MOORE, ESCHER, PENROSE: A CONFORMAL GOLDEN BRAID

Anonymous authors
Paper under double-blind review

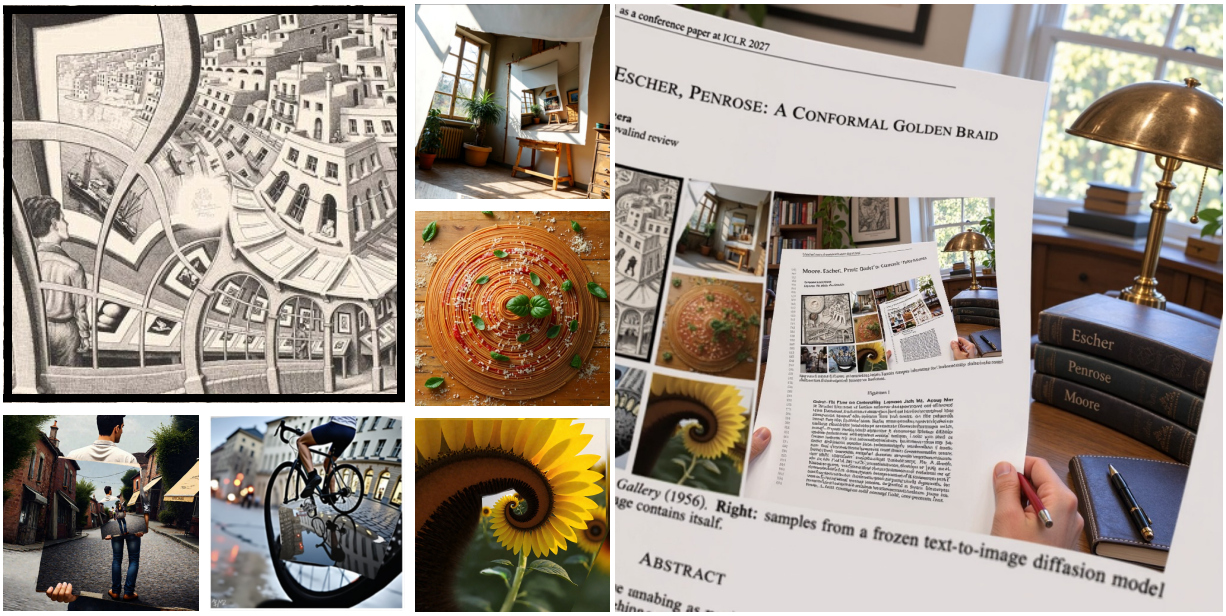

Figure 1: **Top-left:** M. C. Escher, *Print Gallery* (1956). **Surrounding panels:** generated recursive scenes. **Right:** a conditional image of a reader holding this paper, re-noised and refined with our braided sampler. The reader, like you, believes they hold the outermost copy. See the **animated supplementary material**.

ABSTRACT

"I don't think I have ever done anything as peculiar in my life. Among other things, it shows a young man looking with interest at a print on the wall of an exhibition that features himself. How can this be? Perhaps I am not far removed from Einstein's curved universe." So wrote M. C. Escher about his 1956 lithograph *Print Gallery*. Nearly half a century later, a mathematical analysis related its geometry to an untwisted source image through a conformal power map $z \mapsto z^\alpha$, $\alpha \in \mathbb{C}$. Building on this construction, we use a frozen text-to-image diffusion model to generate new self-referential scenes. Prompting alone does not enforce the recursion, while a post-hoc transformation can leave structures poorly connected. Applying the transformation during sampling is also insufficient: the denoiser may "repair" the intended distortion or drift out of the prescribed geometry. We construct a generalized inverse $T^\dagger$ of the non-invertible image transformation T, adapted to its recursive constraint. In the idealized formulation, the Penrose identity $TT^\dagger T = T$ makes $TT^\dagger$ an idempotent projection onto geometrically admissible images. Yet denoising only the transformed image remains an out-of-distribution task, even with projection. We therefore braid denoising steps with T and $T^\dagger$: source-space steps develop the untwisted scene, while transformed-space steps refine its appearance and connections in the final geometry. We generate *Print Gallery*-like compositions and explore further transformations. Rather than distorting a finished image, we let the scene and its distortion develop together.[1]

[1]The title echoes Hofstadter's *Gödel, Escher, Bach: an Eternal Golden Braid* (Hofstadter, 1979).

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

1 INTRODUCTION

"I went there expecting staircases coming out of the windows…"
Roger Penrose, recalling his 1962 visit to M. C. Escher (Wade, 2015)

Maurits Cornelis Escher was not trained as a mathematician. Yet his art both inspired mathematicians and drew on their ideas. After encountering Escher's work, Roger Penrose began devising impossible objects. The triangle and endless staircase he published with his father Lionel in 1958 subsequently inspired Escher's *Waterfall* and *Ascending and Descending* (Penrose & Penrose, 1958; Kersten, 2018; 2019). A different loop appears in the earlier *Print Gallery* (1956). A young man studies a print of a town; following its buildings brings us back to the gallery in which he stands. A picture contains the world that contains it (Museum Escher in The Palace, n.d.).

de Smit & Lenstra (2003) described the print's idealized geometry through a conformal power map connecting a scale-periodic source to its transformed image. It specifies how coordinates correspond, not how buildings, objects, and textures should connect. We seek to generate scenes composed within this geometry, using a pretrained image model.

Prompting for self-reference imposes no exact relation between image regions. Applying the transformation after generation enforces repetition, but leaves the model no opportunity to account for the resulting connections. Introducing it during sampling lets the model respond to the geometry, but subsequent denoising can alter corresponding regions independently or "repair" the intended distortion. Here, distortion is not a defect to remove, but part of the artistic intent.

We use two coupled representations: an untwisted, scale-recursive source and its transformed image. A reverse coordinate warp alone is insufficient: the image transformation selects and repeats content, and a denoised image need not remain in its range. We therefore construct a matched generalized inverse $T^\dagger$ that also acts on such images. Here Penrose returns through generalized inversion (Moore, 1920; Penrose, 1955). In the continuous construction, $TT^\dagger T = T$ ensures that an admissible transformed image survives a round trip unchanged. Thus $TT^\dagger$ is an idempotent projection onto the range of T, restoring the constraint after updates, without a Euclidean nearest-point guarantee.

Geometric consistency alone does not make a convincing composition. Denoising only the transformed image asks the model to interpret unfamiliar arrangements throughout generation. Braiding steps with T and $T^\dagger$ lets the same frozen model develop the source scene and its transformed appearance. For the *Print Gallery* construction, a zero-spin warm-up first encourages scale recursion.

We contribute a matched generalized inverse and a braided sampler for recursive scene generation. The Penrose identities describe continuous operators; resampling and subsequent denoising introduce constraint error. Figure 4 previews the range of geometries and scenes; the **supplementary material** provides animated views.

1.1 RELATED WORK

"Originality is merely an illusion."
— M. C. Escher (Locher, 1982)

Escher's geometry and computational art. de Smit & Lenstra (2003) formalized the geometry of *Print Gallery* and used it to reconstruct the missing central region. Cipolina-Kun et al. (2022) subsequently compared CoModGANs, LaMa, and GLIDE for completing the artwork: they straightened it into several views, inpainted these views, and transformed the completion back. Generation in the straightened representation is a direct precedent. We generate new scenes and repeatedly update both representations during sampling. Generative Escher Meshes (Aigerman & Groueix, 2024) combines diffusion guidance with geometric constraints to synthesize tileable shapes and textures.

Self-similarity and generation across scales. Iterated function systems and fractal image coding represent structure through repeated contractive maps (Hutchinson, 1981; Jacquin, 1992). Neural Collages (Poli et al., 2022) learns contractive operators whose fixed points represent images. Our geometry prescribes admissible images, with content from a frozen model. Generative Powers of Ten (Wang et al., 2024) coordinates diffusion across scales for semantic zooms; ZoomLDM (Yellapragada et al., 2025) learns generation at multiple magnifications. These methods allow content

to change with scale, whereas our construction repeats the same content. Fractal Generative Models (Li et al., 2025) uses architectural recursion without imposing this equality on outputs.

