hexagon:2611.00302v1

TRIANGLES IN SETS OF POSITIVE UPPER DENSITY IN THE PLANE

GENERATED BY GPT-6 ASTRA PRO

Abstract. This note proves that every measurable subset of the plane with positive upper Banach density contains a translate of a rotation of every sufficiently large dilate of any fixed noncollinear triangle.

1. Triangles and the counting estimate

The upper Banach density of a Lebesgue measurable set $E \subset \mathbb{R}^2$ is

$$d^*(E) = \limsup_{L\to\infty} \sup_{z\in\mathbb{R}^2} \frac{|E \cap (z + [0,L]^2)|}{L^2}.$$

Theorem 1.1. *Let $\Delta \subset \mathbb{R}^2$ consist of three noncollinear points, and let $E \subset \mathbb{R}^2$ be measurable with $d^*(E) > 0$. There exists $\lambda_0 > 0$ such that for every $\lambda \geq \lambda_0$ the set E contains a translate of a rotation of $\lambda\Delta$.*

Furstenberg, Katznelson, and Weiss [8, Theorem B] proved that, for all sufficiently large λ, every positive-radius neighborhood of E contains a congruent copy of $\lambda\Delta$. Ziegler [10] extended this neighborhood result to arbitrary finite configurations. Bourgain's theorem for nondegenerate simplices [1] gives exact copies of large triangles in ambient dimension at least three. Falconer, Kovač, and Yavicoli [6, Theorem 1.4] obtained the planar conclusion when $d^*(E) > 1/2$. Castro-Silva [2, Conjecture 1] conjectured the conclusion up to congruence for sets of positive upper density in squares centered at the origin. Theorem 1.1 gives orientation-preserving copies under the upper Banach density hypothesis.

The proof combines the density Hales–Jewett theorem with an operator-norm continuity estimate for the triangle-counting form, derived from the smoothing theorem (Theorem 1.2). If there were an unbounded sequence of missed scales, sufficiently many of them could be chosen far apart. The density Hales–Jewett theorem would then give a positive lower bound for a weighted sum of triangle counts at their nonempty subsums. Each subsum is close to its largest constituent scale relative to that scale, so continuity in operator norm makes the corresponding count close to zero, contradicting the lower bound.

Write $J(x_1, x_2) = (-x_2, x_1)$. After translation, rotation, and a fixed change of scale, a noncollinear triangle has vertex set $\Delta_M = \{0, e_1, Me_1\}$ for $M = \alpha I + \beta J$ with $\beta \neq 0$. Since M commutes with rotations, a copy at scale s has vertices $x, x + su_\theta, x + sMu_\theta$, where $u_\theta = (\cos\theta, \sin\theta)$, with $\theta \in \mathbb{R}/(2\pi\mathbb{Z})$. Put $v_\theta = Ju_\theta$ and let σ be the probability measure $\mathrm{d}\theta/(2\pi)$ on the unit circle. Integrable densities are identified with their measures. For a finite measure ν on $\mathbb{R}^2$ and $s > 0$, let ν_s denote its pushforward by $u \mapsto su$, and put

$$\Lambda_\nu(f_0, f_1, f_2) = \iint f_0(x) f_1(x+u) f_2(x+Mu)\, \mathrm{d}\nu(u)\, \mathrm{d}x. \tag{1.1}$$

Denote its $L^2 \times L^2 \times L^\infty$ norm by $\|\Lambda_\nu\|$. Cauchy–Schwarz gives $\|\Lambda_\nu\| \leq \|\nu\|_{\mathrm{TV}}$, and a change of variables gives $\|\Lambda_{\nu_s}\| = \|\Lambda_\nu\|$.

Date: November 6, 2026.

Theorem 1.2. *For every probability density ρ on $\mathbb{R}^2$,*

$$\limsup_{\varepsilon\downarrow 0}\sup_{s>0}\|\Lambda_{(\sigma*\rho_\varepsilon-\sigma)_s}\| = 0, \qquad \rho_\varepsilon(u)=\varepsilon^{-2}\rho(u/\varepsilon).$$

This includes the Gaussian $\rho(u)=\exp(-\pi|u|^2)$.

Let ξ_j denote the frequency at vertex j. The frequency dual to the displacement u in (1.1) is $b=\xi_1+M^T\xi_2$. Stationary equations play a similar role in the trilinear estimates of Christ [3, 4] and Christ, Durcik, and Roos [5]. Here their exact form forces $b=0$ almost everywhere on a set where measurable frequency fields satisfy them (Section 2). Next, consider translations of the inputs on the scale R^{-1}. Only stationary frequency triples contribute to the limiting correlations, which gives cancellation on $|b|\asymp R$ (Section 3). Finally, after fixed linear changes of coordinates, the three vertex frequencies are pairwise separated by distances comparable to $|b|$. A telescoping estimate then sums the annuli without a rate of decay (Section 4).

1.1. **From smoothing to triangles.** Theorem 1.2 gives

$$\lim_{r\to 1}\|\Lambda_{\sigma_r-\sigma}\| = 0. \tag{1.2}$$

Indeed, for fixed $t>0$ the norm is at most $2\|\Lambda_{\sigma*\rho_t-\sigma}\|+\|(\sigma*\rho_t)_r-\sigma*\rho_t\|_1$. The last term tends to zero by continuity of dilation in L^1. Let $t\downarrow 0$.

