Positive Definite Kernels and Similarities over the Natural Numbers

A First Axiomatic Draft

Orges Leka

March 21, 2026

Contents

1 Introduction

This text is a collection of ideas coming mainly from several of my MathOverflow questions over the last six years around positive definite kernels over the natural numbers. The starting point was a number of observations related to the abc conjecture. Since then, many related examples, constructions, and questions have accumulated in a rather scattered way.

My dream has been to collect these scattered observations into a coherent framework. With the rise of useful large language models in mathematics, this dream is beginning to become reality. The present text is therefore a mixture of my own mathematical observations, written and organized with the help of LLMs.

Any comment on this text is highly appreciated.

2 Motivation and scope

This note proposes a first axiomatic framework for *positive definite kernels and similarities over the natural numbers*. The guiding examples are gcd-based similarities such as

$$\frac{\gcd(a,b)}{\min(a,b)}, \qquad \frac{\gcd(a,b)}{\max(a,b)}, \qquad \frac{2\gcd(a,b)}{a+b}, \qquad \frac{\gcd(a,b)^2}{ab},$$

set-based realizations satisfying $X_a \cap X_b = X_{\gcd(a,b)}$, and projective kernels such as

$$k(a,b) = \frac{2\gcd(a,b)^3}{ab(a+b)}.$$

The long-term goal is not merely to collect formulas, but to isolate named properties that occur independently and then study how they interact. This should make it possible to classify broad families of kernels and to extract arithmetic invariants relevant to radical inequalities and abc-type questions.

The basic philosophy is the following.

- A kernel is first a *Hilbert-space object*: positive definiteness, normalization, Gram representations, and operator models.
- It may additionally carry *arithmetic structure*: scale invariance, rational values, denominator control, multiplicative or divisor-set models.
- Finally, it may admit *arithmetic extractions*: reciprocal integrality, radical transforms, irreducibility, abc-type inequalities.

The definitions below are arranged accordingly.

3 Similarities and positive definite kernels

Definition 3.1 (Similarity)**.** *Let X be a nonempty set. Following the terminology of the* Encyclopedia of Distances*, a function*

$$s : X \times X \to \mathbb{R}$$

is called a similarity *if for all $x, y \in X$ one has:*

(i) $s(x, y) \geq 0$,

(ii) $s(x, y) = s(y, x)$,

(iii) $s(x, y) \leq s(x, x)$,

(iv) $s(x, y) = s(x, x)$ *if and only if* $x = y$.

Definition 3.2 (Positive definite kernel)**.** *A function*

$$k : X \times X \to \mathbb{C}$$

is called positive definite *if it is Hermitian and for every finite choice $x_1, \ldots, x_r \in X$ the matrix*

$$\big(k(x_i, x_j)\big)_{1 \leq i,j \leq r}$$

is positive semidefinite.

Definition 3.3 (Positive definite similarity kernel on $\mathbb{N}$)**.** *A function*

$$k : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$$

is called a positive definite similarity kernel *if it is simultaneously a similarity on $\mathbb{N}$ and a positive definite kernel. If in addition*

$$k(a, a) = 1 \qquad (a \in \mathbb{N}),$$

we call k normalized*.*

Remark 3.4. *A normalized positive definite similarity kernel satisfies*

$$0 < k(a, b) \leq 1, \qquad k(a, a) = 1, \qquad k(a, b) = 1 \iff a = b.$$

By the Moore–Aronszajn theorem, such a kernel is exactly the Gram kernel of a family of distinct unit vectors:

$$k(a, b) = \langle \phi(a), \phi(b) \rangle.$$

Definition 3.5 (Associated Hilbert distance)**.** *Given a positive definite similarity kernel k, define*

$$d_k(a, b)^2 := 2 - 2k(a, b).$$

Then d_k is the Hilbert distance induced by any Gram realization of k.

Definition 3.6 (Associated geodesic distance)**.** *Given a positive definite similarity kernel k, define*

$$\delta_k(a, b) := \arccos(k(a, b)).$$

Since $0 < k(a, b) \leq 1$, this is well-defined and takes values in $[0, \pi/2)$. If

$$k(a, b) = \langle \phi(a), \phi(b) \rangle$$

for unit vectors $\phi(a)$ in a Hilbert space, then $\delta_k(a, b)$ is exactly the geodesic distance between $\phi(a)$ and $\phi(b)$ on the unit sphere.

Remark 3.7. *The two associated distances are related by*

$$d_k(a,b)^2 = 2 - 2\cos(\delta_k(a,b)).$$

Thus d_k is the chordal distance, whereas δ_k is the associated geodesic distance on the unit sphere.

Definition 3.8 (Almost orthogonal kernel)**.** *A normalized positive definite similarity kernel k is called* almost orthogonal *if*

$$\inf_{a \neq b} k(a,b) = 0.$$

Equivalently, the associated unit vectors admit pairs with angles arbitrarily close to $\pi/2$.

4 Measurable set models and induced similarities

Definition 4.1 (Injective measurable set model)**.** *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. An* injective measurable set model over $\mathbb{N}$ *is an injective map*

$$n \longmapsto X_n \in \mathcal{A}$$

such that

$$0 < \mu(X_n) < \infty \qquad (n \in \mathbb{N}).$$

The quantity $\mu(X_n)$ generalizes the finite-cardinality case $|X_n|$.

Remark 4.2. *If μ is counting measure on a finite or countable ground set, then*

$$\mu(X_n) = |X_n|.$$

Thus all formulas below simultaneously cover finite combinatorial models and measure-theoretic models.

Definition 4.3 (GCD-intersection model)**.** *An injective measurable set model $(X_n)_{n\in\mathbb{N}}$ is called a* gcd-intersection model *if*

$$X_a \cap X_b = X_{\gcd(a,b)} \qquad (a,b \in \mathbb{N}).$$

Proposition 4.4 (Standard similarities from measurable set models)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be an injective measurable set model. Then the functions*

$$s_{\mathrm{BB}}(a,b) := \frac{\mu(X_a \cap X_b)}{\max\{\mu(X_a), \mu(X_b)\}},$$

$$s_{\mathrm{S}}(a,b) := \frac{2\,\mu(X_a \cap X_b)}{\mu(X_a) + \mu(X_b)},$$

$$s_{\mathrm{J}}(a,b) := \frac{\mu(X_a \cap X_b)}{\mu(X_a \cup X_b)},$$

$$s_{\cos}(a,b) := \frac{\mu(X_a \cap X_b)}{\sqrt{\mu(X_a)\mu(X_b)}}$$

are similarities on $\mathbb{N}$.

Proof. Symmetry and nonnegativity are immediate. Since $X_a \cap X_b \subseteq X_a$ and $X_a \cap X_b \subseteq X_b$, each quotient is bounded above by its diagonal value. Finally, if for instance

$$s_{\mathrm{BB}}(a,b) = s_{\mathrm{BB}}(a,a) = 1,$$

then

$$\mu(X_a \cap X_b) = \max\{\mu(X_a), \mu(X_b)\}.$$

Hence one of the sets is contained in the other with equal measure, so in each of the four cases one gets $X_a = X_b$, and injectivity yields $a = b$. The converse is obvious. □

Example 4.5 (Two basic gcd-intersection models)**.** *(a) On $\Omega = \mathbb{N} \times \mathbb{N}$ with counting measure, define*

$$X_n := \{(d, i) : d \mid n,\ 0 \le i \le d-1\}.$$

Then the map $n \mapsto X_n$ is injective,

$$\mu(X_n) = |X_n| = \sigma(n),$$

and

$$X_a \cap X_b = X_{\gcd(a,b)}.$$

(b) On $\Omega = \mathbb{Q}_{>0}$ with counting measure, define

$$X_n := \left\{ \frac{k}{n} : 1 \le k \le n \right\}.$$

Then the map $n \mapsto X_n$ is injective,

$$\mu(X_n) = |X_n| = n,$$

and

$$X_a \cap X_b = X_{\gcd(a,b)}.$$

Corollary 4.6 (Explicit similarities in the gcd-intersection case)**.** *If (X_n) is an injective gcd-intersection model, then the above similarities become*

$$s_{\mathrm{BB}}(a,b) = \frac{\mu(X_{\gcd(a,b)})}{\max\{\mu(X_a), \mu(X_b)\}},$$

$$s_{\mathrm{S}}(a,b) = \frac{2\,\mu(X_{\gcd(a,b)})}{\mu(X_a) + \mu(X_b)},$$

$$s_{\mathrm{J}}(a,b) = \frac{\mu(X_{\gcd(a,b)})}{\mu(X_a) + \mu(X_b) - \mu(X_{\gcd(a,b)})},$$

$$s_{\cos}(a,b) = \frac{\mu(X_{\gcd(a,b)})}{\sqrt{\mu(X_a)\mu(X_b)}}.$$

In the counting-measure models this yields, for instance,

$$\frac{\gcd(a,b)}{\max(a,b)}, \qquad \frac{2\gcd(a,b)}{a+b}, \qquad \frac{\gcd(a,b)}{a+b-\gcd(a,b)}, \qquad \frac{\gcd(a,b)}{\sqrt{ab}},$$

and likewise the corresponding σ-analogues.