Constrained diffusion sampling. DDRM (Kawar et al., 2022) exploits the spectral structure of linear inverse problems, while DDNM (Wang et al., 2023) corrects clean-image estimates using a generalized inverse. DPS (Chung et al., 2023) uses approximate likelihood guidance for noisy linear and nonlinear measurements. ΠGDM (Song et al., 2023) also uses pseudoinverse-based guidance, including nonlinear and nondifferentiable measurement operators. Universal Guidance, Self-Guidance, and fast constrained sampling provide further ways to steer a frozen prior (Bansal et al., 2023; Epstein et al., 2023; Graikos et al., 2025). Here the constraint relates regions of the generated image to one another, and sampling alternates between source and transformed representations using the matched operators.

Coupled views and geometric consistency. MultiDiffusion (Bar-Tal et al., 2023) fuses diffusion paths under shared constraints, and SyncTweedies (Kim et al., 2024) synchronizes clean-image estimates through a canonical representation. Visual Anagrams (Geng et al., 2024) combines denoising predictions under orthogonal image transformations. LookingGlass (Chang et al., 2025) extends this approach to latent flow models and more general spatial warps through Laplacian-pyramid image processing. Diffusion Illusions (Burgert et al., 2024) optimizes images whose arrangements reveal different interpretations. Tiled Diffusion (Madar & Fried, 2025) enforces compatibility between tile boundaries, including self-tiling; Refracting Reality (Yin et al., 2026) synchronizes directly and indirectly visible content using optical warps. These methods establish denoising across coupled views; our contribution is a matched operator pair for scale-recursive geometry and its use in sequential source/Escher sampling. The forward operator discards content outside a selected source band, while the generalized inverse selects its Droste extension. Their composition is an idempotent projector onto the twisted image class, allowing the constraint to be restored without compounding the warp. Braided sampling inserts source-space denoising between the inverse and forward maps to develop the scene within this recursive geometry.

Idempotence and generalized inversion. Idempotent Generative Network (Shocher et al., 2024) trains a generator using an idempotence objective; here the geometric operator is specified analytically and the image prior is frozen. Gofer & Gilboa (2024) develop generalized inverses for nonlinear operators, including a norm-based pseudoinverse, while Ehrlich et al. (2026) construct surjective pseudo-invertible neural networks (SPNNs) and define inversion through a bijective completion. These formulations distinguish algebraic consistency from the criterion selecting an inverse.

2 PRELIMINARIES

2.1 FROM SCALE REPETITION TO ESCHER'S GEOMETRY

The *Droste effect* is the repeated appearance of an image within itself.[2] Following de Smit & Lenstra (2003), we start with repetition under scaling alone, then apply a conformal power map to introduce rotation. Place the recursion centre at the origin and represent image locations by complex numbers $z = x + iy = re^{i\theta}$. Choose a scale factor $\lambda > 1$ and write $L = \log \lambda$. Our source images satisfy $s(\lambda z) = s(z)$: zooming by λ reproduces the same image. Denote this source class by $\mathcal{S}$.

Power map. We transform source z into output w using a complex power on a logarithm branch:

$$w = z^{\alpha} = \exp(\alpha \log z).$$

Locally, away from the origin, this map is **conformal**: it preserves angles. Write $\log z = \log r + i\theta$. In these coordinates, scaling becomes a horizontal shift and rotation becomes a vertical shift. Multiplication by α rotates and scales this plane; exponentiation returns us to image coordinates.

Choosing the exponent. An output point must receive the same colour at angles θ and $\theta + 2\pi$. For the *Print Gallery* construction, one full turn in the output corresponds to one full turn plus one

[2]The name comes from Droste cocoa packaging, which depicts a nurse carrying a tray with a smaller package bearing the same image.

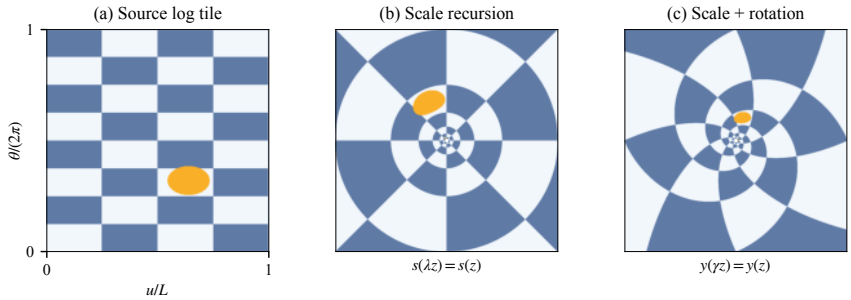

Figure 2: **The Droste effect in three representations.** Left: a tile repeated in logarithmic coordinates, with horizontal period $L = \log \lambda$ and vertical period 2π. Middle: returning to image coordinates gives scale repetition, $s(\lambda z) = s(z)$. Right: the conformal power map adds rotation, producing the Escher geometry with $y(e^{\alpha L} w) = y(w)$. All three panels use the same pattern, with $\lambda = 16$.

scale step in the source. Both leave the source image unchanged. Since the inverse coordinate map multiplies logarithmic coordinates by $\beta = \alpha^{-1}$, this requires $\beta(2\pi i) = 2\pi i + L$, giving

$$\beta = 1 - i\frac{L}{2\pi}, \qquad \alpha = \beta^{-1}. \tag{1}$$

From coordinates to images. A source band $R/\lambda < |z| \le R$ determines the scale-recursive image by repetition inward and outward. Each output location reads its colour from the corresponding source location:

$$y(w) = s(\exp(\beta \log w)), \qquad y(e^{\alpha L} w) = y(w). \tag{2}$$

The choice of β makes this sampling rule independent of the logarithm branch. Multiplication by $e^{\alpha L}$ advances one source scale period, reproducing the same image. In the output, this corresponds to scaling by $e^{L \operatorname{Re} \alpha}$ and rotating by $L \operatorname{Im} \alpha$ radians. These images form the transformed class $\mathcal{E}$. For $\lambda = 256$, the smaller copy has scale approximately $1/22.58$ and rotates clockwise by $157.63°$ (de Smit & Lenstra, 2003). The construction excludes the origin; finite images sample a bounded window around it. Figure 2 compares these three representations.

2.2 Generalized inverses and idempotent projections

For $A \in \mathbb{C}^{m \times n}$, the Moore–Penrose inverse $A^\dagger$ uniquely satisfies (Moore, 1920; Penrose, 1955):

$$\begin{aligned} AA^\dagger A &= A, & A^\dagger A A^\dagger &= A^\dagger, \\ (AA^\dagger)^* &= AA^\dagger, & (A^\dagger A)^* &= A^\dagger A, \end{aligned} \tag{3}$$

where $*$ denotes the conjugate transpose. It returns the minimum-norm least-squares solution.

Keeping only the first two identities defines a *reflexive generalized inverse* $A^\ddagger$, which need not be unique. The round trips are already idempotent: $AA^\ddagger AA^\ddagger = AA^\ddagger, \quad A^\ddagger AA^\ddagger A = A^\ddagger A$. The last two Penrose identities additionally make these projections orthogonal.

Nonlinear inverses. For nonlinear maps, the first two identities still make sense: $gg^\dagger g = g, \quad g^\dagger g g^\dagger = g^\dagger$. The last two rely on a linear adjoint and require a replacement. Gofer & Gilboa (2024) retain reflexive consistency and select $g^\dagger(y)$ by first minimizing $\|g(x) - y\|$, then minimizing $\|x\|$ among those minimizers. This recovers the Moore–Penrose inverse in Euclidean linear spaces.

For surjective g, Ehrlich et al. (2026) fix a bijective completion $G(x) = (g(x), q(x))$ and select

$$g^\dagger(y) = \operatorname*{arg\,min}_{x:\, g(x) = y} \|G(x) - G(0)\|_2. \tag{4}$$

Thus distance is measured in the completed coordinates. For $g = A \circ \phi$, with ϕ bijective and $\phi(0) = 0$, a completion matched to these linearizing coordinates gives $g^\dagger = \phi^{-1} \circ A^\dagger$, recovering all four Penrose identities in those coordinates. Their SPNNs implement this construction.