Proof of Theorem 1.1. It suffices to consider Δ_M. Write $D_s(f)=\Lambda_{\sigma_s}(f,f,f)$ and fix $0<\delta<d^*(E)$. The density Hales–Jewett theorem [7] gives n such that every subset of $\{0,1,2\}^n$ of density at least $\delta/4$ contains a combinatorial line. There are 4^n-3^n lines, each with a nonempty set J of variable coordinates. The words of a line agree off J and take the values $0,1,2$, respectively, throughout J.

Suppose the missed scales are unbounded. Choose missed scales $\lambda_1<\cdots<\lambda_n$ so rapidly increasing that for every nonempty J, with $s_J=\sum_{j\in J}\lambda_j$ and $k=\max J$,

$$\|\Lambda_{\sigma_{s_J}-\sigma_{\lambda_k}}\| = \|\Lambda_{\sigma_{s_J/\lambda_k}-\sigma}\| < \delta/(4\cdot 4^n). \tag{1.3}$$

This is possible by (1.2), since the relative error is at most $(\lambda_1+\cdots+\lambda_{k-1})/\lambda_k$. Put $S=\max(1,\|M\|)\sum_{j=1}^n\lambda_j$. Choose a square Q with $|E\cap Q|\ge\delta|Q|$ and with side so large that $Q'=Q+[-S,S]^2$ satisfies $|Q'|\le 2|Q|$. Take $f=\mathbf{1}_{E\cap Q}$, replacing $E\cap Q$ by a Borel subset of the same measure if necessary.

For fixed θ put $d_0=0$, $d_1=u_\theta$, $d_2=Mu_\theta$. At base point x mark every word w satisfying $f(x+\sum_j\lambda_j d_{w_j})=1$. The marked fraction integrates to $|E\cap Q|$ over Q', by translation in x. Consequently it is at least $\delta/4$ on a set of area at least $\delta|Q|/2$. At every such x at least one whole line is marked. For a line with variable coordinates J, translating x by the contribution of its fixed letters yields the three shifts $0, s_Ju_\theta, s_JMu_\theta$. Integrating over all x and then over θ therefore gives

$$\sum_{\varnothing\ne J\subset\{1,\dots,n\}} 3^{n-|J|}D_{s_J}(f)\ge\delta|Q|/2.$$

But $D_{\lambda_k}(f)=0$ at every chosen scale, so (1.3) bounds every term $D_{s_J}(f)$ by $\delta|Q|/(4\cdot 4^n)$. Summing over the lines gives a contradiction. □

2. Measurable stationarity

Write $p=(x,\theta)$, $\mathrm{d}\mu(p)=\mathrm{d}x\,\mathrm{d}\sigma(\theta)$, $L_0=0$, $L_1=I$, $L_2=M$, and $\pi_jp=x+L_ju_\theta$. For constants $a_j=\xi_j$, differentiating the phase $\sum_j\xi_j\cdot\pi_jp$ in x and θ gives its stationary equations. The same equations for measurable planar vector fields a_j are

$$\sum_{j=0}^{2}a_j(\pi_jp)=0, \qquad v_\theta\cdot(a_1(\pi_1p)+M^Ta_2(\pi_2p))=0. \tag{2.1}$$

The first equation is applied separately to each scalar coordinate. Its finite-dimensional kernel permits passage to analytic fields.

2.1. **Circle operators.** Use the Fourier convention $\widehat{f}(\xi) = \int f(x)\mathrm{e}(-x\cdot\xi)\,\mathrm{d}x$, where $\mathrm{e}(t) = \exp(2\pi i t)$, and write $\check{f}$ for the inverse Fourier transform. For functions on $\mathbb{R}^2$, unadorned L^p norms are with respect to Lebesgue measure.

Lemma 2.1. *Let N be an invertible linear similarity of $\mathbb{R}^2$ and W a measurable weight with $|W| \le 1$. For*

$$Tf(z) = \int W(z,\theta) f(z + Nu_\theta)\,\mathrm{d}\sigma(\theta)$$

there is a nondecreasing function $\omega_N(s) \to 0$ as $s \downarrow 0$ such that $\|\sqrt{e}\,T\|_{2\to 2} \le \omega_N(\|e\|_1)$ whenever $e \in L^1$ and $0 \le e \le 1$. If W vanishes for z outside a bounded set, then T is compact on L^2.

Proof. The difference of two independent uniform points on the circle of radius $|Ne_1|$ has a density $w \in L^1$, since the map $(\theta,\varphi) \mapsto N(u_\theta - u_\varphi)$ has nonzero Jacobian except when the two tangents are parallel. Expanding TT^* shows that its kernel is bounded in absolute value by $w(y-z)$. Where $w \le H$, the kernel of $\sqrt{e}\,TT^*\sqrt{e}$ is dominated by the rank-one kernel $H\sqrt{e(z)e(y)}$. The remainder has norm at most $\|w\mathbf{1}_{\{w>H\}}\|_1$ by Young's inequality. Hence

$$\|\sqrt{e}\,T\|_{2\to 2}^2 \le H\|e\|_1 + \|w\mathbf{1}_{\{w>H\}}\|_1,$$

which gives the claimed modulus by first choosing H large. For bounded z-support of W, the part of the TT^* kernel with $w \le H$ is Hilbert–Schmidt. Thus TT^*, and hence T, is compact. □

When M is a rotation, Greenleaf, Iosevich, Krause, and Liu [9] proved a restricted strong-type estimate at the $L^2 \times L^2 \to L^2$ endpoint for the corresponding bilinear average.

