Research Proposal

Airfoil Generation via SE(3)-Equivariant Diffusion on Dual-Quaternion Frame Graphs

Team: Rishi Shah (solo)

Date: May 24, 2026

Project: Airfoil Generation via SE(3)-Equivariant Diffusion on Dual-Quaternion Frame Graphs

Target Compute: 2x AMD Instinct MI300X, approx. 2 weeks

1. PROBLEM STATEMENT

This project proposes a novel generative framework for aerodynamic shape design that represents an airfoil as a graph of local geometric frames. Each local frame is encoded as a dual quaternion, allowing the airfoil to be modelled as a sequence of rigid transformations that capture local position, orientation, curvature, and surface evolution.

Rather than performing diffusion in Euclidean coordinate space, the model performs diffusion on a structured product space: the Lie group SE(2) governs pose (position and in-plane orientation), while a Euclidean fiber encodes higher-order local geometry (signed curvature, curvature rate, and local half-thickness). This fiber-bundle decomposition keeps group operations mathematically valid while making full use of all six degrees of freedom available in a dual-quaternion parametrisation. The denoising network is designed to be SE(2)-equivariant and operates on two complementary edge features: relative SE(2) frame transformations and Euclidean fiber differences. This allows the model to learn geometric relationships directly from the underlying transformation structure. The central hypothesis is that learning distributions over local geometric transformations and intrinsic curvature will produce smoother, more physically realistic airfoils than learning distributions over raw coordinates. The framework scales naturally to full 3-D wing surfaces.

2. PROPOSED APPROACH

A 2-D airfoil curve lives in the z=0 plane. A local Frenet-Serret frame requires exactly three numbers: $(x, y, \theta) \in SE(2)$. A unit dual quaternion carries six independent degrees of freedom; the three unused dimensions are repurposed as a geometric fiber attached to each surface frame:

Extra DOF	Encoded Quantity	Aerodynamic Significance
Out-of-plane rotation 1	Signed curvature $K(s)$	Pressure gradient, separation onset
Out-of-plane rotation 2	Curvature rate $K'(s)$	Smoothness, transition prediction
z-translation	Local half-thickness $t(s)$	Thickness distribution, drag

The resulting node representation lives in $SE(2) \times R^3$. Message passing employs Frenet-Serret edge transformations. Diffusion is performed in the Lie algebra by linearization around the group identity, admitting closed-form score matching. A geometric consistency constraint loss regularizes the network by penalizing frames whose curvature encoding is inconsistent with the implied frame rotation rate.

The denoising network is implemented as an SE(2)-equivariant graph transformer with dual edge-processing heads (Pose head + Fiber head). The model is conditioned on Reynolds number (Re), angle of attack (α), target lift (C_l), and target drag (C_d) via FiLM or cross-attention layers. Airfoil reconstruction solves a global least-squares integration block to recover positions without accumulating sequential drift, enforcing trailing-edge closure as a hard equality constraint.

3. DATASET

The framework utilizes a combination of open aerodynamic source profiles and large-scale synthetic data augmentation:

Corpus / Dataset	Source	Raw Size / Sample Count	Role
UIUC Airfoil Database	Public Aerodynamic Repo	~1,600 unique airfoils	Base seed configurations
Augmented Perturbations	CST coefficient variations (±20%)	~50,000 - 90,000 shapes	Synthetic training expansion
Aerodynamic Targets	XFOIL simulations over (Re, AoA) grid	Fully Labeled Map	Surrogate network training & conditioning

Preprocessing & Normalization

- **Canonical Frame Normalization:** Chord length scaled to 1.0; leading edge translated to origin; trailing edge anchored at (1, 0).
- Each airfoil is re-sampled along its arc-length into N points (typically N=128) to extract positions, tangents, normals, curvatures, and thickness metrics.
- Samples with self-intersections or non-closing trailing edges are strictly rejected.

4. EVALUATION METRICS

Category	Metric	Description / Goal
Geometry	Curvature smoothness	L^2 norm of K''(s); minimizing non-physical kinks
Geometry	Surface continuity	G1 and G2 continuity checks along the entire perimeter
Geometry	Self-intersection / Closure	Fraction of self-intersections (target 0); Trailing edge distance error
Aerodynamics	Lift & Drag Error	Absolute error vs. target C_l and C_d coefficients
Aerodynamics	Lift-to-drag-ratio	C_l/C_d optimization performance at design operating point
Generative	Diversity & Novelty	Spread in shape space; Fraction of unique samples not in train set

5. COMPUTE REQUIREMENTS

The model leverages multi-accelerator compute to handle graph message passing and coordinate-free diffusion processing blocks efficiently.

Phase	Hardware	Estimated Wall-Clock	Notes
Model Pretraining & Augmentation	2x AMD Instinct MI300X	~5 days	Joint SE(2) and fiber diffusion transformer training

Aerodynamic Surrogate Training	1x AMD Instinct MI300X	~1 day	Separate training of the physics guide network
Guided Denoising Sweeps & Eval	1x AMD Instinct MI300X	~2 days	Generating, filtering, and post-processing candidate airfoils

Software Dependencies

- ROCm 6.x + PyTorch ROCm wheel environment
- NumPy, SciPy, scikit-learn, and custom FFT / Lie algebra utilities
- XFOIL standard bindings for automated aerodynamic validation sweeps
- No Docker required; standard Python venv execution setup

6. EXPECTED DELIVERABLES

- Converged weights for the SE(2)-equivariant graph transformer and the separate neural aerodynamic surrogate.
- An expanded, cleaned, and labeled training database of 50,000–90,000 airfoil profiles complete with XFOIL evaluation logs.
- A set of 10-20 novel high-efficiency candidate airfoil geometries meeting target C_l and C_d conditions.
- A benchmark headline table comparing geometric quality and lift-to-drag performance against coordinate-diffusion and CST baselines.
- A fully compiled conference draft targeted at the International Conference on Machine Learning (ICML).

7. RISK & MITIGATION

Risk	Likelihood	Mitigation
Log-map singularity near high-curvature leading edge	Medium	Arc-length re-sampling to guarantee consecutive frames differ by < $\pi/2$; apply smooth clamping function $\tanh(\|\|\omega\|\|/\pi) \times \pi$ to rotation magnitudes.
Pose and fiber representations become geometrically inconsistent	Medium	Enforce a geometric consistency loss regularizer linking rotational twist to integrated curvature during training.
Surrogate guidance diverges at high noise levels	High	Apply noise annealing—only activate surrogate guidance in the final 30% of denoising steps where the geometry is recognizable.