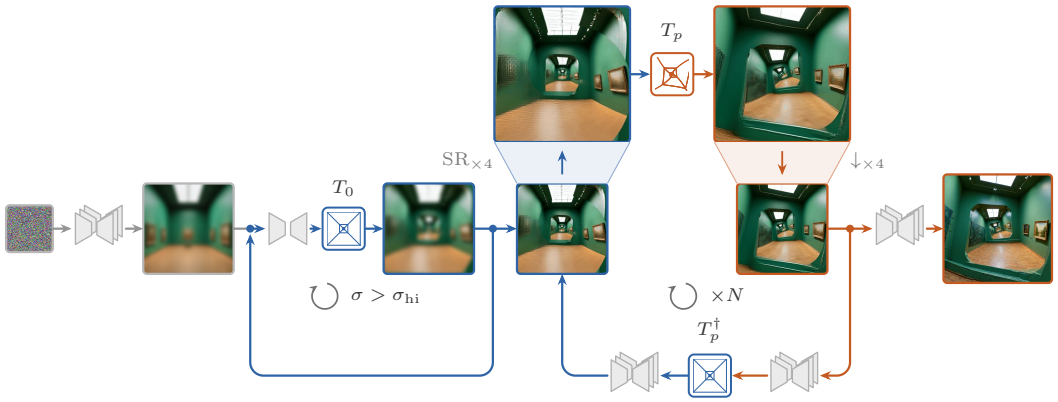

Figure 3: **Braided sampling.** Warm-up interleaves denoising with untwisted repetition T_0. The sampler switches between Droste and Escher representations using T and $T^\dagger$, with denoising blocks $D^{(n)}$ between switches. It ends with refinement in the Escher representation. At each switch, $4\times$ super-resolution precedes the map, then resizing to working resolution. Dashed arrows indicate repetition; the noise-level axis is schematic. The early Droste preview is an illustrative blur of the later panel, not a recorded checkpoint.

Our inverse convention. We use $T^\dagger$ to denote a chosen reflexive generalized inverse, selected by the scale-repeating Droste structure and matched fundamental bands. The dagger does not assert the Moore–Penrose minimum-norm or orthogonality properties. Its round trips restore recursive image structure; here, *projection* means idempotence, without a nearest-point claim.

3 Method

We generate recursive scenes with a frozen text-to-image model by applying geometric maps to its predicted clean images between sampling steps. Alternating denoising with T and its matched generalized inverse $T^\dagger$ develops the untwisted scene and the connections created by its Escher transformation. Figure 3 shows this *braided sampling* process, from scale-repetition warm-up to final refinement in the transformed representation.

3.1 Moving between representations

The maps must accept images that violate recursion, since denoising can change corresponding copies independently. Let B_1 extract the chosen source band and E_1 repeat its content under scaling. Likewise, B_2 reads the corresponding band in the Escher image, and E_2 renders it with the scale and rotation from Sec. 2.1. Both pairs use the same band coordinates. In the continuous construction, extracting a rendered band returns its content unchanged: $B_1E_1 = B_2E_2 = \mathrm{id}$. Define

$$T = E_2B_1, \qquad T^\dagger = E_1B_2. \tag{5}$$

Each map keeps the selected band and regenerates the remaining image. This explains both its loss of information on arbitrary inputs and its ability to impose recursion in the destination representation.

Why the inverse is matched. Cancelling B_1E_1 and B_2E_2 in these compositions gives

$$\begin{aligned} TT^\dagger T &= T, & T^\dagger TT^\dagger &= T^\dagger, \\ TT^\dagger &= E_2B_2, & T^\dagger T &= E_1B_1. \end{aligned} \tag{6}$$

Thus an admissible image survives a geometric round trip unchanged. For an arbitrary image, the round trip restores repetition from its selected band. On the admissible classes $\mathcal{S}$ and $\mathcal{E}$, the maps are mutual inverses; on the ambient image spaces they are reflexive generalized inverses. The band boundaries and angular origin must agree between extraction and rendering.

Restoring the constraint without compounding the warp. Applying T again to an Escher image generally compounds the coordinate warp; T itself is not the projector onto the Escher class. The

matched pair instead gives $P_{\mathcal{E}} = TT^{\dagger} = E_2B_2$, which is idempotent and fixes every admissible Escher image in the continuous construction. The inverse selects a Droste representative: it reads the chosen Escher band, untwists it, and extends it by scale repetition. Appendix C.1 gives an explicit construction. An immediate $T^{\dagger}$–T round trip restores the constraint; our sampler additionally denoises in the Droste representation between these maps to update the scene content.

Two paths for updating the content. The model can update a Droste image directly, or update its Escher rendering and carry the changes back:

$$s \xrightarrow{T} y \xrightarrow{D_{\mathcal{E}}} \widetilde{y} \xrightarrow{T^{\dagger}} s'. \tag{7}$$

Here $D_{\mathcal{E}}$ denotes denoising in the Escher representation. The two paths develop the untwisted scene and its warped connections. Each geometric map ideally restores recursion in its destination, without an additional round-trip projection at the switch.

3.2 Sampling schedule

We use the same frozen model and prompt in both representations. In the schematic sequence

$$\cdots \to D_{\mathcal{S}} \to T \to D_{\mathcal{E}} \to T^{\dagger} \to D_{\mathcal{S}} \to T \to \cdots, \tag{8}$$

each D denotes a block of sampling steps. Geometric operations act on the model's predicted clean image, rather than on its noisy input.

Starting with scale repetition. At high noise levels, we apply the untwisted map $T_0 = E_1B_1$ after each clean prediction, at the working resolution. This introduces nested copies as the scene develops. Despite its identity coordinate warp, T_0 selects and repeats a band. Warm-up ends at σ_{hi}.

Switching representations. The default conformal schedule applies T at $\sigma_{\mathrm{hi}} = 0.87$, $T^{\dagger}$ nine steps later, and a final T at $\sigma_{\mathrm{lo}} = 0.5$, then finishes denoising in the Escher representation. The gap g controls only the first two switches; Appendix B gives the full schedule.

Updating the noisy state. Only when a geometric map is applied, including warm-up, do we re-noise the modified clean prediction with fresh noise:

$$h_{\mathrm{next}} = a_{\mathrm{next}}\widetilde{h}_{\mathrm{clean}} + b_{\mathrm{next}}\epsilon, \qquad \epsilon \sim \mathcal{N}(0, I). \tag{9}$$

Here h is the sampling state and a, b are the schedule's signal and noise coefficients. Other steps use the regular sampler update. Fresh noise avoids warping existing noise: a linear resampler W would change its covariance from I to WW^{*}.

3.3 Resolution and transformation choices

Resolving quality and resolution. At each switch, we apply the geometric map to a $4\times$ super-resolved pixel image, then resize it to the working resolution:

$$M_{\mathrm{pix}}(x) = \mathrm{down}_4(M(\mathrm{SR}_4(x))), \qquad M \in \{T, T^{\dagger}\}. \tag{10}$$

Super-resolution provides more samples for compressed inner copies. For latent backbones, each switch decodes, super-resolves, maps, resizes, and re-encodes the prediction. As an optional final step we re-noise the finished estimate to a moderate level ($\sigma_{\mathrm{tt}} = 0.75$) and denoise it once more without operators, so the prior can smooth the residual seams the warps may leave between repeated copies.

Varying the geometry. We explore five transformation families that arrange the scale-repeating source content in different geometries. For the conformal twist, $\beta_p = 1 - ipL/(2\pi)$ advances p source scale periods per output revolution; changing the sign of p reverses the winding direction. *Poles* combines twists around several prescribed centres, while *Möbius* applies a Möbius coordinate change before the twist, producing two linked recursion centres. *Square* uses square-perimeter coordinates to produce square-shaped nesting, and *Rimrings* adds a radial term that packs repetitions near an outer circle. For Rimrings, we omit inverse steps because inversion is numerically unstable; visible divisions between repeated bands are intentional in this setting. Appendix C gives the map definitions and rendering details.

