Separable Rank of Radial Kernels: Machine-Checked Characterization in Lean 4 and Applications

GitHub Copilot

Source archive (CC0 1.0): https://github.com/OlliNiemitalo/kernel-rank-lean

November 14, 2026

Abstract

Filtering an image or three-dimensional wave field directly with a radially symmetric (round) kernel $K(x) = \varphi(\|x\|^2)$ of width W costs W^d multiply–adds per output sample. When a 2-D kernel is a product $h(x_0)v(x_1)$ of a horizontal and a vertical 1-D filter, it can instead be applied as two cheap 1-D passes; the *separable rank* R is the smallest number of such coordinate-product terms needed to sum to K. On a d-dimensional grid, rank R reduces W^d taps to dRW one-dimensional taps per sample (for example, a 9-tap-wide 3-D stencil drops from $9^3 = 729$ taps to $3 \times 4 \times 9 = 108$ taps when $R = 4$). Single bell-shaped Gaussians $\exp(-\alpha\|x\|^2)$ have rank 1, but cannot form flat-top camera defocus disks or ringed wave stencils, motivating sums of damped radial oscillations $\exp(-\alpha\|x\|^2)\cos(\omega\|x\|^2+\theta)$ (*complex-rate Gaussians*) and polynomials in $\|x\|^2$ times exponentials (*exponential polynomials*).

We determine which continuous radial kernels on $\mathbb{R}^d$ ($d \geq 2$) have finite separable rank and find the exact minimum number of product terms each requires. First, $K(x) = \varphi(\|x\|^2)$ has finite rank if and only if $\varphi(s)$ ($s \geq 0$) is a finite sum of exponential polynomials, whereas decaying complex-rate Gaussians with one shared width and equally spaced frequencies can approximate any continuous radial kernel vanishing at infinity. Second, an alternating angle-sampling identity evaluates each complex-rate Gaussian in d real products (rather than two to the power $d-1$ under naive real/imaginary expansion), yielding a constant $d-2$ real-over-complex overhead across all polynomial degrees. Third, allowing uniform limits on compact sets (*border rank*) never lowers rank in 2-D, whereas in 3-D on $\|x\|^2\cos(\|x\|^2)$ over $\mathbb{R}$, general limits, radial multi-Gaussian limits, and exact decompositions strictly separate as $5 < 6 < 7$. All 32 mathematical theorems are formally verified in the Lean 4 proof assistant.

Analytical operator constructions and illustrative 2-D bokeh and 3-D acoustic wave case studies examine steerable pass bounds, 16-bit channel orthogonalization, GPU crossovers against `cuFFT` and compact FDTD, and spatially varying media $c(\mathbf{x})$ in Section 6, alongside explicit floating-point stability limits for high polynomial degree, high dimension, and border-rank approximations in Section 7. Code, proofs, and data are public domain (CC0 1.0) at https://github.com/OlliNiemitalo/kernel-rank-lean.

1 Introduction

1.1 The problem and a two-dimensional visual example

Convolving a d-dimensional array with a dense $W \times \cdots \times W$ spatial stencil requires W^d multiply–accumulate operations per output sample. When the kernel $K : \mathbb{R}^d \to \mathbb{F}$ admits a rank-R coordinatewise decomposition

$$K(x_0, \ldots, x_{d-1}) = \sum_{n=0}^{R-1} \prod_{j=0}^{d-1} f_{n,j}(x_j), \tag{1}$$

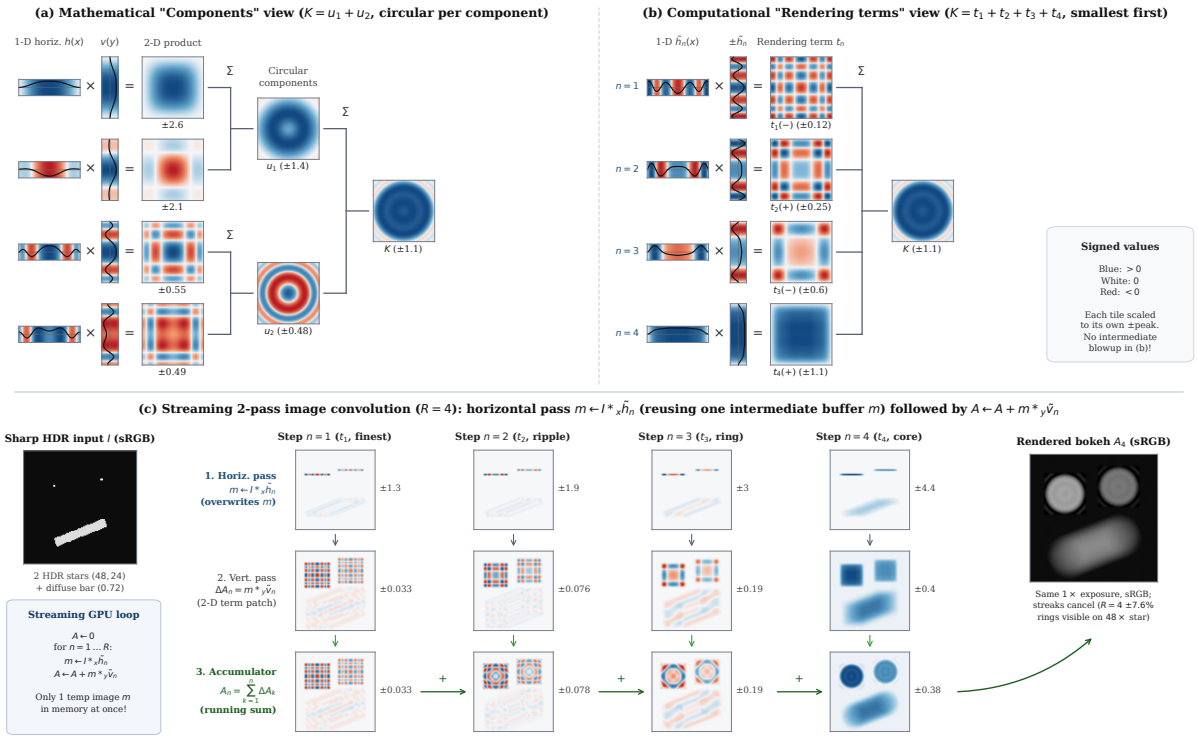


Figure 1: Visual guide to separable radial filtering on a two-complex-Gaussian ($n_{\mathrm{cx}} = 2$, exact rank $R = 4$) circular disc-blur kernel [58]: (a) mathematical *Components* view, where pairs of 1-D outer products $h(x)v(y)$ form circularly symmetric concentric-ring components u_1, u_2 summing to K; (b) computational *Rendering terms* view, where orthogonalizing rowSpan(K) yields four direct real products $t_n(x, y) = \tilde{h}_n(x) \cdot (\pm\tilde{h}_n(y))$ ordered smallest-first so intermediate amplitudes stay bounded; (c) streaming two-pass linear-light HDR blur overwriting *one* horizontal intermediate buffer $m \leftarrow I *_x \tilde{h}_n$ followed by vertical accumulation $A \leftarrow A + m *_y \tilde{v}_n$ (I and A_4 shown at 1× exposure via IEC 61966-2-1 sRGB; axis-aligned streaks cancel, while $R = 4$'s ±7.6% ripple remains visible on the 48× HDR star).

the d-dimensional convolution reduces to R cascades of d one-dimensional passes, costing dRW operations per sample (or $O(dR)$ operations independent of W when the one-dimensional factors admit recursive causal/anti-causal IIR realizations [19, 77]). Over a commutative semiring $\mathbb{F}$, the *separable rank* $\mathrm{srank}_{\mathbb{F}}(K) \in \mathbb{N} \cup \{\infty\}$ is the least $R \in \mathbb{N}$ for which (1) holds: matrix rank [70] when $d = 2$ and $\mathbb{F}$ is a field, and functional canonical polyadic (CP) or *separation rank* [32, 5, 39, 29] when $d \geq 3$.

For circularly or spherically symmetric (*radial*) kernels in dimension $d \geq 2$,

$$K(x) = \varphi\left(\sum_{j=0}^{d-1} x_j^2\right), \qquad x = (x_0, \ldots, x_{d-1}) \in \mathbb{R}^d,$$

classical results of Maxwell [49] and Kannappan and Sahoo [36] show that a continuous nonvanishing radial kernel with $\mathrm{srank}_{\mathbb{R}}(K) = 1$ must be a Gaussian $\varphi(s) = ae^{-\alpha s}$. Because real-rate Gaussians ($\alpha > 0$) are strictly monotone and convex in $s = \|x\|^2$, nonnegative Gaussian mixtures cannot reproduce flat passbands, sharp cutoffs, or diffraction rings [78]. Practical filter and propagator designs [54, 57, 58, 27, 44] therefore sum *complex-rate* Gaussians $\mathrm{Re}(\beta e^{(-\alpha+i\omega)\|x\|^2})$, whose damped radial oscillations $\cos(\omega\|x\|^2 + \theta)$ synthesize sharp transitions while keeping finite separable rank.

Figure 1 illustrates on a 2-D disc blur ($R = 4$) why distinguishing *radial components* from *rank*-1 *rendering terms* matters in both theory and computation:

Table 1: Question-to-result reading map ($d \geq 2$, polynomial degree $k = \deg P$, $N_{\text{naive}}(d) := 2^{d-1}$).

Question	Plain-language answer	Paper ref.	Exported Lean 4 theorem
Which continuous radial kernels on $\mathbb{R}^d$ have finite rank?	Exactly finite sums of exponential polynomials $\sum_u \mathrm{Re}(P_u(s)e^{\lambda_u s})$ on $s = \|x\|^2 \geq 0$	Thm. 3.1	`separableRank_radial_lt_top_iff_exists`
Can pure Gaussians approximate arbitrary radial shapes?	Yes: equal-width harmonic complex Gaussians uniformly approximate any $\psi \in C_0(\mathbb{R}_{\geq 0})$ on $\mathbb{R}^d$	Thm. 3.2	`exists_harmonic_gaussian_uniform_approx`
What is the exact complex rank on $\mathbb{R}^d$?	$\sum_{\mu \in S}((d-1)\deg P_\mu + 1)$ (1 per pure complex exponential, in every d)	Thm. 3.4	`separableRank_radial_complex_eq_iff`
How many *real* products does a complex Gaussian need?	d real products via angle sampling (3 in 3-D), vs. $N_{\text{naive}}(d) = 2^{d-1}$ (4 in 3-D) naively	Thm. 3.5, Prop. 3.6	`separableRank_radial_gaussian_eq_iff`, `re_prod_eq_sum_re`
What is the exact real rank for degree $k = \deg P \geq 0$?	$(d-1)k + 1$ per real rate and $2(d-1)k + d$ per nonreal pair (a constant $d-2$ gap over $\mathbb{C}$)	Thm. 3.5, Prop. 3.7	`separableRank_radial_eq_iff`, `re_coeff_prod_eq_sum_re`
How is 2-D rank characterized by differential equations?	Automatic C^∞ smoothness + unique monic minimal $p_{\min}(D)\varphi_{\mathbb{C}} = 0$ with $\deg p_{\min} = R$	Thm. 3.11	`exists_minimal_differential_annihilator`
Can limits of lower-rank kernels (*border rank*) cost less?	Never in 2-D ($= R$); in $d \geq 3$ multi-Gaussian limits save $(d-2)\sum_u k_u$ terms ($\leq B_d(\Gamma)$)	Thm. 3.13	`exists_rank_le_borderTotalCost_tendstoUniformlyOn_radial_profile`
Is the multi-Gaussian border bound $B_d(\Gamma)$ always exact?	Sharp over $\mathbb{C}$, for all real rates, and on $\cos\|x\|^2$ in $\mathbb{R}^3$; *non-sharp* on $\|x\|^2\cos\|x\|^2$ in 3-D (over $\mathbb{C}$: 4; over $\mathbb{R}$: $5 < 6 < 7$)	Thm. 3.14, Tab. 6	`borderTotalCost_le_of_tendsto_radial_real_profile`, `exists_rank_le_five_tendstoUniformlyOn_radial_sq_mul_cos`

1. **Components view (Figure 1a):** Writing $g(x) = \sqrt{\beta}\, e^{(-\alpha+i\omega)x^2} = g_{\text{re}}(x) + i g_{\text{im}}(x)$, taking the real part of $g(x)g(y)$ uses two real outer products, $\mathrm{Re}(g(x)g(y)) = g_{\text{re}}(x)g_{\text{re}}(y) + g_{\text{im}}(x)(-g_{\text{im}}(y))$. Neither outer product is circular alone, yet their sum cancels the grid corners into a circularly symmetric component (u_1: wide dome; u_2: sharpening rings) with $u_1 + u_2 = K$.
2. **Rendering terms view (Figure 1b):** A filter pipeline evaluates only the total sum $K(x,y) = \sum_{n=1}^{R} \tilde{h}_n(x)\tilde{v}_n(y)$. When many components nearly cancel ($R = 12$), individual u_i grow hundreds of times larger than K, causing severe rounding in 16-bit half-float (`float16`) buffers; orthogonalizing the 1-D filter basis (`separableChannels` [58]; Section 6.2.1) rotates the R pairs into bounded terms $t_n(x,y) = \tilde{h}_n(x) \cdot (\pm\tilde{h}_n(y))$ ordered smallest-first.
3. **Streaming execution (Figure 1c):** Looping $n = 1, \ldots, R$ overwrites a single temporary buffer $m \leftarrow I *_x \tilde{h}_n$ and accumulates $A \leftarrow A + m *_y \tilde{v}_n$ (or tiles strips inside GPU L2 cache so intermediates never touch DRAM; Section 6).

1.2 Guide to questions, answers, and Lean 4 theorems

Table 1 maps each mathematical question studied in this paper directly to its answer, theorem number, and machine-checked Lean 4 declaration.

1.3 Scope: what is proved, what is formalized, and what is illustrative

To keep the evidentiary status of every claim transparent from the outset:

- **Proved and machine-checked in Lean 4 (Sections 2–5, Appendix A):** All 32 mathematical theorems on continuous-domain exact rank, angle-sampling syntheses, 2-D differential annihilators, border-rank upper and lower bounds, and harmonic Gaussian universal approximation are formally verified in Lean 4 (`KernelRankLean`, 6,935 lines, zero `sorry`,

Table 2: Structured comparison of prior work with the present results ($n = \dim \mathcal{W}(g)$ denotes 2-D rank).

Literature thread	Object & assumptions	Dim. & field	Rank / border-rank result	Formal proof
Maxwell [49]; Kannappan–Sahoo [36]; Schoenberg [65]	Radial $K(\|x\|^2)$ on $\mathbb{R}^d$ (rank 1 or pos.-def.)	$d \geq 2$, $\mathbb{R}$	Rank 1 $\iff$ Gaussian; completely monotone $\iff$ positive Gaussian mixture	None
Levi-Civita [46]; Anselone–Korevaar [3]; Laird [43]	Pointwise $g(s_0 + s_1)$ on $\mathbb{R}$ or $\mathbb{R}_{\geq 0}$ (cont.)	$d = 2$, $\mathbb{R}/\mathbb{C}$	Finite 2-way rank $\iff$ exponential polynomial ($R = n$); 2-D rank closed	None
Rochberg [64]; Andersson–Carlsson [2]; Mourrain [51]	Bounded L^2 Hankel / convolution operators	$d = 2$ (or L^2 multi-D), $\mathbb{C}$	Kronecker symbol theorem + Hilbert–Schmidt / AAK SVD tools; requires L^2 or bounded domain	None
Beylkin–Mohlenkamp [5]; Oseledets [59]; Rahimi–Recht [61]	Gaussian-sum CP [12, 6], TT/Tucker [39, 29], pairwise RFF $k(x-y)$	$d \geq 2$, $\mathbb{R}/\mathbb{C}$	Upper-bound approximations; radial TT/Tucker rank $= n$ (n^d or $n \times n$ cores); no CP converse	None
Winograd [75]; Alder–Strassen [1]; Bläser et al. [10]	Multiplication tensor of finite-dim. algebra $\mathcal{A}$	$d = 3$ or $d \geq 3$, field $\mathbb{F}$	$(d-1)\dim \mathcal{A} - (d-2)m_{\text{irr}}$ on $\mathbb{F}[X]/(p)$ ($\|\mathbb{F}\|$ large); discrete border bounds [8, 45, 79, 16]	None
Bentkamp et al. [4]; Jain [35]; Yang et al. [76]	Circuit CP rank; 8^3 octonion; 3×3 $\mathbb{F}_2$ matmul	Finite tensors, $\mathbb{R}/\mathbb{F}_2$	Matricization bounds; fixed-size lower bounds (≥ 18, ≥ 21)	Isabelle/HOL; Lean 4
This work	Continuous radial $K(x) = \varphi(\|x\|^2)$ on $\mathbb{R}^d$	All $d \geq 2$, $\mathbb{R}$ and $\mathbb{C}$	Finite-rank iff + exact $\mathbb{C}/\mathbb{R}$ CP rank + angle-sampling syntheses + $\mathbb{R}^3$ border gap ($5<6<7$) + harmonic Gaussian density	**Lean 4** (32 thms, 0 `sorry`)

depending only on Lean's three standard classical axioms; Section 5). A supplementary Isabelle/HOL theory verifies one alternative nonstandard semigroup proof (Appendix A).

- **Analytically derived outside Lean 4 (Sections 6.1 and 7):** Fourier rank invariance, planar harmonic differential-operator identities, continuous two-stage steerable pass-count bounds, and closed-form floating-point conditioning bounds (showing where high degree k, large dimension d, and $O(h^{-1})$–$O(h^{-3})$ border-rank weights lose precision) are proved analytically in standard mathematical prose.
- **Illustrative numerical and GPU case studies (Section 6.2, Appendix B):** Reproducible Python/SciPy/PyTorch/Triton experiments (checked numerically against `float64` references, outside Lean 4) illustrate finite-grid truncation, `float16` stabilization ablations, and single-GPU crossovers on 2-D bokeh and 3-D acoustics. Analytical resolution-scaling estimates (such as FDTD runtime/memory ratios at matched dispersion) are explicitly labeled as extrapolations and paired with a smaller-grid measured GPU calibration.

1.4 Related work

Table 2 positions our contribution relative to six lines of prior work; three structural comparisons clarify the main trade-offs:

1. **Pointwise $[0,\infty)^d$ formulation vs. L^2 Hankel/convolution operators:** Bounded L^2 operator approaches [64, 2, 69, 74, 71] assume square-integrability or compact truncation, gaining Hilbert–Schmidt singular values and Adamjan–Arov–Krein (AAK) norm-optimal rank-R truncation—tools unavailable in our purely pointwise setting (so Section 6 uses Chebyshev LP optimization instead). Conversely, working pointwise on $[0,\infty)^d$ via Nagumo's Banach-algebra semigroup theorem [52, 22, 66, 42] requires *only* continuity on $[0,\infty)$: it covers undamped and polynomially growing profiles ($\cos\|x\|^2$, $\|x\|^2\cos\|x\|^2$, polynomials) that lie outside $L^2([0,\infty))$, and characterizes exact algebraic CP rank in $d \geq$

Table 3: Master notation table: every symbol has a single fixed role throughout the paper ($d \geq 2$).

Symbol	Lean 4 identifier	Fixed meaning (never overloaded)
$\mathbb{F}$; $\mathbb{R}, \mathbb{C}, \mathbb{R}_{\geq 0}$	$\mathbb{k}$; $\mathbb{R}, \mathbb{C}, \mathbb{R}_{\geq 0}$	Base semiring or field; reals, complexes, $[0, \infty)$
$K : \mathbb{R}^d \to \mathbb{F}$; $K_{\mathbb{C}}$	`K`	d-dimensional kernel; complexification $x \mapsto (K(x) : \mathbb{C})$
$R = \operatorname{srank}_{\mathbb{F}}(K)$	`separableRank K (ENat)`	Exact coordinatewise separable rank in $\mathbb{N} \cup \{\infty\}$ (Def. 2.1)
$\underline{\operatorname{srank}}^{\mathrm{co}}_{\mathbb{F}}(K)$, $\underline{\operatorname{srank}}^{\mathrm{pt}}_{\mathbb{F}}(K)$	See Table 4	Compact-open and pointwise border rank in $\mathbb{N} \cup \{\infty\}$ (Def. 2.4)
$\varphi : \mathbb{R} \to \mathbb{R}$; $\varphi_{\mathbb{C}}$	φ : $\mathbb{R} \to \mathbb{R}$	*Real-valued* radial profile; complexification $s \mapsto (\varphi(s) : \mathbb{C})$
$\psi : \mathbb{R} \to \mathbb{C}$; $g : \mathbb{R} \to \mathbb{F}$	ψ : $\mathbb{R} \to \mathbb{C}$	*Complex-valued* radial profile; general $\mathbb{F}$-valued profile
$s = \|x\|^2$; $W = 2L + 1$	$\sum_j$ `x j ^ 2`; —	Squared Euclidean radius on $\mathbb{R}^d$; 1-D discrete tap count on $[-L, L]$
$\mathcal{W}(g)$	`translateSpan g`	Half-line translate span $\operatorname{span}_{\mathbb{F}}\{u \mapsto g(u+t) : t \geq 0\}$
$\Gamma = [(\lambda_1, P_1), \ldots, (\lambda_m, P_m)]$	`cs : List Component`	Canonical real representation ($m = n_{\mathrm{re}} + n_{\mathrm{cx}} = m_{\mathrm{irr}}$; Def. 2.2)
$c_d(\lambda, P)$, $C_d(\Gamma)$	`multiCost`, `multiTotalCost`	Exact d-D real cost per component and total (C_2 = `totalCost`)
$b_d(\lambda, P)$, $B_d(\Gamma)$	`borderCost`, `borderTotalCost`	Multi-Gaussian border cost $C_d(\Gamma) - (d-2)\sum_u \deg P_u$
$N_{\mathrm{naive}}(d) := 2^{d-1}$	—	Naive real products per k=0 complex Gaussian ($2, 4, 8$ for d=$2, 3, 4$)
$M^{\mathbb{C}}_{d,k} = (d-1)k+1$; $M^{\mathbb{R}}_{d,k}$	—; $2(d-1)k + d$	Root-of-unity node counts over $\mathbb{C}$ and for nonreal rates over $\mathbb{R}$
$H_P(s_0; X)$; $E_k(s_j; X)$	`revHassePoly`; `truncExpPoly`	Reversed Hasse polynomial; truncated exponential $\sum_{q=0}^{k} s_j^q X^q / q!$

3 where no Hilbert–Schmidt SVD exists. Continuity is also minimal for avoiding non-measurable Hamel-basis solutions $e^{A(\|x\|^2)}$ of Cauchy's equation, while discontinuous indicators ($\mathbf{1}_{\|x\| \leq 1}$, rank ∞; $\mathbf{1}_{\|x\|=0}$, rank 1) and singular Green's functions ($\|x\|^{2-d}$ [13, 28, 47, 37]) require separate domain-restricted treatments (Section 7).