5 Operator model and finite sections

Positive definite kernels on $\mathbb{N}$ admit a natural operator-theoretic packaging.

Definition 5.1 (Kernel operator)**.** *Let $k : \mathbb{N} \times \mathbb{N} \to \mathbb{C}$ be Hermitian. A bounded operator T on $\ell^2(\mathbb{N})$ is said to* realize *k if*

$$\langle T e_m, e_n \rangle = k(m,n) \qquad (m, n \in \mathbb{N}),$$

where (e_n) denotes the standard orthonormal basis.

Proposition 5.2 (Positive operators and positive definite kernels)**.** *Let T be a bounded positive self-adjoint operator on $\ell^2(\mathbb{N})$. Then*

$$k_T(m,n) := \langle Te_m, e_n\rangle$$

is a positive definite kernel on $\mathbb{N}$.

Conversely, let k be a positive definite kernel on $\mathbb{N}$, and let

$$G_N = [k(i,j)]_{1\le i,j\le N}$$

be its finite Gram matrices. Then the following are equivalent:

(i) there exists a bounded positive self-adjoint operator T on $\ell^2(\mathbb{N})$ realizing k;

(ii) the quadratic form

$$q(c) := \sum_{m,n\ge 1} k(m,n)c_m\overline{c_n}$$

is bounded on finitely supported vectors;

(iii)

$$\sup_N \|G_N\|_{\ell^2_N\to\ell^2_N} < \infty.$$

In that case,

$$q(c) = \langle Tc, c\rangle$$

on finitely supported vectors and

$$\|T\| = \sup_N \|G_N\|.$$

Proof. If $T \ge 0$, then for every finitely supported family (c_j),

$$\sum_{i,j} c_i\overline{c_j}\, k_T(i,j) = \Big\langle T\sum_i c_ie_i, \sum_j c_je_j\Big\rangle = \Big\|T^{1/2}\sum_i c_ie_i\Big\|^2 \ge 0,$$

so k_T is positive definite.

Now let k be positive definite and define q on $c_{00}\subset\ell^2(\mathbb{N})$ by the displayed formula. Since the matrices G_N are positive semidefinite, one has $q(c)\ge 0$ on c_{00}.

Assume first that (iii) holds. If c is supported in $\{1,\dots,N\}$, then

$$q(c) = c^*G_Nc \le \|G_N\|\,\|c\|_2^2 \le M\|c\|_2^2$$

with $M := \sup_N\|G_N\|$. Hence (ii) holds.

Conversely, if (ii) holds with bound M, then for every $c\in\mathbb{C}^N$ viewed as a finitely supported vector,

$$c^*G_Nc = q(c) \le M\|c\|_2^2.$$

Since each G_N is Hermitian positive semidefinite, this implies $\|G_N\|\le M$ for all N. Hence (ii) and (iii) are equivalent.

Assume now that (ii) holds. By polarization,

$$B(c,d) := \sum_{m,n\ge 1} k(m,n)c_m\overline{d_n}$$

defines a bounded sesquilinear form on c_{00}, because

$$|B(c,d)| \le q(c)^{1/2}q(d)^{1/2} \le M\|c\|_2\|d\|_2.$$

Therefore B extends uniquely to a bounded sesquilinear form on $\ell^2(\mathbb{N}) \times \ell^2(\mathbb{N})$. By the Riesz representation theorem there exists a unique bounded operator T with

$$B(c, d) = \langle Tc, d \rangle.$$

Since $B(c, c) = q(c) \geq 0$, the operator T is positive, hence self-adjoint. Taking $c = e_m$, $d = e_n$ yields

$$\langle Te_m, e_n \rangle = B(e_m, e_n) = k(m, n),$$

so T realizes k. Thus (ii) implies (i).

Finally, if (i) holds, then for finitely supported c,

$$q(c) = \langle Tc, c \rangle \leq \|T\| \, \|c\|_2^2,$$

so (ii) holds. Moreover, $G_N = P_N T P_N$ is the compression of T to the first N coordinates, hence $\|G_N\| \leq \|T\|$. The equalities above show that the least bound in (ii) equals both $\|T\|$ and $\sup_N \|G_N\|$. □

Remark 5.3 (Finite sections and spectrum)**.** *For a bounded self-adjoint realizing operator T, the matrices G_N are the finite sections $P_N T P_N$. Spectral approximation of T by* $\mathrm{spec}(G_N)$ *is delicate in general; however, for compact positive operators the nonzero eigenvalues of the finite sections converge to the nonzero eigenvalues of T.*

6 Named arithmetic properties

We now introduce a list of additional properties. A given positive definite similarity kernel may possess many of them simultaneously.

Definition 6.1 (Projective)**.** *A kernel k is* projective *if*

$$k(ca, cb) = k(a, b) \qquad (a, b, c \in \mathbb{N}).$$

Definition 6.2 (Rational)**.** *A kernel k is* rational *if $k(a, b) \in \mathbb{Q}$ for all $a, b \in \mathbb{N}$.*

Definition 6.3 (Integral-reciprocal)**.** *A rational kernel k is* integral-reciprocal *if*

$$\frac{1}{k(a, b)} \in \mathbb{N} \qquad (a, b \in \mathbb{N}).$$

Definition 6.4 (Bounded denominator)**.** *A rational kernel k is* bounded-denominator *if there exists $S \in \mathbb{N}$ such that*

$$\frac{S}{k(a, b)} \in \mathbb{N} \qquad (a, b \in \mathbb{N}).$$

The least such S, if it exists, is denoted by $S(k)$.

Definition 6.5 (Primitive decay)**.** *A projective kernel k has* primitive decay *if for every sequence of coprime pairs (x_n, y_n) with $x_n + y_n \to \infty$ one has*

$$k(x_n, y_n) \to 0.$$

Definition 6.6 (Height-sensitive)**.** *A projective kernel k is* height-sensitive *if there exist a function h on coprime pairs and a monotone decreasing function F such that*

$$k(a, b) = F\big(h\big(\Big(\frac{a}{\gcd(a, b)}, \frac{b}{\gcd(a, b)}\Big)\big)\big).$$

Definition 6.7 (Reciprocal and radical profiles)**.** *For an integral-reciprocal kernel define*

$$M_k(a,b) := \frac{1}{k(a,b)} \in \mathbb{N}, \qquad R_k(a,b) := \mathrm{rad}(M_k(a,b)).$$

Definition 6.8 (Radical stability)**.** *An integral-reciprocal kernel k is* radically stable at exponent $t > 0$ *if*

$$K_t(a,b) := R_k(a,b)^{-t}$$

is positive definite. It is fully radically stable *if this holds for every $t > 0$.*

Definition 6.9 (Irreducible bounded-denominator similarity)**.** *Let k be a bounded-denominator kernel and write*

$$\widetilde{k}(a,b) := \frac{k(a,b)}{S(k)}.$$

We say that k is multiplicatively reducible *if there exist bounded-denominator kernels k_1, k_2 such that*

$$\widetilde{k}(a,b) = \widetilde{k_1}(a,b)\widetilde{k_2}(a,b) \qquad (a,b \in \mathbb{N}),$$

and irreducible *otherwise.*

7 Rational and integer embeddings

The rational-valued subclass is especially important because rational kernels admit rational Hilbert embeddings.

Theorem 7.1 (Rational embedding theorem)**.** *Let*

$$k : \mathbb{N} \times \mathbb{N} \to \mathbb{Q}$$

be a symmetric positive definite kernel with

$$k(n,n) = 1 \qquad (n \in \mathbb{N}).$$

Then there exists a map

$$\phi : \mathbb{N} \to \ell^2(\mathbb{Q})$$

such that

$$k(a,b) = \langle \phi(a), \phi(b) \rangle_{\ell^2} \qquad (a,b \in \mathbb{N}).$$

Moreover, one may choose the vectors inductively so that

$$\phi(n) \notin \mathrm{span}\{\phi(1), \dots, \phi(n-1)\}$$

at each stage.

Proof. We construct the vectors inductively.

For $n = 1$, set

$$\phi(1) = (1, 0, 0, \dots).$$

Now assume that $\phi(1), \dots, \phi(n-1) \in \ell^2(\mathbb{Q})$ have already been constructed so that

$$\langle \phi(i), \phi(j) \rangle = k(i,j) \qquad (1 \le i, j \le n-1),$$

and are linearly independent. After identifying their span with $\mathbb{Q}^{n-1}$, let

$$G_{n-1} = [k(i,j)]_{1 \le i,j \le n-1} \in M_{n-1}(\mathbb{Q}), \qquad v_{n-1} := (k(1,n), \dots, k(n-1,n))^T \in \mathbb{Q}^{n-1}.$$

Set

$$w_{n-1} := G_{n-1}^{-1} v_{n-1} \in \mathbb{Q}^{n-1}.$$

Because G_{n-1} is positive definite, it is invertible over $\mathbb{Q}$. The vector w_{n-1} prescribes the correct scalar products with the previous vectors.