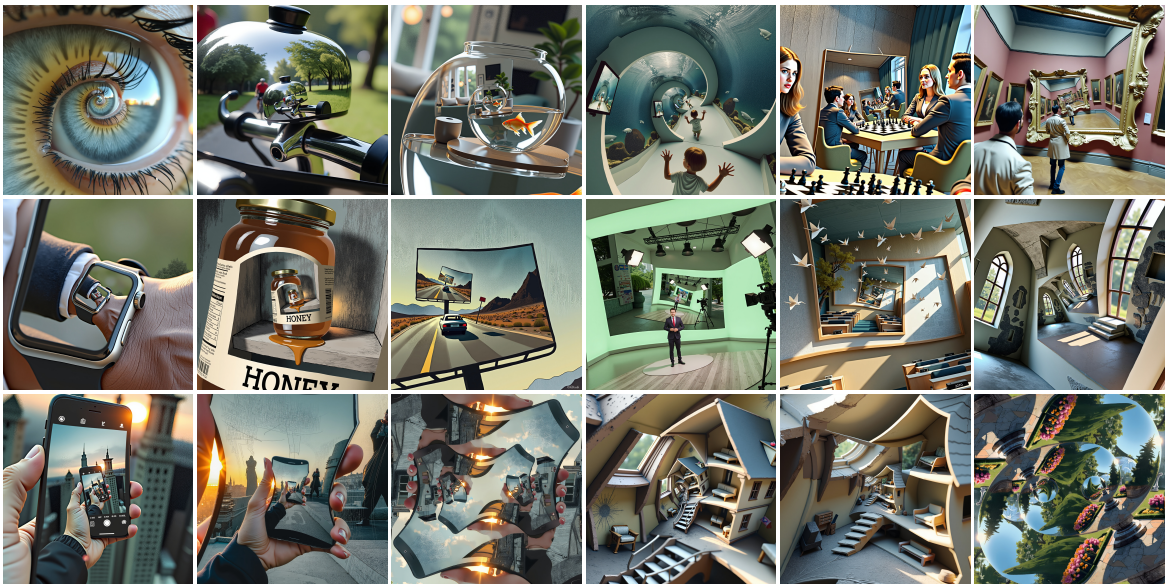

Figure 4: **Overview of generated results.** Diverse scenes with conformal twists, square spirals, and multi-center maps. See the **animated supplementary material**.

4 EXPERIMENTS

Figure 4 opens with results across scenes and transformation families. We evaluate qualitatively, inspecting recursive layout and the content within smaller copies. All experiments use a frozen text-to-image backbone. Baselines and ablations hold the base prompt and initial seed fixed within each row, apart from the stated prompt extension. Appendices A and D give comparison conventions and a geometric round-trip check. The **supplementary material** animates repeated views of the images without generating new content. See Appendix E for further results.

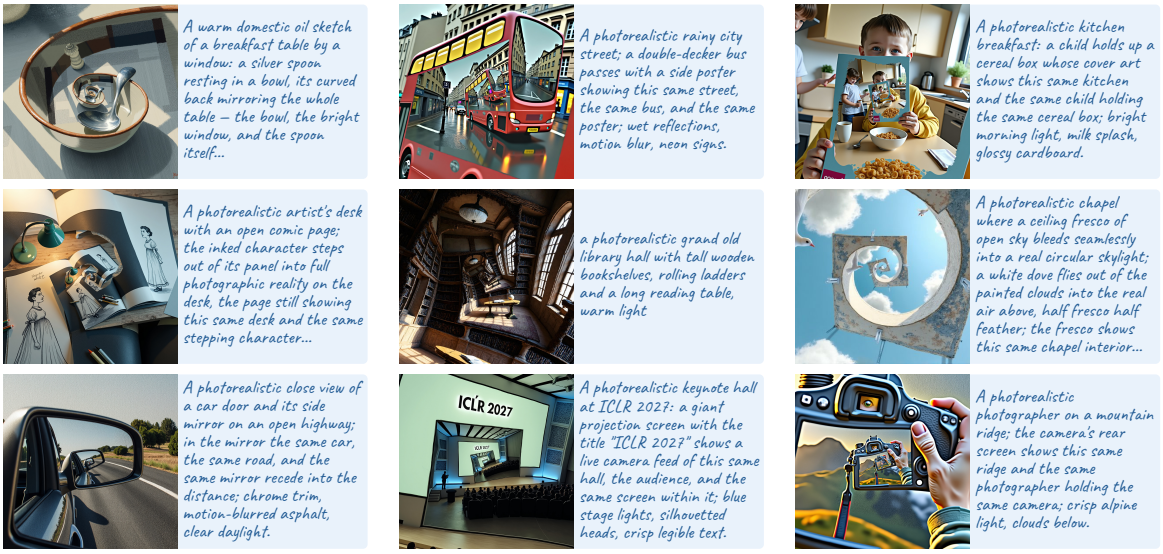

Figure 5: **Conformal results with input prompts.** Longer prompts are excerpted. Animated views appear in the **supplementary material**.

4.1 CLASSICAL PRINT GALLERY TRANSFORMATION

Figure 5 focuses on the conformal construction, with prompts or excerpts. Recursion follows architectural frames, mirrors, camera displays, and screens. The library prompt requests an ordinary

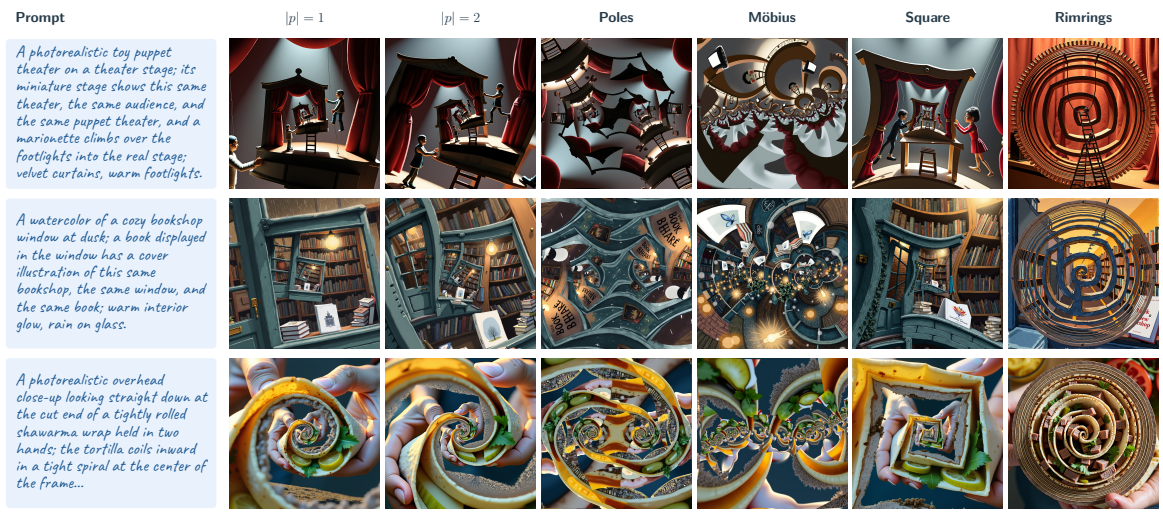

Figure 6: **Twist and transformation choices.** Each row shares a prompt and initial seed. The conformal columns use $|p| = 1, 2$ (negative p for the bookshop), followed by Poles, Möbius, Square, and Rimrings. Longer prompts are excerpted. Rimrings omits inverse steps. See the **supplementary material** for animated views.



Figure 7: **Rimrings: repetition with visible boundaries.** Bands accumulate near an outer rim; their divisions form material layers, ledges, coils, or plate edges. These examples omit the unstable inverse. Animated views appear in the **supplementary material**.

reading room; geometry supplies the nesting. See the **animated supplementary material** and Appendix E for further views and examples.

4.2 Twist and transformation choices

Figure 6 varies the conformal twist and transformation family for three fixed prompt–seed pairs. Poles and Möbius repeat motifs around two centres, while Square produces roughly rectangular nesting. Across these maps, recognizable scene elements persist as their arrangement changes.

Figure 7 shows bands forming tortilla layers, garden ledges, coils, and plate edges. Rimrings omits the unstable inverse: these divisions are intentional, so the aim is not to merge them into a continuous scene. The **supplementary material** animates this accumulation toward the rim.

4.3 Baselines and component ablations

Figure 8 compares prompting and post-processing with cumulative sampler additions. Prompt-only and extended-prompt baselines produce plausible scenes but lack the geometric methods' nested structure. Post-hoc T enforces repetition while carrying over scene boundaries that the denoiser cannot adapt to the new connections.

