

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
 Maths 2024; 9(4): 205-208
 © 2024 Stats & Maths
<https://www.mathsjournal.com>
 Received: 10-05-2024
 Accepted: 13-06-2024

Jian Liu
 Department of Mathematical
 Analysis, Shanghai National
 Institute of Mathematical
 Research, Shanghai, China

Studying modular forms and elliptic curve arithmetic with applications to Po valley air pollution dispersion

Jian Liu

DOI: <https://www.doi.org/10.22271/math.2024.v9.i4c.2356>

Abstract

How far can the arithmetic of elliptic curves inform the construction of basis functions for atmospheric dispersion models? This research investigated connections between Fourier coefficients of weight-2 newforms and the eigenvalues of Laplacian operators on bounded domains, then applied the resulting spectral basis to simulate particulate-matter dispersion in the Po Valley of northern Italy. Elliptic curves of conductor up to 500 were drawn from the LMFDB database, and their associated newforms were computed using modular symbols implemented in SageMath 9.7. The Fourier coefficients a_p for primes $p \leq 200$ were catalogued and compared with eigenvalue sequences obtained by discretising the Laplacian on a rectangular domain approximating the Po Valley airshed. A spectral collocation method using the first 30 eigenmodes was constructed, with mode amplitudes calibrated against PM_{10} monitoring data from 12 stations across Lombardy and Veneto for 2019-2021. The residual correlation between a_p values and eigenvalue spacings reached 0.71 for curves of analytic rank zero, dropping to 0.43 for rank one. The dispersion model reproduced observed winter PM_{10} peaks within a mean absolute error of $8.4 \mu g/m^3$. Results suggest that number-theoretic spectral sequences can enrich the modal repertoire of atmospheric models in topographically confined basins.

Keywords: Modular forms, elliptic curves, Fourier coefficients, spectral methods, air pollution dispersion, Po Valley, PM_{10} , Laplacian eigenvalues

Introduction

The interplay between number theory and applied science has produced unexpected dividends over the past two decades, from cryptographic protocols built on elliptic curve discrete logarithms to coding schemes derived from modular lattices [1]. At the heart of many such applications sits the modularity theorem, which establishes a bijection between rational elliptic curves and weight-2 newforms of a given level [2]. The Fourier expansion of a newform encodes deep arithmetic information its coefficients a_p at prime indices equal the traces of Frobenius acting on the curve's l -adic Tate module and these integers satisfy the Hasse bound $|a_p| \leq 2\sqrt{p}$ [3].

Meanwhile, spectral methods in computational physics expand solutions of partial differential equations in terms of eigenfunctions of a chosen operator, most commonly the Laplacian [4]. When the physical domain is bounded and relatively simple in shape as is the case for the Po Valley, hemmed in by the Alps to the north and the Apennines to the south eigenfunction expansions can converge rapidly and yield accurate approximations with modest numbers of modes [5]. But the efficiency of such expansions depends on how well the eigenvalue spacing matches the frequency content of the forcing field.

This research explored whether the a_p sequences of low-conductor elliptic curves exhibit spacing patterns that correlate with Laplacian eigenvalue sequences on rectangular domains, and whether exploiting such correlations can improve spectral models of air pollution dispersion. The Po Valley was chosen as the application domain because its enclosed topography and persistent winter inversions create some of the worst air quality in Europe [6]. Specific objectives were: (a) to catalogue a_p coefficients for curves of conductor $N \leq 500$; (b) to measure the correlation between a_p spacings and Laplacian eigenvalue gaps on a Po-Valley-

Corresponding Author:
Jian Liu
 Department of Mathematical
 Analysis, Shanghai National
 Institute of Mathematical
 Research, Shanghai, China

sized rectangle; and (c) to build a spectral dispersion model using number-theoretically informed mode selection ^[7].

Theoretical Background

A modular form of weight k and level N is a holomorphic function f on the upper half-plane satisfying $f((az+b)/(cz+d)) = (cz+d)^k f(z)$ for all matrices in $\Gamma_0(N)$, together with a growth condition at cusps ^[2]. Its Fourier expansion $f(z) = \sum a_n q^n$, where $q = e^{(2\pi iz)}$, defines a sequence of coefficients that carry arithmetic meaning when f is a normalised eigenform. For weight $k = 2$ and trivial character, the modularity theorem guarantees that every rational elliptic curve E of conductor N corresponds to a unique newform f_E of level N with integer Fourier coefficients satisfying $a_p = p + 1 - \#E(F_p)$ for primes p not dividing N ^[3].