Consider the Gram matrix

$$G_n = \begin{pmatrix} G_{n-1} & v_{n-1} \\ v_{n-1}^T & 1 \end{pmatrix}.$$

Since G_n is positive definite, the Schur complement is positive:

$$1 - v_{n-1}^T G_{n-1}^{-1} v_{n-1} > 0.$$

Thus

$$1 - \|w_{n-1}\|_2^2 > 0.$$

This is a positive rational number. By Lagrange's four-square theorem, every positive rational number is a sum of four rational squares: if $r = u/v > 0$, write $uv = x_1^2 + \cdots + x_4^2$ with integers x_i, then

$$r = \left(\frac{x_1}{v}\right)^2 + \cdots + \left(\frac{x_4}{v}\right)^2.$$

Hence choose $\beta_1, \ldots, \beta_4 \in \mathbb{Q}$ such that

$$\beta_1^2 + \beta_2^2 + \beta_3^2 + \beta_4^2 = 1 - \|w_{n-1}\|_2^2.$$

Define

$$\phi(n) := (w_{n-1,1}, \ldots, w_{n-1,n-1}, \beta_1, \beta_2, \beta_3, \beta_4, 0, 0, \ldots).$$

Then all coordinates are rational,

$$\langle \phi(i), \phi(n) \rangle = k(i,n) \qquad (1 \le i \le n-1),$$

and

$$\|\phi(n)\|_2^2 = \|w_{n-1}\|_2^2 + \beta_1^2 + \cdots + \beta_4^2 = 1 = k(n,n).$$

Finally, if $\phi(n)$ lay in the span of the previous vectors, its last four coordinates would vanish, forcing $\beta_1 = \cdots = \beta_4 = 0$, hence $\|w_{n-1}\|_2^2 = 1$, contradiction. This completes the induction. □

Theorem 7.2 (Integral conditions versus integer-valued embeddings)**.** *Let*

$$k : \mathbb{N} \times \mathbb{N} \to \mathbb{Q}$$

be a positive definite kernel. For each $n \in \mathbb{N}$, let

$$G_n = [k(i,j)]_{1 \le i,j \le n}, \qquad v_n = (k(1,n+1), \ldots, k(n,n+1))^T \in \mathbb{Q}^n.$$

Then the following are equivalent:

(i) $k(1,1) \in \mathbb{N}$, $k(a,b) \in \mathbb{Z}$ for all $a, b \in \mathbb{N}$, and $G_n^{-1} v_n \in \mathbb{Z}^n$ for every n;

(ii) there exists a map $\phi : \mathbb{N} \to \ell^2(\mathbb{Z})$ such that

$$k(a,b) = \langle \phi(a), \phi(b) \rangle_{\ell^2} \qquad (a, b \in \mathbb{N});$$

(iii) $k(1,1) \in \mathbb{N}$, $k(a,b) \in \mathbb{Z}$ for all $a, b \in \mathbb{N}$, and

$$\det(G_n) \mid [\mathrm{adj}(G_n) v_n]_i \qquad (1 \le i \le n,\ n \in \mathbb{N}).$$

Proof. The equivalence of (i) and (iii) is immediate from

$$G_n^{-1} = \frac{\operatorname{adj}(G_n)}{\det(G_n)}.$$

Assume (ii). Then every scalar product $k(a,b) = \langle \phi(a), \phi(b) \rangle$ is an integer, and in particular $k(1,1) \in \mathbb{N}$. The vectors giving the new scalar products with the first n vectors are integral, so $G_n^{-1} v_n \in \mathbb{Z}^n$.

Assume now (i). We construct $\phi(1), \phi(2), \dots$ inductively in $\ell^2(\mathbb{Z})$. For $n = 1$, write

$$k(1,1) = a^2 + b^2 + c^2 + d^2 \qquad (a,b,c,d \in \mathbb{Z})$$

by Lagrange's theorem, and set

$$\phi(1) = (a, b, c, d, 0, 0, \dots).$$

Suppose $\phi(1), \dots, \phi(n)$ have already been constructed in $\ell^2(\mathbb{Z})$ with the required Gram relations. Let

$$G_n = [k(i,j)]_{1 \le i,j \le n}, \qquad v_n = (k(1,n+1), \dots, k(n,n+1))^T.$$

By assumption, $w_n := G_n^{-1} v_n \in \mathbb{Z}^n$. Exactly as in the rational case, positivity of the Schur complement gives

$$k(n+1,n+1) - \|w_n\|_2^2 > 0.$$

This difference is a positive integer, so by Lagrange's theorem it is a sum of four integer squares:

$$k(n+1,n+1) - \|w_n\|_2^2 = \beta_1^2 + \beta_2^2 + \beta_3^2 + \beta_4^2$$

with $\beta_i \in \mathbb{Z}$. Then

$$\phi(n+1) := (w_{n,1}, \dots, w_{n,n}, \beta_1, \beta_2, \beta_3, \beta_4, 0, 0, \dots)$$

has integer coordinates and satisfies the required Gram relations. This completes the induction. □

8 Projective reduction

Projective kernels are controlled by their values on coprime pairs.

Proposition 8.1 (Reduction to primitive pairs)**.** *A kernel k on $\mathbb{N}$ is projective if and only if there exists a unique function*

$$F : \{(x,y) \in \mathbb{N}^2 : \gcd(x,y) = 1\} \to \mathbb{R}$$

such that

$$k(a,b) = F\left(\left(\frac{a}{\gcd(a,b)}, \frac{b}{\gcd(a,b)}\right)\right) \qquad (a,b \in \mathbb{N}).$$

Proof. If k is projective and $g = \gcd(a,b)$, then

$$k(a,b) = k\left(g\frac{a}{g}, g\frac{b}{g}\right) = k\left(\frac{a}{g}, \frac{b}{g}\right).$$

Hence k depends only on the coprime pair $\left(\frac{a}{\gcd(a,b)}, \frac{b}{\gcd(a,b)}\right)$, which defines F uniquely.

Conversely, any function of the reduced pair is unchanged under common scaling, so it defines a projective kernel. □

This motivates viewing projective kernels as kernels on primitive pairs, or equivalently on positive rational points of $\mathbf{P}^1(\mathbb{Q})$.

9 Core examples

Example 9.1 (Classical gcd-similarities)**.** *Define*

$$s_{\mathrm{Si}}(a,b) := \frac{\gcd(a,b)}{\min(a,b)}, \qquad s_{\mathrm{BB}}(a,b) := \frac{\gcd(a,b)}{\max(a,b)},$$

$$s_{\mathrm{J}}(a,b) := \frac{\gcd(a,b)}{a+b-\gcd(a,b)}, \qquad s_{\mathrm{S}}(a,b) := \frac{2\gcd(a,b)}{a+b},$$

$$s_{\mathrm{Cos}}(a,b) := \frac{\gcd(a,b)^2}{ab}.$$

All of these are normalized similarities. The kernels s_{BB}, s_{S}, and s_{Cos} arise directly from the set models above. The Simpson similarity satisfies

$$s_{\mathrm{Si}}(a,b) = \mathbf{1}_{\gcd(a,b)=\min(a,b)},$$

so it is positive definite as the indicator of the divisibility preorder. The Jaccard similarity is a natural similarity but is recorded here as a candidate rather than a proved positive definite kernel.

Example 9.2 (Divisor-sum kernels)**.** *Let $w : \mathbb{N} \to [0,\infty)$ be any nonnegative weight function. Define*

$$K_w(a,b) := \sum_{d \mid \gcd(a,b)} w(d).$$

Then K_w is positive definite on $\mathbb{N}$.

Proof. For each $d \in \mathbb{N}$, define $f_d(n) := \mathbf{1}_{d|n}$. Then

$$K_w(a,b) = \sum_{d \geq 1} w(d) f_d(a) f_d(b),$$

a nonnegative sum of rank-one positive semidefinite kernels. □

Example 9.3 (The cubic projective kernel)**.** *Define*

$$k_3(a,b) := \frac{2\gcd(a,b)^3}{ab(a+b)}.$$

Then k_3 is a normalized, rational, projective, integral-reciprocal kernel with primitive decay. Indeed, if $a = gx$, $b = gy$, $\gcd(x,y) = 1$, then

$$k_3(a,b) = \frac{2}{xy(x+y)}.$$

Hence

$$M_{k_3}(a,b) = \frac{xy(x+y)}{2} \in \mathbb{N}.$$

10 Closure properties

Proposition 10.1 (Basic closure properties)**.** *Let k, k_1, k_2 be positive definite kernels on $\mathbb{N}$.*

(a) If $\lambda_1, \lambda_2 \geq 0$, then $\lambda_1 k_1 + \lambda_2 k_2$ is positive definite.

(b) The pointwise product $k_1 k_2$ is positive definite.

