Visualizing High-Dimensional Points with Figures

Orges Leka

March 14, 2026

Abstract

This note proposes a visual language for high-dimensional vectors based on canonical planar figures. The guiding analogy is the Cartesian picture: a point (x, y) is represented by its position in the plane, while a high-dimensional vector $x = (x_1, x_2, \dots)$ is represented by a superposition of memorable shapes. The distinguished basis figures are chosen to be a point for the first coordinate, a segment for the second, an equilateral triangle for the third, a square for the fourth, and in general a regular m-gon for the mth coordinate. Negative coefficients are encoded by reflection through the origin. We also describe an alternative *zeta-fan* dictionary, in which the basis indices are arranged as rays on a semicircle according to the cumulative $1/n^2$ law. We then illustrate the construction with two arithmetic kernels on $\{1, \dots, 30\}$: the meet kernel from the divisibility-poset picture and the classical gcd kernel with the Smith feature map.

1 Motivation

The ordinary Cartesian picture identifies a pair $(x, y) \in \mathbb{R}^2$ with a single point in the plane. This is effective because it makes addition, distance, and direction visually accessible. The aim here is to keep that spirit for vectors with many coordinates. Instead of drawing a point in a very high-dimensional ambient space, we assign to each coordinate a canonical planar figure and represent a vector by superposing these figures.

The key design principle is that the indexing of the basis shapes should be canonical and memorable. Regular polygons provide a natural answer: complexity grows with the index, and as $m \to \infty$ the regular m-gon tends toward a circle. Thus the family interpolates from discrete geometry to roundness in a visually meaningful way.

2 A polygonal figure model for ℓ^2_{fin} and ℓ^2

For each $m \geq 1$ we define a canonical figure symbol B_m as follows:

- B_1: a point,
- B_2: a line segment,
- B_3: an equilateral triangle,
- B_4: a square,
- in general, B_m: a regular m-gon.

Each B_m is centered at the origin and given a fixed orientation. These visible figures should be understood as canonical glyphs attached to the indices $m = 1, 2, 3, \dots$.

For a finitely supported vector $x = (x_1, x_2, \dots) \in \ell^2_{\mathrm{fin}}$, we write formally

$$P_x = \sum_{m \geq 1} x_m B_m.$$

At this stage, this is a symbolic expression. Its precise Hilbert-space meaning will be given in Section 9 below.

Remark 2.1. *The visible size of the glyphs in the rendered images is a graphical choice made for readability. It is not part of the Hilbert-space structure itself.*

3 Negative coefficients

Negative coefficients are represented by reflection through the origin. Thus, if $x_m < 0$, the contribution of the mth mode is not $|x_m|B_m$ but

$$|x_m|(-B_m),$$

where $-B_m$ is the image of B_m under the map $z \mapsto -z$, i.e. a rotation by 180° about the origin. This keeps the symbolic representation linear and visually coherent:

$$P_x = \sum_{m\geq 1} x_m B_m.$$

4 An alternative zeta-fan dictionary

The polygon family is not the only possible canonical dictionary of basis figures. A second natural choice is to encode the basis vectors by rays in a semicircular fan, with angular spacing determined by the zeta distribution of exponent 2.

For $n \geq 1$, let

$$p_n := \frac{1/n^2}{\zeta(2)}, \qquad F(n) := \sum_{k=1}^{n} p_k, \qquad F(0) := 0.$$

Thus $(p_n)_{n\geq 1}$ is a probability distribution on $\mathbb{N}$, and F is its cumulative distribution function. We use these masses to partition a semicircle into consecutive arcs of angular widths πp_n. The nth arc is

$$I_n := \big[\pi F(n-1),\ \pi F(n)\big],$$

and we place the nth basis symbol at the midpoint angle

$$\theta_n := \frac{\pi}{2}\big(F(n-1) + F(n)\big).$$

The corresponding visible glyph Z_n is then the oriented radial segment from the origin in direction θ_n; for instance one may take

$$Z_n(t) := t\,(\cos\theta_n, \sin\theta_n), \qquad 0 \leq t \leq 1.$$

This construction has two useful visual features. First, small indices receive noticeably more angular room, since the weights p_n decay like n^{-2}. Second, the higher modes accumulate toward one endpoint of the semicircle, so the fan provides a canonical notion of increasingly fine angular resolution.

Exactly as in Section 9, the zeta-fan symbols may be treated abstractly as basis labels. One may therefore replace the polygon symbols B_n by zeta-fan symbols Z_n and define

$$F_{\mathrm{fin}}^{\mathrm{fan}} := \left\{ \sum_{m=1}^{M} x_m Z_m : M \in \mathbb{N},\ x_m \in \mathbb{R} \right\}, \qquad \langle Z_i, Z_j \rangle_{F^{\mathrm{fan}}} := \delta_{ij},$$

followed by Hilbert-space completion. In this way the exact inner-product formalism is unchanged; only the visible dictionary of basis tokens is replaced.

The zeta-fan viewpoint is especially convenient when a positive definite kernel on a finite set is first factorized by Cholesky decomposition. If

$$G = (K(a_i, a_j))_{1 \le i,j \le N} = LL^\top,$$

then the rows of L provide feature vectors in $\mathbb{R}^N$, and each object a_i may be visualized as the zeta-fan superposition

$$\Gamma^{\mathrm{fan}}_{a_i} = \sum_{m=1}^{N} L_{im} Z_m.$$

Thus the same kernel geometry can be rendered either in the polygon dictionary or in the zeta-fan dictionary.