2. **Coordinatewise CP rank vs. pairwise ML features and TT/Tucker formats:** Random Fourier features [61] factor $k(x-y)$ across *sample pairs* $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ (where even $e^{-\alpha\|x-y\|^2}$ has infinite exact 2-argument rank via $e^{2\alpha x \cdot y}$, and $e^{i\omega \cdot z}$ is non-decaying), whereas across Cartesian coordinates $(z_0, \ldots, z_{d-1}) \in \mathbb{R}^d$ a Gaussian has exact d-way rank 1. Meanwhile, every mode or prefix unfolding of $g(s_0 + \cdots + s_{d-1})$ depends only on partial sums, so Tucker [39, 29] and Tensor-Train [59] unfolding ranks equal the 2-D rank n independent of d, at the price of an n^d core or $n \times n$ inter-axis transfer contractions; uncoupled scalar CP cascades (1) pay only $\operatorname{srank}_{\mathbb{F}} = (d-1)n - (d-2)m_{\mathrm{irr}}$.

3. **Finite-dimensional algebra complexity, border rank, and applied stencils:** Identifying $\mathcal{W}(g)$ with $\mathcal{A} = \mathbb{F}[X]/(p_{\min,\mathbb{F}})$ reduces our multilinear core to $(d-1)$-fold multiplication in $\mathcal{A}$ [75, 23, 1, 9, 15, 10]; we lift this invariant to continuous radial functions on $\mathbb{R}^d$, replace ill-conditioned real Vandermonde inversion on nonreal factors by closed-form angle-sampling syntheses, and establish strict real border-rank non-sharpness ($5 < B_3 = 6 < C_3 = 7$) alongside Hankel/Waring and border-rank theory [17, 34, 41, 53, 18, 20, 63, 11, 14, 79, 16]. Our applied sections connect directly to planar steerable and bokeh filters [70, 26, 38, 60, 7, 62, 72, 54, 27, 55, 56, 57, 58], wideband acoustic stencils [68, 67, 40, 73, 30, 31, 25], and Lean 4/Isabelle formalizations [50, 48, 21, 24, 33, 4, 35, 76].

2 Setup and Notation

Quick reading guide to Sections 2 and 3. Table 3 fixes every symbol once (in particular, R denotes *only* separable rank, Γ denotes a representation list, $\mathbb{F}$ denotes a general semiring/field

so it never clashes with K, and $N_{\text{naive}}(d) := 2^{d-1}$—two to the power $d-1$—names the naive real-product count). Three definitions below set up all theorems in Section 3: **Definition 2.1** ($\operatorname{srank}_{\mathbb{F}}$), **Definition 2.2** ($\Gamma$ and the integer cost functions $C_d(\Gamma), B_d(\Gamma)$), and **Definition 2.4** (border rank, mapped to Lean 4 in Table 4). Across all classification and exact-rank theorems (Theorems 3.1, 3.4, 3.5, and 3.11), the *only* hypotheses are $d \geq 2$ and continuity of φ (or ψ) on $[0,\infty)$—neither differentiability, square-integrability, nor decay at infinity is assumed (only global uniform approximation in Theorem 3.2 assumes $\lim_{s\to\infty}\psi(s) = 0$).

Definition 2.1 (Coordinatewise separable rank; Lean: `IsSumOfProducts`, `separableRank`)**.** Let ι be a finite index set, $(X_i)_{i\in\iota}$ a family of sets, and $\mathbb{F}$ a commutative semiring. A kernel $K : (\prod_{i\in\iota} X_i) \to \mathbb{F}$ is a *sum of R products* ($R \in \mathbb{N}$) if there exist univariate factors $f_{n,i} : X_i \to \mathbb{F}$ ($0 \leq n < R$, $i \in \iota$) such that $K(x) = \sum_{n=0}^{R-1} \prod_{i\in\iota} f_{n,i}(x_i)$ for all $x \in \prod_{i\in\iota} X_i$ (taking the empty sum 0 when $R = 0$). Its *separable rank* over $\mathbb{F}$ is

$$\operatorname{srank}_{\mathbb{F}}(K) := \inf\{R \in \mathbb{N} : K \text{ is a sum of } R \text{ products over } \mathbb{F}\} \in \mathbb{N} \cup \{\infty\},$$

where $\inf\emptyset = \infty$, and writing $\operatorname{srank}_{\mathbb{F}}(K) = R$ with $R \in \mathbb{N}$ denotes equality in $\mathbb{N} \cup \{\infty\}$.

Separable rank is subadditive under finite sums and non-increasing under coordinate pullbacks. Whenever every coordinate map $\sigma_i : Y_i \to X_i$ ($i \in \iota$) is *surjective* (with right inverse τ_i), pulling back and pushing forward preserve rank identically:

$$\operatorname{srank}_{\mathbb{F}}(y \mapsto K(i \mapsto \sigma_i(y_i))) = \operatorname{srank}_{\mathbb{F}}(K). \tag{2}$$

Because $x_j \mapsto x_j^2$ maps $\mathbb{R}$ onto $\mathbb{R}_{\geq 0} = [0,\infty)$ (with right inverse $\sqrt{s_j}$), (2) equates the d-way rank of any radial kernel on $\mathbb{R}^d$ with that of the d-way additive Hankel kernel on $\mathbb{R}^d_{\geq 0}$:

$$\operatorname{srank}_{\mathbb{F}}\left(x \in \mathbb{R}^d \mapsto g\left(\sum_{j=0}^{d-1} x_j^2\right)\right) = \operatorname{srank}_{\mathbb{F}}\left(s \in \mathbb{R}^d_{\geq 0} \mapsto g\left(\sum_{j=0}^{d-1} s_j\right)\right). \tag{3}$$

Consequently, $K(x) = \varphi(\|x\|^2)$ depends only on $\varphi|_{[0,\infty)}$ (leaving $\varphi(s)$ for $s < 0$ unconstrained), and replacing x_j^2 by x_j^2/a_j^2 ($a_j \neq 0$) or $|x_j|^p$ ($p > 0$) extends every theorem below verbatim to axis-aligned ellipsoids and ℓ_p superellipsoids. Setting $x_2 = \cdots = x_{d-1} = 0$ for $d \geq 2$ yields the 2-D slice lower bound:

$$\operatorname{srank}_{\mathbb{F}}(y \in \mathbb{R}^2 \mapsto g(y_0^2 + y_1^2)) \leq \operatorname{srank}_{\mathbb{F}}\left(x \in \mathbb{R}^d \mapsto g\left(\sum_{j=0}^{d-1} x_j^2\right)\right). \tag{4}$$

Definition 2.2 (Canonical real representations and cost functions)**.** A *component* (`Component`) is a pair $(\lambda, P) \in \mathbb{C} \times \mathbb{C}[X]$ of degree $k = \deg P$, evaluating at $s \in \mathbb{R}$ as $\operatorname{Re}(P(s)e^{\lambda s})$. A finite list $\Gamma = [(\lambda_1, P_1), \ldots, (\lambda_m, P_m)]$ is a *canonical representation* (`IsRepresentation`) if the rates λ_u are pairwise distinct and each (λ_u, P_u) is well-formed ($P_u \neq 0$, $\operatorname{Im}\lambda_u \geq 0$, and $\operatorname{Im}\lambda_u = 0 \implies P_u \in \mathbb{R}[X]$; see Remark 2.3); its profile is $\operatorname{profile}(\Gamma, s) := \sum_{u=1}^m \operatorname{Re}(P_u(s)e^{\lambda_u s})$. For $d \geq 2$, define the exact component cost $c_d(\lambda, P)$ and multi-Gaussian border cost $b_d(\lambda, P)$ by

$$c_d(\lambda,P) := \begin{cases} (d-1)\deg P + 1, & \operatorname{Im}\lambda = 0, \\ 2(d-1)\deg P + d, & \operatorname{Im}\lambda > 0, \end{cases} \qquad b_d(\lambda,P) := \begin{cases} \deg P + 1, & \operatorname{Im}\lambda = 0, \\ d(\deg P + 1), & \operatorname{Im}\lambda > 0, \end{cases} \tag{5}$$

with total costs $C_d(\Gamma) := \sum_{u=1}^m c_d(\lambda_u, P_u)$ (`multiTotalCost`) and $B_d(\Gamma) := \sum_{u=1}^m b_d(\lambda_u, P_u)$ (`borderTotalCost`). When every $\deg P_u = 0$ (n_{re} real and n_{cx} nonreal Gaussians), $C_d(\Gamma) = B_d(\Gamma) = n_{\text{re}} + d\, n_{\text{cx}}$.

Table 4: Correspondence between mathematical border-rank clauses (Definition 2.4) and their Lean 4 / Mathlib encoding in `Approximation.lean` and `Main.lean`.

Mathematical clause	Exact Lean 4 / Mathlib encoding	Role of filter & totalized division
Upper bound $\underline{\mathrm{srank}}_{\mathbb{F}}^{\mathrm{co}}(K) \le r$ (uniform on compact sets)	$\exists F : \mathbb{R} \to (\mathrm{Fin}\, d \to \mathbb{R}) \to \mathbb{F}$ with $\forall h$, `separableRank` $(F\, h) \le r$ and $\forall E$, `IsCompact` $E \to$ `TendstoUniformlyOn` $F\, K$ (`nhdsWithin` $0\, \{0\}^{\mathrm{c}})\, E$	Punctured filter $h \to 0$ $(h \neq 0)$; composing with $h_n = \frac{1}{n+1}$ yields sequential $F_{h_n} \to K$ in τ_{co}
Value of F_h at $h = 0$ in difference quotients $\frac{e^{hs}-1}{h}$	Total function $h : \mathbb{R}$ using Lean's convention $0^{-1} = 0$ (so $u_0(s) = 0$)	Keeps $F\, 0$ well-typed with $\mathrm{srank}_{\mathbb{F}}(F\, 0) \le r$ without a subtype $\{h : \mathbb{R} // h \neq 0\}$; $\{0\}^{\mathrm{c}}$ excludes $h = 0$ from the limit
Lower bound $\underline{\mathrm{srank}}_{\mathbb{F}}^{\mathrm{pt}}(K) \ge r$ (pointwise limit)	$\forall (\iota)\, (l : \mathtt{Filter}\, \iota)\, [l.\mathtt{NeBot}]\, (F : \iota \to \dots)$, $(\forall^{\mathrm{f}} a$ in l, `separableRank` $(F\, a) \le r - 1) \to \neg(\forall x$, `Tendsto` $(\lambda a \Rightarrow F\, a\, x)\, l$ (`nhds` $(K\, x)))$	$[l.\mathtt{NeBot}]$ rules out the vacuous bottom filter $\bot$; setting $\iota = \mathbb{N}, l =$ `atTop` recovers sequences $(F_n)_{n \in \mathbb{N}}$
Passing minors / adjugate traces to the limit	Continuity of `Matrix.det` and `Matrix.trace` on $[$`T2Space` $\mathbb{F}]$	Needs *only* pointwise convergence $F_a(x) \to K(x)$; individual $O(h^{-1})$ factors need not converge

Remark 2.3 (Why Γ is unique, and how components map to rendering terms)**.** Because $\mathrm{Re}(\overline{P}(s)e^{\overline{\lambda}s}) = \mathrm{Re}(P(s)e^{\lambda s})$ and $\mathrm{Re}(iae^{\mu s}) = 0$ for $a, \mu \in \mathbb{R}$, restricting to $\mathrm{Im}\,\lambda_u \ge 0$ and $P_u \in \mathbb{R}[X]$ when $\mathrm{Im}\,\lambda_u = 0$ removes sign and conjugate redundancies: each real rate contributes $P_u(s)e^{\lambda_u s}$ over $\mathbb{C}$, each $\mathrm{Im}\,\lambda_u > 0$ rate contributes the conjugate pair $\frac{1}{2}P_u(s)e^{\lambda_u s} + \frac{1}{2}\overline{P}_u(s)e^{\overline{\lambda}_u s}$, and $\mathbb{C}$-linear independence of exponential polynomials on $(0, \infty)$ determines the unordered set of components of $\varphi|_{[0,\infty)}$ uniquely. Geometrically, Γ specifies $m = n_{\mathrm{re}} + n_{\mathrm{cx}}$ radially symmetric *components* $U_u(x) = \mathrm{Re}(P_u(\|x\|^2)e^{\lambda_u \|x\|^2})$ (Figure 1a), whereas real Lagrange interpolation ($\mathrm{Im}\,\lambda_u = 0$) and Propositions 3.6–3.7 ($\mathrm{Im}\,\lambda_u > 0$) expand each U_u into $c_d(\lambda_u, P_u)$ rank-1 coordinate-product (*rendering*) terms totaling $R = C_d(\Gamma)$ terms (related in $d = 2$ by $(HT, VT^{-\top})$ for $T \in \mathrm{GL}_R(\mathbb{R})$; Figure 1b).

Definition 2.4 (Compact-open topology and border rank)**.** Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $d \ge 1$. The *compact-open topology* τ_{co} on $C(\mathbb{R}^d, \mathbb{F})$ is generated by the uniform seminorms $\|F\|_{E,\infty} := \sup_{x \in E} |F(x)|$ over all compact subsets $E \subseteq \mathbb{R}^d$ (metrizable via $\mathbb{R}^d = \bigcup_{M=1}^{\infty} [-M, M]^d$). Writing $\mathcal{F} \neq \bot$ for a nontrivial filter, the *compact-open* and *pointwise border ranks* of $K : \mathbb{R}^d \to \mathbb{F}$ in $\mathbb{N} \cup \{\infty\}$ are

$$\underline{\mathrm{srank}}_{\mathbb{F}}^{\mathrm{co}}(K) := \inf\{r \in \mathbb{N} : \exists (F_n)_{n \in \mathbb{N}},\ \mathrm{srank}_{\mathbb{F}}(F_n) \le r,\ \|F_n - K\|_{E,\infty} \to 0\ (\forall \text{ compact } E)\},$$

$$\underline{\mathrm{srank}}_{\mathbb{F}}^{\mathrm{pt}}(K) := \inf\{r \in \mathbb{N} : \exists \mathcal{F} \neq \bot,\ (F_\iota),\ \mathrm{srank}_{\mathbb{F}}(F_\iota) \le r \text{ along } \mathcal{F},\ F_\iota(x) \xrightarrow{\mathcal{F}} K(x)\ (\forall x)\},$$

satisfying $\underline{\mathrm{srank}}_{\mathbb{F}}^{\mathrm{pt}}(K) \le \underline{\mathrm{srank}}_{\mathbb{F}}^{\mathrm{co}}(K) \le \mathrm{srank}_{\mathbb{F}}(K)$. Table 4 records the exact correspondence between Definition 2.4 and the Lean 4 statements in `KernelRankLean/Approximation.lean`.

3 Main Mathematical Results

3.1 (A) Finite-rank classification and universal approximation

In plain terms, a continuous radial kernel on $\mathbb{R}^d$ ($d \ge 2$) splits into a finite sum of 1-D coordinate products *if and only if* its profile in $s = \|x\|^2$ is a finite sum of polynomials times exponentials (for instance, in $d = 3$, $e^{-\|x\|^2}\cos(2\|x\|^2)$ has finite rank 3, whereas $(1 + \|x\|^2)^{-1}$ and $e^{-\|x\|^4}$ have infinite rank); nevertheless, any continuous radial kernel that fades to zero at infinity can be approximated uniformly across all of $\mathbb{R}^d$ by decaying complex-rate Gaussians sharing one width parameter:

Theorem 3.1 (d-D Finite-Rank Classification; Lean: `separableRank_radial_lt_top_iff_exists`)**.** *Let $d \geq 2$ and let $\varphi : \mathbb{R} \to \mathbb{R}$ be continuous on $[0,\infty)$. Then* $\operatorname{srank}_{\mathbb{R}}(x \in \mathbb{R}^d \mapsto \varphi(\sum_{j=0}^{d-1} x_j^2)) < \infty$ *if and only if there exists a canonical representation $\Gamma = [(\lambda_1, P_1), \dots, (\lambda_m, P_m)]$ such that $\varphi(s) = \sum_{u=1}^{m} \operatorname{Re}(P_u(s) e^{\lambda_u s})$ for all $s \geq 0$.*

Theorem 3.2 (Equal-Width Harmonic Complex-Gaussian Universal Approximation on $\mathbb{R}^d$; Lean: `exists_harmonic_gaussian_uniform_approx`)**.** *Let $d \geq 2$, let $\varphi : \mathbb{R} \to \mathbb{R}$ be continuous on $[0,\infty)$ with $\lim_{s\to\infty} \varphi(s) = 0$, and let $\alpha > 0$ and $\varepsilon > 0$. There exist $\omega_0 > 0$, $M \in \mathbb{N}$, $a_0 \in \mathbb{R}$, and $a_1, \dots, a_M \in \mathbb{C}$ such that the equal-width harmonic multi-Gaussian kernel*

$$K_M(x) := a_0 e^{-\alpha\|x\|^2} + \sum_{m=1}^{M} \operatorname{Re}\Big(a_m e^{(-\alpha + im\omega_0)\|x\|^2}\Big) \tag{6}$$

satisfies $\operatorname{srank}_{\mathbb{R}}(K_M) \leq 1 + dM$ and $|K_M(x) - \varphi(\|x\|^2)| < \varepsilon$ for all $x \in \mathbb{R}^d$.

Remark 3.3 (Envelope compensation vs. free nonlinear rates)**.** Factoring $e^{-\alpha s}$ ($s = \|x\|^2$) out of (6) leaves a T-periodic trigonometric polynomial ($T = 2\pi/\omega_0$) approximating $e^{+\alpha u}\varphi(u)$ on $[0, S_0]$, while $e^{-\alpha(kT+u)} e^{+\alpha u}\varphi(u) = e^{-k\alpha T}\varphi(u)$ damps periodic replicas ($k \geq 1$). Fixing equal widths and harmonic spacing is universal as $M \to \infty$, whereas freeing (α_u, ω_u) achieves much smaller rank R at sharp transitions via near-cancellation [12, 58].

3.2 (B) Exact complex and real ranks and angle-sampling identities

In plain terms, over $\mathbb{C}$ each exponential-polynomial rate μ of degree k costs $(d-1)k+1$ products, whereas over $\mathbb{R}$ each complex-conjugate pair pays a *constant* $d-2$ extra products ($2(d-1)k+d$ vs. $2(d-1)k+2$): in $d=3$, a pure complex Gaussian ($k=0$) costs 3 real products (saving one product over the $N_{\text{naive}}(3) = 4$ naive real/imaginary terms), and $\|x\|^2\cos(\|x\|^2)$ ($k=1$) costs 7 real products. Table 5 summarizes the four counting rules before the formal statements:

1. *Real rate ($\operatorname{Im}\lambda = 0$):* Placing degree-$k$ polynomials on the last $d-1$ coordinates produces degree $(d-1)k$, requiring $M_{d,k}^{\mathbb{C}} = (d-1)k+1$ real products (1 at $k=0$; 3 at $d=3, k=1$).
2. *Nonreal rate ($\operatorname{Im}\lambda > 0$) and the constant $d-2$ gap:* Over $\mathbb{C}$, $\frac{1}{2}Pe^{\lambda s} + \frac{1}{2}\overline{P}e^{\bar{\lambda} s}$ costs $2((d-1)k+1)$. Over $\mathbb{R}$ at $k=0$, every $\mathbb{R}$-linear functional $\ell : \mathbb{C} \to \mathbb{R}$ has a nonzero nullvector, killing one product per mode $j = 0, \dots, d-2$ and forcing d real products ($2 + (d-2)$); each degree step $k \mapsto k+1$ adds $2(d-1)$ real degrees of freedom over both $\mathbb{C}$ and $\mathbb{R}$, keeping the real-over-complex penalty locked at $d-2$ for every $k \geq 0$.
3. *Naive expansion ($N_{\text{naive}}(d) = 2^{d-1}$) vs. angle sampling ($d$):* Writing $z_j = a_j + ib_j$, expanding $\operatorname{Re}(\prod_{j=0}^{d-1}(a_j + ib_j))$ sums $(-1)^{|S|/2}(\prod_{j\notin S} a_j)(\prod_{j\in S} b_j)$ over all even-size subsets $S \subseteq \{0, \dots, d-1\}$, which has $\sum_{2|q}\binom{d}{q} = 2^{d-1} =: N_{\text{naive}}(d)$ real products ($2, 4, 8$ for $d = 2, 3, 4$), whereas antipodal folding $r \leftrightarrow r + d$ in Proposition 3.6 halves the $2d$-angle average to d angles ($2, 3, 4$ products).

Theorem 3.4 (d-D Complex Radial Rank; Lean: `separableRank_radial_complex_eq_iff`)**.** *Let $d \geq 2$ and let $\psi : \mathbb{R} \to \mathbb{C}$ be continuous on $[0,\infty)$. For every $R \in \mathbb{N}$,* $\operatorname{srank}_{\mathbb{C}}(x \in \mathbb{R}^d \mapsto \psi(\sum_{j=0}^{d-1} x_j^2)) = R$ *if and only if there exist a finite set $S \subset \mathbb{C}$ and nonzero polynomials $(P_\mu)_{\mu\in S} \subset \mathbb{C}[X]$ such that $\psi(s) = \sum_{\mu\in S} P_\mu(s) e^{\mu s}$ for all $s \geq 0$ and $\sum_{\mu\in S}((d-1)\deg P_\mu + 1) = R$.*

Theorem 3.5 (Exact d-D Real Radial Rank; Lean: `separableRank_radial_eq_iff`)**.** *Let $d \geq 2$ and let $\varphi : \mathbb{R} \to \mathbb{R}$ be continuous on $[0,\infty)$. For every $R \in \mathbb{N}$,* $\operatorname{srank}_{\mathbb{R}}(x \in \mathbb{R}^d \mapsto \varphi(\sum_{j=0}^{d-1} x_j^2)) = R$ *if and only if there exists a canonical representation $\Gamma = [(\lambda_1, P_1), \dots, (\lambda_m, P_m)]$ (distinct rates*

Table 5: Single-component product counts across $d \in \{2,3,4\}$ and $k = \deg P \in \{0,1\}$ (naive nonreal count expands each complex product into $N_{\text{naive}}(d) = 2^{d-1}$ real products: $N_{\text{naive}}(d)$ at k=0, $d\,N_{\text{naive}}(d)$ at k=1).