Lemma 2.2. *The triangle form satisfies*

$$|\Lambda_\sigma(f_0,f_1,f_2)| \le C_M \prod_{j=0}^{2} \|f_j\|_2. \tag{2.2}$$

On any permutation of $L^2 \times L^2 \times L^\infty$, its norm is at most $\omega_M(s) \to 0$ when one of the L^2 inputs is supported on a set of area at most s. For every finite measure κ, both estimates hold for $\sigma * \kappa$ with an extra factor $\|\kappa\|_{\mathrm{TV}}$.

Proof. For (2.2), partition the circle into small arcs I. Cauchy–Schwarz in the angle gives

$$\begin{aligned}\int \Big|\int_I f_1(x+u_\theta) f_2(x+Mu_\theta)\,\mathrm{d}\sigma(\theta)\Big|^2\,\mathrm{d}x \\ \le \int_{I^2}\int |f_1(y)|^2 |f_2(y+Mu_\varphi-u_\theta)|^2\,\mathrm{d}y\,\mathrm{d}\sigma(\theta)\,\mathrm{d}\sigma(\varphi).\end{aligned}$$

On the diagonal the absolute Jacobian of $(\theta,\varphi) \mapsto Mu_\varphi - u_\theta$ is $|\beta|$. Choose the arcs so that on each I^2 this map is one-to-one with Jacobian bounded below. Changing variables bounds the last integral by $C_M\|f_1\|_2^2\|f_2\|_2^2$. Sum over I and pair with f_0. For the second assertion, bound one input by its supremum. On the opposite edge use Lemma 2.1 with $W = 1$ and e the indicator of the small support. Take ω_M to be the maximum of the moduli for the finitely many edge maps $L_j - L_i$. Convolution by κ integrates translations of f_1, f_2 by w, Mw against $\mathrm{d}\kappa(w)$. □

2.2. **The scalar kernel.** For a set E of configurations with vertices in bounded measurable sets X_j, the degree at $z \in X_j$ is

$$d_{j,E}(z) = \int \mathbf{1}_E(z - L_j u_\theta, \theta)\, \mathrm{d}\sigma(\theta).$$

A *core of degree* c has $d_{j,E} \geq c > 0$ almost everywhere on each of its vertex sets. Every E of positive measure contains a positive-measure core. Indeed, fix $0 < c < \mu(E)/\sum_j |X_j|$. At each step simultaneously delete all vertices of current degree less than c and their incident configurations. The total loss is at most $c\sum_j |X_j|$ since each vertex is deleted once. On the decreasing intersection all remaining degrees are at least c by monotone convergence.

On a core E set $\mathcal{H}_E = \bigoplus_j L^2(X_j; \mathbb{R})$. Define $D_E : \mathcal{H}_E \to L^2(E, \mu)$ by $D_E f = \sum_j f_j \circ \pi_j$, and put $K_E = \ker D_E$.

Lemma 2.3. *The space K_E is finite dimensional and $\operatorname{dist}(f, K_E) \leq C_E \|D_E f\|_{L^2(E,\mu)}$. For each fixed $t > 0$, if $F \subset E$ is a core of degree t on vertex sets $Y_j \subset X_j$ and $\mu(E \setminus F)$ is sufficiently small depending on E, t, restriction is an isomorphism from K_E onto K_F.*

Proof. The diagonal blocks of $D_E^* D_E$ are multiplication by $d_{j,E} \geq c$. Each off-diagonal block is a circle operator with weight bounded by one and shift $(L_i - L_j)u_\theta$. These differences are nonzero similarities, so the blocks are compact by Lemma 2.1. Conjugating by the inverse square root of the diagonal gives the identity plus a compact self-adjoint operator. Its positive spectrum is bounded away from zero. Conjugating back gives finite dimensionality and the distance bound.

If f belongs to the kernel on a core of degree t, the block equations and Lemma 2.1 give

$$\|\sqrt{e}\, f_j\|_2 \leq C_M t^{-1} \omega_M(\|e\|_1) \|f\|_2 \qquad (0 \leq e \leq 1).$$

Indeed, the jth block equation expresses the degree times f_j as a sum of two weighted circle operators applied to the other components, up to sign. With $\eta = \mu(E \setminus F)$, each omitted vertex loses all its incidences, so $|X_j \setminus Y_j| \leq \eta/c$. The displayed estimate with $e = \mathbf{1}_{X_j \setminus Y_j}$ makes restriction injective on K_E for small η. Conversely extend $f \in K_F$ by zero. The degrees e_j of $E \setminus F$ are bounded by one and have integral η. Consequently

$$\|D_E f\|_2^2 \leq 3 \sum_j \|\sqrt{e_j}\, f_j\|_2^2 \leq C_M t^{-2} \omega_M(\eta)^2 \|f\|_2^2.$$

The distance bound makes projection onto K_E injective on this copy of K_F for small η. The two injections give equality of dimensions and hence surjectivity of restriction. □

2.3. **Rigidity of the stationary equations.**

Lemma 2.4. *Every positive-measure core E has a positive-measure subset where each component of a basis of K_E agrees, at its vertex, with a real analytic function defined on all of $\mathbb{R}^2$.*

Proof. Use an orthonormal basis, write $\Phi_j(x)$ for the row of its values at vertex j, and extend by zero outside X_j. Its three Gram matrices $\int \Phi_j^T \Phi_j$ sum to the identity. Let E_h be E translated by $-h$ in the base variable, with basis $\Phi_j(\cdot + h)$. Then $\mu(E \triangle E_h) \to 0$.