5 The meet kernel example

For the arithmetic application we consider the kernel

$$K(a, b) = a \wedge b,$$

where, writing

$$a = p_1 \cdots p_r, \qquad b = q_1 \cdots q_s$$

for the sorted prime-factor lists of a and b, one sets

$$a \wedge b := \prod_{i=1}^{\min(r,s)} \min(p_i, q_i).$$

On the finite set $\{1, \ldots, N\}$ this kernel admits the explicit feature map

$$\phi_d(n) = \sqrt{g(d)}\,\mathbf{1}_{\{d \preceq n\}},$$

where $d \preceq n$ means coordinatewise comparison of the sorted prime-factor lists, and the coefficients $g(d)$ are determined recursively by

$$n = \sum_{d \preceq n} g(d).$$

Hence

$$K(a, b) = \sum_d \phi_d(a)\phi_d(b).$$

The figure attached to an integer n is then

$$P_n = \sum_d \phi_d(n) B_d.$$

Here the symbols B_d denote the canonical figure glyphs in the abstract figure basis introduced later. Because the coefficients are nonnegative in this example, the arithmetic figures are built by adding regular polygons without any reflections.

6 Why prime-number figures become visible

This representation is especially effective for primes. A prime p activates only a sparse collection of modes, so its figure has a small number of dominant polygonal components. In contrast, highly composite integers accumulate many active modes and therefore appear as richer superpositions. In the images below, the signatures of $2, 3, 5, 7, 11, 13, \ldots$ are immediately recognizable: the segment, triangle, pentagon, heptagon, and higher regular polygons remain visible inside composite numbers that depend on them.

7 Summary of the construction

The main point can be stated in one sentence:

> A high-dimensional vector is visualized by replacing its scalar coordinates with weighted copies of canonical planar figures.

Regular polygons are particularly attractive because they are easy to index, easy to recognize, and form a smooth conceptual bridge from a point to a segment to triangles, squares, and eventually circle-like modes.

8 Overview plates

Figure 1 shows the first 30 basis figures. Figure 4 shows the corresponding arithmetic figures for the meet kernel on $\{1, \dots, 30\}$.

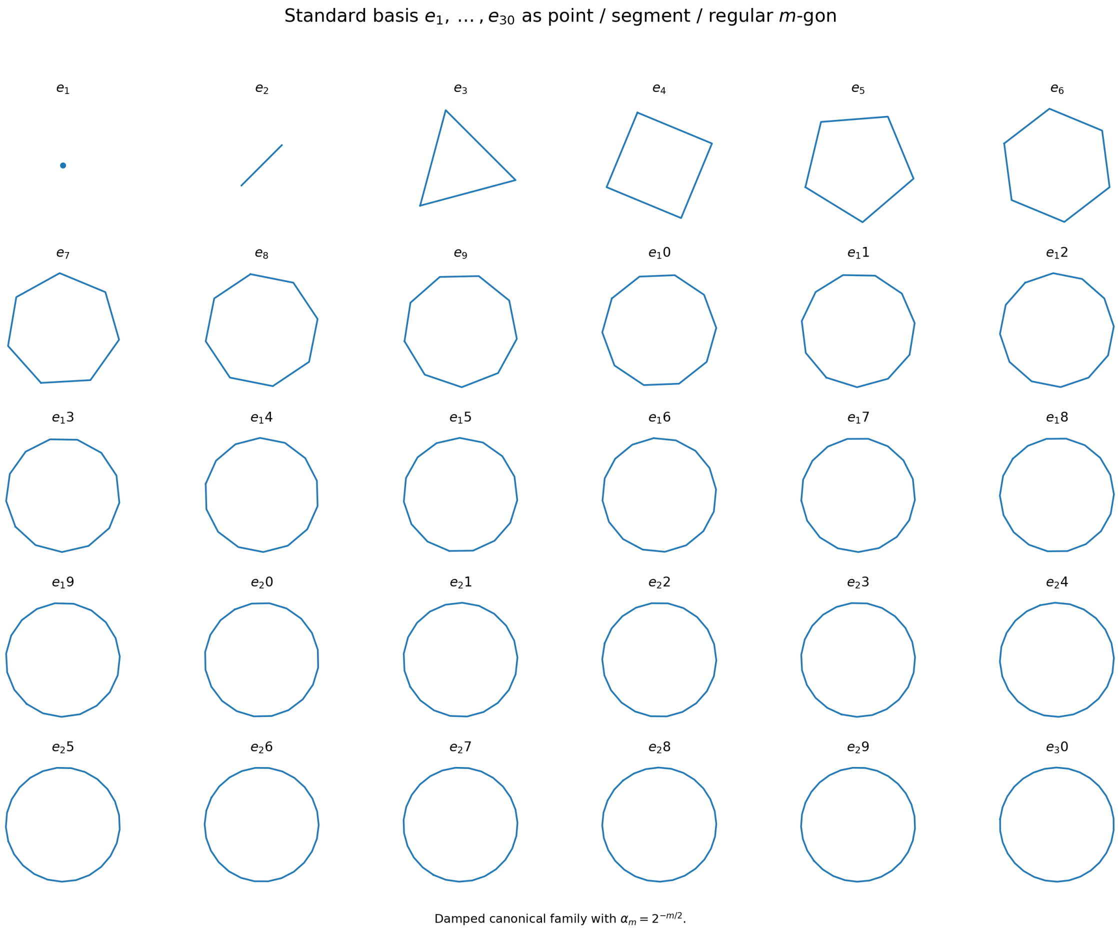


Figure 1: The canonical figures for the first 30 basis vectors: point, segment, triangle, square, and then regular m-gons. In the rendered plate, the visible glyphs are drawn using a fixed-radius plotting convention for readability.