Component type	Profile $\varphi(s)$ ($s = \|x\|^2$)	Dim. d	Complex rank	Exact real c_d	Naive real	Border b_d
Real Gaussian ($k=0$)	$a\,e^{-\alpha s}$	$2,3,4$	$1,1,1$	$1,1,1$	$1,1,1$	$1,1,1$
Real Laguerre ($k=1$)	$(p_0+p_1 s)e^{-\alpha s}$	$2,3,4$	$2,3,4$	$2,3,4$	$2,3,4$	$2,2,2$
Complex Gaussian ($k=0$)	$\mathrm{Re}(\beta\,e^{(-\alpha+i\omega)s})$ ($\omega>0$)	$2,3,4$	$2,2,2$	$2,3,4$	$2,4,8$	$2,3,4$
Complex poly. ($k=1$)	$\mathrm{Re}((p_0+p_1 s)e^{(-\alpha+i\omega)s})$	$2,3,4$	$4,6,8$	$4,7,10$	$4,12,32$	$4,6,8$

with $\operatorname{Im}\lambda_u \geq 0$, $P_u \neq 0$, *and* $P_u \in \mathbb{R}[X]$ *when* $\operatorname{Im}\lambda_u = 0$*) such that* $\varphi(s) = \sum_{u=1}^{m} \mathrm{Re}(P_u(s)e^{\lambda_u s})$ *for all* $s \geq 0$ *and*

$$C_d(\Gamma) = \sum_{\operatorname{Im}\lambda_u = 0} ((d-1)\deg P_u + 1) + \sum_{\operatorname{Im}\lambda_u > 0} (2(d-1)\deg P_u + d) = R.$$

In particular, for pure Gaussians ($\deg P_u = 0$*),* $\mathrm{srank}_{\mathbb{R}} = n_{\mathrm{re}} + d\,n_{\mathrm{cx}}$, *and in* $d = 2$, $\mathrm{srank}_{\mathbb{R}} = C_2(\Gamma) = \mathrm{srank}_{\mathbb{C}}(K_{\mathbb{C}})$.

Both nonreal upper bounds are constructed explicitly by the following two identities (proved in Section 4.2, item 3; $[X^k]F(X)$ denotes the X^k-coefficient of $F \in \mathbb{C}[X]$):

Proposition 3.6 (Alternating angle-sampling ($2d$-th root-of-unity) identity; Lean: `re_prod_eq_sum_re`)**.** *For every integer* $d \geq 1$ *and all* $z_0, \ldots, z_{d-1} \in \mathbb{C}$, *with* $N_{\text{naive}}(d) = 2^{d-1}$,

$$\mathrm{Re}\left(\prod_{j=0}^{d-1} z_j\right) = \frac{N_{\text{naive}}(d)}{d} \sum_{r=0}^{d-1} (-1)^r \prod_{j=0}^{d-1} \mathrm{Re}\big(z_j e^{-i\pi r/d}\big)\,. \tag{7}$$

Proposition 3.7 (Polynomial alternating angle-sampling identity; Lean: `re_coeff_prod_eq_sum_re`)**.** *Let* $d \geq 2$, $k \geq 0$, *and* $M^{\mathbb{R}}_{d,k} = 2(d-1)k + d$. *For any* $U_0, \ldots, U_{d-1} \in \mathbb{C}[X]$ *with* $\deg U_j \leq k$,

$$\mathrm{Re}\left([X^k]\prod_{j=0}^{d-1} U_j(X)\right) = \frac{N_{\text{naive}}(d)}{M^{\mathbb{R}}_{d,k}} \sum_{r=0}^{M^{\mathbb{R}}_{d,k}-1} (-1)^r \prod_{j=0}^{d-1} \mathrm{Re}\left(U_j\left(e^{\frac{2\pi i r}{M^{\mathbb{R}}_{d,k}}}\right) e^{-\frac{i\pi r(2k+1)}{M^{\mathbb{R}}_{d,k}}}\right), \tag{8}$$

reducing at $k = 0$ *(*$M^{\mathbb{R}}_{d,0} = d$, *constant* $U_j(X) = z_j$*) identically to* (7).

Example 3.8 ($d = 3, k = 0$: 3 real products vs. $N_{\text{naive}}(3) = 4$)**.** For $z_j = a_j + ib_j = \beta^{1/3}e^{(-\alpha+i\omega)x_j^2}$ ($j = 0,1,2$), direct expansion $\mathrm{Re}(z_0z_1z_2) = a_0a_1a_2 - a_0b_1b_2 - b_0a_1b_2 - b_0b_1a_2$ uses 4 real 3-way products. Proposition 3.6 evaluates $u_{j,r} = a_j\cos(\frac{\pi r}{3}) + b_j\sin(\frac{\pi r}{3})$ at angles $0, \frac{\pi}{3}, \frac{2\pi}{3}$ ($u_{j,0} = a_j$, $u_{j,1} = \frac{1}{2}a_j + \frac{\sqrt{3}}{2}b_j$, $u_{j,2} = -\frac{1}{2}a_j + \frac{\sqrt{3}}{2}b_j$) to obtain $\mathrm{Re}(z_0z_1z_2) = \frac{4}{3}(u_{0,0}u_{1,0}u_{2,0} - u_{0,1}u_{1,1}u_{2,1} + u_{0,2}u_{1,2}u_{2,2})$ in *three* real products (matching Gauss/Karatsuba multiplication in $\mathbb{C} \cong \mathbb{R}[X]/(X^2+1)$ [75]).

Example 3.9 ($d = 3, k = 1$: 7 real products via $M^{\mathbb{R}}_{3,1} = 7$ vs. 9 or 12)**.** For $K_3(x) = \|x\|^2\cos(\|x\|^2)$ on $\mathbb{R}^3$, splitting $(x_0^2 + x_1^2 + x_2^2)\cos\|x\|^2$ costs $3 \times 4 = 12$ real products naively (Table 5) or $3\times 3 = 9$ via Example 3.8. Setting $U_j(X) = e^{ix_j^2}(1 + x_j^2 X)$ in Proposition 3.7 ($M^{\mathbb{R}}_{3,1} = 7$) yields $K_3(x) = \frac{4}{7}\sum_{r=0}^{6}(-1)^r f_r(x_0)f_r(x_1)f_r(x_2)$ with $f_r(u) = \cos(u^2 - \frac{3\pi r}{7}) + u^2\cos(u^2 - \frac{\pi r}{7})$ in 7 real products (7 is the least modulus keeping all mixed harmonics inside $\{1, \ldots, M-1\}$, proved optimal in Section 4.3).

Remark 3.10 (Commutative-algebra invariant vs. symmetric Waring rank)**.** For $\mathcal{A}_\varphi = \mathbb{R}[X]/(p_{\min,\mathbb{R}}(X))$ ($m_{\text{irr}} = n_{\text{re}} + n_{\text{cx}}$ distinct irreducible factors over $\mathbb{R}$, $\dim_{\mathbb{R}} \mathcal{A}_\varphi = C_2(\Gamma)$), Theorem 3.5 reads

$$\operatorname{srank}_{\mathbb{R}}(K) = C_d(\Gamma) = (d-1)\dim_{\mathbb{R}} \mathcal{A}_\varphi - (d-2)\, m_{\text{irr}} \qquad (d \geq 2), \tag{9}$$

matching Kronecker ($d = 2$), Winograd–Alder–Strassen ($d = 3$ [75, 23, 1, 15]), and Bläser–Mayer–Shringi ($d \geq 4$ [10]). Unlike symmetric Waring rank on $\operatorname{Re}((u+iv)^D)$ (whose real–complex gap $D-2$ grows with degree [18, 63, 11]), coordinatewise decompositions (1) break mode symmetry and incur a degree-independent penalty $(d-2)\Delta m_{\text{irr}} = d-2$ per conjugate pair.

3.3 (C) Two-dimensional minimal differential annihilator

In plain terms, any continuous 2-D radial kernel of finite separable rank R is automatically infinitely differentiable on $[0,\infty)$, and R equals the lowest order of a linear constant-coefficient differential equation $p_{\min}(D)\varphi_{\mathbb{C}} = 0$ satisfied by its profile on $[0,\infty)$:

Theorem 3.11 (2-D Minimal Differential Annihilator; Lean: `exists_minimal_differential_annihilator`)**.** *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be continuous on $[0,\infty)$ with $\operatorname{srank}_{\mathbb{R}}(x \in \mathbb{R}^2 \mapsto \varphi(x_0^2 + x_1^2)) = R < \infty$. For any polynomial $q(X) = \sum_{k=0}^{\deg q} q_k X^k \in \mathbb{C}[X]$ and any function $g : \mathbb{R} \to \mathbb{C}$, define $(q(D)g)(t) := \sum_{k=0}^{\deg q} q_k\, D^k_{[0,\infty)} g(t)$ (`differentialOperator`), where the domain-restricted operator $D^k_{[0,\infty)} : (\mathbb{R} \to \mathbb{C}) \to (\mathbb{R} \to \mathbb{C})$ (`iteratedDerivWithin k g (Set.Ici 0)`) iterates two-sided differentiation at interior points $t > 0$ and* one-sided right differentiation $\lim_{h \to 0^+}(g(h) - g(0))/h$ *at the boundary $t = 0$ (totalized in Mathlib by returning 0 where a derivative fails to exist). Under $\operatorname{srank}_{\mathbb{R}} = R < \infty$, $\varphi_{\mathbb{C}}|_{[0,\infty)}$ belongs to the smooth domain $C^\infty([0,\infty);\mathbb{C})$ on which every classical interior and right derivative exists, and there is a unique monic $p_{\min} \in \mathbb{C}[X]$ of degree R such that $(q(D)\varphi_{\mathbb{C}})(t) = 0$ for all $t \geq 0$ if and only if $p_{\min} \mid q$.*

Remark 3.12 (Automatic smoothness and generator polynomial)**.** No differentiability is assumed in Theorem 3.11: integral smoothing of the shift semigroup (Section 4.1) forces $\varphi_{\mathbb{C}}|_{[0,\infty)}$ to agree with an analytic exponential polynomial, while any non-C^∞ continuous profile (such as $|s-1|$) has $\operatorname{srank}_{\mathbb{R}} = \infty$. Moreover, $\deg p_{\min} = R = \deg \chi_A$ and $p_{\min} \mid \chi_A$ imply $p_{\min} = \chi_A = \mu_A$ for the shift generator $A \in \mathcal{L}(\mathcal{W}(\varphi_{\mathbb{C}}))$.

3.4 (D) Border rank and the 3-D hierarchy (over $\mathbb{C}$: 4; over $\mathbb{R}$: $5 < 6 < 7$)

In plain terms, approximating a polynomial factor $se^{\lambda s}$ by a finite difference $(e^{(\lambda+h)s} - e^{\lambda s})/h$ of two pure Gaussians never saves terms in 2-D, but in $d \geq 3$ dimensions it saves $d-2$ real terms per polynomial degree—and for $K_3(x) = \|x\|^2 \cos(\|x\|^2)$ in 3-D, breaking radial symmetry in the approximants saves *one term further*, giving 4 over $\mathbb{C}$ and $5 < 6 < 7$ over $\mathbb{R}$ (real border rank 5, radial multi-Gaussian border rank 6, and exact real rank 7):

Theorem 3.13 (General d-D Multi-Gaussian Border-Rank Bound and Gap)**.** *Let $d \geq 2$ and let $\Gamma = [(\lambda_1, P_1), \ldots, (\lambda_m, P_m)]$ be any list of well-formed components:*

1. Exact gap identity (`multiTotalCost_eq_borderTotalCost_add`)*:*

$$C_d(\Gamma) = B_d(\Gamma) + (d-2)\sum_{u=1}^{m} \deg P_u. \tag{10}$$

2. Compact-open multi-Gaussian limit: *For $K(x) := \operatorname{profile}(\Gamma, \|x\|^2)$, substituting $u_h(s) = (e^{hs} - 1)/h$ into each $P_u(s)$ gives multi-Gaussian radial kernels $F_h : \mathbb{R}^d \to \mathbb{R}$ satisfying*

$\operatorname{srank}_{\mathbb{R}}(F_h) \le B_d(\Gamma)$ *and* $\|F_h - K\|_{E,\infty} \to 0$ *as* $h \to 0$ $(h \neq 0)$ *on every compact* $E \subset \mathbb{R}^d$, *so* $\underline{\operatorname{srank}}^{\mathrm{co}}_{\mathbb{R}}(K) \le B_d(\Gamma)$ *(strictly below* $C_d(\Gamma)$ *whenever* $d \ge 3$ *and* $\sum_u \deg P_u \ge 1$*; Lean theorem* `exists_rank_le_borderTotalCost_tendstoUniformlyOn_radial_profile`*).*

Theorem 3.14 (Border-Rank Sharpness Regimes and Strict Non-Sharpness on $\mathbb{R}^3$)**.** *Formalized in* `KernelRankLean/Approximation.lean` *(summary in Table 6):*

1. 2-D closure, complex border rank, and all-degree real-rate sharpness: *Over any Hausdorff topological field* $\mathbb{F}$*, 2-way rank is closed under pointwise convergence along nontrivial filters (*`separableRank_fin2_le_of_tendsto`*), implying* $C_2(\Gamma) \le \underline{\operatorname{srank}}^{\mathrm{pt}}_{\mathbb{R}}(K) \le \underline{\operatorname{srank}}^{\mathrm{co}}_{\mathbb{R}}(K) \le B_d(\Gamma)$ *for all* $d \ge 2$ *(*`totalCost_le_of_tendsto_radial_profile`*). Hence* $\underline{\operatorname{srank}}^{\mathrm{pt}}_{\mathbb{C}} = \underline{\operatorname{srank}}^{\mathrm{co}}_{\mathbb{C}} = \sum_{\mu \in S}(\deg P_\mu + 1)$ *for all complex profiles (bounded below by* `sum_natDegree_add_one_le_of_tendsto_radial_complex` *and achieved from above by* `exists_rank_le_sum_natDegree_add_one_tendstoUniformlyOn_radial_complex`*), and* $\underline{\operatorname{srank}}^{\mathrm{pt}}_{\mathbb{R}} = \underline{\operatorname{srank}}^{\mathrm{co}}_{\mathbb{R}} = B_d(\Gamma) = \sum_u(\deg P_u + 1)$ *whenever all rates of* Γ *are real (*$\operatorname{Im}\lambda_u = 0$*, arbitrary* $\deg P_u \ge 0$*;* `borderTotalCost_le_of_tendsto_radial_real_profile`*).*
2. 3-D sharpness at $k = 0$ and strict separation on $K_3(x) = \|x\|^2 \cos(\|x\|^2)$ (over $\mathbb{C}$: 4; over $\mathbb{R}$: $5 < 6 < 7$): *On* $\mathbb{R}^3$, $\underline{\operatorname{srank}}^{\mathrm{pt}}_{\mathbb{R}}(\cos\|\cdot\|^2) = \underline{\operatorname{srank}}^{\mathrm{co}}_{\mathbb{R}}(\cos\|\cdot\|^2) = 3 = B_3$ *(*`not_exists_rank_le_two_tendsto_radial_cos`*), whereas on* $K_3(x) = \|x\|^2\cos(\|x\|^2)$*, complex border rank is* 4*,* $\underline{\operatorname{srank}}^{\mathrm{pt}}_{\mathbb{R}}(K_3) \ge 5$ *(*`not_exists_rank_le_four_tendsto_radial_sq_mul_cos`*), and the* 5*-term non-Gaussian family* (11) *achieves* $\underline{\operatorname{srank}}^{\mathrm{co}}_{\mathbb{R}}(K_3) = 5 < B_3 = 6 < C_3 = 7$ *over* $\mathbb{R}$ *(*`exists_rank_le_five_tendstoUniformlyOn_radial_sq_mul_cos`*).*

Remark 3.15 (Explicit 5-term $\mathbb{R}^3$ limit)**.** Writing $c_j = \cos(x_j^2)$, $s_j = \sin(x_j^2)$, $p_{1,j} = \frac{1}{2}c_j - \frac{\sqrt{3}}{2}s_j$, $p_{2,j} = -\frac{1}{2}c_j - \frac{\sqrt{3}}{2}s_j$, and $q_{m,j} = x_j^2 p_{m,j}$ $(j \in \{0,1,2\})$, the 5-product real family $(h \neq 0)$

$$\begin{aligned} F_h(x) = {} & \frac{1}{3h}\prod_{j=0}^{2}(c_j + 4hx_j^2 c_j) + \frac{1}{h^3}\left(\prod_{j=0}^{2}(s_j + h^2 c_j) - \prod_{j=0}^{2} s_j\right) \\ & - \frac{4}{3h}\prod_{j=0}^{2}(p_{1,j} + hq_{1,j}) + \frac{4}{3h}\prod_{j=0}^{2}(p_{2,j} + hq_{2,j}) \end{aligned} \tag{11}$$

satisfies $F_h(x) - \|x\|^2\cos(\|x\|^2) = O(h)$ uniformly on compact subsets of $\mathbb{R}^3$ (verified algebraically in Section 4.4; its $O(h^{-3})$ floating-point cancellation penalty is analyzed theoretically in Section 7 and measured empirically in Table 20).

4 Mathematical Proofs and Dependency Roadmap

One-paragraph roadmap of the proofs. Figure 2 shows the exact module and hypothesis dependency DAG of our Lean 4 development. Every proof starts from the coordinate surjection $s_j = x_j^2$ (3), which turns $K(x) = g(\|x\|^2)$ on $\mathbb{R}^d$ into the additive Hankel kernel $g(s_0 + \cdots + s_{d-1})$ on $\mathbb{R}^d_{\ge 0}$. In the *necessity* branch, setting $x_2 = \cdots = x_{d-1} = 0$ bounds the 2-D translate space $\mathcal{W}(\varphi_{\mathbb{C}})$ to finite dimension; continuity of φ on $[0,\infty)$ alone makes half-line shifts T_t operator-norm continuous on $\mathcal{W}(\varphi_{\mathbb{C}})$, Nagumo's Bochner-integral trick produces $T_t = \exp(tA)$ without assuming differentiability, and generalized eigenspaces of A yield exponential polynomials (plus $p_{\min} \mid \chi_A$ in $d = 2$, Section 4.1). In the *upper-bound* branch, reversed Hasse polynomials times truncated exponentials turn $P(\sum_j s_j)$ into $[X^k]Q(X)$, extracted over $\mathbb{C}$ by $M^{\mathbb{C}}_{d,k}$ roots of unity, for real rates over $\mathbb{R}$ by Lagrange interpolation, and for nonreal rates over $\mathbb{R}$ by the $M^{\mathbb{R}}_{d,k}$ alternating angle-sampling identity (Section 4.2). In the *lower-bound* branch, linear independence of exponential polynomials turns arbitrary set-theoretic 1-D factors into linear functionals on

Table 6: Regimes in which the multi-Gaussian border bound equals exact pointwise and compact-open border rank.

Field & dim.	Representation / profile class	Exact border rank	Sharp?	Exported Lean 4 theorem(s)
$\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, d=2	All representations $(\forall \lambda_u, \deg P_u)$	$C_2(\Gamma) = B_2(\Gamma) = \mathrm{srank}_{\mathbb{F}}$	**Yes**	`separableRank_fin2_le_of_tendsto`
$\mathbb{C}$, any $d \geq 2$	All complex exp.-polynomials	$\sum_{\mu \in S}(\deg P_\mu + 1)$	**Yes**	`sum_natDegree_add_one_le_of_tendsto_radial_complex`, `exists_rank_le_sum_natDegree_add_one_tendstoUniformlyOn_radial_complex`
$\mathbb{R}$, any $d \geq 2$	All real rates $(\mathrm{Im}\,\lambda_u = 0, \forall \deg P_u)$	$\sum_u(\deg P_u + 1) = B_d(\Gamma)$	**Yes**	`borderTotalCost_le_of_tendsto_radial_real_profile`
$\mathbb{R}$, $d = 3$	Single nonreal Gaussian $\cos(\|x\|^2)$	$3 = B_3 = C_3$	**Yes**	`not_exists_rank_le_two_tendsto_radial_cos`
$\mathbb{R}$, $d = 3$	Degree-1 nonreal $\|x\|^2 \cos(\|x\|^2)$	$5 < B_3 = 6 < C_3 = 7$	**No**	`exists_rank_le_five_tendsto UniformlyOn_radial_sq_mul_cos`, `not_exists_rank_le_four_tendsto_radial_sq_mul_cos`
$\mathbb{R}$, $d \geq 3$	General $\mathrm{Im}\,\lambda_u > 0$ ($d \geq 4$ or $k_u \geq 2$)	In $[C_2(\Gamma),\ B_d(\Gamma)]$	Open	`totalCost_le_of_tendsto_radial_profile`, `exists_rank_le_border TotalCost_tendstoUniformlyOn_radial_profile`

$V = \mathcal{A}_\varphi$, where Alder–Strassen nilpotent peeling on modes $0, \ldots, d-2$ paired with $+1$ index shifting on mode $d-1$ proves optimality (Section 4.3). Finally, finite-difference substitutions $u_h(s)$, Schur-complement minor limits, the polynomial sum-of-squares trace $\mathrm{tr}((\mathrm{adj}(S_0)S_1)^2) \geq 0$, and periodic Fourier density on $\mathbb{R}/T\mathbb{Z}$ establish Theorems 3.2, 3.13, and 3.14 (Section 4.4).