(c) If $k(a,a) > 0$ for all a, then the normalized kernel

$$k^{\mathrm{nor}}(a,b) := \frac{k(a,b)}{\sqrt{k(a,a)k(b,b)}}$$

is positive definite and has unit diagonal.

If the kernels involved are projective, then the resulting kernels are again projective.

Proof. Part (a) is immediate from positivity of Gram matrices. Part (b) is the Schur product theorem. Part (c) follows by conjugating each finite Gram matrix by the positive diagonal matrix with entries $k(a_i,a_i)^{-1/2}$. □

Definition 10.2 (Rationally and integrally embeddable)**.** *A rational positive definite kernel k is* rationally embeddable *if there exists*

$$\phi : \mathbb{N} \to \ell^2(\mathbb{Q})$$

with

$$k(a,b) = \langle \phi(a), \phi(b) \rangle.$$

It is integrally embeddable *if one may choose ϕ with values in $\ell^2(\mathbb{Z})$.*

11 Radical transforms and Schoenberg-type questions

Suppose k is integral-reciprocal and write

$$M_k(a,b) = \frac{1}{k(a,b)}, \qquad R_k(a,b) = \mathrm{rad}(M_k(a,b)).$$

A natural transformed kernel is

$$K_t(a,b) := R_k(a,b)^{-t} \qquad (t > 0).$$

This transform forgets prime powers and keeps only prime support. It is therefore highly nonlinear and should not be expected to preserve positivity in general.

A useful sufficient criterion is conditional negative definiteness of

$$\gamma_k(a,b) := \log R_k(a,b).$$

Indeed, if γ_k is conditionally negative definite, then by Schoenberg's theorem

$$e^{-t\gamma_k(a,b)} = R_k(a,b)^{-t}$$

is positive definite for every $t > 0$.

Question 11.1 (Radical stability problem)**.** *For which integral arithmetic kernels k is the logarithmic radical profile*

$$\gamma_k(a,b) = \log \mathrm{rad}\big(1/k(a,b)\big)$$

conditionally negative definite?

12 Independence examples

The optional properties introduced above are genuinely independent. The examples below were checked numerically with `sympy` on principal sections up to size 12×12 before being written into this note.

Example 12.1 (A rational kernel which is neither projective, nor integral-reciprocal, nor almost orthogonal)**.** *Define*

$$k_{\mathrm{rat}}(a,b) := \begin{cases} 1, & a = b, \\ \dfrac{4ab}{(2a+1)(2b+1)}, & a \neq b. \end{cases}$$

Then k_{rat} is a normalized positive definite similarity kernel. It is rational-valued, but not projective, not integral-reciprocal, and not almost orthogonal.

Proof. Set

$$f(n) := \frac{2n}{2n+1} \in (0,1).$$

For $a \neq b$ one has $k_{\mathrm{rat}}(a,b) = f(a)f(b)$. Hence for every finite set $a_1, \dots, a_r$ the Gram matrix satisfies

$$\big(k_{\mathrm{rat}}(a_i,a_j)\big)_{i,j} = vv^T + D,$$

where $v = (f(a_1), \dots, f(a_r))^T$ and

$$D = \mathrm{diag}(1 - f(a_1)^2, \dots, 1 - f(a_r)^2).$$

Since $1 - f(n)^2 > 0$, the matrix D is positive definite, hence so is $vv^T + D$.

The kernel is rational-valued. It is not projective because

$$k_{\mathrm{rat}}(1,2) = \frac{8}{15}, \qquad k_{\mathrm{rat}}(2,4) = \frac{32}{45}.$$

It is not integral-reciprocal because

$$\frac{1}{k_{\mathrm{rat}}(1,2)} = \frac{15}{8} \notin \mathbb{N}.$$

Finally, since f is increasing,

$$k_{\mathrm{rat}}(a,b) = f(a)f(b) \geq f(1)f(2) = \frac{8}{15} > 0$$

for all distinct a, b, so the kernel is not almost orthogonal. □

Example 12.2 (A projective kernel which is neither rational nor almost orthogonal)**.** *Define*

$$k_{\mathrm{proj}}(a,b) := \frac{1}{2} + \frac{1}{2}\left(\frac{\min(a,b)}{\max(a,b)}\right)^{1/2}.$$

Then k_{proj} is a normalized positive definite similarity kernel. It is projective, but neither rational-valued nor almost orthogonal.

Proof. For $s > 0$ define

$$h_s(a,b) := \left(\frac{\min(a,b)}{\max(a,b)}\right)^s.$$

Writing $x = \log a$ and $y = \log b$, one has

$$h_s(a,b) = e^{-s|x-y|}.$$

The function $x \mapsto e^{-s|x|}$ is positive definite on $\mathbb{R}$ because

$$e^{-s|x-y|} = \frac{1}{\pi}\int_{\mathbb{R}} \frac{s}{s^2+t^2} e^{itx} e^{-ity}\, dt.$$

Hence h_s is positive definite on $\mathbb{N}$ for every $s > 0$. Taking $s = 1/2$ and adding the constant kernel $1/2$ preserves positive definiteness.

Projectivity follows from

$$\frac{\min(ca, cb)}{\max(ca, cb)} = \frac{\min(a,b)}{\max(a,b)}.$$

The value

$$k_{\mathrm{proj}}(1,2) = \frac{1}{2} + \frac{1}{2}\,2^{-1/2}$$

is irrational, so the kernel is not rational-valued. Since $k_{\mathrm{proj}}(a,b) \geq 1/2$ for all a, b, it is not almost orthogonal. □

Example 12.3 (An almost orthogonal kernel which is neither projective nor rational)**.** *Define*

$$k_{\mathrm{ao}}(a,b) := e^{-|a-b|}.$$

Then k_{ao} is a normalized positive definite similarity kernel. It is almost orthogonal, but neither rational-valued nor projective.

Proof. This is the restriction to $\mathbb{N}$ of the Laplace kernel on $\mathbb{R}$:

$$e^{-|x-y|} = \frac{1}{\pi}\int_{\mathbb{R}} \frac{1}{1+t^2} e^{itx} e^{-ity}\, dt.$$

Hence it is positive definite. Also $k_{\mathrm{ao}}(a,a) = 1$ and $0 < e^{-|a-b|} < 1$ for $a \neq b$, so it is a similarity kernel. It is almost orthogonal because for fixed a,

$$e^{-|a-b|} \to 0 \qquad (b \to \infty).$$

The value e^{-1} is irrational, and projectivity fails since

$$k_{\mathrm{ao}}(1,2) = e^{-1}, \qquad k_{\mathrm{ao}}(2,4) = e^{-2}.$$

□

Example 12.4 (An integral-reciprocal kernel which is neither projective nor almost orthogonal)**.** *Define*

$$k_{\mathrm{int}}(a,b) := \begin{cases} 1, & a = b, \\ \dfrac{1}{2}, & a \neq b \text{ and } a \equiv b \pmod 2, \\ \dfrac{1}{3}, & a \not\equiv b \pmod 2. \end{cases}$$

Then k_{int} is a normalized positive definite similarity kernel. It is integral-reciprocal, but neither projective nor almost orthogonal.

Proof. Let J denote the all-ones kernel and let

$$\nu(n) := \begin{cases} 1, & n \text{ odd}, \\ -1, & n \text{ even}. \end{cases}$$

Then one checks directly that

$$k_{\text{int}}(a,b) = \frac{1}{2}\,\delta_{ab} + \frac{5}{12}J(a,b) + \frac{1}{12}\nu(a)\nu(b).$$

Each summand is positive semidefinite, so the kernel is positive definite. The reciprocal takes only the values $1, 2, 3$, hence is integral. It is not projective because

$$k_{\text{int}}(1,2) = \frac{1}{3}, \qquad k_{\text{int}}(2,4) = \frac{1}{2},$$

and it is not almost orthogonal because $k_{\text{int}}(a,b) \geq 1/3$ for all a, b. □

Remark 12.5. *These four examples show that rationality, projectivity, almost orthogonality, and integral-reciprocity are logically independent over the class of normalized positive definite similarity kernels on $\mathbb{N}$, except for the trivial implication*

$$\textit{integral-reciprocal} \Longrightarrow \textit{rational}.$$

13 Kernels on the positive rationals

Projective kernels on $\mathbb{N}$ admit a natural extension to the multiplicative group $\mathbb{Q}_{>0}$ of positive rational numbers. This point of view was already suggested in a MathOverflow discussion on kernels over $\mathbb{Q}_{>0}$ and projective reduction. [1]

Definition 13.1 (Rational extension)**.** *Let $k : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ be a projective kernel. For $r, s \in \mathbb{Q}_{>0}$ choose positive integers a, b, c, d such that*

$$r = \frac{a}{b}, \qquad s = \frac{c}{d},$$

and define

$$K(r,s) := k\left(\frac{ad}{\gcd(ad,bc)}, \frac{bc}{\gcd(ad,bc)}\right).$$

Proposition 13.2 (Well-definedness and positivity of the rational extension)**.** *Let k be a projective positive definite kernel on $\mathbb{N}$. Then its rational extension K is well-defined and is a positive definite kernel on $\mathbb{Q}_{>0}$.*

Proof. First note that

$$\frac{r}{s} = \frac{ad}{bc}.$$

Hence the pair

$$\left(\frac{ad}{\gcd(ad,bc)}, \frac{bc}{\gcd(ad,bc)}\right)$$

is the primitive representative of the ratio r/s. It depends only on r and s, not on the chosen presentations $r = a/b$, $s = c/d$. Therefore K is well-defined.