There is a common core $F_h \subset E \cap E_h$ of degree $c/4$ with $\mu(E \setminus F_h) \to 0$. Indeed, start by deleting $\eta = \mu(E \setminus E_h)$ incidences from E, then prune at $c/4$. If V is the sum of the measures of the pruned vertex sets, at most $cV/4$ further incidences are lost. Their original degrees give $cV \leq 3(\eta + cV/4)$, so the total loss is at most 4η. Since $\mu(E_h) = \mu(E)$ the same bound holds for E_h. Lemma 2.3 applies to both cores for small h, with the same threshold by translation. The surviving vertex sets $Y_{j,h}$ lie in $X_j \cap (X_j - h)$ up to null sets. Since the two bases restrict to bases of K_{F_h}, there is a unique invertible matrix $C(h)$ with

$$\Phi_j(x + h) = \Phi_j(x) C(h), \qquad |X_j \setminus Y_{j,h}| \to 0. \tag{2.3}$$

By absolute continuity the Gram matrices on $Y_{j,h}$ sum to $I + o(1)$. Multiply (2.3) by Φ_j^T, integrate, and use L^2 translation continuity to get $C(h) \to I$. To compose (2.3) at $h, k, h+k$, restrict to $Y_{j,h} \cap (Y_{j,k} - h) \cap Y_{j,h+k}$. The complements in X_j again tend to zero in measure, since $|X_j \triangle (X_j - h)| \to 0$. The summed Gram matrix on the intersections is invertible for small h, k, so $C(h+k) = C(h)C(k)$. By the continuous one-parameter group theorem on the coordinate axes, $C(h) = P(h) := \exp(h_1 T_1 + h_2 T_2)$ near zero for commuting real matrices T_1, T_2.

Write $z_j = \pi_j p$. For every triple of sufficiently small h_j, the equalities $\Phi_j(z_j) = \Phi_j(z_j - h_j)P(h_j)$ hold simultaneously on a subset of E of measure at least $\mu(E)/2$. Indeed, the loss is at most $\sum_j |X_j \setminus (Y_{j,h_j} + h_j)| = o(1)$ since the vertex degrees are at most one. The equality event is jointly measurable in p, h_0, h_1, h_2 by taking Borel versions of the Φ_j. Integrate over the three small shifts, then change variables $b_j = z_j - h_j$ at fixed p. Fubini gives fixed b_0, b_1, b_2 and a positive-measure subset of E where $\Phi_j(z_j) = \Phi_j(b_j)P(z_j - b_j)$ for all j. These expressions are real analytic functions of their vertex variables. □

Lemma 2.5. *On a core E of positive degree, every square-integrable measurable solution of (2.1) satisfies $a_1(\pi_1 p) + M^T a_2(\pi_2 p) = 0$ for almost every $p \in E$.*

Proof. If the asserted equality fails on a set of positive measure, prune that set to a positive-measure core and use this new core in what follows. The coordinates of the vector triple belong to its scalar kernel. On the positive-measure subset given by Lemma 2.4 the three fields therefore agree with global real analytic functions. An analytic function on a connected real analytic manifold that vanishes on a set of positive volume vanishes identically. Thus these functions satisfy (2.1) on all of $\mathbb{R}^2 \times S^1$.

Write $A = a_0(x)$, $B = a_1(x+u)$, $C = a_2(x+Mu)$ for these analytic fields, where $u = u_\theta$ and $v = Ju$. The equations give $B + M^T C = tu$ for a smooth scalar $t = t(x, \theta)$, and

$$(M^T - I)C = A + tu, \qquad -(M^T - I)B = M^T A + tu.$$

Put $K = Da_0(x)$, $X_1 = \partial_\theta - v \cdot \nabla_x$, and $X_2 = \partial_\theta - Mv \cdot \nabla_x$. Since $X_1 B = X_2 C = 0$,

$$(X_1 t)u + tv = M^T K v, \qquad (X_2 t)u + tv = KMv.$$

The v-component of the second identity gives $t = v^T KMv$, which is even under $(u, v) \mapsto (-u, -v)$. The spatial transport terms $(L_j v \cdot \nabla_x t)u$ are therefore even. Taking the odd part of both identities gives $\partial_\theta(tu) = M^T Kv = KMv$. Hence $M^T K = KM$, and integration in θ gives $tu = KMu + c(x)$. Both tu and KMu are odd, so $c = 0$. Thus KM preserves every line and is scalar. In particular K commutes with M. It follows that $(M^T - M)K = 0$. As $M^T - M = -2\beta J$ is invertible, $K = 0$ and $t = 0$. This contradicts the choice of the core. □

3. Cancellation on one frequency annulus

Proposition 3.1. *Let $\Psi \in L^1(\mathbb{R}^2)$ have integral zero and put $\Psi_R(u) = R^2 \Psi(Ru)$. For every bounded box Q and $C < \infty$,*

$$\sup_{|f_j| \le \mathbf{1}_Q,\ \mathrm{Lip}\, f_j \le CR} |\Lambda_{\sigma * \Psi_R}(f_0, f_1, f_2)| \longrightarrow 0 \qquad (R \to \infty).$$

Proof. Take any sequence $R_n \to \infty$ and admissible inputs $f_{j,n}$. Extract local correlations of the translates $f_{j,n}(z + s/R_n)$. All weak-* limits below are in L^∞ of the vertex variable z or the configuration variable $p = (x, \theta)$. First extract a common subsequence for the two-point products at each vertex and the three-point products on configuration space, with all shifts rational. Their uniform Lipschitz bounds in the shift variables give versions of the limits that extend continuously to all real shifts off a single null set in each space. The same bounds give weak-* convergence to these extensions at every real shift.