Starting with T at $\sigma = 0.87$, we add warm-up, $T^\dagger$ followed by a final T at $\sigma = 0.5$, and in-loop SR. The $8\times$ crops expose differences in the inner copies: with warm-up alone, the cereal box contains a conventional illustration and the library mirror contains shelves without another clear mirror. The inverse/forward pair develops another nested scene; SR clarifies small edges and details. Each column shows the cumulative effect of the components included so far. The final column adds optional re-noising and denoising ("time travel"), which is excluded from the main sampler.

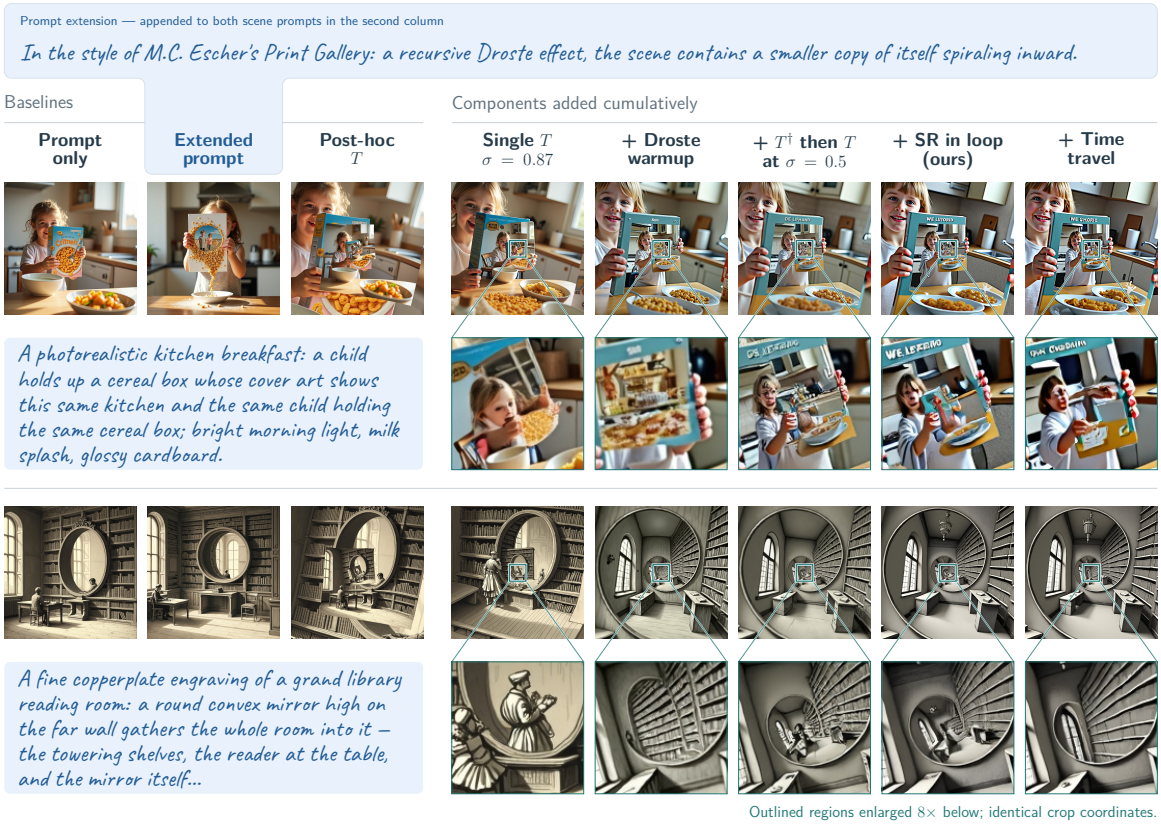

Figure 8: **Baselines and cumulative ablation.** Rows fix the base prompt and seed; the blue callout supplies the prompt extension. The first three columns show prompt-only, extended-prompt, and post-hoc T baselines. Starting from T at $\sigma = 0.87$, subsequent columns add warm-up, $T^\dagger$ followed by a final T at $\sigma = 0.5$, in-loop SR (ours), and optional time travel. Identical central crops are enlarged $8\times$ below.

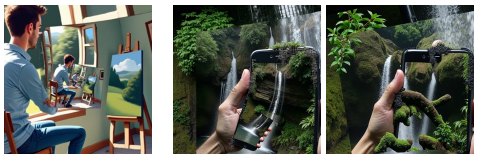

Figure 9: **Failure cases.** Left: recursion anchors on the window rather than the canvas. Right: time travel amplifies the seam (without, then with).

5 Limitations

The fixed recursion centre may miss the intended subject. In Figure 9 (left), warm-up anchors the nest on a window rather than the off-centre painter's canvas. Optional time travel can also amplify seams instead of smoothing them, as the waterfall example shows (right pair). More generally, finite resolution and subsequent denoising limit geometric consistency: the continuous identities do not guarantee exact recursion in the final raster image.

6 Conclusion

Braided sampling combines a frozen image model with matched geometric maps to develop recursive scenes in source and transformed representations. Despite finite resolution and imperfect scene consistency, the results show how geometry can shape a composition throughout generation.

"I never got a pass mark in math ... mathematicians now use my prints to illustrate their books."
—M. C. Escher (Ernst, 1976)

USE OF LARGE LANGUAGE MODELS

LLMs assisted with manuscript editing and drafting, figure preparation, literature search, and checks of the technical exposition. Image results were generated with the text-to-image models described in the paper. The authors are responsible for the final text, claims, references, and artifacts.

REPRODUCIBILITY STATEMENT

Our method adds no training and uses only frozen, publicly available models, so reproducing it needs only the sampler specification. Section 3.2 and Appendix B give the full recipe, with every hyper-parameter and its swept values in Table 4 and the per-step schedule in Table 3; the chart maps T and $T^{\dagger}$ are defined in Appendix C and checked in Appendix D. We use FLUX.1-dev and PixelDiT-1300M with the compute reported in Appendix B. Code will be released upon acceptance.

ETHICS STATEMENT

This work explores artistic image generation using frozen open-source pretrained models. The method may inherit biases and harmful-content risks from these models, and generated images could be misused to mislead viewers. We identify our generated examples as synthetic and attribute the referenced artworks. Use and redistribution of such images should respect applicable licenses and creators' rights.

REFERENCES

Noam Aigerman and Thibault Groueix. Generative Escher Meshes. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*, SIGGRAPH '24, pp. 1–11. ACM, 2024. doi: 10.1145/3641519.3657452. URL http://dx.doi.org/10.1145/3641519.3657452.

Arpit Bansal, Hong-Min Chu, Avi Schwarzschild, Soumyadip Sengupta, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Universal guidance for diffusion models. In *CVPR Workshops*, 2023.

Omer Bar-Tal, Lior Yariv, Yaron Lipman, and Tali Dekel. MultiDiffusion: Fusing diffusion paths for controlled image generation. In *ICML*, 2023.

Black Forest Labs. FLUX.1. https://github.com/black-forest-labs/flux, 2024. Accessed: September 26, 2026.

Ryan Burgert, Xiang Li, Abe Leite, Kanchana Ranasinghe, and Michael Ryoo. Diffusion Illusions: Hiding Images in Plain Sight. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers 24*, SIGGRAPH '24, pp. 1–11. ACM, 2024. doi: 10.1145/3641519.3657500. URL http://dx.doi.org/10.1145/3641519.3657500.

Pascal Chang, Sergio Sancho, Jingwei Tang, Markus Gross, and Vinicius Azevedo. LookingGlass: Generative Anamorphoses via Laplacian Pyramid Warping. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 24–33, 2025.

Hyungjin Chung, Jeongsol Kim, Michael T. McCann, Marc L. Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In *ICLR*, 2023.

Lucia Cipolina-Kun, Simone Caenazzo, and Gaston Mazzei. Comparison of CoModGans, LaMa and GLIDE for Art Inpainting Completing M.C Escher's Print Gallery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 716–724, 2022.

Bart de Smit and Hendrik W. Lenstra, Jr. The Mathematical Structure of Escher's Print Gallery. *Notices of the American Mathematical Society*, 50(4):446–451, 2003. URL https://pub.math.leidenuniv.nl/~smitbde/papers/2003-de_smit-lenstra-escher.pdf. In the feature "Artful Mathematics: The Heritage of M. C. Escher".