To prove positive definiteness, let $r_1, \ldots, r_N \in \mathbb{Q}_{>0}$ and choose for each r_i a primitive pair $(x_i, y_i) \in \mathbb{N}^2$ with

$$r_i = \frac{x_i}{y_i}, \qquad \gcd(x_i, y_i) = 1.$$

Then by construction

$$K(r_i, r_j) = k\left(\frac{x_i y_j}{\gcd(x_i y_j, x_j y_i)}, \frac{x_j y_i}{\gcd(x_i y_j, x_j y_i)}\right).$$

Since k is projective, this is exactly the value of k on the primitive representative of r_i/r_j, hence the matrix $\left(K(r_i, r_j)\right)_{i,j}$ is obtained from values of k on $\mathbb{N}$. Therefore it is positive semidefinite. □

Remark 13.3. *Thus every projective positive definite kernel on $\mathbb{N}$ may equally well be viewed as a positive definite kernel on $\mathbb{Q}_{>0}$. This is often the more natural setting, since $\mathbb{Q}_{>0}$ carries the multiplicative group structure and projective reduction becomes intrinsic.*

14 Frey–Hellegouarch-type kernels

The Frey–Hellegouarch curve attached to a pair (a, b) leads naturally to several arithmetic kernels. Two of them admit particularly clean positivity arguments.

The reciprocal discriminant kernel

For $a, b \in \mathbb{N}$, write

$$a = gx, \qquad b = gy, \qquad \gcd(x, y) = 1.$$

Then the discriminant of the corresponding Frey–Hellegouarch curve is

$$\Delta(a, b) = 16\,(xy(x + y))^2.$$

Hence

$$\frac{1}{\Delta(a, b)} = \frac{1}{16\,(xy(x + y))^2}.$$

Proposition 14.1 (Reciprocal discriminant kernel)**.** *Assume that*

$$k_3(a, b) := \frac{2\,\gcd(a, b)^3}{ab(a + b)}$$

is positive definite. Then

$$k_\Delta(a, b) := \frac{1}{\Delta(a, b)}$$

is also positive definite.

Proof. With $a = gx$, $b = gy$, $\gcd(x, y) = 1$, one has

$$k_3(a, b) = \frac{2}{xy(x + y)}.$$

Therefore

$$k_\Delta(a, b) = \frac{1}{16\,(xy(x + y))^2} = \frac{1}{64}\,k_3(a, b)^2.$$

If k_3 is positive definite, then k_3^2 is positive definite by the Schur product theorem. Multiplication by the positive constant $1/64$ preserves positive definiteness. □

The reciprocal j-invariant kernel

For the same curve, the j-invariant is

$$j(a,b) = 16^2 \, \frac{(a^2+ab+b^2)^3}{a^2b^2(a+b)^2}.$$

Hence

$$\frac{1}{j(a,b)} = \frac{a^2b^2(a+b)^2}{16^2\,(a^2+ab+b^2)^3}.$$

A proof that this defines a positive definite kernel was given in a MathOverflow answer. [2]

Theorem 14.2 (Reciprocal j-invariant kernel)**.** *The kernel*

$$k_j(a,b) := \frac{1}{j(a,b)} = \frac{a^2b^2(a+b)^2}{16^2\,(a^2+ab+b^2)^3}$$

is positive definite on $\mathbb{N}$*. Consequently, by projective extension, it also yields a positive definite kernel on* $\mathbb{Q}_{>0}$*.*

Proof. We decompose k_j into simpler positive definite factors.

First, the Hilbert kernel

$$h(a,b) := \frac{1}{a+b}$$

is positive definite, since

$$\frac{1}{a+b} = \int_0^1 x^{a+b-1}\,dx$$

and therefore

$$h(a,b) = \int_0^1 x^{a-1/2} x^{\,b-1/2}\,\frac{dx}{x}.$$

Second, the rank-one kernel

$$r(a,b) := ab$$

is positive definite, since

$$r(a,b) = u(a)u(b), \qquad u(n) = n.$$

Hence

$$u(a,b) := \frac{ab}{(a+b)^2} = r(a,b)\,h(a,b)^2$$

is positive definite by Schur multiplication.

Now observe that

$$a^2+ab+b^2 = (a+b)^2 - ab,$$

so

$$\frac{(a+b)^2}{(a^2+ab+b^2)^3} = \frac{1}{(a+b)^4}\,\frac{1}{\left(1-\frac{ab}{(a+b)^2}\right)^3} = h(a,b)^4\,\frac{1}{(1-u(a,b))^3}.$$

Since $0 \le u(a,b) < 1$ for all $a,b \in \mathbb{N}$, we have the absolutely convergent power series

$$\frac{1}{(1-u)^3} = \sum_{m=0}^{\infty} \binom{m+2}{2} u^m.$$

Each power $u(a,b)^m$ is positive definite by repeated Schur multiplication, and all coefficients are nonnegative, so

$$(a,b) \longmapsto \frac{1}{(1-u(a,b))^3}$$

is positive definite as a pointwise limit of positive definite kernels.

Therefore

$$v(a,b) := \frac{(a+b)^2}{(a^2+ab+b^2)^3}$$

is positive definite, being the product of the positive definite kernels $h(a,b)^4$ and $(1-u(a,b))^{-3}$.

Finally,

$$k_j(a,b) = \frac{1}{16^2}\, a^2b^2\, v(a,b),$$

and the factor

$$(a,b) \longmapsto a^2b^2$$

is again rank one and positive definite. Hence k_j is positive definite. □

Remark 14.3. *The projective character of the formula implies that k_j descends canonically to a positive definite kernel on $\mathbb{Q}_{>0}$ via the rational extension construction above.*

14.1 The j-invariant of the associated family

Let

$$R(a,b) := \mathrm{rad}\left(\frac{ab}{\gcd(a,b)^2}\right), \qquad E_{a,b} : y^2 = x^3 - R(a,b)^2 x.$$

Since this is a short Weierstrass model of the form

$$y^2 = x^3 + Ax + B$$

with

$$A = -R(a,b)^2, \qquad B = 0,$$

its j-invariant is

$$j(E_{a,b}) = 1728\,\frac{4A^3}{4A^3+27B^2} = 1728.$$

Hence

$$j(a,b) = 1728 \qquad \text{for all } a,b \in \mathbb{N}.$$

Therefore

$$\frac{1}{j(a,b)} = \frac{1}{1728}$$

is a constant kernel, and so it is positive semidefinite on $\mathbb{N}$. Indeed, for any $a_1,\dots,a_n \in \mathbb{N}$ and $c_1,\dots,c_n \in \mathbb{R}$,

$$\sum_{i,j=1}^{n} c_i c_j \frac{1}{j(a_i,a_j)} = \frac{1}{1728}\left(\sum_{i=1}^{n} c_i\right)^2 \geq 0.$$

Thus $1/j(a,b)$ is positive definite in the semidefinite sense, but not strictly positive definite.

15 A radical kernel and an associated elliptic curve

In this section we prove that the kernel

$$k(a,b) := \frac{1}{\operatorname{rad}\left(\frac{ab}{\gcd(a,b)^2}\right)} \qquad (a,b \in \mathbb{N})$$

is positive definite on $\mathbb{N}$, and then attach to it a natural family of elliptic curves whose conductor is governed by the square of the same radical quantity.

15.1 Positive definiteness of the kernel $1/\operatorname{rad}(ab/\gcd(a,b)^2)$

We begin with a useful auxiliary notion.

Definition 15.1. *For $n = \prod_p p^{\nu_p(n)} \in \mathbb{N}$ and $a \in \mathbb{N}$, we write*

$$n \parallel a$$

if for every prime $p \mid n$ one has

$$\nu_p(a) = \nu_p(n).$$

Thus the primes dividing n occur in a with exactly the same exponents as in n, while primes not dividing n are unrestricted.

Definition 15.2. *For $a, b \in \mathbb{N}$, define*

$$\gcd'(a,b) := \prod_{\nu_p(a)=\nu_p(b)\geq 1} p^{\nu_p(a)}.$$

Equivalently, $\gcd'(a,b)$ is the largest positive integer n such that

$$n \parallel a \qquad \text{and} \qquad n \parallel b.$$

Lemma 15.3. *For all $a, b \in \mathbb{N}$ one has*

$$\frac{1}{\operatorname{rad}\left(\frac{ab}{\gcd(a,b)^2}\right)} = \frac{\operatorname{rad}(\gcd(a,b))\,\operatorname{rad}(\gcd'(a,b))}{\operatorname{rad}(a)\operatorname{rad}(b)}.$$

Proof. It is enough to compare the local contribution of each prime p.

