))$ is integrable with average below $1/\zeta + \Psi_J(B_k)$, so Theorem 4.4(b) produces a point x_{k+1} with $U(x_{k+1}) < L^{\delta}_{A_k}(a(x_{k+1}))$. The weight is $w_{k+1} = 1/L^{\delta}_{A_k}(a(x_{k+1}))$, positive because $U \ge 0$, and Theorem 4.3 and Theorem 5.3 restore the invariant. □

The number of steps is arbitrary; only the initial gap condition drives it. What ties the number of points to the frame bounds is the choice of the initial states and of the two shifts, made in the proof of Theorem 1.1 below.

The induction and the read-off use only a few properties of the upper state: that $\Psi_J(B)$ and U^{ζ}_B are positive, Theorem 5.3 and Theorem 5.4, the bound $\Psi_J(B)^{-1} \cdot J \preceq B$ (Theorem 2.7(c)) with the potential of a multiple of the identity, and the crude bound $b(y)b(y)^* \preceq \|b(y)\|^2 \cdot 1$ with $\int_\Omega \|b\|^2\, d\mu = \operatorname{Tr} J$. In Lean they are collected in `Discretization.UpperBarrier`, and the construction, the read-off of the upper frame bound and the theorem with its edge cases are proved once for every such upper state. The operators of Chapter 5 are one instance (`Discretization.Infinite.operatorUpperBarrier`); the matrices over a finite set κ, with the matrix proofs of the upper lemmas, are another (`Discretization.matrixUpperBarrier`).

6.2 Proof of the discretization theorem

Proof of Theorem 1.1. The first family is normalized first. Replace a by $I^{-1/2}a$, whose Gram matrix is $I^{-1/2} I\, I^{-1/2} = 1$ by Theorem 3.1(b). Once points and weights are found for it, conjugating the lower bound by $I^{1/2}$ preserves the Loewner order and turns $(1-r)^2 \cdot 1$, with $r = \sqrt{(m-1)/n}$, into $(1-r)^2 \cdot I$; the upper bound does not involve the first family. From now on $\int_\Omega a(x)a(x)^*\, d\mu(x) = 1$.

By Theorem 3.2(a) and Theorem 3.3(a), J is positive and of finite trace with $\operatorname{Tr} J = \int_\Omega \|b\|^2\, d\mu = M\Lambda$, so the upper potential and verifier of Theorem 5.1 are defined. Put $r = \sqrt{(m-1)/n}$, so that $nr^2 + 1 = m$, and for $M \ge 1$ put $s = \sqrt{(M-1)/n}$, so that $ns^2 + 1 = M$.

Case $m \ge 2$, $M \ge 1 + 1/n$ (`Discretization.bss_generalized_of_gram_eq_one`), the potential argument. Put

$$\delta = \tfrac{1-r}{n}, \quad \zeta = \tfrac{1+s}{n}, \quad A_0 = \tfrac{\delta m}{r} \cdot 1, \quad B_0 = \tfrac{\zeta \operatorname{Tr} J}{s} \cdot 1.$$

Since $2 \le m \le n$ we have $0 < r < 1$, and $M \ge 1 + 1/n$ gives $s \ge 1/n$. Then $\Phi(A_0) = r/\delta$ and $\Psi_J(B_0) = s/\zeta$, so the initial gap closes exactly, $1/\delta - \Phi(A_0) = n = 1/\zeta + \Psi_J(B_0)$. Theorem 6.1 with $k = n$ gives points and weights with $\Phi(A_n) \le r/\delta$ and $\Psi_J(B_n) \le s/\zeta$, and Theorem 4.5 and Theorem 5.5 turn them into the two bounds. The remaining three cases need one potential only: the verifier on the other side is replaced by a constant, whose average is again n. For $m = 1$ the lower constant $(1 - r)^2$ is 1.