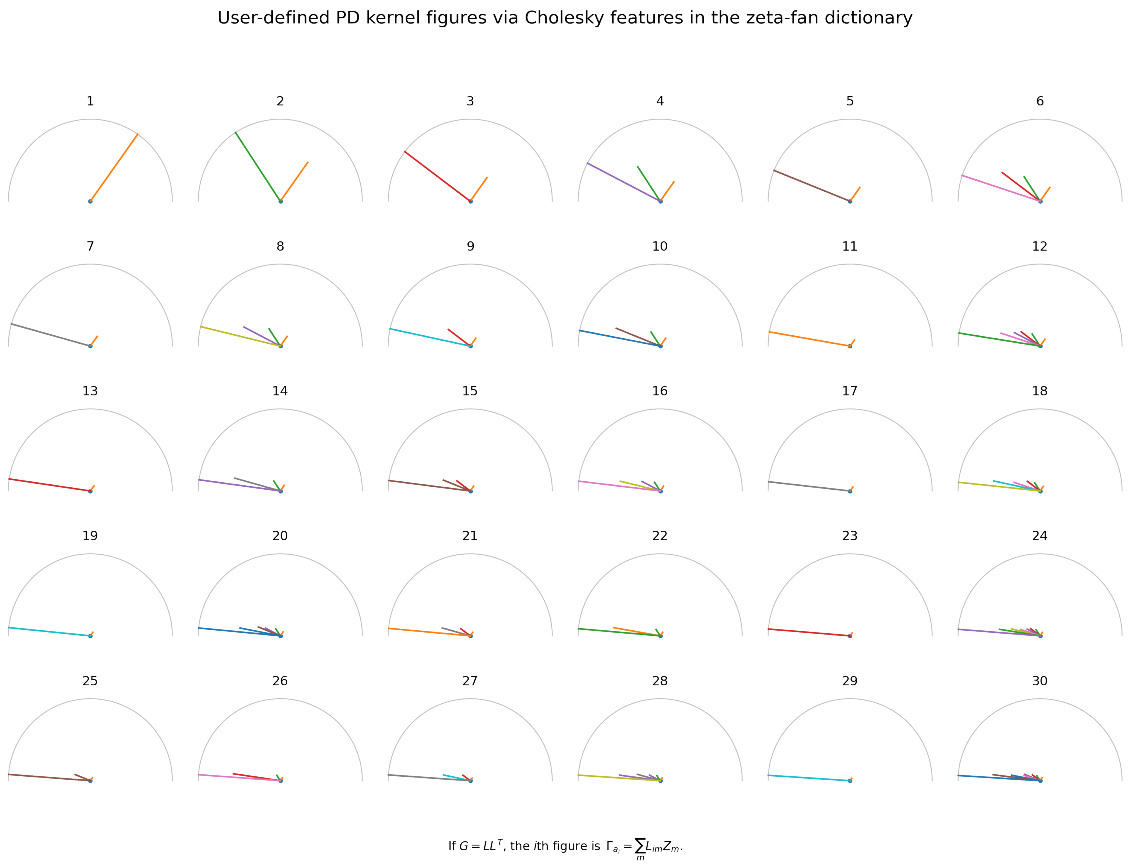


Figure 2: The gcd-kernel figures for the integers $1, \ldots, 30$, using the explicit arithmetic feature map and the zeta-fan.