Write

$$\alpha = \nu_p(a), \qquad \beta = \nu_p(b).$$

Then

$$\nu_p\left(\frac{ab}{\gcd(a,b)^2}\right) = \alpha + \beta - 2\min(\alpha,\beta) = |\alpha - \beta|.$$

Hence the p-factor of the left-hand side is

$$\begin{cases} 1, & \alpha = \beta, \\ p^{-1}, & \alpha \neq \beta. \end{cases}$$

Now consider the right-hand side. The p-factor of $\operatorname{rad}(a)\operatorname{rad}(b)$ is

$$p^{\mathbf{1}_{\alpha\geq 1}+\mathbf{1}_{\beta\geq 1}}.$$

The p-factor of $\mathrm{rad}(\gcd(a,b))$ is p if and only if $\min(\alpha,\beta)\geq 1$, and otherwise 1. The p-factor of $\mathrm{rad}(\gcd'(a,b))$ is p if and only if $\alpha=\beta\geq 1$, and otherwise 1.

Thus the p-factor of the right-hand side is:

$$\begin{cases} 1, & \alpha=\beta=0,\\ p^{-1}, & \text{exactly one of } \alpha,\beta \text{ is positive},\\ 1, & \alpha=\beta\geq 1,\\ p^{-1}, & \alpha\neq\beta,\ \alpha,\beta\geq 1. \end{cases}$$

This is exactly the same as the local factor on the left-hand side. Therefore the two expressions are equal. □

Theorem 15.4. *The kernel*

$$k(a,b):=\frac{1}{\mathrm{rad}\left(\frac{ab}{\gcd(a,b)^2}\right)}$$

is positive definite on $\mathbb{N}$. *In fact, it is strictly positive definite.*

Proof. By Lemma 15.3, we may write

$$k(a,b)=\frac{f(a,b)}{\mathrm{rad}(a)\,\mathrm{rad}(b)},\qquad \text{where}\qquad f(a,b):=\mathrm{rad}(\gcd(a,b))\ \mathrm{rad}(\gcd'(a,b)).$$

Since multiplication by a factor of the form $u(a)u(b)$ preserves positive definiteness, it is enough to prove that f is positive definite.

We now decompose f into a sum of rank-one positive semidefinite kernels.

Let D denote the set of squarefree positive integers. For each $d\in D$, define

$$g_d(a,b):=\varphi(d)\,\mathbf{1}_{d\mid a}\,\mathbf{1}_{d\mid b}.$$

Then

$$\sum_{d\in D} g_d(a,b)=\mathrm{rad}(\gcd(a,b)).$$

Indeed, if $m=\mathrm{rad}(\gcd(a,b))$, then the squarefree divisors d with $d\mid a$ and $d\mid b$ are exactly the divisors of m, so

$$\sum_{d\in D} g_d(a,b)=\sum_{d\mid m}\varphi(d)=m=\mathrm{rad}(\gcd(a,b)).$$

Similarly, for each $n\in\mathbb{N}$, define

$$h_n(a,b):=\varphi(\mathrm{rad}(n))\,\mathbf{1}_{n\parallel a}\,\mathbf{1}_{n\parallel b}.$$

Then

$$\sum_{n\geq 1} h_n(a,b)=\mathrm{rad}(\gcd'(a,b)).$$

To see this, write

$$\gcd'(a,b)=\prod_{i=1}^{r} p_i^{e_i}.$$

The integers n satisfying $n\parallel a$ and $n\parallel b$ are exactly the numbers of the form

$$n=\prod_{i\in I} p_i^{e_i}\qquad (I\subseteq\{1,\dots,r\}),$$

and for such n one has

$$\varphi(\mathrm{rad}(n)) = \prod_{i \in I} (p_i - 1).$$

Hence

$$\sum_{n \| a,\ n \| b} \varphi(\mathrm{rad}(n)) = \prod_{i=1}^{r} (1 + (p_i - 1)) = \prod_{i=1}^{r} p_i = \mathrm{rad}(\gcd(a,b)).$$

Therefore

$$f(a,b) = \left(\sum_{d \in D} g_d(a,b) \right) \left(\sum_{n \geq 1} h_n(a,b) \right) = \sum_{d \in D} \sum_{n \geq 1} f_{d,n}(a,b),$$

where

$$f_{d,n}(a,b) := g_d(a,b)\, h_n(a,b) = \varphi(d)\varphi(\mathrm{rad}(n))\, \mathbf{1}_{d|a} \mathbf{1}_{d|b} \mathbf{1}_{n \| a} \mathbf{1}_{n \| b}.$$

Now each $f_{d,n}$ is positive semidefinite, because it has rank one: if we define

$$u_{d,n}(a) := \sqrt{\varphi(d)\varphi(\mathrm{rad}(n))}\ \mathbf{1}_{d|a} \mathbf{1}_{n \| a},$$

then

$$f_{d,n}(a,b) = u_{d,n}(a)\, u_{d,n}(b).$$

Hence f is a sum of positive semidefinite kernels, and is therefore positive semidefinite. It follows that k is positive semidefinite as well.

It remains to prove strict positive definiteness. Let $a_1, \ldots, a_r \in \mathbb{N}$ be pairwise distinct, and let $c_1, \ldots, c_r \in \mathbb{R}$ not all be zero. Choose a_j maximal among the a_i with $c_i \neq 0$, and set

$$n := a_j, \qquad d := \mathrm{rad}(n).$$

Consider the summand $f_{d,n}$.

If $f_{d,n}(a_i, a_i) \neq 0$, then $d \mid a_i$ and $n \parallel a_i$. The second condition means that every prime dividing n occurs in a_i with exactly the same exponent as in n. Since also $d \mid a_i$, this forces $n \mid a_i$. If a_i had any further prime factor not dividing n, then $a_i > n$, contradicting the maximality of $n = a_j$. Thus $a_i = n = a_j$.

Therefore, among the numbers $a_1, \ldots, a_r$, the kernel $f_{d,n}$ is supported only at the point a_j, and hence

$$\sum_{i,\ell=1}^{r} c_i c_\ell\, f_{d,n}(a_i, a_\ell) = \varphi(d)\varphi(\mathrm{rad}(n))\, c_j^2 > 0.$$

Since f is the sum of the kernels $f_{d,n}$, we obtain

$$\sum_{i,\ell=1}^{r} c_i c_\ell\, f(a_i, a_\ell) > 0.$$

Thus f is strictly positive definite, and so is k. □

Corollary 15.5. *The kernel*

$$k_2(a,b) := \frac{1}{\mathrm{rad}\left(\frac{ab}{\gcd(a,b)^2}\right)^2}$$

is also positive definite.

Proof. This follows from the Schur product theorem, since

$$k_2(a,b) = k(a,b)^2$$

is the pointwise product of the positive definite kernel k with itself. □

15.2 An associated family of elliptic curves

The previous theorem suggests introducing the squarefree quantity

$$R(a,b) := \operatorname{rad}\left(\frac{ab}{\gcd(a,b)^2}\right).$$

We then define the associated elliptic curve

$$E_{a,b}: \qquad y^2 = x^3 - R(a,b)^2 x.$$

Proposition 15.6. *For every $a, b \in \mathbb{N}$, the curve $E_{a,b}$ is an elliptic curve over $\mathbb{Q}$, with*

$$\Delta(E_{a,b}) = 64\, R(a,b)^6.$$

Proof. This is the special case $A = -R(a,b)^2$, $B = 0$ of the short Weierstrass model

$$y^2 = x^3 + Ax + B.$$

For such a model one has

$$\Delta = -16(4A^3 + 27B^2).$$

Substituting $A = -R(a,b)^2$ and $B = 0$ gives

$$\Delta = -16\big(4(-R(a,b)^2)^3\big) = 64\, R(a,b)^6.$$

Since $R(a,b) \geq 1$, the discriminant is nonzero, so $E_{a,b}$ is indeed an elliptic curve. □

Proposition 15.7. *Let $R(a,b) = \operatorname{rad}\left(\frac{ab}{\gcd(a,b)^2}\right)$, and note that $R(a,b)$ is squarefree. Then*

$$E_{a,b} : y^2 = x^3 - R(a,b)^2 x$$

is the congruent-number curve attached to $R(a,b)$, and its conductor is

$$N(E_{a,b}) = \begin{cases} 32\, R(a,b)^2, & \text{if } R(a,b) \text{ is odd,} \\ 16\, R(a,b)^2, & \text{if } R(a,b) \text{ is even.} \end{cases}$$

Proof. This is the standard conductor formula for the congruent-number family

$$E_n : \; y^2 = x^3 - n^2 x$$

with squarefree parameter n, applied to $n = R(a,b)$. □

Corollary 15.8. *Up to the constant factor 16 or 32, the conductor of $E_{a,b}$ is exactly*

$$\operatorname{rad}\left(\frac{ab}{\gcd(a,b)^2}\right)^2.$$

Consequently,

$$\frac{1}{N(E_{a,b})} = \begin{cases} \dfrac{1}{32}\, \dfrac{1}{\operatorname{rad}\left(\frac{ab}{\gcd(a,b)^2}\right)^2}, & \text{if } R(a,b) \text{ is odd,} \\ \dfrac{1}{16}\, \dfrac{1}{\operatorname{rad}\left(\frac{ab}{\gcd(a,b)^2}\right)^2}, & \text{if } R(a,b) \text{ is even,} \end{cases}$$

so the reciprocal conductor is, piecewise up to a constant, governed by the positive definite kernel from the previous subsection.