Case $m = 1$, $M \ge 1{+}1/n$ (`Discretization.bss_generalized_of_unique`). The lower verifier is replaced by $L(x) = n|a(x)|^2$, whose average is n because the single function is normalized. The induction then tracks only the upper state, using Theorem 5.3 and the average of Theorem 5.4 with Theorem 3.5(a), and takes the largest weight the constant allows, $1/w_i = L(x_i)$. Each point therefore contributes exactly $1/n$ to the lower bound, which becomes the identity $\sum_i w_i |a(x_i)|^2 = 1$. The upper bound is Theorem 5.5.

Case $m \ge 2$, $M \le 1{+}1/n$ (`Discretization.bss_generalized_of_small_dim`). The upper verifier is replaced by $U(x) = n\|b(x)\|^2 / \operatorname{Tr} J$, whose average is n. The induction tracks only the lower matrix, with Theorem 4.3, and the lower bound is Theorem 4.5. The weights it produces satisfy $w_i U(x_i) \le 1$, that is $w_i \|b(x_i)\|^2 \le \operatorname{Tr} J/n$. Summing over the n points and using Theorem 2.8(c) gives $\sum_i w_i\, b(x_i) b(x_i)^* \preceq \operatorname{Tr} J \cdot 1 = M\Lambda \cdot 1$, and $M \le 1 + 1/n$ forces $s = \sqrt{(M-1)/n} \le 1/n$ and hence $M = 1 + ns^2 \le 1 + s \le (1+s)^2$.

Case $m = 1$, $M \le 1{+}1/n$ (`Discretization.bss_generalized_of_unique_of_small_dim`). Both verifiers are now the constants $L(x) = n|a(x)|^2$ and $U(x) = n\|b(x)\|^2 / \operatorname{Tr} J$, of equal average n, and neither depends on the state of the construction. Theorem 3.5(b) produces a single point y with $U(y) \le L(y)$ and $L(y) > 0$; using it n times with the weight $w = 1/L(y)$ makes the lower bound the identity $\sum_i w_i |a(x_i)|^2 = 1$, while $w\, U(y) \le 1$ and Theorem 2.8(c) give the upper bound as in the previous case. No induction is needed. □

In Lean the normalisation is `Discretization.bss_of_gram_eq_one`, stated once for an arbitrary conclusion about the second family, and the four cases are proved once for every upper state of `Discretization.UpperBarrier`. A finite second family, $H = \mathbb{C}^\kappa$, is proved with the matrix instance, which keeps Theorem 1.3 free of operators.

Chapter 7

Discretization of the uniform norm and the L_p-norms

Let Ω be an arbitrary set, ι a finite nonempty index set with $m = \#\iota$ elements, and $a = (a_k)_{k\in\iota}$ a family of bounded functions $\Omega \to \mathbb{C}$ whose coordinate functions are linearly independent. The theorem of Kiefer and Wolfowitz [4] gives a finitely supported probability measure on Ω whose L_2 norm dominates the uniform norm on the span with the constant $\sqrt{m+\varepsilon}$ (Theorem 7.1); the complex and non-compact case is treated in [5]. Thinned with the one-sided discretization, the measure becomes $n \geq m$ equally weighted points (Theorem 7.2), and its mixture with a probability measure μ, thinned in the same way, discretizes all $L_p(\mu)$-norms at once (Theorem 7.4).

Theorem 7.1 (Optimal design (Kiefer–Wolfowitz))**.** *Let ι be a finite nonempty index set with $m = \#\iota$ elements, let Ω carry a measurable structure, let $a = (a_k)_{k\in\iota}$ be bounded measurable functions on it with linearly independent coordinate functions, and let $\varepsilon > 0$. Then there is a finitely supported probability measure $\varrho = \sum_{k=1}^{N} w_k \delta_{x_k}$ on Ω whose Gram matrix $\int_\Omega a(y)a(y)^*\, d\varrho(y) = \sum_k w_k\, a(x_k)a(x_k)^*$ is positive definite and such that*

$$|f(y)|^2 \;\leq\; (m+\varepsilon)\int_\Omega |f|^2\, d\varrho$$

for every $y \in \Omega$ and every $f \in \operatorname{span}\{a_k\}$. If Ω is a compact space whose σ-algebra contains the open sets and the a_k are continuous, the same holds with $\varepsilon = 0$. In both cases ϱ can be taken with at most $N = m^2 + 1$ atoms.