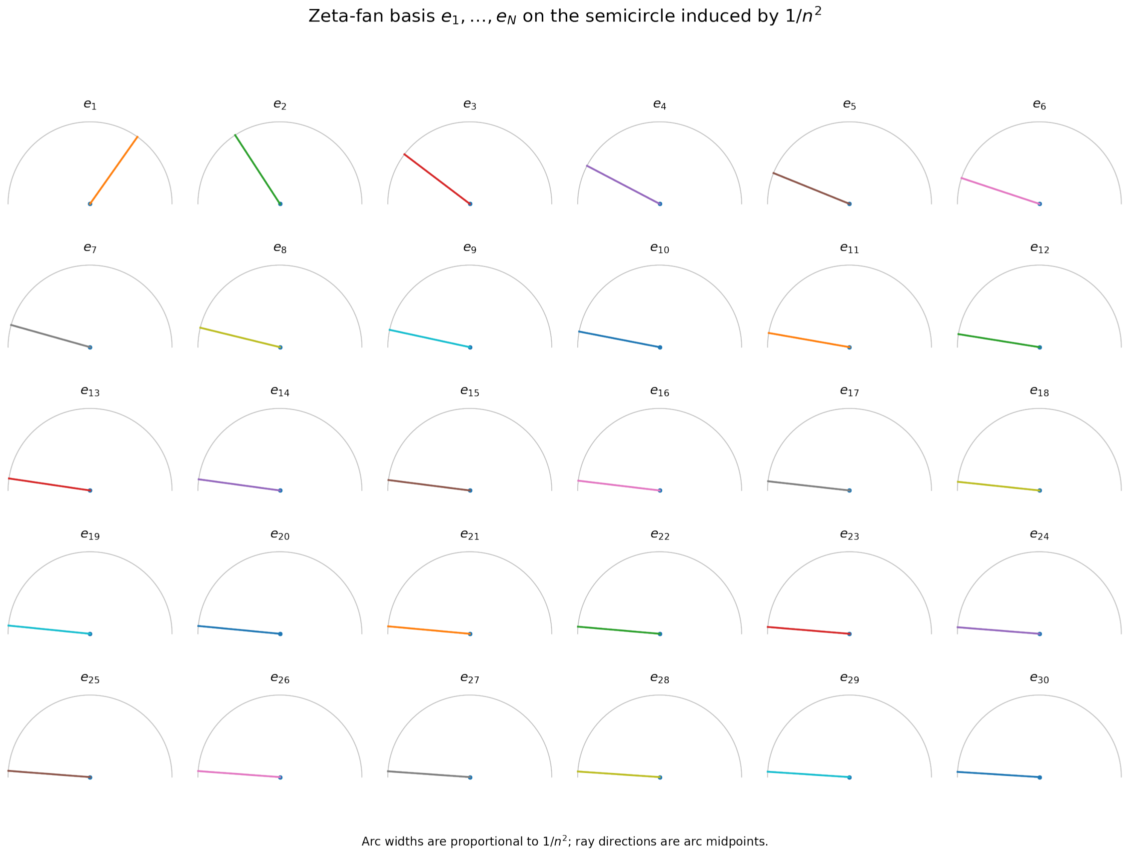


Figure 3: The canonical figures for the first 30 basis vectors with zeta-fan as described above.

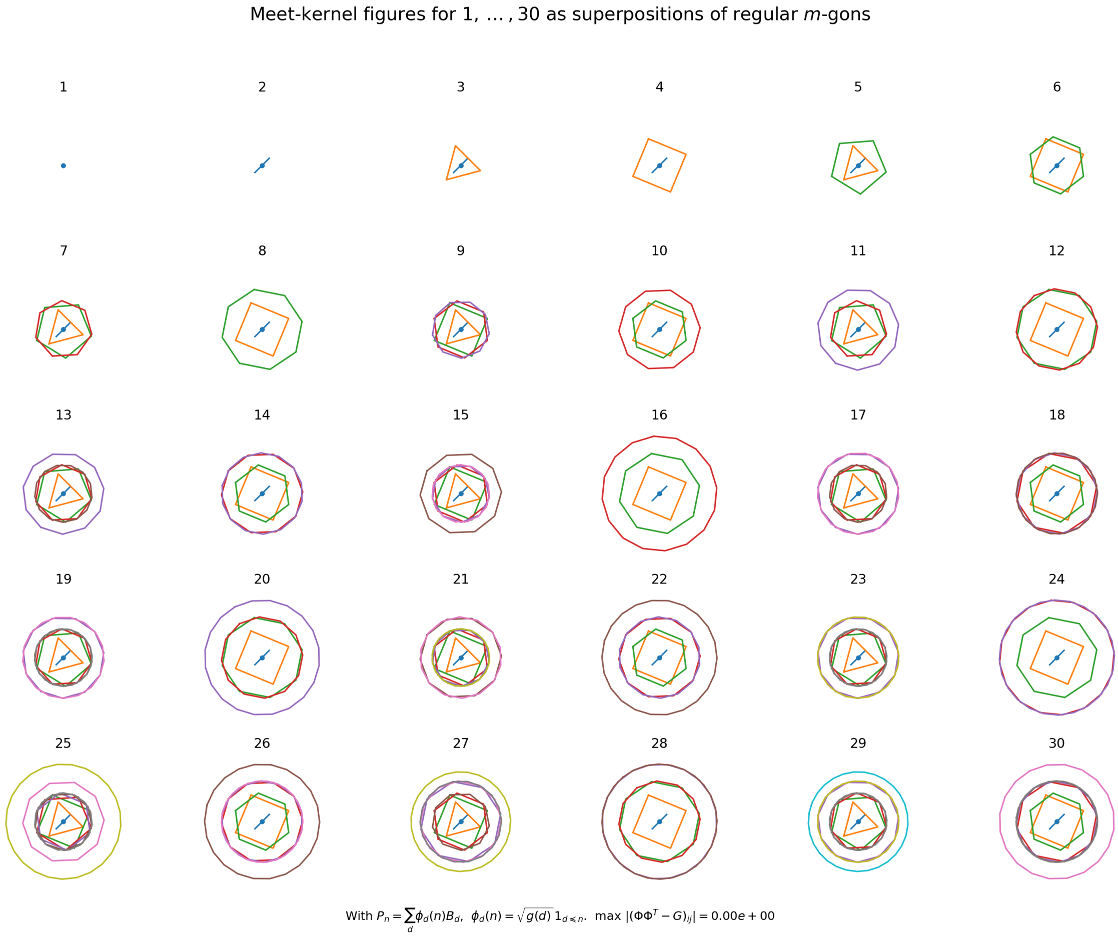


Figure 4: The meet-kernel figures for the integers $1, \ldots, 30$, using the explicit arithmetic feature map and the regular-polygon dictionary.

9 Preservation of the Inner Product

The exact preservation of inner products is most cleanly formulated in an abstract Hilbert space whose basis vectors are indexed by the canonical figure symbols.

9.1 An abstract Hilbert space of figures determined by the basis symbols

A key point of the present construction is that, for the purpose of preserving inner products, the concrete geometric realization of the basic figures is secondary. What matters first is that the family of symbols

$$B_1, B_2, B_3, \ldots$$

is canonically indexed and that each B_i encodes the integer i unambiguously.

