Polyhedra, Moduli, and Quantum Geometry

Variational principles, spectral methods, and computation

30 October 2026

Abstract

Hyperbolic and Euclidean polyhedra, Weil–Petersson theory, mapping class groups, Dehn surgery, and quantum invariants admit a common approach through geometric variational principles. This monograph develops the connections between them. Volume and length functions determine geometric equations; their Hessians govern boundary response, numerical solution, and semiclassical amplitudes. Two quantum and semiclassical viewpoints enter: three-manifold invariants built from tetrahedral data, and Colin de Verdière's finite-dimensional approach to surface trace formulas.

The comparisons yield several explicit results. In a prescribed final cone-angle chamber, the leading small-area coefficient of an exact Weil–Petersson volume polynomial is 2^{-d} times the Thurston–Deligne–Mostow volume, where $d = n-3$. At the flat weights, the same complex hyperbolic form is the logarithmic curvature of a continued unit-volume Liouville correlation; its four-point fusion and Hermitian area pairing are calculated explicitly. A normalized quantum tetrahedral integral has a uniform Airy transition on the regular one-parameter family, including the first correction.

The computational strand connects the Casson–Rivin program to finite elements and resistor networks. For the planar prescribed-angle realization problem, the negative volume Hessian is twice the cotangent stiffness matrix. Its Schur complement describes boundary response, and its Gaussian determinant has a rooted-forest expansion. A compatible barrier supplies a path-following iteration bound with Laplacian linear subproblems. On a geometric Dehn-filling branch, a holomorphic Legendre critical value recovers the filling volume to all orders; the inverse volume Hessian determines peripheral response, and the Gaussian surgery factor agrees with the peripheral torsion factor.

On surfaces, character-twisted trace data organize multicurves by length. A length-valued Euler product has isolated unit leading atoms, while a trace identity obstructs a perfect fixed linear selector of simple curves. Weighted counts imply mixing under precise endpoint hypotheses; a marked-torus model gives an explicit global correlation bound on the nonsquare genus-two $(2, 2)$ stratum. Flat-point lower bounds and arithmetic transfer estimates separate the geometric and Möbius obstructions to sharp counting.

These results lead to a conjectural program: uniform quantum gluing and Dehn filling, effective extraction of connected-curve coefficients, weighted counting beyond the marked-torus model, and sharp norm-ball discrepancies. Established inputs, deductions proved here, conditional implications, and numerical evidence are distinguished throughout.

Contents

Introduction: the connections and the results **8**
The principal results . . . 9
The conjectural program . . . 10
Structure and routes through the monograph . . . 10

I Polyhedra and moduli measures 11

1 The volume problem and its normalization **11**

2 Ideal polyhedra, angle cells, and a partition function **11**
2.1 The geometric parametrization . . . 11
2.2 A normalized cell integral . . . 12
2.3 The string-theoretic connection . . . 12

3 The ideal tetrahedron: an explicit volume integral **13**
3.1 Lambda lengths and the ambient cross-ratio . . . 13
3.2 Integration in angle coordinates . . . 14
3.3 What happens at the punctures of the parameter plane? . . . 14

4 Hyperideal polyhedra: an actual local integrand **15**
4.1 From vertices to orthogeodesic lengths . . . 15
4.2 The WP density on a fixed-boundary leaf . . . 15
4.3 General cells and what remains to integrate . . . 16
4.4 A further tree-coordinate result . . . 17

5 Boundary lengths, cone angles, and continuation **17**
5.1 Fixed signature and the polynomial chamber . . . 17
5.2 Pure cone/cusp volumes beyond the first chamber . . . 18
5.3 What continuation means for compact polyhedra . . . 19

6 Equal parameters and the torus–sphere correspondence **19**

7 Euclidean shapes and Deligne–Mostow periods **20**
7.1 The fixed-curvature leaf at the Euclidean end . . . 20
7.2 The same periods describe hypergeometric functions and developed edges . . . 20
7.3 Collision angles and the lattice conditions . . . 22
7.4 Four equal deficits and elliptic periods . . . 22

8 From ideal polyhedra to the Euclidean limit of WP geometry **23**
8.1 A prescribed intrinsic interpolation . . . 23
8.2 The normalized local WP form . . . 23
8.3 The class calculation and an exact final-chamber polynomial . . . 24
8.4 Examples and nonuniformity at balanced weights . . . 25
8.5 Why the complex hyperbolic volumes look different . . . 26

II Variational geometry and computation 27

9 Volume Hessians, finite elements, and computational complexity **27**
9.1 The cotangent form and the correct deformation coordinates . . . 27
9.2 Boundary response, Schur complements, and a Gaussian factor . . . 28
9.3 Why the nonlinear iteration needs a barrier . . . 30
9.4 Flips connect the local matrix to mapping-class geometry . . . 31
9.5 The three-manifold extension and the remaining questions . . . 32