Remark 15.9. *The key point is that the arithmetic support of the conductor is exactly the square of the squarefree quantity*

$$R(a,b) = \mathrm{rad}\left(\frac{ab}{\gcd(a,b)^2}\right),$$

which is the same radical expression that appears in the kernel

$$\frac{1}{R(a,b)} \qquad \textit{and} \qquad \frac{1}{R(a,b)^2}.$$

Thus the family $E_{a,b}$ provides a natural bridge between positive definite kernels on $\mathbb{N}$ and arithmetic invariants of elliptic curves.

16 A conditional radical inequality from a projective kernel candidate

For $a, b \in \mathbb{N}$, write

$$a = gx, \qquad b = gy, \qquad \gcd(x,y) = 1.$$

Consider the projective arithmetic kernel candidate

$$f(a,b) := \frac{x+y}{\mathrm{rad}(xy(x+y))^3}.$$

The positive definiteness of f has not been proved so far. However, if it were true, then one would obtain a strong radical inequality immediately from the Cauchy–Schwarz inequality for Gram kernels.

Proposition 16.1 (Conditional radical inequality)**.** *Assume that the kernel*

$$f(a,b) = \frac{x+y}{\mathrm{rad}(xy(x+y))^3}, \qquad x = \frac{a}{\gcd(a,b)}, \quad y = \frac{b}{\gcd(a,b)},$$

is positive definite on $\mathbb{N}$. Then for every coprime pair $x, y \in \mathbb{N}$ one has

$$x + y \le \mathrm{rad}(xy(x+y))^3.$$

Equivalently, for coprime positive integers a, b, c with $a + b = c$ one has

$$c \le \mathrm{rad}(abc)^3.$$

Proof. Assume that f is positive definite. For $a = b$ one has

$$x = y = 1,$$

hence

$$f(a,a) = \frac{1+1}{\mathrm{rad}(1 \cdot 1 \cdot 2)^3} = \frac{2}{2^3} = \frac{1}{4}.$$

Thus the rescaled kernel

$$\widetilde{f}(a,b) := 4f(a,b)$$

is again positive definite and satisfies

$$\widetilde{f}(a,a) = 1 \qquad (a \in \mathbb{N}).$$

Therefore $\widetilde{f}$ is the Gram kernel of a family of unit vectors in some Hilbert space. By the Cauchy–Schwarz inequality,

$$\widetilde{f}(a,b) \le 1 \qquad (a,b \in \mathbb{N}).$$

Substituting the definition of $\widetilde{f}$ gives

$$4\,\frac{x+y}{\operatorname{rad}(xy(x+y))^3} \le 1$$

if we keep the kernel exactly as stated above. Hence, if one wants the sharp constant 1, one should instead work with the normalized candidate

$$F(a,b) := \frac{4(x+y)}{\operatorname{rad}(xy(x+y))^3},$$

which has diagonal value 1. In that normalized form, positive definiteness yields

$$\frac{4(x+y)}{\operatorname{rad}(xy(x+y))^3} \le 1,$$

that is,

$$x+y \le \operatorname{rad}(xy(x+y))^3.$$

Equivalently, if x, y are coprime and $c = x+y$, then

$$\operatorname{rad}(xy(x+y)) = \operatorname{rad}(xyc),$$

so

$$c \le \operatorname{rad}(xyc)^3.$$

Renaming (x,y,c) as (a,b,c) gives the stated form

$$c \le \operatorname{rad}(abc)^3 \qquad \text{whenever } a+b=c,\ \gcd(a,b)=1.$$

□

Remark 16.2 (Computational evidence)**.** *The normalized kernel*

$$F(a,b) := \frac{4(x+y)}{\operatorname{rad}(xy(x+y))^3}$$

has been checked numerically on principal sections up to size $N = 8000$, and no violation of positive definiteness has been found.

Remark 16.3. *At present this is only computational evidence. The positive definiteness of F — and hence the resulting radical inequality above — remains unproved.*

17 Positive definite kernels over the naturals and the Riemann Hypothesis

Proposition 17.1 (The Redheffer Gram kernel)**.** *For $n \ge 1$, let $R_n = (r_{ij})_{1\le i,j\le n}$ be the Redheffer matrix,*

$$r_{ij} = \begin{cases} 1, & \text{if } j = 1 \text{ or } i \mid j, \\ 0, & \text{otherwise.} \end{cases}$$

Let

$$G_n := R_n^{\top} R_n.$$

Then G_n is positive semidefinite, and its entries are given by

$$(G_n)_{jk} = \begin{cases} n, & j = k = 1, \\ \tau(k), & j = 1,\ k \geq 2, \\ \tau(j), & k = 1,\ j \geq 2, \\ \tau(\gcd(j,k)), & j, k \geq 2. \end{cases}$$

In particular, away from the first row and column, G_n is the divisor kernel

$$(j,k) \longmapsto \tau(\gcd(j,k)).$$

Moreover,

$$\det(G_n) = M(n)^2,$$

where $M(n) = \sum_{m \leq n} \mu(m)$ is the Mertens function. Hence

$$\text{RH} \iff \det(G_n) = O_\varepsilon(n^{1+\varepsilon}) \qquad (\forall \varepsilon > 0).$$

Proof. Since $G_n = R_n^\top R_n$, it is a Gram matrix of the column vectors of R_n, hence positive semidefinite.

Now compute the entries. For $j, k \geq 2$,

$$(G_n)_{jk} = \sum_{i=1}^{n} r_{ij} r_{ik} = \sum_{i=1}^{n} 1_{i|j}\, 1_{i|k}.$$

Because $j, k \leq n$, every common divisor i of j and k lies in $\{1, \ldots, n\}$, so this sum counts the common divisors of j and k. Therefore

$$(G_n)_{jk} = \tau(\gcd(j,k)).$$

For $k \geq 2$,

$$(G_n)_{1k} = \sum_{i=1}^{n} r_{i1} r_{ik} = \sum_{i=1}^{n} 1 \cdot 1_{i|k} = \tau(k),$$

and similarly $(G_n)_{j1} = \tau(j)$ for $j \geq 2$. Finally,

$$(G_n)_{11} = \sum_{i=1}^{n} r_{i1}^2 = \sum_{i=1}^{n} 1 = n.$$

For the determinant, use

$$\det(G_n) = \det(R_n^\top R_n) = \det(R_n)^2.$$

Wilf recalls Redheffer's identity

$$\det(R_n) = M(n),$$

so

$$\det(G_n) = M(n)^2.$$

By the classical equivalence

$$\text{RH} \iff M(n) = O_\varepsilon(n^{1/2+\varepsilon}) \quad (\forall \varepsilon > 0),$$

we get

$$\text{RH} \iff M(n)^2 = O_\varepsilon(n^{1+\varepsilon}) \iff \det(G_n) = O_\varepsilon(n^{1+\varepsilon}),$$

after renaming $\varepsilon/2$ as ε. □

Remark 17.2 (Set model for the divisor kernel)**.** *Let*

$$X_m := \{d \in \mathbb{N} : \ d \mid m\},$$

viewed inside $(\mathbb{N}, \mu)$ *with counting measure. Then*

$$X_a \cap X_b = X_{\gcd(a,b)}, \qquad \mu(X_m) = |X_m| = \tau(m),$$

and therefore

$$\tau(\gcd(a,b)) = |X_a \cap X_b| = \langle 1_{X_a}, 1_{X_b} \rangle_{L^2(\mathbb{N},\mu)}.$$

Thus the divisor kernel is a natural arithmetic Gram kernel arising from the divisor-set model. The Redheffer Gram kernel is obtained from this kernel by a distinguished deformation of the first row and column.