4.1 Necessity and the two-dimensional core

Row-span dimension and 2-D complexification. In two coordinates over a field $\mathbb{F}$, let $\mathrm{rowSpan}(K) \leq (X \to \mathbb{F})$ be the $\mathbb{F}$-linear span of the row sections $\{x \mapsto K(x,y) : y \in X\}$. Then $\mathrm{srank}_{\mathbb{F}}(K) = \dim_{\mathbb{F}} \mathrm{rowSpan}(K) \in \mathbb{N} \cup \{\infty\}$: if $K(x,y) = \sum_{n=0}^{R-1} f_n(x) g_n(y)$, every row $x \mapsto K(x,y)$ lies in $\mathrm{span}\{f_0, \ldots, f_{R-1}\}$, whereas any basis $b_0(x), \ldots, b_{R-1}(x)$ of $\mathrm{rowSpan}(K)$ expands $K(x,y) = \sum_{n=0}^{R-1} c_n(y) b_n(x)$. For the two-variable Hankel kernel $(s_0, s_1) \mapsto g(s_0 + s_1)$ on $\mathbb{R}^2_{\geq 0}$, the row span is the *translate span*

$$\mathcal{W}(g) = \mathrm{span}_{\mathbb{F}}\{u \in \mathbb{R}_{\geq 0} \mapsto g(u+t) : t \in \mathbb{R}_{\geq 0}\} \leq (\mathbb{R}_{\geq 0} \to \mathbb{F}).$$

Scalar extension $\mathbb{R} \hookrightarrow \mathbb{C}$ satisfies $\mathrm{srank}_{\mathbb{C}}(K_{\mathbb{C}}) \leq \mathrm{srank}_{\mathbb{R}}(K)$ in every dimension d, with equality in $d = 2$ because any complex dependence $\sum_n (\alpha_n + i\beta_n) b_n(x) = 0$ among real functions $b_n : X \to \mathbb{R}$ forces $\sum_n \alpha_n b_n = 0$ and $\sum_n \beta_n b_n = 0$ separately.

Banach shift semigroup without differentiability. For any $d \geq 2$, if $\mathrm{srank}_{\mathbb{R}}(x \in \mathbb{R}^d \mapsto \varphi(\|x\|^2)) < \infty$, the slice inequality (4) and 2-D complexification give $\dim_{\mathbb{C}} \mathcal{W}(\varphi_{\mathbb{C}}) < \infty$. For any continuous $\psi : \mathbb{R} \to \mathbb{C}$ with $\dim_{\mathbb{C}} \mathcal{W}(\psi) = n < \infty$:

1. Equip $\mathcal{W}(\psi) \leq (\mathbb{R}_{\geq 0} \to \mathbb{C})$ with the Hausdorff product topology of pointwise convergence (avoiding boundedness assumptions on ψ). Since $\dim_{\mathbb{C}} \mathcal{W}(\psi) = n < \infty$, all Hausdorff vector topologies on $\mathcal{W}(\psi)$ coincide (`isNormableSpace_of_finiteDimensional`), making $\mathcal{W}(\psi)$ a complex Banach space (`FiniteDimensional.complete`).
2. Non-negative shifts $(T_t h)(u) = h(u+t)$ ($t \geq 0$) preserve $\mathcal{W}(\psi)$ without derivatives; continuity of ψ on $[0, \infty)$ gives pointwise and hence operator-norm continuity of $t \mapsto T_t \in \mathcal{L}(\mathcal{W}(\psi))$.
3. Following Nagumo [52] (Engel and Nagel [22, Thm. I.3.7]; cf. Anselone and Korevaar [3, p. 750]), continuity at 0 with $T_0 = I$ yields $\delta > 0$ such that the Bochner integral $I_\delta =$

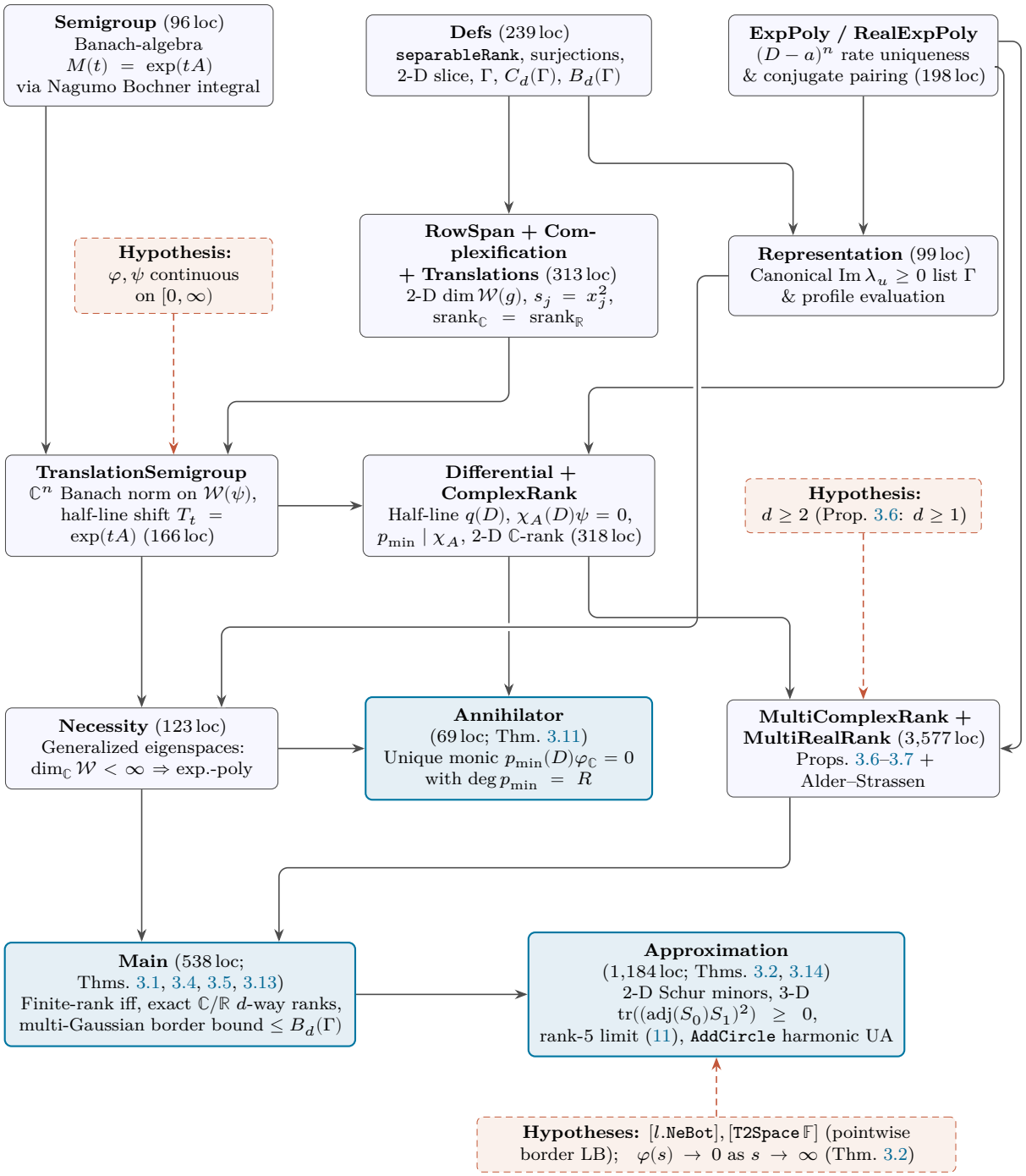


Figure 2: Module and hypothesis dependency graph of `KernelRankLean` (17 files, 27 `KernelRank.*` theorems; module breakdown and complete 32-theorem audit map in Appendix A). Solid arrows trace Lean `import` dependencies; dashed red boxes mark the exact module where each analytic, dimensional, or topological hypothesis enters.

$\int_0^\delta T_u\,du$ satisfies $\|I_\delta - \delta I\| < \delta$ and is invertible by Neumann series (`Semigroup.lean`). Changing variables in $T_t I_\delta = \int_t^{t+\delta} T_u\,du$ shows $t \mapsto T_t$ is right-differentiable on $[0,\infty)$ with commuting generator $A = (T_\delta - I)I_\delta^{-1}$, so $T_t = \exp(tA)$ and $\psi(t) = \mathrm{eval}_0(\exp(tA)\psi|_{\mathbb{R}_{\geq 0}}) \in C^\infty([0,\infty))$.

4. Decomposing $\mathcal{W}(\psi) = \bigoplus_{\mu\in S} \ker((A-\mu I)^n)$ over $\mathbb{C}$ gives $\exp(tA)v = e^{\mu t}\sum_{q=0}^{n-1} \frac{t^q}{q!}(A-\mu I)^q v$ on each generalized eigenspace, so $\psi(t) = \sum_{\mu\in S} P_\mu(t)e^{\mu t}$ on $[0,\infty)$; pairing conjugates $(\mu, \bar\mu)$ on $\varphi = \frac{1}{2}(\varphi_{\mathbb{C}} + \overline{\varphi_{\mathbb{C}}})$ yields a canonical representation Γ.

Cayley–Hamilton divisibility for Theorem 3.11. Differentiating $t \mapsto \exp(tA)\psi|_{\mathbb{R}_{\geq 0}}$ on $[0,\infty)$ gives $q(D)\psi(t) = \mathrm{eval}_0(q(A)\exp(tA)\psi|_{\mathbb{R}_{\geq 0}})$, so Cayley–Hamilton ($\chi_A(A) = 0$) yields $\chi_A(D)\psi = 0$ on $[0,\infty)$ with $\deg\chi_A = \dim_{\mathbb{C}}\mathcal{W}(\psi)$. Conversely, $(D-\mu)^{\deg P_\mu+1}$ kills $P_\mu(t)e^{\mu t}$ while multiplying leading terms at $\nu \neq \mu$ by $(\nu-\mu)^{\deg P_\mu+1} \neq 0$, so $q(D)\psi = 0$ iff $p_{\min}(X) = \prod_{\mu\in S}(X-\mu)^{\deg P_\mu+1}$ divides $q(X)$. Thus $p_{\min} \mid \chi_A$ gives $\sum_{\mu\in S}(\deg P_\mu + 1) = \deg p_{\min} \leq \dim_{\mathbb{C}}\mathcal{W}(\psi)$, matched from above by Taylor expansion of $P_\mu(s_0+s_1)e^{\mu(s_0+s_1)}$.

4.2 Constructive multi-way upper bounds via Hasse polynomials and roots of unity

Warm-up ($d = 3, k = 1$): extracting a degree-1 radial polynomial as an $[X^1]$ coefficient. Consider $P(s) = p_0 + p_1 s$ on $s = s_0 + s_1 + s_2$. Putting $U_0(X) = p_1 + (p_0 + p_1 s_0)X$, $U_1(X) = 1 + s_1 X$, and $U_2(X) = 1 + s_2 X$, the X^1-coefficient of $U_0(X)U_1(X)U_2(X)$ is

$$[X^1]((p_1+(p_0+p_1s_0)X)(1+s_1X)(1+s_2X)) = (p_0+p_1s_0)\cdot 1\cdot 1 + p_1 s_1\cdot 1 + p_1\cdot 1\cdot s_2 = P(s_0+s_1+s_2),$$

a product of three univariate polynomials whose j-th factor depends *only on* s_j. Since $\deg(U_0U_1U_2) \leq 3$, over $\mathbb{C}$ averaging $\zeta^{-r}U_0(\zeta^r)U_1(\zeta^r)U_2(\zeta^r)$ over the $M_{3,1}^{\mathbb{C}} = 3$ cube roots of unity ($\zeta = e^{2\pi i/3}$) isolates X^1, whereas over $\mathbb{R}$ for a nonreal rate $e^{\lambda(s_0+s_1+s_2)}$ ($\mathrm{Im}\,\lambda > 0$), Proposition 3.7 extracts $\mathrm{Re}([X^1]\prod_{j=0}^2 e^{\lambda s_j}U_j(X))$ using $M_{3,1}^{\mathbb{R}} = 2(3-1)(1) + 3 = 7$ real 3-way products.

General degree $k = \deg P$ ($d \geq 2$). Using the q-th Hasse derivative $D^{[q]}P(s_0) = \frac{1}{q!}P^{(q)}(s_0)$ defined by $P(s_0+t) = \sum_{q=0}^k (D^{[q]}P)(s_0)\,t^q$, define the *reversed Hasse polynomial* $H_P(s_0;X)$ on mode 0 and *truncated exponential polynomials* $E_k(s_j;X)$ on modes $j = 1,\dots,d-1$:

$$H_P(s_0;X) = \sum_{q=0}^k q!\,(D^{[q]}P)(s_0)X^{k-q},\ E_k(s_j;X) = \sum_{q=0}^k \frac{s_j^q}{q!}X^q,\ Q(X) = H_P(s_0;X)\prod_{j=1}^{d-1}E_k(s_j;X). \tag{12}$$

Binomial convolution gives $[X^q]\prod_{j=1}^{d-1}E_k(s_j;X) = (s_1+\cdots+s_{d-1})^q/q!$ for $0 \leq q \leq k$, so $[X^k]Q(X) = P(s_0+\cdots+s_{d-1})$ with $\deg Q \leq dk$. Three extraction schemes yield all upper bounds:

1. *Complex rates over $\mathbb{C}$:* Put $M_{d,k}^{\mathbb{C}} = (d-1)k+1$ and $\zeta = e^{2\pi i/M_{d,k}^{\mathbb{C}}}$. Since $d \geq 2$, every $n \in \{0,\dots,dk\}$ satisfies $-M_{d,k}^{\mathbb{C}} < -k \leq n-k \leq (d-1)k < M_{d,k}^{\mathbb{C}}$, so $M_{d,k}^{\mathbb{C}} \mid (n-k) \iff n = k$, giving $[X^k]Q(X) = \frac{1}{M_{d,k}^{\mathbb{C}}}\sum_{r=0}^{M_{d,k}^{\mathbb{C}}-1}\zeta^{-rk}Q(\zeta^r)$ in $M_{d,k}^{\mathbb{C}} = (d-1)k+1$ complex products.

2. *Real rates over $\mathbb{R}$ ($\mathrm{Im}\,\lambda = 0$, $P \in \mathbb{R}[X]$):* Since $A(X) = \prod_{j=1}^{d-1}E_k(s_j;X) \in \mathbb{R}[X]$ has $\deg A \leq (d-1)k < M_{d,k}^{\mathbb{C}}$, real Lagrange interpolation at $M_{d,k}^{\mathbb{C}}$ distinct real nodes expresses $P(\sum_j s_j)e^{\lambda\sum_j s_j}$ in $M_{d,k}^{\mathbb{C}} = (d-1)k+1$ real d-way products.

3. *Nonreal rates over* $\mathbb{R}$ *(*$\operatorname{Im}\lambda > 0$*, Proof of Propositions 3.6 and 3.7):* Put $M = M_{d,k}^{\mathbb{R}} = 2(d-1)k + d = (2k+1)d - 2k$ ($M = d$ at $k = 0$). Expanding $\operatorname{Re} w = \frac{1}{2}(\bar{w} + w)$ in (8) over subsets $S \subseteq \{0, \dots, d-1\}$ of size $m = |S|$ and tuples $q \in \{0, \dots, k\}^d$ with $(-1)^r = e^{i\pi r M/M}$ yields geometric phases $\exp(2\pi i r K(S,q)/M)$ with integer frequency

$$K(S,q) = (2k+1)m + \sum_{j \notin S} q_j - \sum_{j \in S} q_j - k \in \mathbb{Z}.$$

For $m = 0$, $-M < -k \le K \le (d-1)k < M$, so $M \mid K \iff \sum_j q_j = k$; for $m = d$, $K = M + (k - \sum_j q_j)$, so $M \mid K \iff \sum_j q_j = k$; and for $1 \le m \le d-1$ ($d \ge 2$), $1 \le (k+1)m - k \le K(S,q) \le (k+1)m + (d-1)k \le M-1$, so $\sum_{r=0}^{M-1} e^{2\pi i r K(S,q)/M} = 0$. Setting $U_0(X) = e^{\lambda s_0} H_P(s_0; X)$ and $U_j(X) = e^{\lambda s_j} E_k(s_j; X)$ ($j \ge 1$) gives $M_{d,k}^{\mathbb{R}} = 2(d-1)k + d$ real products.

4.3 Multilinear lower bounds via Alder–Strassen degree peeling

Step 1: From set-theoretic 1-D functions to multilinear dual functionals. Let V be a finite-dimensional $\mathbb{F}$-vector space ($\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$), let $v : \mathbb{R}_{\ge 0} \to V$ satisfy $\operatorname{span}_{\mathbb{F}}(v(\mathbb{R}_{\ge 0})) = V$, and let $\mathcal{M} : V^d \to \mathbb{F}$ be d-multilinear. Any rank-R decomposition (1) of $(s_0, \dots, s_{d-1}) \mapsto \mathcal{M}(v(s_0), \dots, v(s_{d-1}))$ defines unique linear functionals $\ell_{n,j} \in V^*$ on a basis $\{v(t_b)\}_{b=1}^{\dim V}$ with

$$\mathcal{M}(u^{(0)}, \dots, u^{(d-1)}) = \sum_{n=0}^{R-1} \prod_{j=0}^{d-1} \ell_{n,j}(u^{(j)}) \qquad \text{for all } u^{(0)}, \dots, u^{(d-1)} \in V. \tag{13}$$

Step 2: Intuition on $d = 3, k = 1$ ($K_3(x) = \|x\|^2 \cos(\|x\|^2)$ requires 7 real terms). Here $v(s) = (e^{is}, se^{is}) \in \mathbb{C}^2$ spans $V = \mathbb{C} \oplus \mathbb{C}\varepsilon \cong \mathbb{C}[T]/(T^2)$ ($\varepsilon^2 = 0$) over $\mathbb{R}$, with $\mathcal{M}((u_0,u_1),(v_0,v_1),(w_0,w_1)) = \operatorname{Re}(u_1 v_0 w_0 + u_0 v_1 w_0 + u_0 v_0 w_1)$. On $E = 0 \oplus \mathbb{C}\varepsilon \cong \mathbb{R}^2$, $\{\ell_{n,0}|_E\}$ must span E^* ($\dim_{\mathbb{R}} E^* = 2$), so a graph map $\psi_0(u_0) = (u_0, \gamma_0(u_0))$ kills *two* terms ($n = 0, 1$). Because $\varepsilon^2 = 0$, $\mathcal{M}(\psi_0(u_0), (0, v_1), (w_0, 0)) = \operatorname{Re}(u_0 v_1 w_0)$ is unaffected by γ_0, so a second map $\psi_1(v_0) = (v_0, \gamma_1(v_0))$ on mode 1 kills *two more* terms ($n = 2, 3$). Evaluating $\mathcal{M}(\psi_0(u_0), \psi_1(v_0), (0, w_1)) = \operatorname{Re}(u_0 v_0 w_1)$ on the remaining $R - 4$ terms requires ≥ 3 real terms (since every $\Lambda : \mathbb{C} \to \mathbb{R}$ has $\dim_{\mathbb{R}} \ker \Lambda \ge 1$), forcing $R \ge 7$.

Step 3: General multi-mode degree peeling over $\mathbb{C}$ and over $\mathbb{R}$. Over $\mathbb{C}$, $v(s)_{i,q} = \frac{s^q}{q!} e^{\mu_i s}$ spans $V = \prod_{i \in \iota} \mathbb{C}^{\tau_i^0}$ ($\tau_i^0 = \deg P_i + 1$), mapping the Hankel kernel to

$$\mathcal{T}_\tau(u^{(0)}, \dots, u^{(d-1)}) = \sum_{i \in \iota} \sum_{q_0 + \cdots + q_{d-1} < \tau_i} w_i(q_0 + \cdots + q_{d-1}) \prod_{j=0}^{d-1} u_{i,q_j}^{(j)}, \qquad w_i(\tau_i - 1) \ne 0 \ (\tau_i > 0).$$

Strong induction on $\sum_i \tau_i$ peels 1 term when $\tau_{i_0} = 1$, and when $\tau_{i_0} \ge 2$ uses $2(\tau_{i_0} - 1) \ge \tau_{i_0}$ to kill $d-1$ non-interacting terms along $e_{i_0, \tau_{i_0} - 1}$ on modes $0, \dots, d-2$ before shifting mode $d-1$ up by $+1$ ($\tau_{i_0} \mapsto \tau_{i_0} - 1$). Over $\mathbb{R}$, the parallel induction on $(\prod_{i \in \iota_{\mathbb{R}}} \mathbb{R}^{\tau_{\mathbb{R},i}^0}) \times (\prod_{u \in \iota_{\mathbb{C}}} \mathbb{C}^{\tau_{\mathbb{C},u}^0})$ peels 1 or $d-1$ terms on real blocks, d terms on nonreal $\tau_{\mathbb{C},u_0} = 1$ blocks (via $\dim_{\mathbb{R}} \ker(\Lambda : \mathbb{C} \to \mathbb{R}) \ge 1$), and $2(d-1)$ terms on nonreal $\tau_{\mathbb{C},u_0} \ge 2$ blocks, proving Theorem 3.5.

4.4 Proofs of border-rank bounds, adjugate obstruction, and universal approximation

Multi-Gaussian finite-difference limit (Theorem 3.13). Identity (10) follows from $(k+1) + (d-2)k = (d-1)k + 1$ and $d(k+1) + (d-2)k = 2(d-1)k + d$. Substituting $u_h(s) = (e^{hs} - 1)/h = L_h(e^{hs})$ ($\deg L_h = 1$) into $P(s)$ gives $Q_h = P \circ L_h$ of degree $\le k$, so $P(u_h(s))e^{\lambda s} =$

$\sum_{m=0}^{k}([Y^m]Q_h)\prod_{j=0}^{d-1}e^{(\lambda+mh)x_j^2}$ has real rank $\le k+1 = b_d(\lambda,P)$ when $\operatorname{Im}\lambda = 0$ and $\le d(k+1) = b_d(\lambda,P)$ by Proposition 3.6 when $\operatorname{Im}\lambda > 0$ (realizing the W-tensor degeneration [8, 14, 79] at $k=1,\lambda=0$). Since $|e^y - 1 - y| \le y^2 e^{|y|}$, $u_h(s) \to s$ uniformly on $[0,B]$ for $B = \sup_{x\in E}\|x\|^2$.