[Lean: `Discretization.KieferWolfowitz.exists_optimal_measure`
`Discretization.KieferWolfowitz.exists_optimal_measure_of_compact`

`Discretization.KieferWolfowitz.exists_optimal_design_card_le`
`Discretization.KieferWolfowitz.exists_optimal_design_of_compact_card_le]`

The proof is given in Section 7.1. Thinning the measure with Theorem 1.5 gives the bound with any number $n \geq m$ of equally weighted points, the case $p = \infty$ of Proposition 8 of [3]. For $n = 2m$ points Krieg, Pozharska, Ullrich and Ullrich [5, Theorem 2] prove $\|f\|_\infty \leq 42\big(\sum_{i=1}^{2m} |f(x_i)|^2\big)^{1/2}$, and Proposition 8 of [3] improves the constant.

Theorem 7.2 (Discretization of the uniform norm with $n \geq m$ points)**.** *Let ι be a finite nonempty index set with $m = \#\iota$ elements, let $a = (a_k)_{k\in\iota}$ be bounded functions on a set Ω with linearly independent coordinate functions, and let $\varepsilon > 0$. Then for every $n \geq m$ there are points $x_1, \dots, x_n \in \Omega$, not necessarily distinct, such that*

$$|f(y)|^2 \;\leq\; \frac{m+\varepsilon}{\big(1-\sqrt{(m-1)/n}\big)^2} \cdot \frac{1}{n}\sum_{i=1}^{n} |f(x_i)|^2$$

for every $y \in \Omega$ and every $f \in \operatorname{span}\{a_k\}$. If Ω is a compact space and the a_k are continuous, the same holds with $\varepsilon = 0$.

[Lean: `Discretization.exists_uniform_discretization_by_l2`
`Discretization.exists_uniform_discretization_of_compact_by_l2`]

Proof. Give Ω the σ-algebra of all its subsets, or the Borel sets in the compact case, and take the measure $\varrho = \sum_k w'_k \delta_{x'_k}$ of Theorem 7.1, with the constant $K = m + \varepsilon$, where $\varepsilon = 0$ in the compact case, and positive definite Gram matrix G. On the index set $\{1, \dots, N\}$ the weights define the probability measure $\sum_k w'_k \delta_k$, with respect to which the family $k \mapsto a(x'_k)$ has Gram matrix G and $|f(x'_k)|^2$ has integral $\sum_k w'_k |f(x'_k)|^2 = \int_\Omega |f|^2 \, d\varrho$. Theorem 1.5 for this family gives n indices, that is points $x_1, \dots, x_n$ of ϱ, with

$$(1-r)^2 \int_\Omega |f|^2 \, d\varrho \;\leq\; \frac{1}{n}\sum_{i=1}^{n} |f(x_i)|^2, \qquad r = \sqrt{(m-1)/n},$$

for every $f \in \operatorname{span}\{a_k\}$, and $r < 1$ because $n \geq m$. The bound of Theorem 7.1 finishes: $|f(y)|^2 \leq K \int_\Omega |f|^2 \, d\varrho \leq K(1-r)^{-2} \frac{1}{n} \sum_i |f(x_i)|^2$. □

For $n = 2m$ one has $r < 1/\sqrt{2}$, hence $(1-r)^{-2}/(2m) < (1+\sqrt{2})^2/m$, and

$$\|f\|_\infty \;\le\; (1+\sqrt{2})\sqrt{1+\varepsilon/m}\,\Big(\sum_{i=1}^{2m} |f(x_i)|^2\Big)^{1/2} \qquad \text{for every } f \in \operatorname{span}\{a_k\}:$$

the uniform norm on the span is dominated by the Euclidean norm of $2m$ sample values. Since $r < 1/\sqrt{2}$ is strict, a small enough ε gives the constant $1+\sqrt{2}$ of [3] itself.