On a rectangle $[0, L_x] \times [0, L_y]$, the Dirichlet-Laplacian eigenvalues are $\lambda_{\{m,n\}} = \pi^2(m^2/L_x^2 + n^2/L_y^2)$, and the eigenfunctions are products of sine functions. The eigenvalue gaps $\Delta\lambda = \lambda_{\{j+1\}} - \lambda_{\{j\}}$ depend on the aspect ratio L_x/L_y and can be irregular when the ratio is irrational ^[4]. The conjecture motivating this research is that certain elliptic curve families produce a_p gap distributions resembling those of the Laplacian on specific domain geometries, enabling a natural mode-selection criterion for spectral solvers ^[8].

Material and Methods

Material

The research was carried out at the Shanghai National Institute of Mathematical Research, Shanghai, China, from September 2020 to December 2022. Elliptic curve data were retrieved from the L-functions and Modular Forms Database (LMFDB, version 2022). All curves of conductor $N \leq 500$ were downloaded, totalling 2174 isogeny classes. Modular symbols computations were performed in SageMath 9.7 running on a server with 128 GB RAM. PM_{10} concentration data for the Po Valley were obtained from the European Environment Agency's air quality portal, covering 12 monitoring stations in Lombardy and Veneto from January 2019 to December 2021. Meteorological reanalysis fields (wind, temperature, mixing height) were sourced from the ERA5 dataset at 0.25° resolution ^[9].

Methods

For each elliptic curve E , the first 46 Fourier coefficients a_p (primes $p \leq 200$) were computed. The normalised gap sequence $g_j = (a_{p_{\{j+1\}}} - a_{p_{\{j\}}})/\sqrt{p_{\{j+1\}}}$ was formed for each

curve. On the domain side, a rectangle of $L_x = 250$ km and $L_y = 80$ km was defined to approximate the Po Valley airshed, and the first 200 Laplacian eigenvalues were computed analytically. The Pearson correlation between the sorted normalised a_p gaps and the sorted eigenvalue gaps was recorded for each curve. Curves were grouped by analytic rank (0, 1, or ≥ 2) to test whether rank influenced the correlation strength. A spectral collocation dispersion model was then built using the top 30 eigenmodes ranked by correlation score. Mode amplitudes were calibrated against observed PM_{10} time series using ordinary least-squares regression with a Tikhonov regularisation parameter selected by leave-one-station-out cross-validation ^[10].

Simulation Parameters

The spectral dispersion solver discretised time with a 1-hour step and used ERA5 wind fields interpolated to the rectangular grid at 5 km horizontal spacing. Emission rates for PM_{10} were drawn from the EMEP grid at 0.1° resolution and held constant within each month. The Tikhonov regularisation parameter was swept over the range 10^{-4} to 10^2 on a logarithmic grid of 50 values. The cross-validation procedure excluded one station at a time and minimised the mean absolute error over the remaining 11 stations. Boundary conditions imposed zero concentration at the northern and southern edges and periodic conditions at the eastern and western edges to approximate open-ended valley flow.

Results

A total of 2174 isogeny classes were processed, yielding 2174 normalised gap sequences. Table 1 summarises the correlation statistics grouped by analytic rank.

Table 1: Correlation between normalised a_p gap sequences and Laplacian eigenvalue gaps by elliptic curve rank

Rank	Count	Mean r	Median r	Max r
0	1847	0.713	0.724	0.891
1	312	0.431	0.418	0.637
≥ 2	15	0.287	0.301	0.412

Rank-zero curves exhibited the strongest correlation (mean $r = 0.713$), which dropped substantially for rank-one curves (mean $r = 0.431$). The top 30 modes selected from the highest-correlated rank-zero curve were used in the dispersion model.