2-D minor closure and 3-D adjugate sum-of-squares (Theorem 3.14). If $\operatorname{srank}_{\mathbb{F}}(K) > r$ on $X_0 \times X_1$, Schur-complement induction finds an $(r+1)\times(r+1)$ evaluation submatrix $M(K)$ with $\det M(K) \neq 0$, whereas $\det M(F_\iota) = 0$ whenever $\operatorname{srank}_{\mathbb{F}}(F_\iota) \le r$; when all $\operatorname{Im}\lambda_u = 0$, $B_d(\Gamma) = C_2(\Gamma)$ gives sharpness immediately. In $d=3$, any $G : X^3 \to \mathbb{R}$ of rank $\le n$ has mode-0 slices $S_0 = V\operatorname{diag}(d_0)W$, $S_1 = V\operatorname{diag}(d_1)W$ satisfying the everywhere-defined sum-of-squares identity

$$\operatorname{tr}\Big(\big(\operatorname{adj}(S_0)\,S_1\big)^2\Big) = (\det V\ \det W)^2 \sum_{k=0}^{n-1}\Big(d_{1,k}\prod_{j\neq k} d_{0,j}\Big)^2 \ge 0.$$

Evaluating $\cos(\|x\|^2)$ on $(0,\sqrt{\pi/2})^2$ gives $\operatorname{tr} = -2 < 0$ (ruling out rank ≤ 2); evaluating $\|x\|^2\cos(\|x\|^2)$ on $(0,\sqrt{\pi/2},\sqrt{\pi},\sqrt{3\pi/2})^2$ gives $\operatorname{tr} = -1024(\pi/2)^8 < 0$ (ruling out rank ≤ 4; cf. [20, 45]). Conversely, expanding (11) in h cancels $-\frac{1}{3h}c_0c_1c_2$ and $\frac{1}{h}(c_0s_1s_2 + s_0c_1s_2 + s_0s_1c_2)$, leaving $\|x\|^2\cos(\|x\|^2) + O(h)$ uniformly on compact sets.

Periodic Fourier density for universal approximation (Theorem 3.2). Clamping $e^{+\alpha u}\varphi(u)$ on $[0, S_0+1]$ into a T-periodic continuous function $G \in C(\mathbb{R}/T\mathbb{Z},\mathbb{C})$ with $e^{-\alpha(T-1)}M_{\mathrm{b}} < \varepsilon/4$ and applying Fourier density (`span_fourier_closure_eq_top`) plus Proposition 3.6 yields (6) with global error $< \varepsilon$ on $\mathbb{R}^d$.

5 Lean 4 Formalization and Axiom Audit

The `KernelRankLean` library in Lean 4 [50, 48] comprises 17 files totaling 6,935 physical lines (6,586 nonblank; full module breakdown, the 32-theorem declaration map, and a controlled comparison of five standard and nonstandard semigroup formalizations appear in Appendix A, Tables 11–13). Note that source line counts reflect proof style, definition granularity, and Mathlib library coverage rather than intrinsic mathematical complexity. Roughly two-thirds of the library ($\approx$ 4,400 lines) consists of reusable domain-general components—coordinatewise `separableRank` on arbitrary finite Pi types (`Defs`), 2-way row-span rank and pointwise border closure (`RowSpan`, `Complexification`, `Approximation`), Banach-algebra semigroups (`Semigroup`, 96 lines via Mathlib's Bochner integral [21]), and multilinear Alder–Strassen degree peeling (`MultiComplexRank`, `MultiRealRank`)—extending formalized tensor-rank work from discrete circuits and fixed-size tensors [4, 35, 76] to infinite-dimensional function spaces on $\mathbb{R}^d$ across all $d \ge 2$ and $k \ge 0$. Four proof-engineering choices kept build times short and avoided unification bottlenecks:

1. *Dependent multi-way definition with homogeneous 2-D row span:* Defining `separableRank` on $\prod_{i\in\iota} X_i$ unifies d-D and 2-D theorems, while keeping `rowSpan` on homogeneous $\operatorname{Fin} 2 \to \alpha$ avoids 200,000-heartbeat higher-order unification timeouts on $\operatorname{Fin} 2 \to \mathbb{R}_{\ge 0}$.
2. *Explicit uniformity on `translateSpan`:* Installing the $\mathbb{C}^n$-transferred norm uniformity explicitly avoids a typeclass diamond with the subtype uniformity of $\mathbb{R}_{\ge 0} \to \mathbb{C}$.
3. *Cayley–Hamilton divisibility ($p_{\min} \mid \chi_A$):* Proving $\dim_{\mathbb{C}} \mathcal{W}(\psi) \ge \deg p_{\min}$ via polynomial divisibility avoids cyclic-vector constructions and yields Theorem 3.11 directly.
4. *Local `Classical.decEq` and global adjugate polynomials:* Local `let := Classical.decEq` avoids shadowing $\operatorname{Fin} d$'s decidable equality in `Finset.mul_prod_erase`, and $\operatorname{tr}((\operatorname{adj}(S_0)S_1)^2) \ge 0$ avoids slice-invertibility assumptions along limiting sequences.

Classical axioms, constructive witnesses, and mechanical verification audit. Although `separableRank` is marked `noncomputable` (as an infimum over arbitrary functions $X_i \to \mathbb{F}$ over classical Cauchy-quotient reals), every upper-bound and border-rank proof constructs explicit closed-form factor formulas. Across all 32 audited Lean 4 theorems (27 in `KernelRank.*` and 5 in `Experiments.*`; Tables 12–13), `#print axioms` in `scripts/check` reports *exclusively* Lean 4's three standard classical axioms: `propext` (propositional extensionality), `Quot.sound` (quotient soundness), and `Classical.choice` (Hilbert's choice operator, entailing excluded middle via Diaconescu's theorem and entering via $\mathbb{R}/\mathbb{C}$ arithmetic, surjection right inverses (2), and finite-dimensional basis extraction). Calculus of Inductive Constructions with `propext`, `Quot.sound`, and `Classical.choice` is equiconsistent with ZFC plus countably many inaccessible cardinals [50, 48]; `scripts/check` mechanically excludes `sorryAx`, `Lean.ofReduceBool`/`native_decide`, `unsafe`, and custom `axioms`, while positive and negative test theorems (`separableRank_radial_sq_mul_cos` $= 3d-2$ and `not_exists_rank_le_four_tendsto_radial_sq_mul_cos`) confirm non-vacuity of the 239-line specification in `Defs.lean`.

6 Applications: Analytical Constructions and Illustrative Case Studies

Following the scope boundary in Section 1.3, Section 6.1 derives exact continuous-domain operator identities that follow directly from Theorems 3.4–3.5 and Proposition 3.6, while Section 6.2 presents reproducible illustrative numerical and GPU case studies on finite pixel/voxel grids (with full numerical benchmark tables and 9-decimal coefficients collected in Appendix B, Tables 14–22).

6.1 (A) Analytical constructions from the exact-rank theory

Fourier rank invariance on $\mathbb{R}^d$. For $\operatorname{Re}\lambda < 0$, differentiating $\mathcal{F}_{\mathbb{R}^d}[e^{\lambda\|x\|^2}](\xi) = (\pi/(-\lambda))^{d/2} \exp(\|\xi\|^2/(4\lambda))$ k times with respect to λ maps $P(\|x\|^2)e^{\lambda\|x\|^2}$ ($\deg P = k$) to $\widetilde{P}(\|\xi\|^2)e^{\|\xi\|^2/(4\lambda)}$ with $\deg\widetilde{P} = k$ and $\operatorname{Re}(1/(4\lambda)) < 0$. By Theorems 3.4 and 3.5, the d-dimensional Fourier transform preserves exact complex and real d-way separable rank across all integrable radial exponential-polynomial kernels.

Harmonic "beach-ball" differential operators and analytical radial derivatives. Complex partial-derivative operators $(\partial_x + i\partial_y)^m$ provide a rotation-equivariant basis for 2-D steerable filters [26, 38, 7, 62]. Specializing Proposition 3.6 to equal symbolic factors $z_0 = \cdots = z_{m-1} = \partial_x + i\partial_y$ ($D_\theta := \cos\theta\,\partial_x + \sin\theta\,\partial_y$) gives, for every $m \geq 1$ (odd or even),

$$\operatorname{Re}\big((\partial_x + i\partial_y)^m\big) = \frac{N_{\text{naive}}(m)}{m} \sum_{r=0}^{m-1} (-1)^r D^m_{\pi r/m}. \tag{14}$$

On a regular $2m$-point ring $\theta_n = \pi n/m$ of radius $h \to 0$, $D_{\theta_{r+m}} = -D_{\theta_r}$ and Fourier orthogonality of $(-1)^r$ unify the odd-order "beach-ball" operators of Niemitalo [56] and the even-order operators of Perona [60]:

$$\lim_{h\to 0} \frac{1}{h^m} \sum_{n=0}^{2m-1} (-1)^n f\Big(x + h\cos\frac{\pi n}{m},\, y + h\sin\frac{\pi n}{m}\Big) = \frac{1}{(m-1)!\,2^{m-2}} \operatorname{Re}\big((\partial_x + i\partial_y)^m\big) f(x,y).$$

Moreover, since $\partial_{\bar z}(z) = 0$ for $\partial_{\bar z} = \frac{1}{2}(\partial_x + i\partial_y)$, induction yields $\big(\frac{\partial_x + i\partial_y}{2}\big)^m \varphi(x^2+y^2) = (x+iy)^m \varphi^{(m)}(x^2+y^2)$ with no lower-order radial derivatives [7, 62]. When φ is a multi-Gaussian profile of 2-D rank $R = n_{\text{re}} + 2n_{\text{cx}}$ (all $\lambda_u \neq 0$, e.g. [57]), differentiating analytically on $\mathbb{R}^2$ before grid sampling avoids discrete finite-difference Nyquist distortion [55] and gives exact continuous

Table 7: Overview of the illustrative numerical and GPU case studies (full tables in Appendix B).

Theory reference	Finite-grid / precision question	Illustrative outcome	Section / Table
Thm. 3.5 ($d = 2$)	Sampled SVD rank vs. $C_2(\Gamma)$; 2-D vs. 3-D lifting	Sampled SVD rank equals $C_2(\Gamma)$ across all 18 blur families; $2c$ beats $1c_2, 1r_4$ in both ripple and C_3	Sec. 6.2.1; Tab. 14
Sec. 7 & $\mathrm{GL}_R(\mathbb{R})$	Orthogonalization vs. magnitude reordering & `fp32` outer sum	Reordering (8.18/255) and `fp32` outer sum (6.76/255) fail on $6c$; $\mathrm{GL}_R(\mathbb{R})$ cuts 20-scene `fp16` error to 0.30–0.31/255	Sec. 6.2.1; Tab. 16
Thm. 3.1	Discrete rank-R stencils vs. continuous profiles	Converges onto $1c/2c/4c/6c$ for $R_{\mathrm{px}} \geq 16$ (warm-starts $1.3\times$–$2.2\times$ lower at k=0); hits 0 when $L{+}1 \leq R$	Sec. 6.2.1; Tab. 15
Props. 3.6–3.7 ($d = 3$)	Exact R vs. naive split, 4 FDTD baselines, and `cuFFT`	24.0% GPU (1.98 → 1.51 ms, +28 MiB) and 21.2% CPU savings; 30.9× measured speedup vs. 3×-refined 25-pt FDTD	Sec. 6.2.2; Tabs. 18–19
Thm. 3.14 & Rem. 3.15	Finite-h precision of 5-, 6-, 7-term K_3 formulas	Empirical `fp16`/`fp32`/`fp64` minima match $\Theta(\varepsilon_{\mathrm{mach}}^{1/4})$, $\Theta(\varepsilon_{\mathrm{mach}}^{1/2})$, and $\Theta(\varepsilon_{\mathrm{mach}})$ (20.4%, 4.8%, 0.28% in `fp16`)	Sec. 6.2.2; Tab. 20
Thm. 3.5 (shared λ_u)	Spatially varying $c(\mathbf{x})$ via local weights $\gamma_u(\nu(\mathbf{x}))$	$\leq 0.185\%$ dispersion across $\nu \in [0.35, 0.80]$ and 7.8% 3-D lens error where 1-step `cuFFT` is inapplicable	Sec. 6.2.2; Fig. 7

2-way real ranks $(\lfloor m/2 \rfloor + 1)R$ for $\cos(m\theta)$ and $\lceil m/2 \rceil R$ for $\sin(m\theta)$ (reducing a 161×161 gradient pair from 51,842 taps/pixel to 2,576 at $R = 4$, a 20× reduction or 40× with even/odd folding).

Continuous pass-count bounds for simultaneous steerable pairs. Splitting $(C_m, S_m) = (A_c + B_c, A_s + B_s)$ into P_y shared vertical passes plus horizontal passes and P_x shared horizontal passes plus vertical passes satisfies $\dim \mathrm{rowSpan}_x(C_m, S_m) \leq \mathrm{rank}(A_c) + \mathrm{rank}(A_s) + P_x$ and $\dim \mathrm{colSpan}_y(C_m, S_m) \leq P_y + \mathrm{rank}(B_c) + \mathrm{rank}(B_s)$; since C_m, S_m span $\{x^0, \ldots, x^m\}g_n(x)$ and $\{y^0, \ldots, y^m\}g_n(y)$, *every* two-stage continuous 1-D schedule requires $\geq 2(m+1)R$ passes ($4R$ for (h_x, h_y)). Savings instead come from even/odd tap-parity folding, sharing passes across multiple harmonic orders ($5R$ vs. $6R$ for $m = 0, 1$), and recursive coordinate halving $T(d) = T(\lfloor d/2 \rfloor) + T(\lceil d/2 \rceil) + d \leq d(\lceil \log_2 d \rceil + 1)$ for d-D multi-Gaussian gradients ($8R_3$ vs. $9R_3$ in $d = 3$).

6.2 (B) Illustrative numerical and GPU case studies

Hardware platform and timing protocol. All illustrative experiments in `Applications/` (`blur_2d.py`, `acoustics_3d.py`, `gpu_kernels.py`; Table 7) ran on Linux `7.0.0-43-generic` (`x86_64`) with an Intel Core Ultra 5 125H CPU (14 cores / 18 threads, 61 GiB DDR5) and an external NVIDIA GeForce RTX 4070 Ti Super eGPU (16,376 MiB GDDR6X, 48 MB L2 cache, 672 GB/s peak VRAM bandwidth, PCIe Gen3 ×4, driver `615.71.09`) under Python `3.14.4`, NumPy `2.5.4`, SciPy `1.18.1` (HiGHS), PyTorch `2.15.0` (CUDA `13.2`), and Triton `3.9.0`. Every GPU timing uses on-device `torch.cuda.Event` barriers around VRAM-resident buffers (excluding PCIe transfers), reports mean ± std across $N = 10$ batches of 15 calls (150 launches; 6×2 on `DirectStep3D`; inter-run autotuning shift $\leq 6\%$ in 2-D, $\leq 3\%$ in 3-D), and asserts agreement with a `float64` reference (max `float32` error $\leq 6.6 \times 10^{-6}$). Table 8 defines the accuracy and isotropy metrics used below.

6.2.1 Case study I: 2-D circular bokeh blur

Approximating the discontinuous bokeh disc $\mathbf{1}_{r \leq 1}$ ($\mathrm{srank}_{\mathbb{R}} = \infty$) across a transition band $[1, 1.2]$ via LP/Nelder–Mead search yields three practical findings (full numbers in Appendix B, Tables 14–17):

Table 8: Error, isotropy, and dispersion metrics used across the 2-D and 3-D case studies.

Metric	Symbol	Definition	Domain / units
2-D Chebyshev ripple	$\delta(\varphi)$	$\max_{r\in[0,1]\cup[1.2,\,1.2\sqrt{2}]} \lvert\varphi(r^2)-\mathbf{1}_{r\le1}\rvert$	Pass/stopband L^∞
Discrete shell anisotropy	$\Delta_{\mathrm{sh}}(K)$	$\max_{N,\,\lvert S_N\rvert>1}(\max_{S_N} K - \min_{S_N} K)$ on $S_N=\{\lVert\mathbf{m}\rVert_2^2=N\}$	Integer shells
`float16` sRGB / linear err.	$E_{\mathrm{sRGB}}, E_{\mathrm{lin}}$	$255\max\lvert\mathrm{sRGB}(\hat{I}_{16}^{[0,1]})-\mathrm{sRGB}(I_{64}^{[0,1]})\rvert$ and $255\max\lvert\hat{I}_{16}-I_{64}\rvert$	/255 (N=20 HDR scenes)
3-D phase-velocity err.	$\lvert\delta c/c_0\rvert, k_\tau$	$\max_{0<\lVert\mathbf{k}\rVert_2\le k_c}\lvert\arccos(\widehat{M}(\mathbf{k})/2)/(\nu\lVert\mathbf{k}\rVert_2)-1\rvert$ (232,620 octant pts)	% error; cutoff k_τ
3-D pulse waveform err.	E_{pulse}	$\lVert p_{\mathrm{num}}(\mathbf{x}_r,\cdot)-p_{\mathrm{ref}}(\mathbf{x}_r,\cdot)\rVert_{\ell^2}/\lVert p_{\mathrm{ref}}(\mathbf{x}_r,\cdot)\rVert_{\ell^2}$ vs. D'Alembert	Relative ℓ^2 (%)

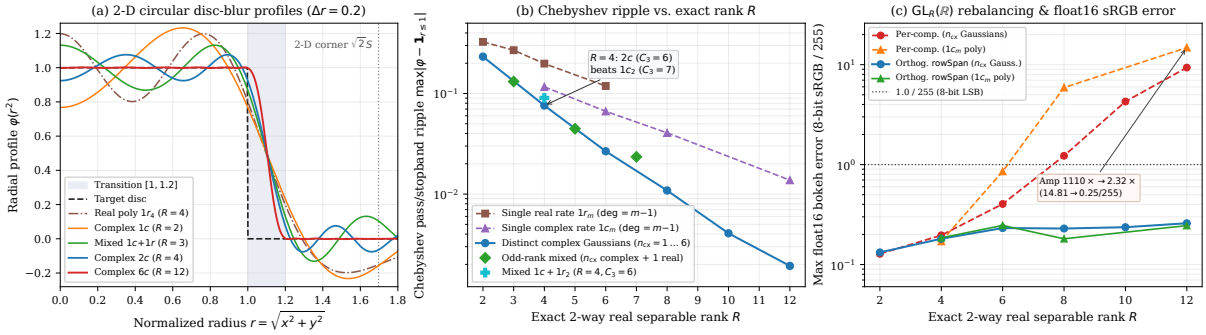


Figure 3: Two-dimensional circular disc blur ($\Delta r = 0.2$, `Applications/blur_2d.py`): (a) minimax radial profiles $\varphi(r^2)$ against the unit disc; (b) minimax pass/stopband ripple across exact 2-way real ranks $R \in [2, 12]$; (c) max IEEE `float16` bokeh error (/255, 512×512 HDR scene, $48\times$ glints, $R_{\mathrm{px}} = 25, W = 61$) before and after nuclear-norm-optimal $\mathrm{GL}_R(\mathbb{R})$ orthogonalization.

1. **Orthogonalization vs. non-orthogonalized `float16` baselines (Table 16):** On $R = 12$ ($6c$ and $1c_6$), individual raw components exceed K by $\kappa_\infty^{\mathrm{raw}} = 595$–$1{,}110\times$. Across $N = 20$ HDR scenes ($R_{\mathrm{px}} = 25, W = 61, 48\times$ glints), norm balancing ($7.90\pm0.92/255$ on $6c$), magnitude reordering alone ($8.18 \pm 0.92/255$), and even a `float32` outer accumulator over `float16` 1-D passes ($6.76\pm1.10/255$) all leave severe rounding inside each 1-D pass, whereas nuclear-norm-optimal $\mathrm{GL}_R(\mathbb{R})$ QR/eigenvalue orthogonalization ($\kappa_\infty \le 2.49$) drops sRGB error to $0.30 \pm 0.06/255$ (and ascending $|e_n|$ order cuts linear highlight error from 0.40 to 0.15/255).

2. **GPU crossovers vs. dense W^2 Triton and `cuFFT` (Table 17):** On 1920×1080 RGBA, L2-strip `SeparableBlur2D` is $28.0\times$ (`fp32`) / $29.5\times$ (`fp16`) faster than tuned dense $W \times W$ Triton at $R_{\mathrm{px}} = 64, R = 4$, and beats 2-D `cuFFT` (0.45–0.56 ms) at $R = 4, R_{\mathrm{px}} \in \{8, 16\}$ (0.33–0.35 ms), whereas `cuFFT` wins for $R \ge 8$ or $R_{\mathrm{px}} \ge 32$ (and `PocketFFT` wins $4\times$–$6\times$ on CPU).

3. **Discrete rank-R convergence and warm-starting (Figures 4–5, Table 15):** Direct ℓ_2 SVD incurs 17.7%–19.2% cross-streak anisotropy, whereas shell-isotropic discrete minimax ($\Delta_{\mathrm{sh}} \to 0$) hits 0 error when $L+1 \le R$ and converges onto our continuous $1c/2c/4c/6c$ profiles once $R_{\mathrm{px}} \ge 16$ (0.07582 vs. 0.07593 at $R_{\mathrm{px}} = 64$), where continuous warm-starts start $1.3\times$–$2.2\times$ below cold SVD.