Accordingly, we define the Hilbert space of figures abstractly, without starting from a function space such as $L^2([0,1]; \mathbb{R}^2)$.

Step 1: The abstract figure space. Let $\mathcal{F}_{\text{fin}}$ be the real vector space of all finite linear combinations of the formal figure symbols $B_1, B_2, \ldots$:

$$\mathcal{F}_{\text{fin}} := \left\{ \sum_{m=1}^{M} x_m B_m \ : \ M \in \mathbb{N},\ x_m \in \mathbb{R} \right\}.$$

Here the B_m are regarded as formal basis elements. At this stage, they need not be functions or curves; they are simply the canonical figure labels:

$$B_1 = \text{point}, \qquad B_2 = \text{segment}, \qquad B_3 = \text{equilateral triangle}, \qquad B_4 = \text{square},$$

and in general B_m is the regular m-gon symbol.

Step 2: Inner product by declaration on the basis. We define an inner product on $\mathcal{F}_{\text{fin}}$ by declaring the basis $(B_m)_{m\geq 1}$ to be orthonormal:

$$\langle B_i, B_j \rangle_{\mathcal{F}} := \delta_{ij},$$

and then extending bilinearly. Thus, for

$$X = \sum_m x_m B_m, \qquad Y = \sum_m y_m B_m,$$

where only finitely many coefficients are nonzero, we set

$$\left\langle \sum_m x_m B_m, \sum_m y_m B_m \right\rangle_{\mathcal{F}} = \sum_{m=1}^{\infty} x_m y_m.$$

It follows immediately that

$$\left\| \sum_m x_m B_m \right\|_{\mathcal{F}}^2 = \sum_m x_m^2.$$

Step 3: Completion. The completion of $\mathcal{F}_{\text{fin}}$ with respect to this norm is a Hilbert space, denoted by

$$\mathcal{F}.$$

By construction, $\mathcal{F}$ is canonically isometric to ℓ^2, with the correspondence

$$(x_1, x_2, \dots) \longleftrightarrow \sum_{m=1}^{\infty} x_m B_m.$$

Thus the family $(B_m)_{m\geq 1}$ is an orthonormal basis of $\mathcal{F}$.

Why this is enough. From this viewpoint, the role of the visible figures is symbolic and indexing-theoretic: each B_m is a uniquely recognizable geometric token for the integer m. The exact preservation of inner products is then built into the abstract Hilbert-space structure itself, namely through the declaration

$$\langle B_i, B_j \rangle_{\mathcal{F}} = \delta_{ij}.$$

In particular, one does not need the visible polygons themselves to be orthogonal as planar curves in L^2 or in any other concrete function space.

9.2 The figure map

Given $x = (x_m)_{m\geq 1} \in \ell^2$, define

$$T(x) := \sum_{m=1}^{\infty} x_m B_m \in \mathcal{F}.$$

Since (B_m) is an orthonormal basis of $\mathcal{F}$, the map

$$T : \ell^2 \to \mathcal{F}$$

is an isometric isomorphism, and for all $x, y \in \ell^2$,

$$\langle T(x), T(y) \rangle_{\mathcal{F}} = \sum_{m=1}^{\infty} x_m y_m = \langle x, y \rangle_{\ell^2}.$$

9.3 Negative coefficients

If a coefficient is negative, then

$$x_m B_m = |x_m|(-B_m),$$

so the sign may be visualized by replacing the figure B_m with a reflected or otherwise sign-marked version. This visual convention does not affect the Hilbert-space formalism, since the linear structure is already fixed abstractly.

9.4 Application to the meet-kernel construction

Suppose now that a family of feature vectors is given by

$$\Phi(a) = (\phi_d(a))_{d \geq 1} \in \ell^2,$$

and that the kernel satisfies

$$K(a, b) = \sum_{d \geq 1} \phi_d(a)\phi_d(b) = \langle \Phi(a), \Phi(b) \rangle_{\ell^2}.$$

Defining the figure associated with a by

$$\Gamma_a := \sum_{d \geq 1} \phi_d(a) B_d \in \mathcal{F},$$

we obtain

$$\langle \Gamma_a, \Gamma_b \rangle_{\mathcal{F}} = \sum_{d \geq 1} \phi_d(a)\phi_d(b) = K(a, b).$$

Hence the figure representation preserves the kernel geometry exactly.

In particular, for the meet-kernel feature map

$$\phi_d(n) = \sqrt{g(d)}\, \mathbf{1}_{\{d \preceq n\}},$$

we obtain

$$\langle \Gamma_m, \Gamma_n \rangle_{\mathcal{F}} = K(m, n).$$

Therefore the passage from high-dimensional points to figures is not merely illustrative: it is an isometric realization of the original inner-product geometry.

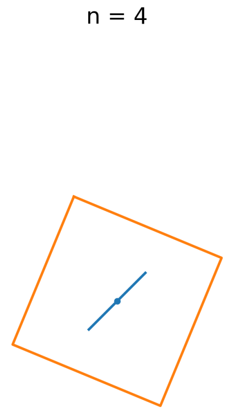
Appendix: Detail views for $n = 1, \ldots, 30$

The next pages display the arithmetic figures four at a time, using a common plotting window so that size and complexity can be compared directly.