At each vertex, the weak-* limit of $f_{j,n}(z + s/R_n)\overline{f_{j,n}(z + t/R_n)}$ depends only on $s - t$, by translating z by t/R_n in a compactly supported test. The limits of squared absolute values of finite

linear combinations of translates show that for almost every z this continuous function of $s-t$ is positive definite. By Bochner's theorem it equals $\int \mathrm{e}((s-t)\cdot\xi)\,\mathrm{d}\nu_{j,z}(\xi)$ for a positive measure of mass at most one. The measures form measurable kernels, as follows by testing against smooth functions and Fourier inversion.

For $e_s(\xi)=\mathrm{e}(s\cdot\xi)$ and a trigonometric polynomial $a=\sum_s c_s e_s$, set $F_{j,n}[a](z)=\sum_s c_s f_{j,n}(z+s/R_n)$. Write $\nu_j=\nu_{j,\pi_j p}$. For $G=\sum_d b_d\otimes c_d$, with b_d, c_d trigonometric polynomials, put

$$G_n(p)=\sum_d F_{1,n}[b_d](\pi_1 p)F_{2,n}[c_d](\pi_2 p).$$

Denote the weak-* limit of $F_{0,n}[a](\pi_0 p)G_n(p)$ by $\ell_p(a,G)$. This defines a bilinear form satisfying

$$|\ell_p(a,G)|\le \|a\|_{L^2(\nu_0)}\|G\|_{L^2(\nu_1\otimes\nu_2)}. \tag{3.1}$$

Indeed, the weak-* limit of $|G_n|^2$ is the square of the second norm in (3.1). To see this, expand the square and test against a compactly supported function φ of p with $|\varphi|\le 1$. In the coordinates $z=x+u_\theta$, each term pairs a product of two translates at vertex 1 with such a product at vertex 2 through the operator of Lemma 2.1, with shift $(M-I)u_\theta$ and weight $\varphi(z-u_\theta,\theta)$. This operator is compact. The one-vertex correlation formula gives the respective weak L^2 limits on bounded vertex boxes. Likewise $|F_{0,n}[a](\pi_0 p)|^2$ has weak-* limit $\|a\|^2_{L^2(\nu_0)}$. Cauchy–Schwarz on shrinking configuration balls gives (3.1) by Lebesgue differentiation. Rational shifts and rational complex coefficients give a single exceptional null set. Continuity then extends the assertions to real shifts and all coefficients. Characters span a dense subspace of L^2 of each finite measure by uniqueness of Fourier transforms. Thus (3.1) extends ℓ_p to the two indicated L^2 spaces.

For $s\in\mathbb{R}^2$ and $h\in\mathbb{R}$, changing variables by $(x,\theta)\mapsto(x+s/R_n,\theta+h/R_n)$ in a continuous test gives

$$\ell_p(a,G)=\ell_p(e_s a,(e_{s+hv_\theta}\otimes e_{s+hMv_\theta})G). \tag{3.2}$$

Indeed the displacement at vertex j differs from $(s+hL_j v_\theta)/R_n$ by $O_M(h^2/R_n^2)$, so the input error tends to zero. The weak-* limit at these angle-dependent shifts follows by approximating them by simple functions, using the uniform Lipschitz bound. Rational parameters and continuity again give simultaneous invariance for almost every p. At each such p, density and (3.1) extend this identity to arbitrary a and G, since multiplication by each character preserves the L^2 norm.

In (3.2) set $s=0$ and average over $0\le h\le H$. Since $H^{-1}\int_0^H \mathrm{e}(hc)\,\mathrm{d}h\to \mathbf{1}_{\{0\}}(c)$, dominated convergence in $L^2(\nu_1\otimes\nu_2)$ and (3.1) restrict G to the hyperplane

$$v_\theta\cdot\xi_1+Mv_\theta\cdot\xi_2=0.$$

For almost every p the product measure on this hyperplane is concentrated on pairs of atoms. Indeed, for $j=1,2$ put $d_j(\theta)=L_j v_\theta$. With $z=\pi_j p$ fixed, Fubini gives

$$\int\!\!\iint \mathbf{1}_{\{\xi\neq\xi',\ d_j(\theta)\cdot(\xi-\xi')=0\}}\,\mathrm{d}\nu_{j,z}(\xi)\,\mathrm{d}\nu_{j,z}(\xi')\,\mathrm{d}\sigma(\theta)=0,$$

since distinct points contribute at most two angles. As $(x,\theta)\mapsto(\pi_j p,\theta)$ preserves measure, for almost every p each positive-mass fiber of $\xi\mapsto d_j(\theta)\cdot\xi$ is concentrated at a single point. Write η_j for the pushforward of ν_j by this projection. By Fubini the product mass of the hyperplane is $\sum_{t:\,\eta_1(\{t\})>0}\eta_1(\{t\})\eta_2(\{-t\})$. Each positive term therefore comes from a pair of point atoms.