6.2.2 Case study II: 3-D acoustic wave propagation and border-rank precision

Discretizing $\partial_{tt}p = c_0^2\nabla^2 p$ by $p^{n+1}[\mathbf{m}]+p^{n-1}[\mathbf{m}] = (M*p^n)[\mathbf{m}]$ ($\nu = c_0\Delta t/\Delta x$) requires $\widehat{M}(\mathbf{k}) \approx 2\cos(\nu\lVert\mathbf{k}\rVert_2)$ [67] with $|\widehat{M}(\mathbf{k})| \le 2$; purely axial stencils leave an irreducible $O(\nu^4 k^4)$ diagonal gap,

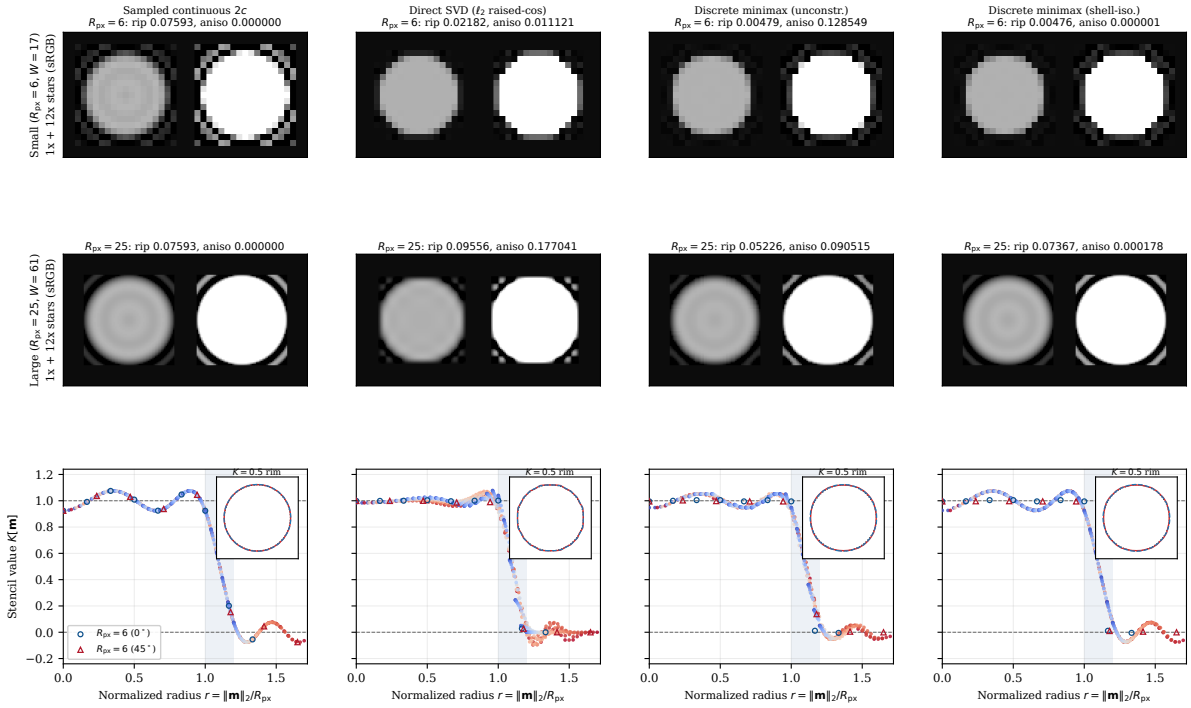


Figure 4: Comparison at $R = 4$ between sampled continuous $2c$ Gaussians, direct ℓ_2 SVD of a raised-cosine disc, unconstrained discrete rank-4 minimax, and shell-isotropic discrete rank-4 minimax (`blur_2d.py`): pixelated small ($R_{\text{px}} = 6, W = 17$, top) and large ($R_{\text{px}} = 25, W = 61$, middle) sRGB bokeh on 1× and 12× HDR stars, and radial pixel scatter / $K = 0.5$ contours (bottom).

while 27-point `IWB` [40] is identically the $R = 1, L = 1$ Gaussian stencil $(2 - 4\nu^2)\delta_{\mathbf{0}} + \frac{\nu^2}{2}2^{-\|\mathbf{m}\|^2}$ (9.38% error by 0.80π). Cutting-plane LP optimization of wider radial stencils ($\leq 7.0 \times 10^{-7}$ 1,000-step `DCT-II` Neumann energy drift) yields four takeaways (Table 9 and Appendix B, Tables 18–20):

1. **Exact 3-per-pair vs. naive 4-per-pair on 256^3 (Table 9, top):** Replacing 9 naive real terms by 7 exact angle-sampled terms cuts 256^3 step time by 24.0% on GPU (1.98 → 1.51 ms, +28 MiB slab scratch; 21.2% on CPU) while holding 3-D dispersion to 0.049% up to 0.80π (2.5% pulse error vs. 78.2%–109.3% for unrefined FDTD).

2. **Analytical FDTD scaling vs. measured multi-grid calibration (Table 9, bottom; Table 19):** Analytically extrapolating voxel spacing as $\Delta x \propto k_\tau$ at $\tau = 0.1\%$ ($k_{0.1\%} = 0.81\pi$ vs. 0.236π, ratio 3.43×) predicts 40× less field memory and 52 × –71× less runtime for `Rank-7` than 25-pt FDTD (15× memory and 13 × –19× time at $\tau = 0.5\%$). Direct multi-grid GPU runs on the same physical pulse calibrate both regimes: 25-pt FDTD needs 3.0× refinement (384^3) to match 128^3 `Rank-7`'s 1.68% pulse error (*measured* 30.9× *speedup,* 14.4× *less VRAM*), and at 3.5× refinement (448^3) takes 57.4× (`fp32`) to 70.3× (`fp16`) longer.

3. **Empirical border-rank precision sweep (Table 20):** Evaluating 7-, 6-, and 5-term formulas for $K_3(x) = \|x\|^2 \cos(\|x\|^2)$ across `fp16`/`fp32`/`fp64` confirms $h \to 0$ convergence alongside the $\Theta(\varepsilon_{\text{mach}})$, $\Theta(\varepsilon_{\text{mach}}^{1/2})$, and $\Theta(\varepsilon_{\text{mach}}^{1/4})$ precision floors derived in Section 7 (0.28%, 4.8%, and 20.4% min error in `fp16`).

4. **Spatially varying sound speed $c(\mathbf{x})$ (Figure 7):** Fixing shared Gaussian rates ($2c +$

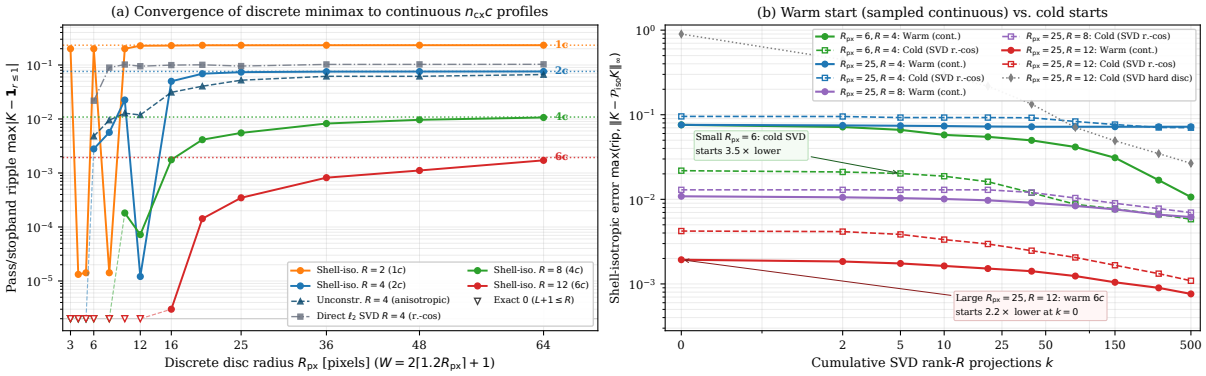


Figure 5: Discrete-to-continuous convergence and warm-starting (`blur_2d.py`): (a) pass/stopband Chebyshev ripple across $R_{\text{px}} \in [3, 64]$ (∇: exact 0 error when $L + 1 \le R$); (b) shell-isotropic minimax objective $\max(\text{rip}, \Delta_{\text{iso}})$ across SVD steps $k \in [0, 500]$ comparing continuous-Gaussian warm starts against raised-cosine and hard-disc SVD cold starts.

Table 9: Headline 3-D acoustic comparison on the RTX 4070 Ti Super (full details in Tables 18–19): fixed 256^3 single step (top) and equal-problem broadband pulse $L_{\text{phys}} = 128, \sigma_0 = 1.40, c_0 T_{\text{phys}} = 42$ (bottom).

Experiment	Scheme (ν)	Grid	GPU time	Total / scratch VRAM	Dispersion / pulse err.	Ratio vs. `Rank-7`
Fixed 256^3 step	`Rank-9` naive (4/pair, 1.00)	256^3	1.98 ± 0.02 ms	128 + 28 MiB	0.049% / 2.5%	1.31× time (−24%)
Fixed 256^3 step	`Rank-7` exact (3/pair, 1.00)	256^3	1.51 ± 0.03 ms	128 + 28 MiB	0.049% / 2.5%	1.00× (`fp16`: 1.09 ms)
Fixed 256^3 step	Periodic `cuFFT` k-space (1.00)	256^3	1.44 ± 0.05 ms	128 + 101 MiB	0% (periodic only)	0.95× time, 3.6× scratch
Fixed 256^3 step	25-pt $O(\Delta x^4, \Delta t^4)$ FDTD ($1/\sqrt{3}$)	256^3	0.325 ± 0.0002 ms	128 + 0 MiB	10.0% / 78.2%	0.22× time (dispersed)
Equal pulse (meas.)	`Rank-7` exact (`fp32`, 1.00)	128^3	7.90 ms (42 steps)	30 MiB total	1.68% max pulse err.	1.00× (`fp16`: 6.45 ms)
Equal pulse (meas.)	25-pt FDTD (1× coarse, $1/\sqrt{3}$)	128^3	3.15 ms (73 steps)	16 MiB total	73.0% max pulse err.	Unresolved pulse
Equal pulse (meas.)	25-pt FDTD (3.0× refined, $1/\sqrt{3}$)	384^3	244.5 ms (219 steps)	432 MiB total	1.60% max pulse err.	30.9× **time,** 14.4× **VRAM**
Equal pulse (meas.)	25-pt FDTD (3.5× refined, $1/\sqrt{3}$)	448^3	453.2 ms (255 steps)	686 MiB total	0.87% max pulse err.	57.4× (`fp32`) / 70.3× (`fp16`)
Equal τ=0.1% (extrap.)	25-pt FDTD (3.43× analytic)	$(3.43N)^3$	Analytic $(k_\tau)^4/\nu$ law	40.4× field mem.	Matched $k_{0.1\%}$ cutoff	Est. 52 × –71× time

$2r, R = 8, W = 13$) keeps all 8 separable passes shift-invariant while local weights $\gamma_u(\nu(\mathbf{x}))$ track $c(\mathbf{x})$, giving $\le 0.185\%$ dispersion across $\nu \in [0.35, 0.80]$ and 7.8% 3-D lens refraction error where 1-step `cuFFT` does not apply.

7 Limitations and Numerical Stability

We collect the theoretical floating-point conditioning bounds and practical scope boundaries of our decompositions in one place:

Theoretical floating-point stability across dimensions d, degrees k, and border steps h. In IEEE arithmetic with unit roundoff $\varepsilon_{\text{mach}}$ (`float16`: $2^{-11} \approx 4.9 \times 10^{-4}$; `float32`: $2^{-24} \approx 6.0 \times 10^{-8}$; `float64`: $2^{-53} \approx 1.1 \times 10^{-16}$):

1. **Degree-0 angle sampling vs. inter-component cancellation:** For $z_j = \beta^{1/d} e^{(-\alpha+i\omega)x_j^2}$ in Proposition 3.6, $|\operatorname{Re}(z_j e^{-i\pi r/d})| \le |z_j|$ bounds the absolute term sum by

$$\frac{N_{\text{naive}}(d)}{d} \sum_{r=0}^{d-1} \prod_{j=0}^{d-1} \left|\operatorname{Re}(z_j e^{-i\pi r/d})\right| \le N_{\text{naive}}(d)\, |\beta|\, e^{-\alpha \|x\|^2}, \qquad N_{\text{naive}}(d) = 2^{d-1}. \tag{15}$$

For $d \in \{2, 3, 4\}$ ($N_{\text{naive}}(d) = 2, 4, 8$), single-component angle sampling is well-conditioned, whereas for $d \gg 1$ ($d \ge 20$) $N_{\text{naive}}(d)$ grows exponentially. In low dimensions, large amplification ($\kappa_\infty^{\text{raw}} = 595$ on 6c) arises instead from *inter-component cancellation* $\sum_u |\beta_u| e^{-\alpha_u s} \gg |\varphi(s)|$: curable in $d = 2$ by $\text{GL}_R(\mathbb{R})$ orthogonalization (Table 16), whereas rigid $d \ge 3$ CP decompositions [41] require keeping $\sum_u |\beta_u|$ bounded during optimization.

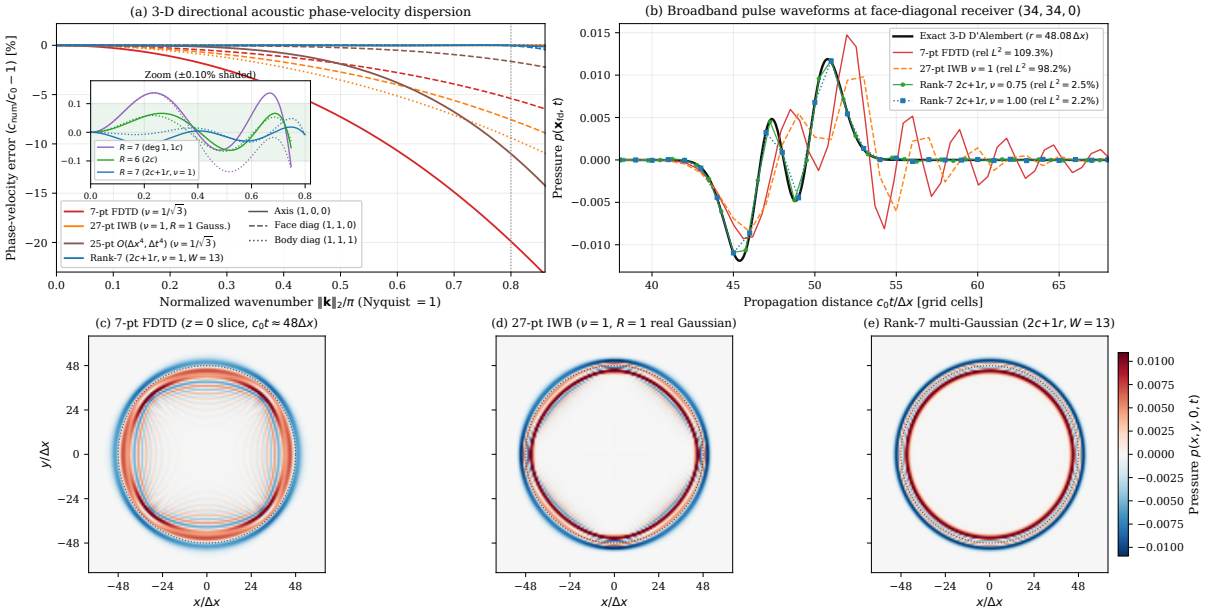


Figure 6: Three-dimensional acoustic wave propagation (`Applications/acoustics_3d.py`): (a) phase-velocity error ($c_{\mathrm{num}}(\mathbf{k})/c_0 - 1$) along $(1,0,0)$, $(1,1,0)$, and $(1,1,1)$; (b) broadband pulse ($\sigma_0 = 1.40\,\Delta x$, 256^3) at $(34,34,0)$ vs. exact 3-D D'Alembert; (c–e) slices $p(x,y,0,t)$ at $c_0 t \approx 48\,\Delta x$.

2. **High-degree polynomials ($k = \deg P \geq 1$):** Rescaling $X \mapsto \rho X$ ($\rho > 0$) in (12) bounds the intermediate sum on $P(s) = s^k/k!$ ($s = \|x\|^2 > 0$) by $N_{\mathrm{naive}}(d)\rho^{-k}e^{\rho s}|e^{\lambda s}|$; picking the saddle-point radius $\rho_\star = k/s$ yields mild Stirling growth in k,

$$\frac{N_{\mathrm{naive}}(d)\rho_\star^{-k}e^{\rho_\star s}}{s^k/k!} = N_{\mathrm{naive}}(d)\,k!\left(\frac{e}{k}\right)^k \sim N_{\mathrm{naive}}(d)\sqrt{2\pi k}, \tag{16}$$

   whereas unscaled $\rho = 1$ grows as $k!e^s/s^k$, and real Lagrange interpolation on $M = (d-1)k+1$ equispaced nodes suffers exponential Lebesgue growth $\Lambda_M \sim \frac{2^M}{eM\log M}$ unless evaluated on Chebyshev nodes ($\Lambda_M = \frac{2}{\pi}\log M + O(1)$).

3. **Finite-h border-rank approximations:** Approximating $s^k e^{\lambda s}$ via $u_h(s)^k e^{\lambda s}$ balances $O(ks^{k+1}|h|)$ truncation error against $O(2^k|h|^{-k}\varepsilon_{\mathrm{mach}})$ cancellation at $|h_\star| \asymp \varepsilon_{\mathrm{mach}}^{1/(k+1)}$, imposing a relative-error floor of $\Theta(\varepsilon_{\mathrm{mach}}^{1/(k+1)})$—and $\Theta(\varepsilon_{\mathrm{mach}}^{1/4})$ for the $O(h^{-3})$ 5-term $\mathbb{R}^3$ limit (11), matching the 20.4% (`float16`), 1.62% (`float32`), and 0.014% (`float64`) minima measured in Table 20. Border-rank degenerations in $d \geq 3$ are therefore complexity barriers and models of coalescing-rate ill-conditioning rather than finite-precision evaluators; practical $d \geq 3$ stencils should use well-separated $k = 0$ Gaussians.

Domain, discretization, rendering, boundary, and verification boundaries. (i) *Domain & discretization:* Theorems 3.1–3.5 apply to continuous profiles on $[0,\infty)$ (leaving $s < 0$ unconstrained); discontinuous indicators ($\mathbf{1}_{r\leq 1}$), non-L^2 singular kernels ($1/\|x\|$), and aliased integer-grid samples ($\mathrm{Im}\,\lambda_u \in \pi\mathbb{Z}$ or $W < R$) lie outside exact continuous rank, and outer rate search (α_u, ω_u) and Brillouin-octant stability audits are numerical rather than interval-certified. (ii) *Rendering & FFT crossovers:* $[-S,S]^2$ truncation has square support, oscillatory flat-top approximations can ring negatively around extreme ($10^4\times$) HDR impulses without clamping, spatially varying bokeh $R_{\mathrm{px}}(x,y)$ needs depth-layering [27], and 2-D `cuFFT` beats FIR cascades for shift-invariant blurs with $R \geq 8$ or $R_{\mathrm{px}} \geq 32$. (iii) *Wide acoustic stencils & hardware specificity:* Wide stencils ($L = 4$–8) cross thin baffles and lack passive closures at curved/impedance

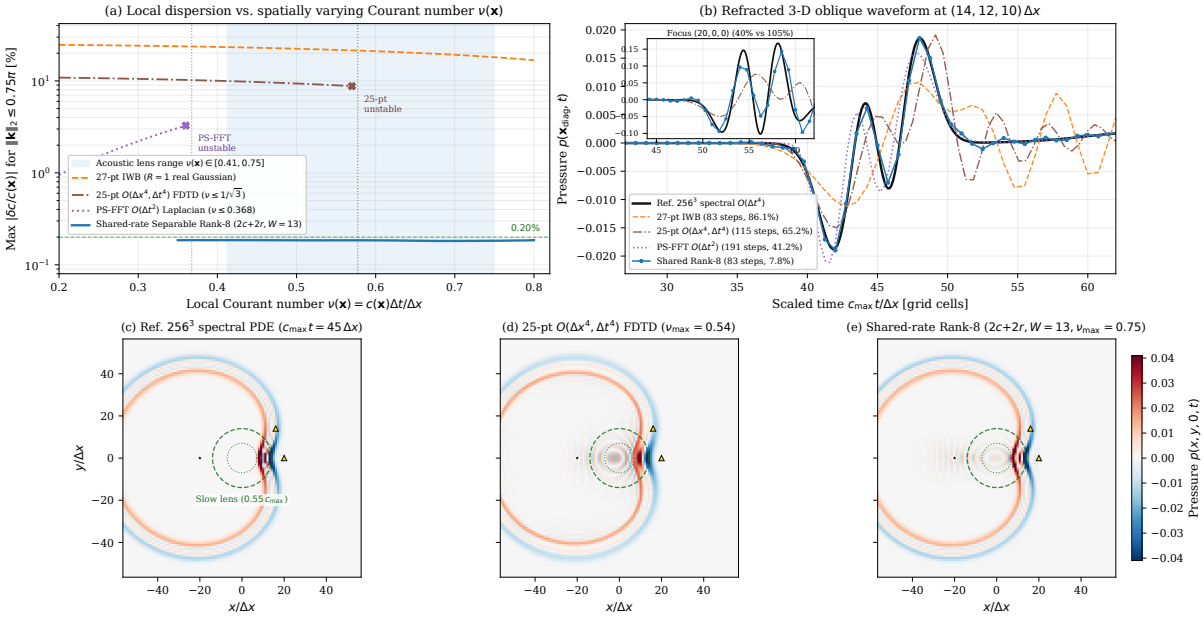


Figure 7: Spatially varying acoustic lens $c(\mathbf{x})/c_{\max} = 1 - 0.45e^{-\|\mathbf{x}\|^2/(14\Delta x)^2}$ where 1-step `cuFFT` k-space does not apply (`acoustics_3d.py --hetero-only`): (a) worst-case 3-D dispersion up to 0.75π across $\nu(\mathbf{x})$; (b) refracted waveforms (128^3) at $(14, 12, 10)\,\Delta x$ and caustic $(20, 0, 0)\,\Delta x$ vs. 256^3 fourth-order spectral PDE reference; (c–e) mid-plane slices at $c_{\max} t = 45\,\Delta x$.

Table 10: Hierarchy of exact and border ranks, illustrated on $K_3(x) = \|x\|^2 \cos(\|x\|^2)$ in $d = 3$ ($\mathrm{Re}(se^{is})$).

Rank notion	Plain-language meaning (for $K : \mathbb{R}^d \to \mathbb{R}$)	General $d \ge 2$ formula / bound	On K_3
$\mathbb{C}$-border / 2-D slice	Least complex terms in a compact-open limit (equals 2-D slice rank)	$\sum_{\mu \in S}(\deg P_\mu + 1) = C_2(\Gamma)$	4
Real border $\underline{\mathrm{srank}}^{\mathrm{co}}_{\mathbb{R}}$	Least *real* product terms in a compact-open limit	$\in [C_2(\Gamma),\ B_d(\Gamma)]$ ($= B_d$ if $\mathrm{Im}\,\lambda_u = 0$)	5
Multi-Gaussian border	Least real terms in a compact-open limit of radial Gaussians	$\le B_d(\Gamma) = C_d(\Gamma) - (d-2)\sum_u k_u$	6
Exact complex $\mathrm{srank}_{\mathbb{C}}$	Least exact product terms with complex factors $f_{n,j} : \mathbb{R} \to \mathbb{C}$	$\sum_{\mu \in S}((d-1)\deg P_\mu + 1)$	6
Exact real $\mathrm{srank}_{\mathbb{R}}$	Least exact product terms with real factors $f_{n,j} : \mathbb{R} \to \mathbb{R}$	$C_d(\Gamma)$ ($+(d-2)$ per nonreal pair)	7

walls, periodic `cuFFT` matches `Rank-7 fp32` on homogeneous boxes, and exact GPU crossovers depend on the RTX 4070 Ti Super's 48 MB L2 cache and register file. (iv) *Verification boundary:* Lean 4 verifies the 32 mathematical theorems in `KernelRankLean` relative to their definitions in `Defs.lean`; human review of those definitions remains essential, and Sections 6–7 are not formally verified in Lean.