$n = 1$

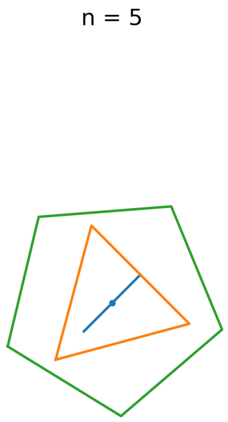
$n = 2$

$n = 3$

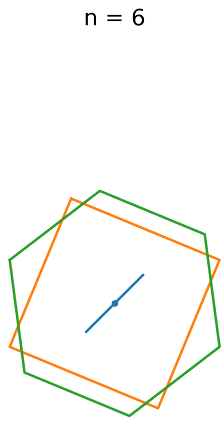


$n = 4$

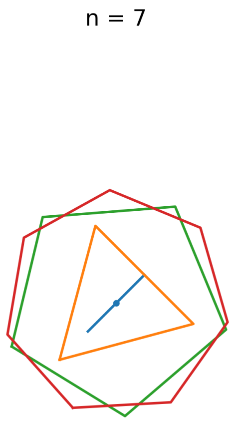
Detail views for $n = 1, \ldots, 4$.



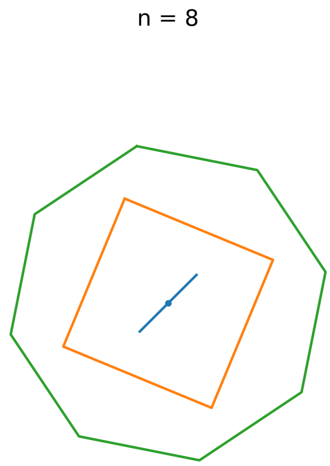
$n = 5$



$n = 6$

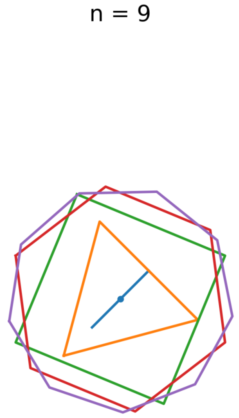


$n = 7$

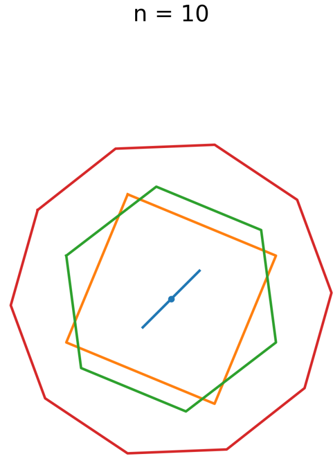


$n = 8$

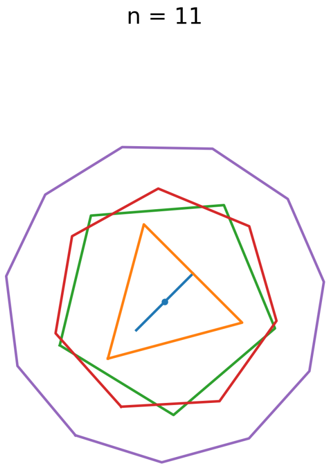
Detail views for $n = 5, \ldots, 8$.



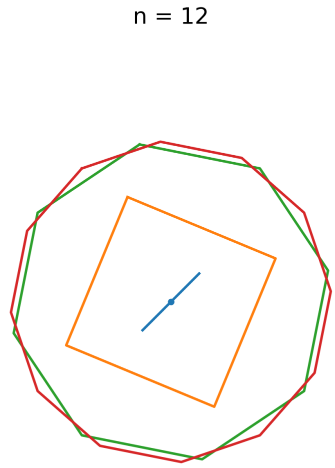
$n = 9$



$n = 10$

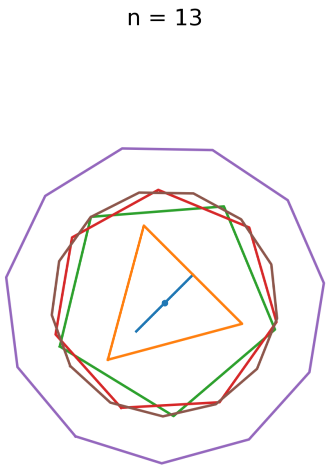


$n = 11$

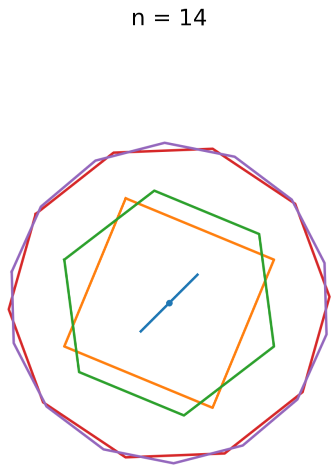


$n = 12$

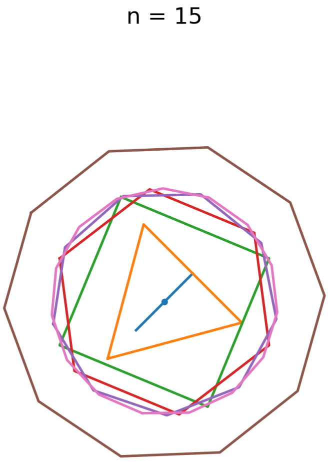
Detail views for $n = 9, \ldots, 12$.