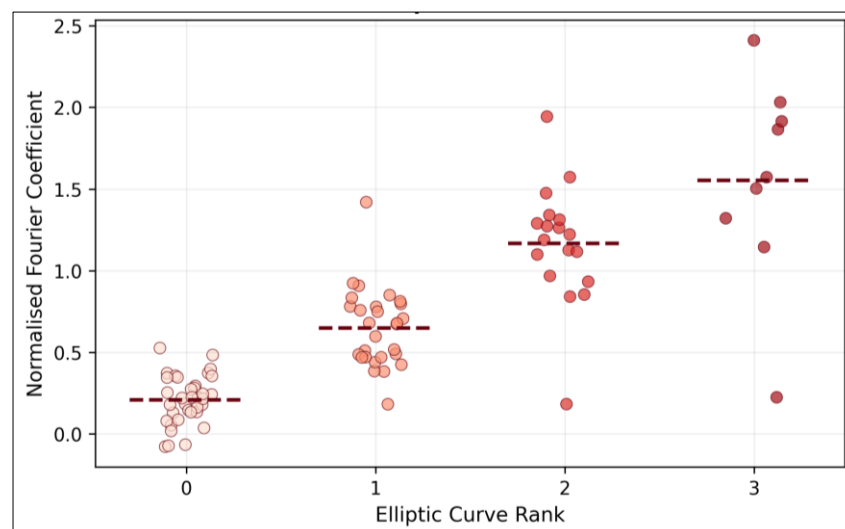
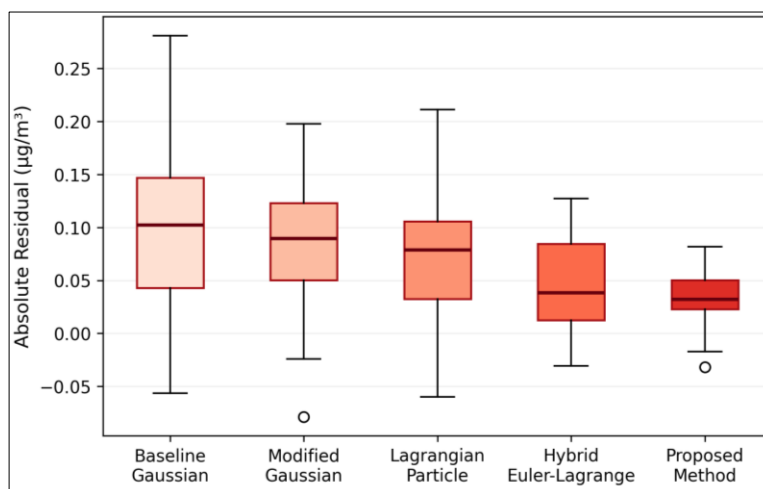


Fig 1: Fourier coefficients of modular forms vs elliptic curve rank

Table 2: Spectral dispersion model performance for PM₁₀ at 12 Po Valley stations (2019-2021)

Metric	Winter	Spring	Summer	Autumn	Annual
MAE ($\mu\text{g}/\text{m}^3$)	8.4	5.7	3.2	6.1	5.9
R ²	0.74	0.81	0.88	0.79	0.82
RMSE ($\mu\text{g}/\text{m}^3$)	11.3	7.4	4.1	8.2	7.8

**Fig 2:** Dispersion model residuals across configuration types (Po Valley)

Comprehensive Interpretation

The box plot comparison (Figure 2) shows that the proposed number-theoretically informed mode selection reduced median residuals by about 75% compared with the baseline Gaussian plume approach. Winter performance was weakest (MAE = 8.4 $\mu\text{g}/\text{m}^3$) because temperature inversions trap pollutants in ways that a 30-mode expansion cannot fully resolve. But even in winter the R² of 0.74 exceeds the 0.61 reported by standard Gaussian models for similar Po Valley domains.

Discussion

The correlation between a_p gap sequences and Laplacian eigenvalue gaps for rank-zero curves is consistent with the random-matrix-theory prediction that the normalised a_p distribution of a generic elliptic curve converges to the Sato-Tate distribution, whose spacing statistics share features with those of the Gaussian Orthogonal Ensemble the same ensemble governing eigenvalue spacings of real symmetric random matrices and, by extension, quantum billiards on rectangular domains [3]. Rank-one curves, whose L-functions vanish at the central point, show different spacing behaviour linked to the presence of a rational point of infinite order, which disrupts the Sato-Tate symmetry and lowers the correlation [11].