For a pair of atoms (ξ_1,ξ_2) on the hyperplane, set $G=\mathbf{1}_{\{(\xi_1,\xi_2)\}}$. With $h=0$, average (3.2) over $s\in[0,H]^2$ and let $H\to\infty$. The characters in s restrict a in $L^2(\nu_0)$ to the single point $\xi_0=-\xi_1-\xi_2$, again by (3.1). This gives zero unless ξ_0 is an atom. Thus the only contributing triples of atoms satisfy both equations

$$\xi_0+\xi_1+\xi_2=0,\qquad v_\theta\cdot(\xi_1+M^T\xi_2)=0.$$

For almost every p, all such atom triples satisfy $\xi_1+M^T\xi_2=0$. To cover all atoms by countably many measurable fields, normalize each nonzero $\nu_{j,z}$ to a probability measure and take all pairs of

quantiles of its two coordinate marginals at levels in $\mathbb{Q} \cap (0,1)$. The quantiles are measurable in z by the distribution function formula. Each coordinate of an atom is an atom of the corresponding marginal and hence is attained on an interval of quantile levels of positive length. Set the fields to zero where $\nu_{j,z} = 0$. If the two displayed equations and $\xi_1 + M^T\xi_2 \neq 0$ held on a positive-measure set for any triple of these fields, restrict to a positive-measure subset on which the vertices and the fields are bounded, then prune to a positive-measure core. Lemma 2.5 gives a contradiction.

For $|G| \leq 1$, truncation on the hyperplane to finitely many atom pairs has $L^2(\nu_1 \otimes \nu_2)$ error at most the square root of the omitted mass. On each contributing pair $\mathrm{e}(y \cdot (\xi_1 + M^T\xi_2)) = 1$. Thus (3.1) gives $\ell_p(1, e_y \otimes e_{My}) = \ell_p(1,1)$ for every $y \in \mathbb{R}^2$. Take $w \in C_c(\mathbb{R}^2)$ equal to one on Q. Every starting sequence has a subsequence along which dominated convergence and $\int \Psi = 0$ give

$$\lim_{n\to\infty} \Lambda_{\sigma * \Psi_{R_n}}(f_{0,n}, f_{1,n}, f_{2,n}) = \int \Psi(y) \int w(x) \ell_p(1, e_y \otimes e_{My}) \, \mathrm{d}\mu(p) \, \mathrm{d}y = 0. \qquad \square$$

3.1. **Localization of the inputs.** The following observation will also give the endpoint in Theorem 1.2. Let $\mu_n = \sigma * \kappa_n$, where $\sup_n \|\kappa_n\|_{\mathrm{TV}} < \infty$ and $|\mu_n|(\{|u| > U\}) \to 0$ for some fixed U. Let P_n be convolution by measures supported in the unit ball and uniformly bounded in total variation. This includes the identity operator.

Lemma 3.2. *Suppose that for each fixed triple of smooth compactly supported cutoffs χ_l,*

$$\sup_{\|f_l\|_\infty \leq 1} |\Lambda_{\mu_n}(\chi_0 P_n f_0, \chi_1 P_n f_1, \chi_2 P_n f_2)| \to 0.$$

Then $\Lambda_{\mu_n}(P_n f_0, P_n f_1, P_n f_2)$ has norm tending to zero on every permutation of $L^2 \times L^2 \times L^\infty$.

Proof. Take a smooth nonnegative partition $\sum_q \chi_{0,q}(x) = 1$ of the base plane by translates of one cutoff on the unit lattice. For $l = 1, 2$ take translated cutoffs $0 \leq \chi_{l,q} \leq 1$ equal to one at $x + L_l u$ whenever $x \in \operatorname{supp} \chi_{0,q}$ and $|u| \leq U$. The sum of the products lies in $[0,1]$ and equals one for $|u| \leq U$. Thus localization costs $o(1)$ in the desired norm by the tail assumption and Cauchy–Schwarz in x.

For a localized form restrict the two L^2 inputs to a fixed box translated by q, containing the unit neighborhoods of the cutoff supports. This is exact since P_n uses only that neighborhood. Divide by their local L^2 norms, omitting zero terms. Truncate the normalized inputs at magnitude $T \geq 1$. Each discarded part has L^2 norm at most one and is supported on a set of area at most T^{-2}. Its form error is $O(\omega_M(T^{-2}))$ for a unit bounded input. Indeed, write its convolution under P_n as an integral of translates and apply Lemma 2.2 to each translate after multiplication by the cutoff. For the truncated inputs the hypothesis gives $o_n(1)T^2$, uniformly in q by translation invariance. Restore the local norms and sum their products by Cauchy–Schwarz and bounded overlap. First let $n \to \infty$ and then $T \to \infty$. $\square$

Corollary 3.3. *Let ψ be an even smooth function supported in a compact annulus away from zero and set $K_R = \sigma * (R^2 \check{\psi}(R\,\cdot))$. For inputs with Fourier supports in $B(0, C_0 R)$, where C_0 is fixed and support for a bounded input is understood distributionally,*

$$|\Lambda_{K_R}(f_0, f_1, f_2)| \leq a(R) \|f_i\|_\infty \prod_{j \neq i} \|f_j\|_2, \qquad i = 0, 1, 2,$$

where $a(R) \to 0$ as $R \to \infty$.