Yamit Ehrlich, Nimrod Berman, and Assaf Shocher. Pseudo-Invertible Neural Networks, 2026. URL https://arxiv.org/abs/2602.06042.

Dave Epstein, Allan Jabri, Ben Poole, Alexei A. Efros, and Aleksander Holynski. Diffusion self-guidance for controllable image generation. In *NeurIPS*, 2023.

Bruno Ernst. *The Magic Mirror of M. C. Escher*. Random House, 1976.

Daniel Geng, Inbum Park, and Andrew Owens. Visual Anagrams: Generating Multi-View Optical Illusions with Diffusion Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 24154–24163, 2024.

Eyal Gofer and Guy Gilboa. Generalized Inversion of Nonlinear Operators. *Journal of Mathematical Imaging and Vision*, 66(4):478–503, 2024. doi: 10.1007/s10851-024-01179-w. URL http://dx.doi.org/10.1007/s10851-024-01179-w.

Alexandros Graikos, Nebojsa Jojic, and Dimitris Samaras. Fast constrained sampling in pre-trained diffusion models. In *Advances in Neural Information Processing Systems*, volume 38, Main Conference, pp. 48205–48240. Curran Associates, Inc., 2025. doi: 10.52202/085713-1609. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/44f85a8ac5151e17cecea24ca33a9725-Paper-Conference.pdf.

Douglas R. Hofstadter. *Gödel, Escher, Bach: An Eternal Golden Braid*. Basic Books, 1979.

John Hutchinson. Fractals and Self-Similarity. *Indiana University Mathematics Journal*, 30(5):713–747, 1981. doi: 10.1512/iumj.1981.30.30055. URL https://iumj.org/article/2972/.

A.E. Jacquin. Image coding based on a fractal theory of iterated contractive image transformations. *IEEE Transactions on Image Processing*, 1(1):18–30, 1992. doi: 10.1109/83.128028. URL http://dx.doi.org/10.1109/83.128028.

Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models. In *NeurIPS*, 2022.

Erik Kersten. Ascending and Descending. Museum Escher in The Palace, 2018. URL https://escherinhetpaleis.nl/en/escher-today/ascending-and-descending.

Erik Kersten. The OMG moment of Waterfall. Museum Escher in The Palace, 2019. URL https://escherinhetpaleis.nl/en/escher-today/waterfall.

Jaihoon Kim, Juil Koo, Kyeongmin Yeo, and Minhyuk Sung. SyncTweedies: A General Generative Framework Based on Synchronized Diffusions. In *Advances in Neural Information Processing Systems*, volume 37, pp. 95198–95237. Curran Associates, Inc., 2024. doi: 10.52202/079017-3016. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/ad1efab57a04d93f097e7fbb2d4fc054-Paper-Conference.pdf.

Tianhong Li, Qinyi Sun, Lijie Fan, and Kaiming He. Fractal Generative Models, 2025. URL https://arxiv.org/abs/2502.17437.

J. L. Locher (ed.). *M. C. Escher: His Life and Complete Graphic Work*. Harry N. Abrams, New York, 1982. ISBN 9780810908581.

Or Madar and Ohad Fried. Tiled Diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7795–7804, 2025.

E. H. Moore. On the reciprocal of the general algebraic matrix. *Bulletin of the American Mathematical Society*, 26:394–395, 1920.

Museum Escher in The Palace. Print Gallery, n.d. URL https://escherinhetpaleis.nl/en/about-escher/masterpieces/print-gallery. Accessed September 24, 2026.

L. S. Penrose and R. Penrose. Impossible objects: A special type of visual illusion. *British Journal of Psychology*, 49(1):31–33, 1958. doi: 10.1111/j.2044-8295.1958.tb00634.x. URL http://dx.doi.org/10.1111/j.2044-8295.1958.tb00634.x.

R. Penrose. A generalized inverse for matrices. *Mathematical Proceedings of the Cambridge Philosophical Society*, 51(3):406–413, 1955. doi: 10.1017/s0305004100030401. URL http://dx.doi.org/10.1017/S0305004100030401.

Michael Poli, Winnie Xu, Stefano Massaroli, Chenlin Meng, Kuno Kim, and Stefano Ermon. Self-Similarity Priors: Neural Collages as Differentiable Fractal Representations. In *Advances in Neural Information Processing Systems*, volume 35, pp. 30393–30405. Curran Associates, Inc., 2022. doi: 10.52202/068431-2203. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/c40bed606c51c8e827c1ba75aa2da054-Paper-Conference.pdf.

Assaf Shocher, Amil Dravid, Yossi Gandelsman, Inbar Mosseri, Michael Rubinstein, and Alexei A. Efros. Idempotent generative network. In *ICLR*, 2024.

Jiaming Song, Arash Vahdat, Morteza Mardani, and Jan Kautz. Pseudoinverse-Guided Diffusion Models for Inverse Problems. In *International Conference on Learning Representations (ICLR)*, 2023. URL https://openreview.net/forum?id=9_gsMA8MRKQ.

Mike Wade. Roger Penrose on his friend MC Escher: the genius that galleries ignored. *The Times*, june 2015. URL https://www.thetimes.com/travel/destinations/uk-travel/england/london-travel/roger-penrose-on-his-friend-mc-escher-the-genius-that-galleries-ignored-90nhp8gsd0l. Accessed: September 26, 2026.

Xiaojuan Wang, Janne Kontkanen, Brian Curless, Steven M. Seitz, Ira Kemelmacher-Shlizerman, Ben Mildenhall, Pratul Srinivasan, Dor Verbin, and Aleksander Holynski. Generative Powers of Ten. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7173–7182, 2024.

Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. In *ICCV Workshops*, 2021.

Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-shot image restoration using denoising diffusion null-space model. In *ICLR*, 2023.

Srikar Yellapragada, Alexandros Graikos, Kostas Triaridis, Prateek Prasanna, Rajarsi Gupta, Joel Saltz, and Dimitris Samaras. ZoomLDM: Latent Diffusion Model for Multi-scale Image Generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 23453–23463, 2025.

Yue Yin, Enze Tao, and Dylan Campbell. Refracting Reality: Generating Images with Realistic Transparent Objects. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4312–4321, 2026.

A IMPLEMENTATION DETAILS

The baseline comparison in Figure 8 uses the displayed base prompts and a fixed initial seed within each row. The extended-prompt baseline appends: "In the style of M.C. Escher's Print Gallery: a recursive Droste effect, the scene contains a smaller copy of itself spiraling inward." The five cumulative ablation columns each include the same central 128×128-pixel crop from the 1024×1024 output: columns and rows 448 through 575, using zero-based coordinates. Displaying that crop at the same size as the full image gives an $8\times$ linear enlargement. Crop locations are identical across the compared conditions; the crops receive no separate enhancement.

B SAMPLER AND HYPER-PARAMETERS

We build on FLUX.1-dev (Black Forest Labs, 2024) in half precision with its distilled guidance at scale 3.5 (no classifier-free guidance pass), sampling 1024×1024 images with 128 flow-matching Euler steps and the default resolution-dependent timestep shift. On a non-operator step we take the standard Euler update. On an operator step we form the clean prediction $\hat{x}_0 = x_t - \sigma_t v_\theta(x_t, t)$, decode it with the VAE, apply the operator in pixel space, re-encode, and re-noise to the next level, $x_{t+1} = (1 - \sigma_{t+1})\hat{x}'_0 + \sigma_{t+1}\epsilon$ with a fresh $\epsilon \sim \mathcal{N}(0, I)$.

Unless stated otherwise we use inset scale $s_{\text{in}} = 1/\lambda = 1/4$ and $p = 1$ source scale period per output revolution. From $\sigma_w = 0.95$ to $\sigma_{\text{hi}} = 0.87$ (about 22 steps) we apply the twist-free map ($p = 0$) on every step to plant the nested copy. We then apply T at σ_{hi}, its generalized inverse $T^\dagger$ nine steps later ($\sigma \approx 0.83$), and a final T at $\sigma_{\text{lo}} = 0.5$; the op gap sets only the $T \to T^\dagger$ spacing, while σ_{lo} places the last T. Before each of these three window warps we upscale $\hat{x}_0$ by $\times 4$ with Real-ESRGAN (Wang et al., 2021) and area-downsample back afterwards, which reduces resampling blur. We omit SR during warm-up because repeated enhancement of noisy predictions can darken them. In-loop warps use bilinear resampling. Optional time travel re-noises the result to $\sigma_{\text{tt}} = 0.75$ and denoises without operators. It may reduce joins but can also amplify seams (Figure 9); it is off in the main sampler. The nonlinear SR and denoising stages lie outside the continuous inverse identities of Sec. 3.1. A 1024^2 image takes about two minutes on a single 96 GB NVIDIA Blackwell GPU, with peak memory about 45 GB when in-loop super-resolution is active.