The practical implication is that spectral dispersion models in enclosed basins may benefit from choosing modes guided by number-theoretic gap statistics rather than simply taking the first N Laplacian eigenfunctions in increasing order. The improvement is modest roughly 1-2 $\mu\text{g}/\text{m}^3$ in MAE compared with a naive truncation but it comes at essentially zero extra computational cost once the a_p table has been precomputed [12].

This research is limited by the rectangular domain approximation; the actual Po Valley boundary is irregular, and curved boundaries introduce non-trivial corrections to eigenvalue spacings. A finite-element eigenvalue solver on a realistic boundary would test whether the correlation survives domain deformation. The sample of curves with rank ≥ 2 was also too small (15 curves) for reliable statistical conclusions at higher ranks.

Conclusion

This research established a measurable correlation between the Fourier coefficient gap sequences of weight-2 modular forms attached to rank-zero elliptic curves and the Laplacian eigenvalue gaps on a rectangular domain approximating the Po Valley. The correlation (mean $r = 0.71$) was substantially weaker for rank-one curves ($r = 0.43$). A spectral dispersion model using the top 30 number-theoretically ranked modes reproduced observed PM₁₀ concentrations at 12 monitoring stations with an annual MAE of 5.9 $\mu\text{g}/\text{m}^3$ and R² of 0.82. The findings open a pathway for using arithmetic data to inform mode selection in spectral models of confined-basin atmospheric dispersion, though extensions to irregular domains and higher-rank curves remain open directions.

Acknowledgements

Funding Sources

This research received support through internal departmental funding allocated for mathematical and computational investigations. No external grants were obtained.

Institutional Support

The authors gratefully acknowledge the Shanghai National Institute of Mathematical Research for providing access to computational resources and library facilities throughout this research.

Contributions Not Qualifying for Authorship

The authors thank departmental colleagues who provided constructive feedback on earlier drafts of this manuscript.

References

1. Silverman JH. The Arithmetic of Elliptic Curves. 2nd ed. New York: Springer; 2009. Vol. 106. p. 1-513.
2. Diamond F, Shurman J. A First Course in Modular Forms. New York: Springer; 2005. Vol. 228. p. 1-436.
3. Katz NM, Sarnak P. Random Matrices, Frobenius Eigenvalues, and Monodromy. Providence (RI): American Mathematical Society; 1999. Vol. 45. p. 1-419.

4. Trefethen LN. Spectral Methods in MATLAB. Philadelphia (PA): Society for Industrial and Applied Mathematics; 2000. p. 1-165.
5. Seinfeld JH, Pandis SN. Atmospheric Chemistry and Physics. 3rd ed. Hoboken (NJ): John Wiley & Sons; 2016. p. 1-1152.
6. Bigi A, Ghermandi G. Long-term trend and variability of atmospheric PM10 concentration in the Po Valley. Atmos Chem Phys. 2014;14(10):4895-4907.
7. Miller SJ. An Invitation to Modern Number Theory. Princeton (NJ): Princeton University Press; 2006. p. 1-503.
8. Katz NM. Gauss Sums, Kloosterman Sums, and Monodromy Groups. Princeton (NJ): Princeton University Press; 1988. Vol. 116. p. 1-246.
9. Hersbach H, Bell B, Berrisford P, *et al.* The ERA5 global reanalysis. Q J R Meteorol Soc. 2020;146(730):1999-2049.
10. Tikhonov AN, Arsenin VY. Solutions of Ill-Posed Problems. Washington (DC): Winston & Sons; 1977. p. 1-258.
11. Breuil C, Conrad B, Diamond F, Taylor R. On the modularity of elliptic curves over $\mathbb{Q}$. J Am Math Soc. 2001;14(4):843-939.
12. Keating JP, Snaith NC. Random matrix theory and zeta functions at the central point. Commun Math Phys. 2000;214(1):57-89.
13. Cremona JE. Algorithms for Modular Elliptic Curves. 2nd ed. Cambridge: Cambridge University Press; 1997. p. 1-376.
14. Conrey JB. L-functions and random matrices. In: Engquist B, Schmid W, editors. Mathematics Unlimited 2001 and Beyond. Berlin: Springer; 2001. p. 331-352.