Proof. Choose p Schwartz with $\widehat{p} = 1$ on a neighborhood of $B(0, C_0)$, so $p_R = R^2 p(R\,\cdot)$ reproduces the given inputs. Multiply p_R by a fixed smooth cutoff supported in the unit ball and equal to one near zero, obtaining a kernel q_R with $\|q_R - p_R\|_1 = o(1)$ and $\|\nabla q_R\|_1 = O(R)$. Apply Lemma 3.2 along any sequence $R \to \infty$ with $P_n f = q_R * f$. The total variation of K_R outside $B(0,2)$ tends to zero because $\check{\psi}$ is Schwartz. For bounded inputs the localized projections have fixed supports, bounded magnitudes, and Lipschitz constants $O(R)$, so Proposition 3.1 gives the hypothesis. Replacing q_R by p_R costs $o(1)$ in form norm by Young's inequality and the uniform total variation bound for K_R. $\square$

4. Summing the frequency annuli

Choose a smooth radial χ equal to one on $B(0,1)$ and supported in $B(0,2)$. Put $\psi(b) = \chi(b/2) - \chi(b)$ and write Λ_r for the form with kernel $\sigma * (r^2 \check{\psi}(r\,\cdot))$. For Schwartz inputs it has the Fourier representation

$$\iint \widehat{f}_0(-\xi_1 - \xi_2)\widehat{f}_1(\xi_1)\widehat{f}_2(\xi_2)\widehat{\sigma}(b)\psi(b/r)\,\mathrm{d}\xi_1\,\mathrm{d}\xi_2, \qquad b = \xi_1 + M^T\xi_2.$$

Lemma 4.1. *For dyadic $R \geq 1$ and $\|f_j\|_2, \|f_j\|_\infty \leq B$, one has*

$$\sum_{r=2^k \geq R} |\Lambda_r(f_0, f_1, f_2)| \leq C_{M,B}\big(\sup_{r \geq R} a(r)\big)^{1/2},$$

with $a(r)$ from Corollary 3.3 for a sufficiently large fixed $C_0 = C_0(M)$.

Proof. Set $A = M^T$ and $\xi_0 = -\xi_1 - \xi_2$. In the coordinates

$$z_0 = (A - I)^{-1}\xi_0, \qquad z_1 = -A^{-1}\xi_1, \qquad z_2 = \xi_2$$

the three differences are

$$z_0 - z_2 = -(A - I)^{-1}b, \quad z_1 - z_2 = -A^{-1}b, \quad z_1 - z_0 = A^{-1}(A - I)^{-1}b.$$

Thus all pairwise distances are comparable to r on the support of $\psi(b/r)$. Translating all three frequency supports by w in these coordinates amounts to modulating the inputs by $((A-I)w, -Aw, w)$. The factors cancel in (1.1), so the form is unchanged.

Let μ be the sum of the three pushforwards of $|\widehat{f}_j(\xi_j)|^2\,\mathrm{d}\xi_j$ to the z-plane, so that $\mu(\mathbb{R}^2) \leq 3B^2$. At each r, partition each input smoothly in z-frequency on a square lattice of mesh hr, using translates and dilates of a single bump. The projections have uniformly bounded convolution L^1 norms and their support squares have side at most $3hr$. Choose $h > 0$ so small that the distances between squares in every contributing triple are comparable to r. Each square has only $O_M(1)$ possible partners. The expansion converges absolutely by (2.2), Cauchy–Schwarz on two families of projections, and bounded overlap. The identity follows by taking finite projection sums and then limits in L^2.

For a contributing triple of squares Q_0, Q_1, Q_2, write $m \leq n \leq p$ for their μ-masses in increasing order. The convolved version of (2.2), with convolution factor of L^1 norm $\|\check{\psi}\|_1$, bounds its contribution by $C(mnp)^{1/2}$. For a second bound translate one square to the origin in the common frequency coordinates. All three supports now lie in $B(0, C_M r)$ in their original frequency planes. Apply Corollary 3.3 with the largest-mass input in the bounded slot to obtain $C_B a(r)(mn)^{1/2}$. Taking the geometric mean bounds the contribution by $C_B a(r)^{1/2} n p^{1/4}$.

To sum this bound, put

$$H(t) = \int \mu(B(z,t))^{1/4}\,\mathrm{d}\mu(z), \qquad 0 \leq H(t) \leq \mu(\mathbb{R}^2)^{5/4}.$$

Fix $c > 0$ small enough that the cr-neighborhoods of Q_0, Q_1, Q_2 are disjoint. They lie in a ball B_* which in turn lies in $B(z, Cr)$ for every $z \in \bigcup_j Q_j$, with C depending only on M. At least two of the neighborhoods have mass at most $\mu(B_*)/2$. For z in either corresponding square,

$$\mu(B(z,Cr))^{1/4} - \mu(B(z,cr))^{1/4} \geq (1 - 2^{-1/4})\mu(B_*)^{1/4} \geq (1 - 2^{-1/4})p^{1/4}.$$