8 Conclusions and Open Problems

- **Theory and formalization:** Continuous radial kernels on $\mathbb{R}^d$ ($d \ge 2$) have finite separable rank iff their profile on $[0, \infty)$ is a finite sum of exponential polynomials (while shared-width harmonic complex Gaussians uniformly approximate all $C_0([0, \infty))$ radial profiles), with exact real rank $C_d(\Gamma)$ achieved by alternating angle sampling (d real products per complex Gaussian vs. $N_{\text{naive}}(d) = 2^{d-1}$) and a four-level 3-D hierarchy on $\|x\|^2 \cos(\|x\|^2)$ (over $\mathbb{C}$: 4; over $\mathbb{R}$: $5 < 6 < 7$; Table 10)—all verified across 32 Lean 4 theorems.
- **Numerics and stability:** In 2-D, $\mathrm{GL}_R(\mathbb{R})$ orthogonalization eliminates 600 × –1,100×

inter-component `float16` rounding ($7.90 \pm 0.92 \to 0.31 \pm 0.13/255$ on $R = 12$, where magnitude reordering alone gives $8.18 \pm 0.92/255$), while in $d \geq 3$ finite-h border limits hit an $\Theta(\varepsilon_{\text{mach}}^{1/(k+1)})$ precision floor that favors well-separated $k = 0$ Gaussians.

- **Illustrative GPU regimes:** L2-cached separable cascades beat `cuFFT` on small/medium 2-D blurs at $R = 4$ (0.33–0.35 ms vs. 0.45–0.46 ms at 1080p) and in 3-D acoustics cut step time by 24% via 3-vs.-4 angle sampling (1.51 ms vs. 1.98 ms on 256^3), achieve a measured 30.9× speedup over 3×-refined 25-point FDTD at matched 1.6% broadband pulse error ($57\times$ –$70\times$ at 3.5× refinement), and handle $> 2:1$ spatially varying $c(\mathbf{x})$ contrasts where 1-step `cuFFT` is inapplicable.

Open problems.

1. Determine a closed-form formula for the exact real compact-open border rank $\underline{\text{srank}}_{\mathbb{R}}^{\text{co}}$ of general nonreal representations Γ with degree $k_u \geq 2$ or dimension $d \geq 4$ (between $C_2(\Gamma)$ and $B_d(\Gamma)$; cf. [79, 16]).
2. Derive commutator/gradient corrections across steep $\nabla(c^2)$ caustics and provably passive one-sided boundary closures for wide signed multi-Gaussian wave stencils ($L \geq 4$).
3. Certify discrete 3-D Brillouin-zone von Neumann stability $|\widehat{M}(\mathbf{k})| < 2$ formally in Lean 4 via interval arithmetic or multivariate sum-of-squares certificates.

A Formalization Breakdown, Theorem Map, and Semigroup Comparison

Tables 11 and 12 list the module line counts and exported theorem identifiers of `KernelRankLean`. Table 13 compares the 96-line Banach-algebra Bochner-integral proof of $M(t) = \exp(tA)$ (`Semigroup.lean`) against four alternative finite-dimensional formalizations in `Experiments/` and `Isabelle/Semigroup_NSA.thy` [24, 33] (using homogeneous infinitesimal scale $\eta = \varepsilon + \|M(\varepsilon) - 1\|$ to avoid difference-quotient bounds in the three nonstandard proofs; line counts reflect these specific formalizations and library APIs rather than intrinsic mathematical difficulty).

Table 11: Module breakdown and source line counts of `KernelRankLean` (earlier milestones: 5,750 loc at `94a5b4a`, 5,324 loc at `9722ff5`, 1,587 loc at `cae13d6`, 1,340 loc at `240dea0`).

Module(s)	Mathematical content	Phys.	Nonbl.
`Defs`, `KernelRankLean`	Multi-way rank, surjections, slices, exact & border costs	254	216
`RowSpan`, `Complexification`	2-way row-span rank; $\mathbb{R} \to \mathbb{C}$ invariance	221	189
`Translations`	d-D radial-to-Hankel reduction; `translateSpan`	92	80
`ExpPoly`, `RealExpPoly`	Exponential-polynomial uniqueness & conjugates	198	170
`Semigroup`, `TranslationSemigroup`	Banach semigroups; half-line shift semigroup	262	219
`Necessity`, `Representation`	Generalized eigenvectors; canonical $\operatorname{Im}\lambda \geq 0$ pairing	222	198
`Differential`, `ComplexRank`, `Annihilator`	Half-line $q(D)$, 2-D complex rank, minimal $p_{\min}$	387	346
`MultiComplexRank`	Exact d-way complex rank $\sum_\mu((d-1)\deg P_\mu + 1)$	1,015	978
`MultiRealRank`	Angle-sampling syntheses & real Alder–Strassen peeling	2,562	2,521
`Main`	Unified d-D/2-D theorems & multi-Gaussian border bound	538	516
`Approximation`	Exact complex/real border ranks & harmonic approx.	1,184	1,153
Total (`KernelRankLean`)		**6,935**	**6,586**

B Detailed Numerical Tables and Explicit Coefficients

Tables 14–20 present the complete numerical and GPU benchmark results referenced in Sections 6.2–7. To allow standalone reproduction of the main filter and stencil designs from this

Table 12: Map from paper results to the 27 `KernelRank.*` theorems audited by `scripts/check` (plus 6 foundational lemmas; the remaining 5 `Experiments.*` audited theorems appear in Table 13, totaling 32 audited Lean 4 theorems).

Paper result	Exported Lean 4 theorem identifier(s) (`KernelRank.*`)	File(s)	Loc
Eq. (2)–(4)	`separableRank_comp_surjective`, `separableRank_padTwo_le`	`Defs.lean`	239
Sec. 4.1 (2-D core)	`separableRank_eq_finrank_rowSpan`, `separableRank_ofReal`	`RowSpan`, `Complexification`	221
Sec. 4.1 (semigroup)	`exists_eq_exp_of_semigroup`, `exists_translation_generator`	`Semigroup`, `TranslationSemigroup`	262
Thm. 3.1	`separableRank_radial_lt_top_iff_exists`	`Main.lean`	538
Thm. 3.2	`exists_harmonic_gaussian_uniform_approx`	`Approximation.lean`	1,184
Thm. 3.4	`separableRank_radial_complex_eq_iff`, `separableRank_sum_expPoly`, `separableRank_radial_sum_expPoly`	`Main`, `MultiComplexRank`	1,015
Thm. 3.5	`separableRank_radial_eq_iff`, `separableRank_radial_two_eq_iff`, `separableRank_radial_profile`, `separableRank_radial_gaussian_profile`, `separableRank_radial_gaussian_eq_iff`	`Main`, `MultiRealRank`	2,562
Props. 3.6–3.7	`re_prod_eq_sum_re`, `re_coeff_prod_eq_sum_re`	`MultiRealRank.lean`	2,562
Thm. 3.11	`exists_minimal_differential_annihilator`	`Annihilator.lean`	69
Thm. 3.13	`multiTotalCost_eq_borderTotalCost_add`, `exists_rank_le_borderTotalCost_tendstoUniformlyOn_radial_profile`, `separableRank_radial_sq_mul_cos`, `separableRank_radial_sq_mul_cos_complex`, `exists_rank_le_two_mul_tendstoUniformlyOn_radial_sq_mul_cos`	`Defs`, `Main`	538
Thm. 3.14(1)	`separableRank_fin2_le_of_tendsto`, `totalCost_le_of_tendsto_radial_profile`, `borderTotalCost_eq_totalCost_of_forall_im_eq_zero`, `borderTotalCost_le_of_tendsto_radial_real_profile`, `sum_natDegree_add_one_le_of_tendsto_radial_complex`, `exists_rank_le_sum_natDegree_add_one_tendstoUniformlyOn_radial_complex`	`Approximation.lean`	1,184
Thm. 3.14(2)	`not_exists_rank_le_two_tendsto_radial_cos`, `not_exists_rank_le_four_tendsto_radial_sq_mul_cos`, `exists_rank_le_five_tendstoUniformlyOn_radial_sq_mul_cos`	`Approximation.lean`	1,184

Table 13: Comparison of the five machine-checked proofs of $M(t) = \exp(tA)$ in the repository (line counts depend on proof style and existing library support in Mathlib vs. Isabelle `HOL-NSA`; audited `Experiments.*` theorems: `exists_eq_exp_of_semigroup_finiteDifference`, `exists_translation_generator_finiteDifference`, `exists_eq_exp_of_semigroup_nonstandard`, `exists_translation_generator_nonstandard`, `NSA.exists_eq_exp_of_semigroup`).

Proof method	System / files	Algebra scope	Phys.	Nonbl.
Bochner integral smoothing (main)	Lean 4: `Semigroup.lean`	Any Banach algebra	96	87
Classical finite differences	Lean 4: `GeneratorBound` + `FiniteDifferenceSemigroup`	Finite-dimensional	256	233
Nonstandard (ultrafilter sequences)	Lean 4: `NonstandardSemigroup.lean`	Finite-dimensional	206	190
Nonstandard (`Hyp` / transfer)	Lean 4: `Ultrapower` + `UltrapowerSemigroup` + step est.	Finite-dimensional	399	344
Nonstandard (`transfer` + power)	Isabelle2026: `Semigroup_NSA.thy` (0 oracles)	Finite-dim., $\|1\|=1$	275	253

PDF (complete JSON archives for all 18 blur families and all 6 acoustic propagators live in `Applications/`), Tables 21–22 tabulate every radial profile in the canonical convention of Def-

inition 2.2:

$$\begin{aligned}\varphi(s) &= \sum_{u=1}^{m} \mathrm{Re}\left(\left(\sum_{q=0}^{k_u} (p_{u,q} + i q_{u,q}) s^q \right) e^{(-\alpha_u + i\omega_u)s} \right) \\ &= \sum_{u=1}^{m} \sum_{q=0}^{k_u} s^q (p_{u,q}\cos(\omega_u s) - q_{u,q}\sin(\omega_u s)) e^{-\alpha_u s},\end{aligned} \tag{17}$$

with $\omega_u = q_{u,q} = 0$ on real rates. In 2-D blur (Table 21), $s = (x^2+y^2)/R_{\mathrm{px}}^2$ on $|x|, |y| \le 1.2\,R_{\mathrm{px}}$ normalized so $\sum_{\mathbf{m}} K[\mathbf{m}] = 1$; in 3-D acoustics (Table 22), $M[\mathbf{m}] = \alpha_0 \delta_{\mathbf{0}}[\mathbf{m}] + \varphi(\|\mathbf{m}\|^2)$ on $[-L, L]^3$ with $\alpha_0 := 2 - \sum_{\mathbf{m}\in[-L,L]^3} \varphi(\|\mathbf{m}\|^2)$ so $\widehat{M}(\mathbf{0}) = 2$ to machine precision.

Table 14: Representative 2-D circular disc-blur designs ($\Delta r = 0.2$, $R_{\mathrm{px}} = 25$, $W = 61$; explicit parameters in Table 21). $\kappa_\infty = (\sum_n \|h_n\|_\infty \|v_n\|_\infty)/\|K\|_\infty$ and max `float16` sRGB error (/255, single scene `seed=7`) before (raw) and after (orth) $\mathrm{GL}_R(\mathbb{R})$ rebalancing.

Structure	Degrees	R (d=2)	SVD	C_3 (d=3)	Ripple $\pm\delta$	$\kappa_\infty^{\mathrm{raw}}$	$\kappa_\infty^{\mathrm{orth}}$	sRGB raw	sRGB orth
$1c$ (Gaussian)	0	2	2	3	0.2324	1.65	1.30	0.13	0.13
$1c + 1r$ (mixed)	0, 0	3	3	4	0.1315	3.45	1.65	0.16	0.20
$2c$ (Gaussians)	0, 0	4	4	6	0.0759	5.28	1.95	0.20	0.18
$1c + 1r_2$ (mixed poly)	0, 1	4	4	6	0.0905	4.50	1.92	0.12	0.23
$1c_2$ (complex poly)	1	4	4	7	0.1165	5.68	1.88	0.17	0.19
$1r_4$ (real Laguerre)	3	4	4	7	0.1982	4.74	1.56	0.21	0.23
$2c + 1r$ (mixed)	0, 0, 0	5	5	7	0.0446	8.73	2.08	0.26	0.20
$4c$ (Gaussians)	0^4	8	8	12	0.0109	70.06	2.31	1.22	0.23
$6c$ (Gaussians)	0^6	12	12	18	0.0019	595.01	2.49	9.35	0.26
$1c_6$ (complex poly)	5	12	12	23	0.0138	1,109.75	2.32	14.81	0.25

Table 15: Sampled continuous complex Gaussians vs. direct discrete rank-R optimization across $R_{\mathrm{px}} \in \{3, 6, 12, 25, 64\}$ ($L + 1 = \lceil 1.2\,R_{\mathrm{px}} \rceil + 1$): pass/stopband Chebyshev ripple $\pm\delta$, with shell anisotropy Δ_{sh} in parentheses.

R_{px}	W	L+1	R	Cont. Gaussian	Direct ℓ_2 SVD (Δ_{sh})	Disc. unconstr. (Δ_{sh})	Disc. shell-iso. (Δ_{sh})
3	9	5	4	0.07593 (0)	0.00000 (0)	0.00000 (0)	0.00000 (0)
6	17	9	4	0.07593 (0)	0.02182 (0.0111)	0.00479 (0.1285)	0.00476 (1×10^{-6})
12	31	16	4	0.07593 (0)	0.09496 (0.1066)	0.01899 (0.1577)	0.04273 (4×10^{-7})
25	61	31	4	0.07593 (0)	0.09556 (0.1770)	0.05226 (0.0905)	0.07367 (1.8×10^{-4})
64	155	78	4	0.07593 (0)	0.10295 (0.1923)	0.06647 (0.0323)	0.07582 (2.0×10^{-5})
12	31	16	8	0.01086 (0)	0.00978 (0.0140)	0.00137 (0.0240)	0.00079 (1×10^{-6})
25	61	31	8	0.01086 (0)	0.01295 (0.0151)	0.00486 (0.0228)	0.00552 (4.3×10^{-5})
64	155	78	8	0.01086 (0)	0.01625 (0.0189)	0.00755 (0.0113)	0.01065 (2.4×10^{-5})
64	155	78	12	0.00193 (0)	0.00480 (0.0064)	0.00100 (0.0066)	0.00171 (1.3×10^{-4})

Code and Data Availability

All source code, machine-checked proofs, numerical optimization scripts, GPU benchmarks, JSON parameter/result archives, and figure-generation scripts for this work are openly available in the public domain (CC0 1.0 Universal) at https://github.com/OlliNiemitalo/kernel-rank-lean:

- `KernelRankLean/` and `Experiments/`: Lean 4 (`v4.36.0-rc2` with Mathlib `v4.36.0-rc2`) formalizations of all d-D and 2-D separable-rank, root-of-unity, differential-annihilator, border-rank, and alternative semigroup theorems;
- `Isabelle/`: supplementary Isabelle2026 (`HOL-Nonstandard_Analysis`) semigroup formalization `Semigroup_NSA.thy`;

Table 16: Six-stage IEEE `float16` conditioning ablation ($R_{\text{px}} = 25, W = 61$, $N = 20$ random 512×512 HDR scenes, peak 48×): Stage 1 unbalanced; Stage 2 pair-balanced $\|h_n\|_2 = \|v_n\|_2$; Stage 2b reordered smallest-$\|h_n\|_2\|v_n\|_2$-first *without* orthogonalization; Stage 2c `float16` 1-D passes summed in `float32` *without* orthogonalization; Stages 3–4 $\mathrm{GL}_R(\mathbb{R})$-orthogonalized (descending vs. ascending $|e_n|$). Rightmost column: Stage 3 → Stage 4 linear-light max error (`seed=7`).

	20-scene mean ± std max 8-bit sRGB error (/255, peak 48×)						Linear (/255)
Design (R)	1. Unbal.	2. Pair-bal.	2b. Reord.	2c. `fp32`-acc.	3. Orth. (desc)	4. Orth. (asc)	48×; 256×
$2c$ (R=4)	0.12 ± 0.02	0.11 ± 0.01	0.11 ± 0.02	0.09 ± 0.02	0.20 ± 0.04	0.24 ± 0.08	0.24→**0.16**; 1.00→**0.60**
$4c$ (R=8)	0.58 ± 0.07	0.54 ± 0.06	0.64 ± 0.10	0.46 ± 0.05	0.25 ± 0.07	0.26 ± 0.10	0.36→**0.16**; 1.23→**0.69**
$6c$ (R=12)	7.80 ± 1.16	7.90 ± 0.92	8.18 ± 0.92	6.76 ± 1.10	$\mathbf{0.30 \pm 0.06}$	$\mathbf{0.31 \pm 0.13}$	0.40→**0.15**; 1.41→**0.60**
$1c_6$ (R=12)	`ovf` (20/20)	15.17 ± 1.39	15.59 ± 1.84	13.33 ± 1.30	$\mathbf{0.28 \pm 0.05}$	$\mathbf{0.27 \pm 0.10}$	0.40→**0.16**; 1.45→**0.62**

Table 17: GPU execution times in ms (mean ± std across $N = 10$ CUDA-Event batches of 15 calls = 150 launches; 20–50 for `DirectBlur2D`; inter-run shift ≤ 6%) for 1920×1080 RGBA 2-D blur on the RTX 4070 Ti Super eGPU.

R_{px}	W	Direct W^2	`cuFFT` (zero)	`cuFFT` (clamp)	Sep. $R = 4$		Sep. $R = 8$		Sep. $R = 12$	
		`fp32`	`fp32`	`fp32`	`fp32`	`fp16`	`fp32`	`fp16`	`fp32`	`fp16`
8	21	$0.516_{\pm.001}$	$0.465_{\pm.0002}$	$0.457_{\pm.001}$	$\mathbf{0.331}_{\pm.0003}$	$0.417_{\pm.0004}$	$0.665_{\pm.001}$	$0.670_{\pm.001}$	$0.860_{\pm.001}$	$1.033_{\pm.032}$
16	41	$1.750_{\pm.033}$	$0.451_{\pm.0005}$	$0.454_{\pm.001}$	$\mathbf{0.354}_{\pm.005}$	$0.396_{\pm.001}$	$0.719_{\pm.001}$	$0.682_{\pm.0003}$	$1.076_{\pm.005}$	$1.143_{\pm.028}$
32	79	$6.449_{\pm.132}$	$\mathbf{0.452}_{\pm.001}$	$0.481_{\pm.001}$	$0.528_{\pm.004}$	$0.496_{\pm.001}$	$1.121_{\pm.001}$	$0.937_{\pm.005}$	$1.661_{\pm.013}$	$1.522_{\pm.012}$
64	155	$24.38_{\pm.110}$	$\mathbf{0.473}_{\pm.002}$	$0.555_{\pm.001}$	$0.871_{\pm.003}$	$0.825_{\pm.003}$	$1.755_{\pm.007}$	$1.672_{\pm.001}$	$2.641_{\pm.007}$	$2.594_{\pm.005}$

Table 18: Fixed-grid (256^3) 3-D acoustic propagators vs. four compact FDTD baselines and periodic k-space `cuFFT` (mean ± std in ms over $N = 10$ batches of 15 steps = 150 steps; 6×2 for `DirectStep3D`; inter-run shift ≤ 3%; `+MiB` is extra scratch VRAM atop 128 MiB in-place / 192 MiB out-of-place fields).

Scheme	ν	W	R (nv)	$\lvert\delta c/c_0\rvert$	$k_{0.1\%}$	$k_{0.5\%}$	Wave L^2	Direct	Naive	`fp32`	`fp16`	+MiB
7-pt FDTD	0.58	3	—	19.88%	$.060\pi$	$.135\pi$	109.3%	—	—	$0.324_{\pm.0003}$	—	0
13-pt $O(\Delta x^4, \Delta t^2)$	0.50	5	—	8.74%	$.099\pi$	$.222\pi$	100.8%	—	—	$0.324_{\pm.0001}$	—	0
27-pt IWB ($1r$)	1.00	3	1 (1)	9.38%	$.085\pi$	$.191\pi$	98.2%	—	—	$0.324_{\pm.0001}$	—	0
25-pt $O(\Delta x^4, \Delta t^4)$	0.58	5	—	11.00%	$.236\pi$	$.355\pi$	78.2%	—	—	$0.325_{\pm.0002}$	—	0
k-space `cuFFT`	1.00	256	—	0.000%	1.00π	1.00π	0.35%	—	—	$1.44_{\pm.05}$	—	+101
Rank-4 ($1c{+}1r, .75\pi$)	0.75	9	4 (5)	0.293%	$.14\pi$	$.76\pi$	11.6%	$2.68_{\pm.04}$	$1.02_{\pm.02}$	$0.857_{\pm.003}$	$0.627_{\pm.009}$	+16
Rank-6 ($2c, .75\pi$)	0.75	11	6 (8)	0.066%	$.76\pi$	$.80\pi$	5.2%	$4.63_{\pm.05}$	$1.66_{\pm.02}$	$1.27_{\pm.02}$	$0.920_{\pm.016}$	+24
Rank-7 (deg 1 $1c, .75\pi$)	0.75	13	7 (9)	0.138%	$.16\pi$	$.78\pi$	7.1%	$7.07_{\pm.01}$	$1.98_{\pm.02}$	$1.51_{\pm.03}$	$1.09_{\pm.02}$	+28
Rank-7 ($2c{+}1r, .80\pi$)	0.75	13	7 (9)	0.045%	$.81\pi$	$.85\pi$	2.8%	$7.07_{\pm.01}$	$1.98_{\pm.02}$	$1.51_{\pm.03}$	$1.09_{\pm.02}$	+28
Rank-7 ($2c{+}1r, .80\pi$)	1.00	13	7 (9)	0.049%	$.81\pi$	$.87\pi$	2.5%	$7.07_{\pm.01}$	$1.98_{\pm.02}$	$1.51_{\pm.03}$	$1.09_{\pm.02}$	+28
Rank-9 ($3c, .85\pi$)	0.75	17	9 (12)	0.044%	$.86\pi$	$.89\pi$	2.0%	$14.51_{\pm.03}$	$3.21_{\pm.02}$	$2.21_{\pm.02}$	$1.57_{\pm.01}$	+36/18

- `Applications/`: Python, SciPy, PyTorch, and Triton optimization, GPU-kernel, and figure scripts (`blur_2d.py`, `acoustics_3d.py`, `gpu_kernels.py`, `generate_figures.py`) plus JSON result archives (`results_blur_2d.json`, `results_acoustics_3d.json`);
- `scripts/`: hermetic setup (`scripts/setup`), formal verification and axiom/oracle audit (`scripts/check`), isolated GPU Python wrapper (`scripts/python`), and PDF build (`scripts/build-paper`).