Tables 1–4 give the exact conformal-sampler settings; the number of window ops (3: T, $T^\dagger$, T), the final level $\sigma_{\text{lo}} = 0.5$, the guidance scale 3.5 and the 128 steps are held fixed. Table 4 lists the remaining knobs with their defaults and the values we swept.

Table 1: Model and sampler.

Component	Setting
Base model	FLUX.1-dev (12B rectified-flow DiT), `FluxPipeline`, fp16
Guidance	distilled guidance embedding, scale 3.5 (no CFG pass)
Resolution	1024×1024 (latent $128 \times 128 \times 16$, VAE factor 8)
Scheduler	FlowMatchEulerDiscrete, 128 steps, resolution-dependent shift
Non-op step	Euler flow step, $\hat{x}_0 = x_t - \sigma_t v_\theta$
Op step	decode $\hat{x}_0$, warp, re-encode, re-noise to the next level
Hardware	single NVIDIA Blackwell GPU (96 GB); ≈2 min/image; peak ≈45 GB

Table 2: Conformal transform T.

Item	Definition
Sampling coordinates	$z = \exp(\beta_p \log w)$ about the focus, then source-band folding
Exponents	$\beta_p = 1 - \mathrm{i}p \log \lambda/(2\pi)$; forward $\alpha_p = \beta_p^{-1}$
Inset scale	$s_{\text{in}} = 1/\lambda = 1/4$ default
Periods per turn	$p = 1$ default; $p = 0$ gives untwisted Droste repetition
Generalized inverse	$T^\dagger = E_1 B_2$ with matched bands (Sec. 3.1)
Resampling	bilinear `grid_sample`, border padding, supersampling 1
In-loop upscaling	$\times 4$ Real-ESRGAN (RRDBNet x4plus, fp16), window ops only

Table 3: Operator schedule over the 128 steps ($\sigma : 1 \to 0$).

Phase	σ range (step)	Operator	Upscaling
Free denoising	$1.0 \to 0.95$ (0–18)	none	–
Warm-up	$0.95 \to 0.87$ (19–40)	T_0, every step	none
T-cycle window	$0.87 \to 0.5$	T (41) $\to T^\dagger$ (50) $\to T$ at σ_{lo} (97)	$\times 4$ SR
Free denoising	$\to 0$	none	–
Time travel (optional)	post-hoc	re-noise to σ_{tt}=0.75, denoise, no ops	–

Table 4: Swept hyper-parameters and their defaults.

Name	Symbol	Default	Swept
Inset scale	s_{in}	$1/4$	$1/8,\ 1/16$
Periods per turn	p	1	$-1,\ 2$
Warm-up start	σ_w	0.95	off, 0.99
First op	σ_{hi}	0.87	$0.80,\ 0.92$
Op gap	steps	9	$5,\ 13$
Time travel	σ_{tt}	off	$0.60,\ 0.75$
In-loop upscaling	–	SR $\times 4$	off, bicubic

B.1 Transfer to PixelDiT

To test whether the recipe depends on latent-space denoising, we rerun it on PixelDiT-1300M (`nvidia/PixelDiT-1300M-1024px`), a 1.3B-parameter pixel-space diffusion transformer with no VAE. It shares FLUX's rectified-flow contract, so the operator step carries over unchanged, except that the VAE round trip becomes the identity and every T and $T^\dagger$ is a pure pixel-space warp of $\hat{x}_0$. We keep the FLUX recipe above and match it step for step, changing only the noise schedule: at equal σ PixelDiT is further along, so FLUX's values give hazy outputs. Re-matching the levels by correlating the running estimate with the final image (warm-up $0.95 \to 0.96$, $\sigma_{\mathrm{hi}}\, 0.87 \to 0.92$, $\sigma_{\mathrm{lo}}\, 0.5 \to 0.69$, time travel $0.75 \to 0.82$) restores contrast and keeps the recursion. A 1024^2 image takes about 20 seconds on one RTX 5090 (versus $\sim$2 minutes for FLUX). The resulting gallery is shown in Appendix E.

C Transformation definitions

We use the source scale factor $\lambda > 1$ and $L = \log \lambda$ from Sec. 2.1. After centering source coordinates, write the output-to-source sampling rule as

$$y(w) = s(\exp Q(w)) .$$

Here Q is the source log-position, and a positive factor η sets its overall image scale. Source scale repetition and angular periodicity identify log-positions that differ by an element of

$$\Lambda = L\mathbb{Z} + 2\pi i\mathbb{Z}.$$

Consequently, a branch change that adds an element of Λ does not change the sampled source colour. Define $\beta_{k,l} = l - ikL/(2\pi)$ for integers k, l.

Conformal twist. For integer p we use

$$Q(w) = \beta_{p,1} \log w + \log \eta, \qquad \beta_{p,1} = 1 - i\frac{pL}{2\pi}.$$

A full turn changes Q by $pL + 2\pi i \in \Lambda$. Thus p counts signed source scale periods per output revolution. For $w = re^{i\theta}$, a fixed source radius satisfies $\log r + (pL/2\pi)\theta = \mathrm{const}$, giving logarithmic spirals. Increasing $|p|$ changes their winding; the map is locally conformal away from the origin.

Poles. For prescribed centres c_j and integer pairs (k_j, l_j),

$$Q(w) = \sum_j \beta_{k_j,l_j} \log(w - c_j) + \log \eta.$$

Each centre contributes a local recursion structure. A turn around c_j changes Q by $k_j L + 2\pi i l_j \in \Lambda$, so branch consistency holds separately at every centre. Away from the centres, the map is locally conformal wherever $Q'(w) = \sum_j \beta_{k_j,l_j}/(w - c_j)$ is nonzero.

Möbius. For distinct points a, b, we precompose the twist with $M(w) = (w - a)/(w - b)$. The renderer uses

$$Q(w) = \beta_{k,l}\big[\log(w - a) - \log(w - b)\big] + \log \eta.$$

This is the two-centre case of Poles with opposite coefficients, linking the recursion around a and b. Different logarithm branches differ by source periods. This coordinate construction does not by itself specify a matched image-space inverse.

Square. For $w = x + iy$, let $\rho_\square = \max(|x|, |y|)$, and let $\tau \in [0, 1)$ be normalized arclength around the square perimeter, starting on the positive real axis. We use

$$Q(w) = l \log \rho_\square + kL\tau + 2\pi i l\tau + \log \eta.$$

One perimeter traversal adds $kL + 2\pi i l \in \Lambda$. The square chart replaces circular nesting with square-shaped nesting. It is non-conformal, and its derivatives change across the rays through the square's corners.

Rimrings. The ideal map adds a real radial term to the twist:

$$Q(w) = \beta_{k,l} \log w + h(|w|) + \log \eta, \qquad h(r) = \frac{Ar}{R_0 - r}, \quad A > 0.$$

The radial term leaves angular branch consistency unchanged. As $r \to R_0^-$, it traverses increasingly many source scale periods, crowding repeated bands near the outer circle. The implementation evaluates h at $\min(|w|, R_0 - 10^{-3})$ to regularize the divergence. We do not apply an inverse for this transform: inversion is numerically unstable, and the visible divisions between bands are part of the desired appearance.

Rendering. All five maps specify output-to-source coordinates. Writing $Q = u + i\theta$, the renderer chooses the largest-radius scale-equivalent sample that remains inside the source canvas. If $r_{\max}(\theta)$ is the available source radius in direction θ, including a 0.995 safety factor, it uses

$$m = \left\lfloor \frac{\log r_{\max}(\theta) - u}{L} \right\rfloor, \qquad r_{\mathrm{src}} = \exp(u + mL).$$

The standalone renderer samples these source positions bicubically on a supersampled output grid and averages to the final resolution. The in-loop implementation uses bilinear sampling (Table 2); the coordinate rules are shared. Replacing exact zero logarithm arguments by 10^{-9} avoids evaluating the logarithm at singular centres. The square radius is likewise floored at 10^{-9}. These formulas define the coordinate renderer; the matched-band inverse identities in Sec. 3.1 require the corresponding band extraction and extension construction.