Those two squares have total mass at least n. Integrate the difference over them and sum over all triples. Bounded overlap and the bounded number of partners give

$$|\Lambda_r(f_0, f_1, f_2)| \leq C_{M,B} a(r)^{1/2}\big(H(Cr) - H(cr)\big).$$

Choose an integer D with $2^D c \geq C$. Since H is increasing, the sum of the differences over dyadic $r \geq R$ telescopes to at most $D\mu(\mathbb{R}^2)^{5/4}$. This proves the lemma. □

Corollary 4.2. *Uniformly for* $\|f_j\|_2, \|f_j\|_\infty \le B$,

$$\Lambda_{\sigma * \rho_\varepsilon - \sigma}(f_0, f_1, f_2) \to 0 \quad (\varepsilon \downarrow 0).$$

Proof. For a finite measure κ, smoothing its u variable by ρ_ε averages translates of the two inputs at vertices $1, 2$, by v, Mv respectively. These translations preserve both norms. Thus Lemma 4.1 bounds the tails $r \ge R$ for both σ and $\sigma * \rho_\varepsilon$ by the same quantity. The remaining kernel of the difference is $\sigma * (\rho_\varepsilon * \check{\chi}_R - \check{\chi}_R)$, where $\check{\chi}_R(u) = R^2 \check{\chi}(Ru)$. For fixed R its total variation tends to zero by translation continuity in L^1. The low and high parts exhaust the form, since translation continuity in L^2 and (2.2) give $\Lambda_{\sigma * \check{\chi}_S}(f_0, f_1, f_2) \to \Lambda_\sigma(f_0, f_1, f_2)$ as $S \to \infty$, and likewise after convolution with ρ_ε. First fix R large and then let $\varepsilon \downarrow 0$. □

Proof of Theorem 1.2. Take P_n to be the identity in Lemma 3.2 and apply it to $\mu_n = \sigma * \rho_{\varepsilon_n} - \sigma$ for an arbitrary $\varepsilon_n \downarrow 0$. The hypothesis follows from Corollary 4.2, since bounded inputs times fixed cutoffs have bounded L^2 norms. The tail outside $B(0, 2)$ is at most $\int_{|w|>1} \rho_{\varepsilon_n}(w)\, dw \to 0$. This proves the estimate at unit scale. Dilation invariance gives uniformity for all scales. □

Remark. One also recovers the Fourier inverse estimate. For $g \in L^2 \cap L^\infty$ with bounded Fourier transform, fix $t > 0$ and put $k = \sigma * \rho_t$ with ρ a smooth, even, compactly supported probability density. Then

$$\sup_{\|f_0\|_2, \|f_1\|_2 \le 1} |\Lambda_\sigma(f_0, f_1, g)| \le \|\Lambda_{\sigma - k}\|\, \|g\|_\infty + C_M \|\widehat{k}\|_1 \|\widehat{g}\|_\infty.$$

For Schwartz f_0, f_1, the function paired with g in Λ_k is $z \mapsto \int k(u) f_0(z - Mu) f_1(z + (I - M)u)\, du$, which is Schwartz since $k \in C_c^\infty$. Fourier pairing is thus valid for $g \in L^2$. At fixed $b = \xi_1 + M^T \xi_2$, apply Cauchy–Schwarz in ξ_2 to $\widehat{f}_0(-b + (M^T - I)\xi_2)$ and $\widehat{f}_1(b - M^T \xi_2)$. Both linear parts are invertible. Integration against $|\widehat{k}(b)|$ gives the second term, and density extends the bound to L^2 tests. Consequently the left side tends to zero for any sequence $g_n \in L^2 \cap L^\infty$ with $\sup_n \|g_n\|_\infty < \infty$ and $\|\widehat{g}_n\|_\infty \to 0$, by first taking the upper limit in n and then letting $t \downarrow 0$.

References

[1] J. Bourgain, *A Szemerédi type theorem for sets of positive density in $\mathbb{R}^k$*, Israel J. Math. **54** (1986), 307–316.

[2] D. Castro-Silva, *Geometrical sets with forbidden configurations*, Forum Math. Sigma **11** (2023), e44.

[3] M. Christ, *On certain trilinear oscillatory integral inequalities*, Adv. Nonlinear Stud. **25** (2025), 502–553.

[4] M. Christ, *A three term sublevel set inequality*, arXiv:2204.04346 (2022).

[5] M. Christ, P. Durcik, and J. Roos, *Trilinear smoothing inequalities and a variant of the triangular Hilbert transform*, Adv. Math. **390** (2021), 107863.

[6] K. Falconer, V. Kovač, and A. Yavicoli, *The density of sets containing large similar copies of finite sets*, J. Anal. Math. **148** (2022), 339–359.

[7] H. Furstenberg and Y. Katznelson, *A density version of the Hales–Jewett theorem*, J. Anal. Math. **57** (1991), 64–119.

[8] H. Furstenberg, Y. Katznelson, and B. Weiss, *Ergodic theory and configurations in sets of positive density*, in *Mathematics of Ramsey Theory*, Algorithms Combin. **5**, Springer, Berlin, 1990, 184–198.

[9] A. Greenleaf, A. Iosevich, B. Krause, and A. Liu, *L^p estimates for bilinear generalized Radon transforms in the plane*, in *Combinatorial and Additive Number Theory V*, Springer Proc. Math. Stat. **395**, Springer, Cham, 2022, 179–198.

[10] T. Ziegler, *Nilfactors of $\mathbb{R}^m$-actions and configurations in sets of positive upper density in $\mathbb{R}^m$*, J. Anal. Math. **99** (2006), 249–266.