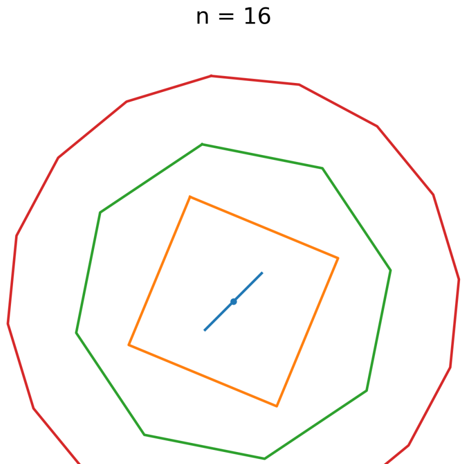
$n = 13$



$n = 14$

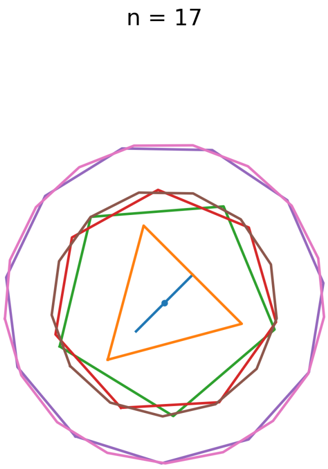


$n = 15$

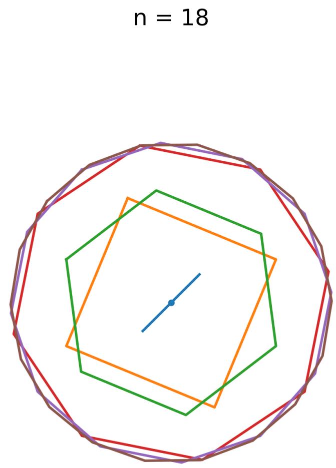


$n = 16$

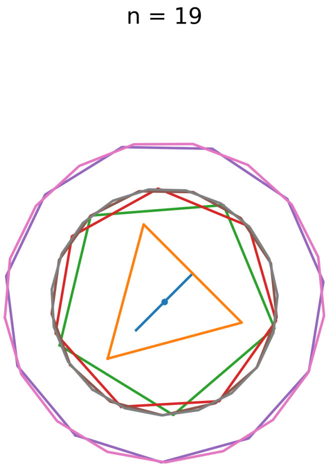
Detail views for $n = 13, \ldots, 16$.



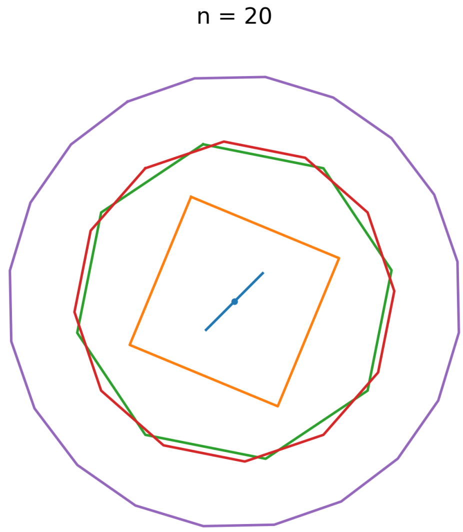
$n = 17$



$n = 18$

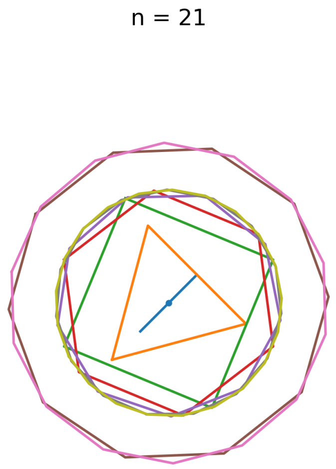


$n = 19$

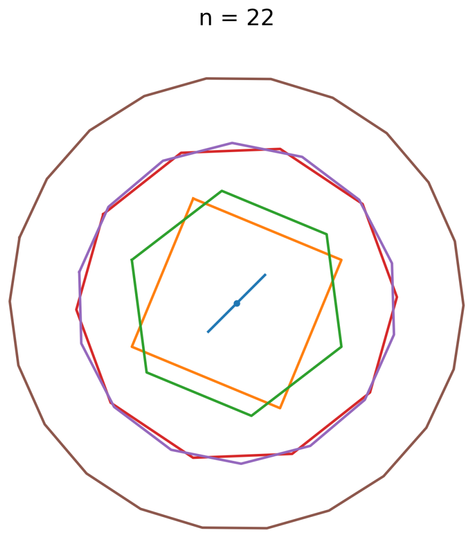


$n = 20$

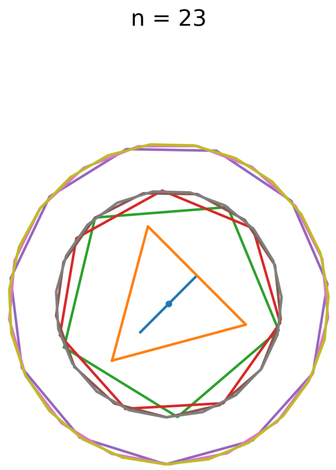
Detail views for $n = 17, \dots, 20$.