An average is at most the largest of its terms, so the same points control the uniform norm by the largest sample value. For $n = m$ points such a bound is due to Novak [7, Lemma 1.2.2], a form of Auerbach's lemma, which gives points with $\|f\|_\infty \le (m+\varepsilon)\max_{1\le i\le m}|f(x_i)|$.

Corollary 7.3 (The uniform norm by the largest sample value)**.** *Let ι be a finite nonempty index set with $m = \#\iota$ elements, let $a = (a_k)_{k\in\iota}$ be bounded functions on a set Ω with linearly independent coordinate functions, and let $\varepsilon > 0$. Then for every $n \ge m$ there are points $x_1, \dots, x_n \in \Omega$, not necessarily distinct, such that*

$$|f(y)|^2 \;\le\; \frac{m+\varepsilon}{\big(1-\sqrt{(m-1)/n}\big)^2}\cdot \max_{1\le i\le n}|f(x_i)|^2$$

for every $y \in \Omega$ and every $f \in \operatorname{span}\{a_k\}$. If Ω is a compact space and the a_k are continuous, the same holds with $\varepsilon = 0$.

[Lean: `Discretization.exists_uniform_discretization_by_max`
`Discretization.exists_uniform_discretization_of_compact_by_max`]

Proof. Take the points of Theorem 7.2: the average $\frac{1}{n}\sum_{i=1}^n |f(x_i)|^2$ is at most its largest term (`Discretization.inv_mul_sum_le_iSup`). □

For $n = 2m$ points and a small enough ε this gives the bound of [3],

$$\|f\|_\infty \;\le\; (2+\sqrt{2})\sqrt{m}\,\max_{1\le i\le 2m}|f(x_i)| \qquad \text{for every } f \in \operatorname{span}\{a_k\}.$$

Between $p = 2$, Theorem 1.5, and $p = \infty$ lie the L_p-norms, Proposition 8 of [3]. The paper takes n points for μ and n for the measure of Theorem 7.1; the mixture $(\mu+\varrho)/2$ serves both with n points.

Corollary 7.4 (Discretization of the L_p-norms)**.** *Let (Ω,μ) be a probability space, let ι be a finite nonempty set with $m=\#\iota$ elements, and let $a=(a_k)_{k\in\iota}$ be bounded measurable functions on Ω with positive definite Gram matrix $\int_\Omega a(x)a(x)^*\,d\mu(x)$. Then for every $n\ge m$ there are points $x_1,\dots,x_n\in\Omega$, not necessarily distinct, with*

$$\kappa\,\|f\|_{L_p(\mu)}\;\le\;m^{1/2-1/p}\Big(\frac{2}{n}\sum_{i=1}^{n}|f(x_i)|^2\Big)^{1/2},\qquad \kappa=\max\Big(1-\sqrt{m/n},\,\frac{1}{2m}\Big),$$

for every $f\in\operatorname{span}\{a_k\}$ and every $2\le p\le\infty$.
[Lean: `Discretization.exists_lp_discretization`]

Proof. Put $r=\sqrt{(m-1)/n}$ and $s=1-r$. First, $0<\kappa<s$: the term $1-\sqrt{m/n}$ lies below s because $m-1<m$, and $1/(2m)<s$ because $r\le\sqrt{(m-1)/m}$ and $(1-\frac{1}{2m})^2=\frac{m-1}{m}+\frac{1}{4m^2}$. Put $\varepsilon=m(s^2/\kappa^2-1)>0$, so that $\kappa^2(m+\varepsilon)=ms^2$.