C.1 The Droste-selected generalized inverse

We make the continuous construction explicit for the conformal family. Fix $p \in \mathbb{Z}$, $\beta_p = 1 - ipL/(2\pi)$, $\alpha_p = \beta_p^{-1}$, and a logarithm branch $\operatorname{Arg} z \in [-\pi, \pi)$. Let $R(\theta) = r_{\max}(\theta)$ be the source radius used by the renderer, including its safety factor. Define the source band and its Droste folding map by

$$A = \{re^{i\theta} : R(\theta)/\lambda < r \le R(\theta)\}, \tag{11}$$

$$d(z) = \lambda^{m(z)} z, \qquad m(z) = \left\lfloor \frac{\log R(\operatorname{Arg} z) - \log |z|}{L} \right\rfloor. \tag{12}$$

Thus $d(z) \in A$, $d(q) = q$ for $q \in A$, and $d(\lambda z) = d(z)$. The forward image operator is

$$F(w) = d(\exp(\beta_p \log w + \log \eta)), \qquad (Tx)(w) = x(F(w)).$$

Here F returns a source-band coordinate, not an image value.

To read this band from an Escher image, fix one output representative for each $q \in A$:

$$\psi(q) = \exp(\alpha_p[\log q - \log \eta + n(q)L]), \qquad n(q) \in \mathbb{Z}.$$

The integer $n(q)$ fixes which scale-rotated copy is read; it is chosen from coordinates alone and held fixed for all images. In the unbounded continuous construction one may take $n(q) = 0$. For a bounded output canvas containing a neighbourhood of the recursion centre, an explicit available-band convention is to choose the largest integer $n(q)$ for which $\psi(q)$ remains inside the canvas, with the rendering safety margin. This specifies the selected Escher band $\psi(A)$.

Since $2\pi i\beta_p = pL + 2\pi i$, logarithm branch changes alter the recovered source coordinate only by an integer power of λ. Consequently $F(\psi(q)) = q$. Define

$$(T^\dagger y)(z) = y(\psi(d(z))).$$

Operationally, this folds z into the source band, finds its matched Escher representative, and reads the colour there. In particular, $T^\dagger y$ is scale-periodic even when y violates recursion. The extraction and extension maps are now explicit:

$$B_1 x = x|_A, \quad (E_1 b)(z) = b(d(z)), \qquad (B_2 y)(q) = y(\psi(q)), \quad (E_2 b)(w) = b(F(w)).$$

Hence $B_1 E_1 = B_2 E_2 = \mathrm{id}$, $T = E_2 B_1$, and $T^\dagger = E_1 B_2$, giving both reflexive identities in Eq. 6.

For $y \in \mathcal{E}$, $T^\dagger y$ is the unique member of $\mathcal{S}$ mapped to y: among all preimages, it selects the Droste extension of the recovered band. For arbitrary y, the selected band also matters because its repeated copies may disagree; $TT^\dagger y$ regenerates an admissible Escher image from that band. This is an algebraic selection, without a minimum-norm or nearest-point claim. The construction excludes the recursion centre. Finite sampling, interpolation, and the subsequent SR and denoising operations are outside these exact continuous identities.

D EMPIRICAL CHECK OF IDEMPOTENCE

We check the geometric maps separately from denoising and SR. For $P = TT^{\dagger}$, an image Tx is a fixed point in the continuous construction, so $P(Tx) = P^2(Tx) = Tx$. Figure 10 compares these images on a checkerboard and displays the residual $|P^2(Tx) - P(Tx)|$. The transformed panels retain the same recursive layout, while blur and edge differences remain visible after resampling. This is a visual check of the discretised geometric round trip; it does not establish exact consistency of the final denoised images.

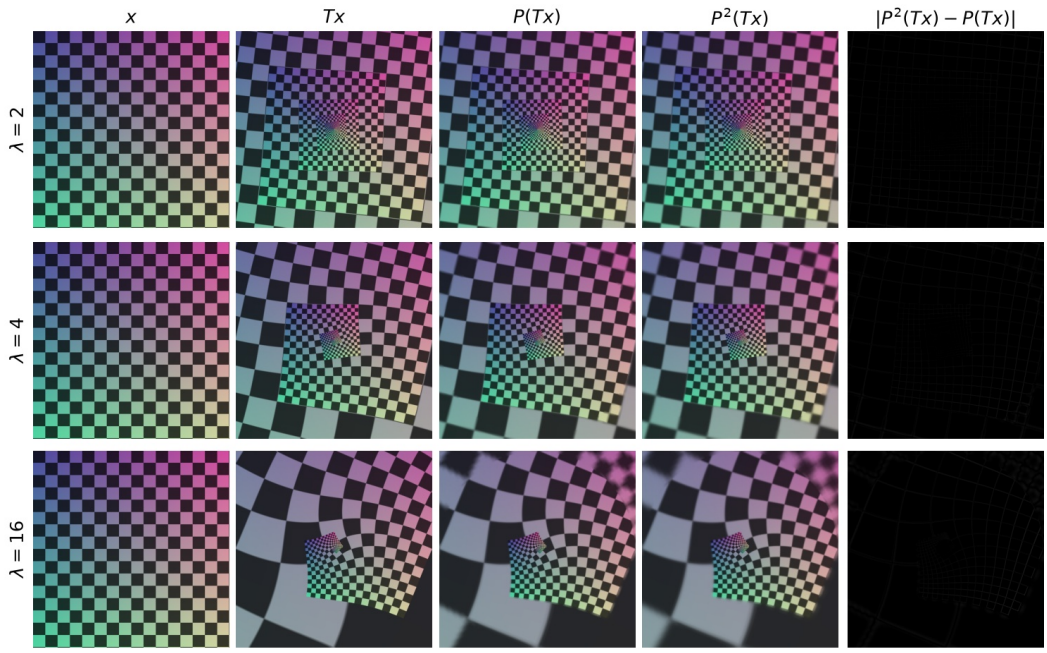

Figure 10: **Checking geometric round trips.** From left to right: x, Tx, $P(Tx)$, $P^2(Tx)$, and the residual $|P^2(Tx) - P(Tx)|$, where $P = TT^{\dagger}$. Rows use deck ratios $\lambda = 2, 4, 16$. The three transformed images coincide in the continuous construction; the rasterized comparison shows blur and residual differences near checkerboard edges. The measured residual $|P^2(Tx) - P(Tx)|$ is 0.011, 0.012, 0.014 (top to bottom), each at the bilinear-resampling floor.



Figure 11: **Transfer to a pixel-space backbone.** The conformal T-cycle recipe applied to PixelDiT-1300M, with the same prompts, numerical seeds, transform, and scale as the FLUX configuration, and adapted intervention noise levels. The warm-up start changes from 0.95 to 0.96, σ_{hi} from 0.87 to 0.92, and σ_{lo} from 0.5 to 0.69. Where time travel is used, its noise level changes from 0.75 to 0.82. Per-cell tags retain the FLUX naming convention (e.g. `tt75`) and do not denote the adapted noise levels.

E Additional results

E.1 Transfer to a pixel-space model

Our sampler applies its geometric operators in pixel space, so it shouldn't be dependent on the backbone denoising in a latent space. To test this we swap FLUX.1-dev, a latent-diffusion model, for PixelDiT-1300M, which denoises directly in pixels, and rerun the same recipe with only the model changed. Figure 11 shows the recursion still forms, so the technique transfers across the backbone and across the latent-to-pixel change; the implementation is described in Appendix B.

E.2 Curated FLUX gallery

Figure 12, shown across the next two pages, collects a larger set of FLUX.1-dev results, distinct from the examples in the main text.



Figure 12: **Additional Escher-recursion results (FLUX.1-dev).** Curated gallery across subjects, transforms (conformal, square, poles, Möbius), inset scales and spiral orders; (Page 1 of 2.)



Figure 12: (Continued, page 2 of 2.)