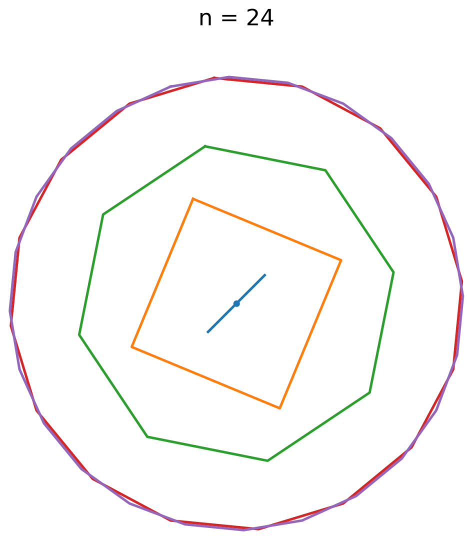
$n = 21$



$n = 22$

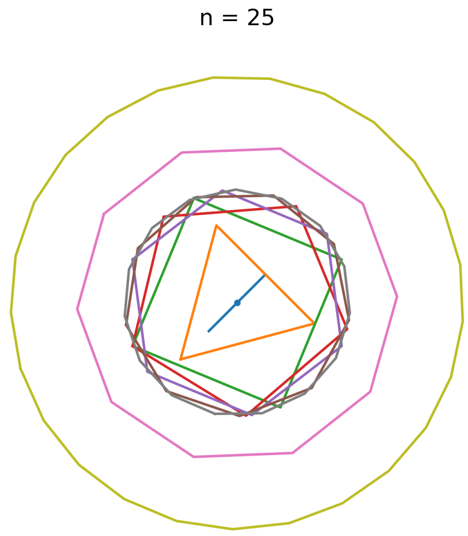


$n = 23$

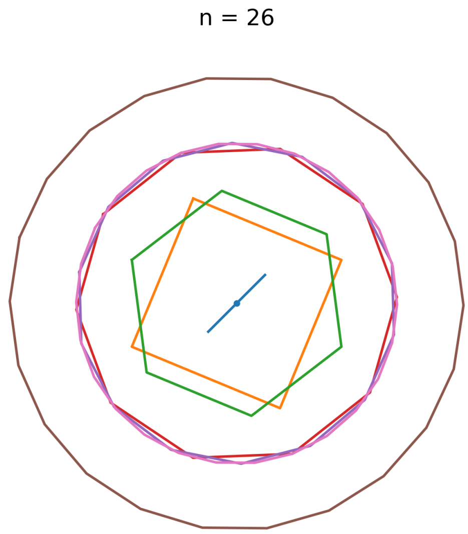


$n = 24$

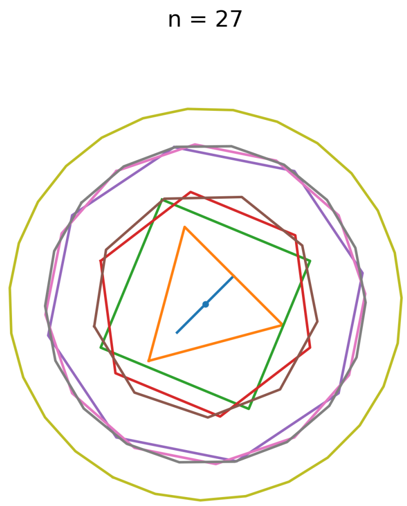
Detail views for $n = 21, \ldots, 24$.



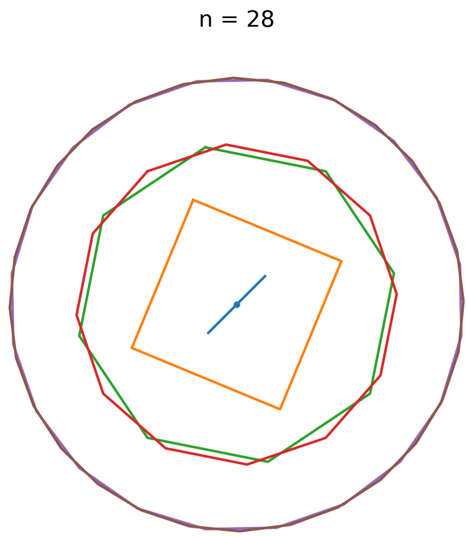
$n = 25$



$n = 26$

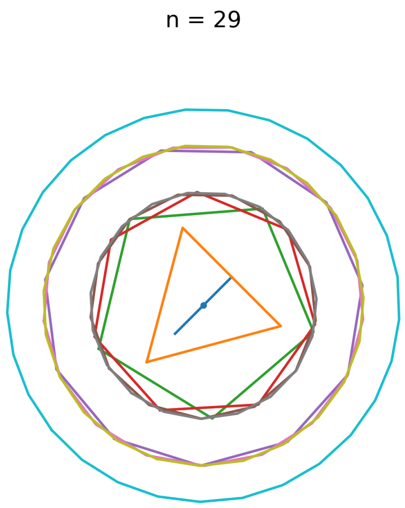


$n = 27$

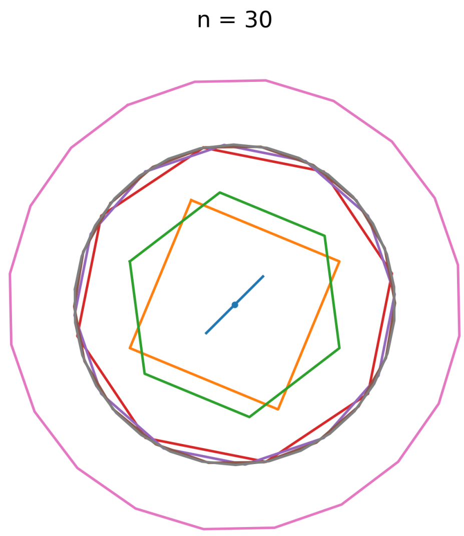


$n = 28$

Detail views for $n = 25, \ldots, 28$.



$n = 29$



$n = 30$

Detail views for $n = 29, \ldots, 30$.