The functions a_k are linearly independent on Ω: if $f=0$ everywhere, then $\int_\Omega|f|^2\,d\mu=c^*Ic=0$ by Theorem 3.1(a), so $c=0$. Let ϱ be the measure of Theorem 7.1 for this ε, with positive definite Gram matrix and $|f(y)|^2\le(m+\varepsilon)\int_\Omega|f|^2\,d\varrho$ at every point. The mixture $\nu=(\mu+\varrho)/2$ is a probability measure whose Gram matrix, the mean of the two, is positive definite, and the bounded functions a_k are square-integrable for it. Theorem 1.5 for ν gives points $x_1,\dots,x_n$ with

$$s^2(A{+}B)\;\le\;X,\qquad A=\int_\Omega|f|^2\,d\mu,\quad B=\int_\Omega|f|^2\,d\varrho,\quad X=\frac{2}{n}\sum_{i=1}^{n}|f(x_i)|^2.$$

For $p=\infty$: $\kappa^2|f(y)|^2\le\kappa^2(m+\varepsilon)B\le ms^2(A+B)\le mX$. For $2\le p<\infty$ put $q=p/2-1\ge0$. Since $|f(y)|^2\le(m+\varepsilon)B$ at every point, $|f|^p\le((m+\varepsilon)B)^q|f|^2$, and integration against μ gives $\int_\Omega|f|^p\,d\mu\le((m+\varepsilon)B)^qA$. With $B^qA\le(A+B)^{q+1}$ and $\kappa\le s$,

$$\kappa^p\int_\Omega|f|^p\,d\mu\;\le\;\big(\kappa^2(m{+}\varepsilon)\big)^q\kappa^2(A{+}B)^{q+1}\;\le\;m^q\big(s^2(A{+}B)\big)^{q+1}\;\le\;m^qX^{q+1},$$

and $q+1=p/2$. In Lean the finite p are stated for the p-th powers and $p=\infty$ at every point. □

For $n=2m$ one has $\kappa\ge1-1/\sqrt{2}$, hence

$$\|f\|_{L_p(\mu)}\;\le\;(2+\sqrt{2})\,m^{1/2-1/p}\Big(\frac{1}{m}\sum_{i=1}^{2m}|f(x_i)|^2\Big)^{1/2}\qquad\text{for every } 2\le p\le\infty.$$

7.1 The proof of the Kiefer–Wolfowitz theorem

The measure of Theorem 7.1 is one whose Gram matrix has an almost maximal determinant. Mixing in a further point cannot increase that determinant by much, and the matrix determinant lemma turns this into the bound $a(y)^*G^{-1}a(y) \le m+\varepsilon$ of Theorem 7.5, from which the theorem follows. No topology and no measurable structure on Ω are used until the very end: a finitely supported measure is a finite list of points with weights.

A *design* consists of finitely many points $x_1,\dots,x_N \in \Omega$ and weights $w_1,\dots,w_N \ge 0$ with $\sum_k w_k = 1$, and its *Gram matrix* is $G = \sum_k w_k\, a(x_k)a(x_k)^*$, the Gram matrix of a in L_2 of the measure $\sum_k w_k\delta_{x_k}$; in Lean it is `Discretization.KieferWolfowitz.designGram`. Its quadratic form is the weighted sum of squares,

$$c^*Gc = \sum_k w_k\,|\langle c, a(x_k)\rangle|^2 \qquad \text{for every } c \in \mathbb{C}^\iota,$$

since the quadratic form of a rank-one matrix uu^* at c is $|\langle c,u\rangle|^2$ (`Discretization.KieferWolfowitz.re_quadForm_designGram`). The statements below spell out points, weights and Gram matrix.

Proposition 7.5 (Almost maximal determinant)**.** *Let ι be a finite nonempty index set with $m = \#\iota$ elements, let $a = (a_k)_{k\in\iota}$ be bounded functions on a set Ω whose coordinate functions are linearly independent, and let $\varepsilon > 0$. Then there are points $x_1,\dots,x_N \in \Omega$ and weights $w_1,\dots,w_N \ge 0$ with $\sum_k w_k = 1$ such that $G = \sum_k w_k\, a(x_k)a(x_k)^*$ is positive definite and*

$$a(y)^*G^{-1}a(y) \;\le\; m+\varepsilon \qquad \textit{for all } y \in \Omega.$$

[Lean: `Discretization.KieferWolfowitz.exists_design_quadForm_inv_le`]

Proof. First, some design has a positive definite Gram matrix (`Discretization.KieferWolfowitz.exi`
Linear independence of the coordinate functions says that $c = 0$ is the only coefficient vector with $\langle c, a(y)\rangle = 0$ for all $y \in \Omega$. It therefore suffices to find points annihilating no nonzero coefficient vector, for then the uniform distribution on them has $c^*Gc > 0$ for every $c \ne 0$ by the formula for c^*Gc. Descending induction on dimension supplies them: while some nonzero vector survives the points chosen so far, the hypothesis gives a point at which it does not vanish, and adding that point lowers the dimension.