Acknowledgments

This work builds directly on problems, mathematical formulations, and software developed by Olli Niemitalo, including `kernel-optimizer` [58] for complex-Gaussian approximations of 2-D

Table 19: Measured multi-grid GPU calibration on an identical physical 3-D wave problem ($L_{\text{phys}} = 128, \sigma_0 = 1.40, c_0 T_{\text{phys}} = 42$, exact grid receivers $(30,0,0),(22,22,0),(18,18,18)$; `acoustics_3d.py --calibration-only`): coarse 128^3 `Rank-7` ($\nu = 1.00$) vs. refined fourth-order 25-pt $O(\Delta x^4, \Delta t^4)$ FDTD ($\nu = 1/\sqrt{3}$).

Scheme & grid G^3	Ref. m	Steps	Step ms ($\pm$std)	Total sim. ms	In-place VRAM	Axial L^2	Max L^2
Rank-7 ($2c{+}1r$, `fp32`), 128^3	1.0×	42	0.188 ± 0.0022	**7.90** ms	30.0 MiB	1.33%	**1.68**%
Rank-7 ($2c{+}1r$, `fp16` slab), 128^3	1.0×	42	0.154 ± 0.0002	**6.45** ms	23.0 MiB	2.78%	2.78%
25-pt $O(\Delta x^4, \Delta t^4)$ FDTD, 128^3	1.0×	73	0.040 ± 0.0005	2.92 ms	16.0 MiB	65.09%	65.09%
25-pt $O(\Delta x^4, \Delta t^4)$ FDTD, 256^3	2.0×	145	0.324 ± 0.0003	47.01 ms	128.0 MiB	7.83%	7.83%
25-pt $O(\Delta x^4, \Delta t^4)$ FDTD, 320^3	2.5×	182	0.751 ± 0.0025	136.64 ms	250.0 MiB	3.28%	3.28%
25-pt $O(\Delta x^4, \Delta t^4)$ FDTD, 384^3	3.0×	218	1.122 ± 0.0005	244.49 ms	432.0 MiB	1.60%	**1.60**%
25-pt $O(\Delta x^4, \Delta t^4)$ FDTD, 448^3	3.5×	255	1.777 ± 0.0027	453.17 ms	686.0 MiB	0.87%	0.87%
25-pt $O(\Delta x^4, \Delta t^4)$ FDTD, 512^3	4.0×	291	2.620 ± 0.0008	762.37 ms	1024.0 MiB	0.51%	0.51%

Table 20: Empirical floating-point accuracy on $K_3(x) = \|x\|^2 \cos(\|x\|^2)$ over $[-1.2, 1.2]^3$ (31^3 grid; `acoustics_3d.py --calibration-only`), comparing exact 7-term angle sampling (Example 3.9), the 6-term radial Gaussian border limit (Theorem 3.13, minimized over $h \in [10^{-9}, 10^{-0.2}]$), and the 5-term non-Gaussian real border limit (11).

Decomposition of $K_3(x)$	Weights / floor	`float16` ($h_\star$)	`float32` ($h_\star$)	`float64` ($h_\star$)
Exact 7-term (Prop. 3.7)	$O(1)$; $\Theta(\varepsilon_{\text{mach}})$	$\mathbf{2.76{\times}10^{-3}}$	$\mathbf{4.27{\times}10^{-7}}$	$\mathbf{1.76{\times}10^{-15}}$
Gaussian border 6-term ($B_3{=}6$)	$O(h^{-1})$; $\Theta(\varepsilon_{\text{mach}}^{1/2})$	$4.83{\times}10^{-2}$ (0.019)	$7.03{\times}10^{-4}$ ($2{\times}10^{-4}$)	$2.93{\times}10^{-8}$ ($8{\times}10^{-9}$)
Non-Gaussian border 5-term (11)	$O(h^{-3})$; $\Theta(\varepsilon_{\text{mach}}^{1/4})$	$2.04{\times}10^{-1}$ (0.15)	$1.62{\times}10^{-2}$ (0.014)	$1.42{\times}10^{-4}$ ($1{\times}10^{-4}$)

circularly symmetric blur kernels and his investigations of isotropic low-pass filters, steerable gradient orientation estimation, and planar harmonic differential operators [54, 55, 56, 57]. Olli Niemitalo initiated and guided the `kernel-rank-lean` repository, proposed the equal-width harmonic Gaussian universal approximation construction, suggested extending the formalization to d dimensions alongside the visual pipeline exposition, proof dependency graph, discrete-vs.-continuous warm-start comparisons, and 2-D/3-D GPU and heterogeneous-medium applications, and shared independent related-work and structural reviews that sharpened both the mathematical positioning and manuscript organization. We also thank the aiXiv Official Agent reviewers on versions 1.0–1.6, whose critiques motivated Theorems 3.5, 3.13, and 3.14, Propositions 3.6–3.7, Definition 2.4, Section 7, the six-stage `separableChannels` ablation, the multi-grid FDTD calibration, and Appendices A–B. Both the repository and this manuscript are released under CC0 1.0 Universal.

References

[1] A. Alder and V. Strassen, "On the algorithmic complexity of associative algebras," *Theoretical Computer Science* **15**(2), 201–211 (1981).

[2] F. Andersson and M. Carlsson, "On general domain truncated correlation and convolution operators with finite rank," *Integral Equations and Operator Theory* **82**(3), 339–370 (2015).

[3] P. M. Anselone and J. Korevaar, "Translation invariant subspaces of finite dimension," *Proceedings of the American Mathematical Society* **15**(5), 747–752 (1964).

[4] A. Bentkamp, J. C. Blanchette, and D. Klakow, "A formal proof of the expressiveness of deep learning," *Journal of Automated Reasoning* **63**(2), 347–368 (2019).

Table 21: Explicit canonical coefficients (17) for representative 2-D circular disc-blur designs ($\Delta r = 0.2$, Table 14).

Design ($R = C_2$, C_3, $\pm\delta$)	u (q)	α_u	ω_u	$p_{u,q}$	$q_{u,q}$
$1c$ (R=2, C_3=3, $\pm$0.2324)	1 (0)	0.862715	1.624792	+0.767553	−1.862917
$2c$ (R=4, C_3=6, $\pm$0.0759)	1 (0)	0.894640	5.265923	+0.413699	+0.555222
	2 (0)	1.973219	1.554245	+0.510373	−4.611414
$1c_2$ (R=4, C_3=7, $\pm$0.1165)	1 (0)	2.878934	3.152485	+0.883521	−7.976371
	1 (1)	2.878934	3.152485	−16.607444	+7.931853
$2c$+$1r$ (R=5, C_3=7, $\pm$0.0446)	1 (0)	0.962816	7.154625	+0.091824	−0.456790
	2 (0)	2.075279	3.285822	−3.434558	+0.588375
	3 (0)	2.494577	0	+4.387353	0
$4c$ (R=8, C_3=12, $\pm$0.0109)	1 (0)	4.338835	1.553577	−5.754013	−46.161778
	2 (0)	3.838091	4.692138	+9.776587	+15.218676
	3 (0)	2.790788	8.181910	−3.042741	−0.293993
	4 (0)	1.345593	12.330086	+0.009312	−0.245279
$6c$ (R=12, C_3=18, $\pm$0.0019)	1 (0)	5.143786	2.080541	−82.448179	−111.252655
	2 (0)	5.613157	6.153062	+113.961587	−58.149957
	3 (0)	5.982623	9.803250	+39.766911	+161.948149
	4 (0)	6.506533	11.060675	−71.545624	−94.857314
	5 (0)	3.871248	14.810010	+1.416029	+3.711177
	6 (0)	2.202364	19.033594	−0.152648	+0.108514

Table 22: Explicit 9-decimal canonical coefficients (17) for all six 3-D acoustic leapfrog propagators in Table 18 (set $\alpha_0 := 2 - \sum_{\mathbf{m}\in[-L,L]^3} \varphi(\|\mathbf{m}\|^2)$; reference α_0 shown for verification).

Scheme (ν, W, α_0)	u (q)	α_u	ω_u	$p_{u,q}$	$q_{u,q}$
Rank-4 ($1c$+$1r$, ν=.75, W=9, α_0=−1.440427676)	1 (0)	0.439979390	0.502750129	+0.137876459	+0.024929224
	2 (0)	0.888101329	0	+0.492695925	0
Rank-6 ($2c$, ν=.75, W=11, α_0=−1.419560209)	1 (0)	0.606539114	0.240982835	+0.597581203	+0.180225818
	2 (0)	0.267859019	0.385016560	−0.007190817	+0.050272529
Rank-7 (deg 1 $1c$, ν=.75, W=13, α_0=−1.434649594)	1 (0)	0.429753880	0.229296784	+0.605530664	+0.525638174
	1 (1)	0.429753880	0.229296784	−0.061586407	−0.061209821
Rank-7 ($2c$+$1r$, ν=.75, W=13, α_0=−1.439243312)	1 (0)	0.381717837	0.311635272	+0.110431944	+0.263079509
	2 (0)	0.163094730	0.300649461	−0.015761751	−0.008519990
	3 (0)	0.599866965	0	+0.491386225	0
Rank-7 ($2c$+$1r$, ν=1.0, W=13, α_0=−1.959018821)	1 (0)	0.321809087	0.262377242	+0.159170237	+0.129231489
	2 (0)	0.126172334	0.278598443	−0.010612383	+0.004174471
	3 (0)	0.438442015	0	+0.241253606	0
Rank-9 ($3c$, ν=.75, W=17, α_0=−1.447076570)	1 (0)	0.301382846	0.246772855	−0.286641369	+0.234406754
	2 (0)	0.398975160	0.094363260	+0.866865360	+0.865008109
	3 (0)	0.111967878	0.275076161	−0.002296910	−0.011770692

[5] G. Beylkin and M. J. Mohlenkamp, "Algorithms for numerical analysis in high dimensions," *SIAM Journal on Scientific Computing* **26**(6), 2133–2159 (2005).

[6] G. Beylkin and L. Monzón, "On approximation of functions by exponential sums," *Applied and Computational Harmonic Analysis* **19**(1), 17–48 (2005).

[7] J. Bigun, T. Bigun, and K. Nilsson, "Recognition by symmetry derivatives and the generalized structure tensor," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(12), 1590–1605 (2004).

[8] D. Bini, M. Capovani, F. Romani, and G. Lotti, "$O(n^{2.7799})$ complexity for $n \times n$ approximate matrix multiplication," *Information Processing Letters* **8**(5), 234–235 (1979).

[9] M. Bläser, "Lower bounds for the bilinear complexity of associative algebras," *Computational Complexity* **9**(2), 73–112 (2000).

[10] M. Bläser, H. Mayer, and D. Shringi, "On the multilinear complexity of associative algebras," in *40th International Symposium on Theoretical Aspects of Computer Science (STACS 2023)*, LIPIcs **254**, pp. 12:1–12:18 (2023).

[11] G. Blekherman, "Typical real ranks of binary forms," *Foundations of Computational Mathematics* **15**(3), 793–798 (2015).

[12] D. Braess, *Nonlinear Approximation Theory*, Springer Series in Computational Mathematics **7**, Springer, Berlin (1986).

[13] D. Braess and W. Hackbusch, "Approximation of $1/x$ by exponential sums in $[1, \infty)$," *IMA Journal of Numerical Analysis* **25**(4), 685–697 (2005).

[14] J. Buczyński and J. M. Landsberg, "Ranks of tensors and a generalization of secant varieties," *Linear Algebra and its Applications* **438**(2), 668–689 (2013).

[15] P. Bürgisser, M. Clausen, and M. A. Shokrollahi, *Algebraic Complexity Theory*, Grundlehren der mathematischen Wissenschaften **315**, Springer, Berlin (1997).

[16] C.-Y. Chang, F. Gesmundo, and J. Zuiddam, "Border subrank of higher order tensors and algebras," arXiv:2604.19872 (2026).

[17] P. Comon and B. Mourrain, "Decomposition of quantics in sums of powers of linear forms," *Signal Processing* **53**(2–3), 93–107 (1996).

[18] P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, "Symmetric tensors and symmetric tensor rank," *SIAM Journal on Matrix Analysis and Applications* **30**(3), 1254–1279 (2008).

[19] R. Deriche, "Using Canny's criteria to derive a recursively implemented optimal edge detector," *International Journal of Computer Vision* **1**(2), 167–187 (1987).

[20] V. de Silva and L.-H. Lim, "Tensor rank and the ill-posedness of the best low-rank approximation problem," *SIAM Journal on Matrix Analysis and Applications* **30**(3), 1084–1127 (2008).

[21] F. Dupuis, R. Y. Lewis, and H. Macbeth, "Formalized functional analysis with semilinear maps," in *13th International Conference on Interactive Theorem Proving (ITP 2022)*, LIPIcs **237**, pp. 10:1–10:19 (2022).

[22] K.-J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Graduate Texts in Mathematics **194**, Springer (2000).

[23] C. M. Fiduccia and Y. Zalcstein, "Algebras having linear multiplicative complexities," *Journal of the ACM* **24**(2), 311–331 (1977).

[24] J. D. Fleuriot, "On the mechanization of real analysis in Isabelle/HOL," in *Theorem Proving in Higher Order Logics (TPHOLs 2000)*, LNCS **1869**, pp. 145–161 (2000).

[25] S. Fomel, L. Ying, and X. Song, "Seismic wave extrapolation using lowrank symbol approximation," *Geophysical Prospecting* **61**(3), 526–536 (2013).

[26] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(9), 891–906 (1991).

[27] K. Garcia, "Circular separable convolution depth of field," in *ACM SIGGRAPH 2017 Talks*, Article 29, pp. 1–2 (2017).

[28] L. Greengard and J. Strain, "The fast Gauss transform," *SIAM Journal on Scientific and Statistical Computing* **12**(1), 79–94 (1991).

[29] W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics **42**, Springer, Heidelberg (2012).

[30] B. Hamilton and S. Bilbao, "Optimised 25-point finite difference schemes for the three-dimensional wave equation," *Proceedings of Meetings on Acoustics* **28**, 015022 (2016).

[31] B. Hamilton and S. Bilbao, "FDTD methods for 3-D room acoustics simulation with high-order accuracy in space and time," *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **25**(11), 2112–2124 (2017).

[32] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *Journal of Mathematics and Physics* **6**(1–4), 164–189 (1927).

[33] J. Hölzl, F. Immler, and B. Huffman, "Type classes and filters for mathematical analysis in Isabelle/HOL," in *Interactive Theorem Proving (ITP 2013)*, LNCS **7998**, pp. 279–294 (2013).

[34] A. Iarrobino and V. Kanev, *Power Sums, Gorenstein Algebras, and Determinantal Loci*, Lecture Notes in Mathematics **1721**, Springer, Berlin (1999).

[35] H. Jain, "Bounds on the real tensor rank of octonion multiplication," arXiv:2608.16649 (2026).

[36] Pl. Kannappan and P. K. Sahoo, "Rotation invariant separable functions are Gaussian," *SIAM Journal on Mathematical Analysis* **23**(5), 1342–1351 (1992).

[37] B. N. Khoromskij, *Tensor Numerical Methods in Scientific Computing*, De Gruyter, Berlin (2018).

[38] J. J. Koenderink and A. J. van Doorn, "Generic neighborhood operators," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**(6), 597–605 (1992).

[39] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review* **51**(3), 455–500 (2009).

[40] K. Kowalczyk and M. van Walstijn, "Room acoustics simulation using 3-D compact explicit FDTD schemes," *IEEE Transactions on Audio, Speech, and Language Processing* **19**(1), 34–46 (2011).

[41] J. B. Kruskal, "Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra and its Applications* **18**(2), 95–138 (1977).

[42] M. Laczkovich, "The Levi-Civita equation in function classes," *Aequationes Mathematicae* **94**(4), 689–701 (2020).

[43] P. G. Laird, "On characterizations of exponential polynomials," *Pacific Journal of Mathematics* **80**(2), 503–507 (1979).

[44] C. H. Lampert and O. Wirjadi, "An optimal nonorthogonal separation of the anisotropic Gaussian convolution filter," *IEEE Transactions on Image Processing* **15**(11), 3501–3513 (2006).

[45] J. M. Landsberg, *Tensors: Geometry and Applications*, Graduate Studies in Mathematics **128**, American Mathematical Society (2012).

[46] T. Levi-Civita, "Sulle funzioni che ammettono una formula d'addizione del tipo $f(x+y) = \sum_{i=1}^{n} X_i(x)Y_i(y)$," *Rendiconti della Reale Accademia dei Lincei* (5) **22** (2º Sem., Fasc. 5), 181–183 (1913).

[47] J. Liang, Z. Gao, and Z. Xu, "A kernel-independent sum-of-Gaussians method by de la Vallée-Poussin sums," *Advances in Applied Mathematics and Mechanics* **13**(5), 1126–1141 (2021).

[48] The mathlib Community, "The Lean mathematical library," in *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020)*, pp. 367–381 (2020).

[49] J. C. Maxwell, "Illustrations of the dynamical theory of gases.—Part I. On the motions and collisions of perfectly elastic spheres," *Philosophical Magazine* (4) **19**(124), 19–32 (1860).

[50] L. de Moura and S. Ullrich, "The Lean 4 theorem prover and programming language," in *Automated Deduction—CADE 28*, LNCS **12699**, pp. 625–635 (2021).

[51] B. Mourrain, "Polynomial–exponential decomposition from moments," *Foundations of Computational Mathematics* **18**(6), 1435–1492 (2018).

[52] M. Nagumo, "Einige analytische Untersuchungen in linearen, metrischen Ringen," *Japanese Journal of Mathematics* **13**, 61–80 (1936).

[53] J. Nie and K. Ye, "Hankel tensor decompositions and ranks," *SIAM Journal on Matrix Analysis and Applications* **40**(2), 486–516 (2019).

[54] O. Niemitalo, "Circularly symmetric convolution and lens blur," technical blog post (2010), https://yehar.com/blog/?p=1495.

[55] O. Niemitalo, "Auto detection of rotation angle on arbitrary image with orthogonal features" (answer), *Signal Processing Stack Exchange* (2019), https://dsp.stackexchange.com/a/58287/15347.

[56] O. Niemitalo, "Beach-ball like differential operators of a two-dimensional function" (question; with answer by K. B. Dave), *Mathematics Stack Exchange* (2019), https://math.stackexchange.com/q/3232122.

[57] O. Niemitalo, "Efficient implementation of 2-D circularly symmetric low-pass filter," *Signal Processing Stack Exchange* (2019), https://dsp.stackexchange.com/q/58449.

[58] O. Niemitalo, *kernel-optimizer: Circularly symmetric 2-D blur kernel approximation by complex Gaussians*, GitHub repository and interactive WebGPU app, https://github.com/OlliNiemitalo/kernel-optimizer (https://yehar.com/kernel-optimizer/).

[59] I. V. Oseledets, "Tensor-train decomposition," *SIAM Journal on Scientific Computing* **33**(5), 2295–2317 (2011).

[60] P. Perona, "Deformable kernels for early vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**(5), 488–499 (1995).

[61] A. Rahimi and B. Recht, "Random features for large-scale kernel machines," in *Advances in Neural Information Processing Systems* **20** (NIPS 2007), pp. 1177–1184 (2008).

[62] M. Reisert and H. Burkhardt, "Complex derivative filters," *IEEE Transactions on Image Processing* **17**(12), 2265–2274 (2008).

[63] B. Reznick, "On the length of binary forms," in *Quadratic and Higher Degree Forms*, Developments in Mathematics **31**, pp. 207–232 (2013).

[64] R. Rochberg, "Toeplitz and Hankel operators on the Paley–Wiener space," *Integral Equations and Operator Theory* **10**(2), 187–235 (1987).

[65] I. J. Schoenberg, "Metric spaces and completely monotone functions," *Annals of Mathematics* **39**(4), 811–841 (1938).

[66] J. J. Stone, *Exponential Polynomials on Commutative Semigroups*, Applied Mathematics and Statistics Laboratories Technical Note 14, Stanford University (1960).

[67] M. Tabei, T. D. Mast, and R. C. Waag, "A k-space method for coupled first-order acoustic propagation equations," *Journal of the Acoustical Society of America* **111**(1), 53–63 (2002).

[68] C. K. W. Tam and J. C. Webb, "Dispersion-relation-preserving finite difference schemes for computational acoustics," *Journal of Computational Physics* **107**(2), 262–281 (1993).

[69] A. Townsend and L. N. Trefethen, "Continuous analogues of matrix factorizations," *Proceedings of the Royal Society A* **471**(2173), 20140585 (2015).

[70] S. Treitel and J. L. Shanks, "The design of multistage separable planar filters," *IEEE Transactions on Geoscience Electronics* **9**(1), 10–27 (1971).

[71] E. E. Tyrtyshnikov, "Incomplete cross approximation in the mosaic-skeleton method," *Computing* **64**(4), 367–380 (2000).

[72] M. Unser and N. Chenouard, "A unifying parametric framework for 2D steerable wavelet transforms," *SIAM Journal on Imaging Sciences* **6**(1), 102–135 (2013).

[73] J. van Mourik and D. Murphy, "Explicit higher-order FDTD schemes for 3D room acoustic simulation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **22**(12), 2003–2011 (2014).

[74] A. Wathen and S. Zhu, "On spectral distribution of kernel matrices related to radial basis functions," *Numerical Algorithms* **70**(4), 709–726 (2015).

[75] S. Winograd, "Some bilinear forms whose multiplicative complexity depends on the field of constants," *Mathematical Systems Theory* **10**(1), 169–180 (1977).

[76] S. Yang, R. Zhao, J. Wu, Y. Wang, W. Li, F. Chen, T. Deng, S. Hong, Y. Li, Z. Li, J. Mi, Y. Pan, K. Zhu, J. Yang, H. Chen, and Y. Yang, "A structural proof of the lower bound 21 for 3×3 matrix multiplication over $\mathbb{F}_2$," arXiv:2609.18722 (2026).

[77] I. T. Young and L. J. van Vliet, "Recursive implementation of the Gaussian filter," *Signal Processing* **44**(2), 139–151 (1995).

[78] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Applied Optics* **46**(10), 1819–1829 (2007).

[79] J. Zuiddam, "A note on the gap between rank and border rank," *Linear Algebra and its Applications* **525**, 33–44 (2017).