Proposition 17.3 (Grommer kernel and positive definiteness)**.** *Let*

$$\zeta^*(s) := \pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s), \qquad \xi(s) := \frac{1}{2}\,s(s-1)\zeta^*(s), \qquad \Xi(t) := \xi\left(\frac{1}{2} + it\right).$$

Further set

$$Y(t) := \Xi(\sqrt{t}), \qquad -\frac{Y'(t)}{Y(t)} = \sum_{m \geq 1} s_m t^{m-1},$$

and define the Hankel kernel:

$$k(a,b) := s_{a+b} \qquad (a, b \in \mathbb{N}),$$

and for each $N \geq 1$ *let*

$$K_N := \big(k(a,b)\big)_{1 \leq a,b \leq N} = \big(s_{a+b}\big)_{1 \leq a,b \leq N}.$$

Then the following are equivalent:

1. *The Riemann hypothesis holds.*
2. K_N *is positive definite for every* $N \geq 1$*.*
3. *The kernel* k *is positive definite on* $\mathbb{N}$*.*

Proof. By Grommer's criterion, RH is equivalent to the positivity of all Hankel determinants

$$\det K_N > 0 \qquad (N \geq 1).$$

Since each K_N is real symmetric, Sylvester's criterion implies that

$$\det K_m > 0 \text{ for all leading principal minors } m \leq N$$

is equivalent to K_N being positive definite. Applied for every N, this gives

$$\text{RH} \iff K_N > 0 \quad \text{for all } N \geq 1.$$

So it remains to prove that

$$K_N > 0 \ \forall N \iff k \text{ is positive definite on } \mathbb{N}.$$

The implication

$$k \text{ positive definite on } \mathbb{N} \Longrightarrow K_N > 0 \ \forall N$$

is immediate: by definition, every finite Gram section

$$\big(k(a_i, a_j)\big)_{1\le i,j\le r}$$

must be positive definite; in particular this holds for $a_i = i$, $1 \le i \le N$.

For the converse, assume that every K_N is positive definite. We construct vectors

$$\phi(1), \phi(2), \cdots \in \ell^2$$

such that

$$\langle \phi(a), \phi(b) \rangle = k(a, b) \qquad (a, b \in \mathbb{N}).$$

This proves that k is a positive definite kernel.

We proceed inductively.

For $N = 1$, since $K_1 = (k(1, 1))$ is positive definite, we have $k(1, 1) > 0$. Set

$$\phi(1) := \sqrt{k(1, 1)}\, e_1.$$

Assume now that for some $n \ge 1$ we have already constructed vectors

$$\phi(1), \ldots, \phi(n) \in \operatorname{span}\{e_1, \ldots, e_n\} \subset \ell^2$$

such that

$$\langle \phi(i), \phi(j) \rangle = k(i, j) \qquad (1 \le i, j \le n).$$

Let C_n be the $n \times n$ matrix whose i-th row is $\phi(i)^T$. Then

$$K_n = C_n C_n^T.$$

Since K_n is positive definite, C_n is invertible.

Now define

$$v_n := \begin{pmatrix} k(1, n+1) \\ k(2, n+1) \\ \vdots \\ k(n, n+1) \end{pmatrix}.$$

We seek $\phi(n + 1)$ in the form

$$\phi(n+1) = \begin{pmatrix} x \\ \alpha \end{pmatrix} \in \mathbb{R}^{n+1} \subset \ell^2,$$

where $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, such that

$$\langle \phi(i), \phi(n+1) \rangle = k(i, n+1) \qquad (1 \le i \le n).$$

Since the first n coordinates of $\phi(i)$ are encoded by C_n, this condition is

$$C_n x = v_n.$$

Hence necessarily

$$x = C_n^{-1} v_n.$$

We now choose α so that

$$\|\phi(n+1)\|^2 = k(n+1, n+1).$$

This requires

$$\alpha^2 = k(n+1,n+1) - |C_n^{-1} v_n|^2.$$

It remains to show that the right-hand side is strictly positive.

But the block form of K_{n+1} is

$$K_{n+1} = \begin{pmatrix} K_n & v_n \\ v_n^T & k(n+1,n+1) \end{pmatrix}.$$

Since K_{n+1} is positive definite and $K_n = C_n C_n^T$, the Schur complement of K_n in K_{n+1} is strictly positive:

$$k(n+1,n+1) - v_n^T K_n^{-1} v_n > 0.$$

Using $K_n^{-1} = (C_n^{-1})^T C_n^{-1}$, we get

$$v_n^T K_n^{-1} v_n = v_n^T (C_n^{-1})^T C_n^{-1} v_n = |C_n^{-1} v_n|^2.$$

Therefore

$$k(n+1,n+1) - |C_n^{-1} v_n|^2 > 0,$$

so we may define

$$\phi(n+1) := \begin{pmatrix} C_n^{-1} v_n \\ \sqrt{k(n+1,n+1) - |C_n^{-1} v_n|^2} \end{pmatrix}.$$

Then $\phi(1), \dots, \phi(n+1)$ have Gram matrix K_{n+1}, completing the induction.

Thus there exists a map $\phi : \mathbb{N} \to \ell^2$ with

$$k(a,b) = \langle \phi(a), \phi(b) \rangle \qquad (a, b \in \mathbb{N}),$$

so k is positive definite on $\mathbb{N}$.

This proves

$$K_N > 0 \ \forall N \iff k \text{ is positive definite on } \mathbb{N}.$$

Combining this with Grommer's criterion completes the proof. □

18 A first list of programmatic problems

For the moment, the main geometric theme should be the existence and structure of *almost orthogonal realizations*, i.e. normalized positive definite similarity kernels whose Gram vectors admit angles arbitrarily close to $\pi/2$.

Question 18.1 (Almost orthogonality problem)**.** *Which projective positive definite similarity kernels on $\mathbb{N}$ are almost orthogonal? More concretely, for kernels of the form*

$$k(a,b) = F\left(\left(\frac{a}{\gcd(a,b)}, \frac{b}{\gcd(a,b)}\right)\right),$$

what conditions on F imply

$$\inf_{a \neq b} k(a,b) = 0?$$

Can one characterize this in terms of primitive decay along coprime pairs of large height?

Question 18.2 (Quantitative decay problem)**.** *Among almost orthogonal projective kernels, how fast can the similarity decay along primitive pairs? For instance, which decay profiles are compatible with positive definiteness:*

$$k(a,b) \asymp H\Big(\Big(\frac{a}{\gcd(a,b)}, \frac{b}{\gcd(a,b)}\Big)\Big)^{-\alpha}, \qquad k(a,b) \asymp (\log H\Big(\Big(\frac{a}{\gcd(a,b)}, \frac{b}{\gcd(a,b)}\Big)\Big))^{-\beta},$$

or variants defined by arithmetic data such as $\max\{a', b'\}$ *for* $\left(\frac{a}{\gcd(a,b)}, \frac{b}{\gcd(a,b)}\right) = (a', b')$*?*

Question 18.3 (Classification near orthogonality)**.** *Classify, or at least organize, rational-valued projective positive definite similarity kernels according to their almost orthogonality behaviour. Which kernels are:*

(i) not almost orthogonal,

(ii) almost orthogonal but with slow decay,

(iii) almost orthogonal with uniform quantitative decay along primitive pairs?

Question 18.4 (Embedding problem for almost orthogonal kernels)**.** *Give intrinsic criteria for a rational positive definite kernel to admit a realization by unit vectors in* $\ell^2(\mathbb{Q})$ *that are almost orthogonal. Likewise, when does an integer-valued or integral-reciprocal kernel admit an integral realization in which the vectors become asymptotically orthogonal along primitive directions?*

Question 18.5 (Integral almost orthogonality)**.** *Classify integral arithmetic kernels, i.e. projective positive definite similarity kernels with*

$$1/k(a,b) \in \mathbb{N},$$

that are almost orthogonal. What arithmetic shapes can the reciprocal profile

$$M_k(a,b) = \frac{1}{k(a,b)}$$

have when $M_k(a,b) \to \infty$ *along primitive pairs?*

Question 18.6 (Irreducibility problem)**.** *Which bounded-denominator similarities are multiplicatively irreducible? Among almost orthogonal kernels, when is the decay mechanism itself multiplicatively irreducible? How does irreducibility interact with positive definiteness?*

Question 18.7 (Set-model problem)**.** *Which arithmetic similarity kernels arise from an injective measurable set model satisfying*

$$X_a \cap X_b = X_{\gcd(a,b)}?$$

When such a model exists, does almost orthogonality correspond to a transparent geometric or measure-theoretic sparsity property of the family (X_n)*?*

Question 18.8 (Operator realization problem)**.** *Which almost orthogonal arithmetic kernels are operator-bounded on* $\ell^2(\mathbb{N})$*? For kernels with bounded realization, how much of the asymptotic near-orthogonality is visible in the finite Gram sections and in the spectrum of the associated operator?*

References

[1] O. Leka, *The abc-conjecture over the positive rationals and Levy-Schoenberg kernels?*, MathOverflow, 2020. https://mathoverflow.net/questions/352880/the-abc-conjecture-over-the-positive-rationals-and-levy-schoenberg-kernels

[2] O. Leka, *Discriminant of elliptic curve (Frey-Hellegouarch), j-invariant and positive definite kernels, similarities?*, MathOverflow, 2020; answer updated 2024. https://mathoverflow.net/questions/361894/discriminant-of-elliptic-curve-frey-hellegouarch-j-invariant-and-positive-def

[3] O. Leka, *Freys elliptic curves and Hilbert spaces?*, MathOverflow, 2020. https://mathoverflow.net/questions/364731/freys-elliptic-curves-and-hilbert-spaces