The determinants of the Gram matrices of designs form a set of real numbers which is bounded above, since the entries of every Gram matrix

are bounded by C^2 and the Leibniz formula bounds a determinant by its entries, and which contains a positive number by the first step. Let S be its supremum and take a design whose determinant exceeds $S/(1+\eta)$, where $\eta = \varepsilon^2/(4m(m+\varepsilon-1))$; its Gram matrix G is positive definite because its determinant is positive.

Appending a point y with weight $\alpha \in [0,1)$ and scaling the old weights by $1-\alpha$ is again a design, with Gram matrix $(1-\alpha)G + \alpha\, a(y)a(y)^*$. By Theorem 2.9 its determinant is $\det G$ multiplied by $(1-\alpha)^{m-1}(1+\alpha(t-1))$ with $t = a(y)^* G^{-1} a(y)$, and that multiple cannot exceed $S/\det G < 1+\eta$. If t were larger than $m+\varepsilon$, the weight of Theorem 7.6 would make it at least $1+\eta$. □

The proof above rests on the following estimate for the gain from mixing in one point.

Lemma 7.6 (Gain from mixing in one point)**.** *Let $m \geq 1$ and $t > m$ be real. Then $\alpha = (t-m)/(2m(t-1))$ lies in $[0,1)$ and*

$$1 + \frac{(t-m)^2}{4m(t-1)} \;\leq\; (1-\alpha)^{m-1}\big(1+\alpha(t-1)\big).$$

[Lean: `Discretization.KieferWolfowitz.exists_mix_ge`]

Proof. By Bernoulli's inequality $(1-\alpha)^{m-1} \geq 1-(m-1)\alpha$, and the difference of the two sides of the claim is then the explicit nonnegative quantity $(t-m)^2/(4m^2(t-1))$. The weight α is the one for which the linear gain $\alpha(t-m)$ is twice the quadratic loss $(m-1)(t-1)\alpha^2$, so half of the gain survives. □

The number of points can be cut down without changing the Gram matrix.

Lemma 7.7 (Few points suffice)**.** *Let Ω be a set, let ι be a finite set with $m = \#\iota$ elements and let $a : \Omega \to \mathbb{C}^\iota$. For all points $x_1, \dots, x_N \in \Omega$ and weights $w_1, \dots, w_N \geq 0$ with $\sum_k w_k = 1$ there are at most m^2+1 points $y_j \in \Omega$ and weights $v_j \geq 0$ with $\sum_j v_j = 1$ and $\sum_j v_j\, a(y_j)a(y_j)^* = \sum_k w_k\, a(x_k)a(x_k)^*$.*

[Lean: `Discretization.KieferWolfowitz.exists_design_card_le`]

Proof. The Gram matrix lies in the convex hull of the matrices $a(y)a(y)^*$, so by Carathéodory's theorem it is a convex combination of affinely independent ones. These matrices are Hermitian, and the Hermitian matrices in $\mathbb{C}^{\iota\times\iota}$ form a real vector space of dimension m^2, since every complex matrix is uniquely $H+iK$ with H and K Hermitian. An affinely independent family in it has at most m^2+1 members. □

Proof of Theorem 7.1. Take the design of Theorem 7.5 for $\varepsilon/2$. For $s = m+\varepsilon$ the bound $s^{-1}a(y)^*G^{-1}a(y) < 1$ is strict, so subtracting $s^{-1}a(y)a(y)^*$ from G leaves a positive definite matrix by Theorem 2.8(b), that is $a(y)a(y)^* \preceq s\cdot G$. Evaluating this at the coefficient vector c of f, with $c^*Gc = \sum_k w_k|f(x_k)|^2$, gives $|f(y)|^2 \le s\sum_k w_k|f(x_k)|^2$. This inequality depends on the design only through its Gram matrix, which Theorem 7.7 preserves while cutting the design down to m^2+1 points. Read as the measure $\varrho = \sum_k w_k\delta_{x_k}$, the design gives the theorem: integration against ϱ is the weighted sum over the points, for every measurable integrand.

For $\varepsilon = 0$ on a compact space, the matrices $a(y)a(y)^*$ form a compact set, whose convex hull is compact as well and consists exactly of the Gram matrices of designs. The determinant is continuous, so it attains its maximum at a design. No mixture can then increase the determinant at all, and Theorem 7.6 applied with the factor 1 gives $a(y)^*G^{-1}a(y) \le m$ without slack. The passage to the norm inequality is the one above for every constant $s > m$, followed by the limit $s \to m$, which is a limit of real numbers, and the point count is again Theorem 7.7.

In Lean the design form is `Discretization.KieferWolfowitz.exists_optimal_design_card_le` and its compact counterpart, and `Discretization.KieferWolfowitz.exists_measure_of_design` turns a design into the measure. □

Remark 7.8 (What the count is good for)**.** *The sharper count $r+1$, with r the dimension of the span of the products $a_k\overline{a_l}$, is not formalised, and nothing needs it.*

Remark 7.9 (The identity decomposed over the design)**.** *For a design with points $x_1,\dots,x_n$, weights $w_1,\dots,w_n$ and invertible Gram matrix G, and with $m = \#\iota$,*

$$\sum_{k=1}^{n} w_k\, a(x_k)^*G^{-1}a(x_k) = m \qquad \text{and} \qquad \sum_{k=1}^{n} w_k\big(a(x_k)^*G^{-1}z\big)a(x_k) = z$$

for every $z \in \mathbb{C}^\iota$, the first being $\mathrm{Tr}(G^{-1}G)$ and the second $GG^{-1}z = z$ written out (`Discretization.KieferWolfowitz.sum_weight_quadForm_inv_eq_card`, `Discretization.KieferWolfowitz.sum_weight_smul_quadForm_inv_eq_self`). With $c_k = w_k\, a(x_k)^* G^{-1} a(x_k)$ this is John's decomposition of the identity over the design points, in the inner product $\langle u, z\rangle = u^* G^{-1} z$ of the ellipsoid of G, and it is the condition dual to the bound of Theorem 7.5. Under Theorem 7.1 with $\varepsilon = 0$ the average forces $a(x_k)^* G^{-1} a(x_k) = m$ at every design point, which makes the points contact points. For an arbitrary convex body the theorem is proved in the companion project Lean-SNumbers (PDF).

References

[1] J. Batson, D. A. Spielman, N. Srivastava, *Twice-Ramanujan sparsifiers*, SIAM Review **56** (2014), no. 2, 315–334, doi, arxiv.

[2] F. Bartel, M. Schäfer, T. Ullrich, *Constructive subsampling of finite frames with applications in optimal function recovery*, Appl. Comput. Harmon. Anal. **65** (2023), 209–248, doi.

[3] A. Chkifa, M. Dolbeault, D. Krieg, M. Ullrich, *Constructive discretization and approximation in reproducing kernel Hilbert spaces*, preprint, 2026, arxiv.

[4] J. Kiefer, J. Wolfowitz, *The equivalence of two extremum problems*, Canad. J. Math. **12** (1960), 363–366, doi.

[5] D. Krieg, K. Pozharska, M. Ullrich, T. Ullrich, *Sampling projections in the uniform norm*, J. Math. Anal. Appl. **553** (2026), no. 2, Paper No. 129873, arxiv.

[6] I. Limonova, V. Temlyakov, *On sampling discretization in L_2*, J. Math. Anal. Appl. **515** (2022), no. 2, Paper No. 126457, arxiv.

[7] E. Novak, *Deterministic and stochastic error bounds in numerical analysis*, Lecture Notes in Mathematics **1349**, Springer-Verlag, Berlin, 1988.