III Quantum geometry, degeneration, and surgery 33

10 A geometric introduction to the quantum constructions **33**
10.1 Six labels: recoupling and tetrahedral weights . . . 33
10.2 The invariants and their volume limits . . . 34
10.3 A complete example: the figure-eight knot . . . 35
10.4 Why dilogarithms occur . . . 36
10.5 Stationary phase, loops, torsion, and degenerate saddles . . . 37
10.6 Flat connections, Chern–Simons theory, and boundary states . . . 38
10.7 Quantizing peripheral holonomy and performing surgery . . . 39
10.8 Liouville theory, conformal blocks, and continuous fusion . . . 40
10.9 Why a screening residue meets flat cone geometry . . . 42
10.10 Parameters, limits, and a route through the calculations . . . 43

11 Polyhedral degeneration and quantum volume asymptotics **44**
11.1 Three different geometric quantities . . . 44
11.2 An explicit small-polyhedron expansion . . . 44
11.3 Schläfli, doubling, and degeneration/regeneration . . . 45
11.4 The established tetrahedral quantum connection . . . 46
11.5 A proposed joint scale and a model integral . . . 46

12 An exact quantum comparison and the boundary measure **47**
12.1 A first-moment identity at the flat weights . . . 47
12.2 Analytic continuation and the equivalent residue . . . 48
12.3 Four points: the exact quantum potential and its measure . . . 49
12.4 An actual boundary torsion measure . . . 49
12.5 The four-point character surface, including a correction . . . 50
12.6 What a normalized quantum kernel would have to produce . . . 51
12.7 The auxiliary tetrahedron and the order of limits . . . 51

13 The four-point residue: fusion, area, and an explicit density **52**
13.1 The two spectral pinches and their coefficients . . . 52
13.2 Why the screening integrals are the Virasoro blocks . . . 55
13.3 The finite fusion matrix . . . 55
13.4 The Hermitian pairing and its direct verification . . . 56
13.5 A closed formula for the four-point density and total volume . . . 57
13.6 Four equal weights: the logarithmic fusion calculation . . . 58
13.7 What the calculation establishes . . . 59

14 The normalized quantum-group comparison and the compact limit **59**
14.1 Fixing the Teschner–Vartanov normalization . . . 59
14.2 The double-Gamma factors and the finite sine matrix . . . 60
14.3 Symmetric normalization and the residue prescription . . . 61
14.4 The actual screened Gram matrix has rank two at the limit . . . 62
14.5 The screened classical saddle also degenerates . . . 63
14.6 A genuine compact family: the regular tetrahedron . . . 63
14.7 The regular classical fold and its scale . . . 64
14.8 Correction of the integration-coordinate dictionary . . . 65

15 A uniform quantum Airy transition for the regular family **66**
15.1 The exact normalization and the result . . . 66
15.2 A uniform double-sine estimate from its real integral . . . 67
15.3 The smooth quantum phase and amplitude . . . 68
15.4 The full contour and its tail estimate . . . 68
15.5 Airy localization and the first correction . . . 69
15.6 A numerical check of the exact integral . . . 70
15.7 Interpretation and scope . . . 71

16 Dehn filling, variational response, and quantum surgery **71**
16.1 Peripheral conventions and the two small parameters . . . 71
16.2 What the linear term does to real volume optimization . . . 72

16.3 The cusp shape in the inverse volume Hessian . . . 73
16.4 An exact holomorphic critical-value identity . . . 73
16.5 The Neumann–Zagier coefficients through sixth order . . . 74
16.6 Computing the quartic coefficient from angle-volume derivatives . . . 75
16.7 The closed quantum-volume statement and its local consequence . . . 76
16.8 The Gaussian factor and the torsion surgery formula . . . 77
16.9 A figure-eight benchmark and the scope of the conclusion . . . 77
16.10 The symplectic pairing behind the calculation . . . 79

IV Curves, trace formulas, and quotients 84

17 Unfolding and the problem of finding constant series 84
17.1 What a volume computation determines . . . 84
17.2 Operations that preserve constancy . . . 85
17.3 Other established constant series . . . 85

18 Figure-eight curves and simple lifts 86
18.1 The pants trace relation . . . 86
18.2 What the local double cover accomplishes . . . 86
18.3 The two types on a punctured torus . . . 87
18.4 A positive single-length kernel for all figure-eights . . . 87

19 Cutting into pants and varying twists 88
19.1 What data a cut curve retains . . . 88
19.2 An exact local distance function . . . 88
19.3 Holonomy and first variation . . . 89
19.4 Unfolding a twist lattice . . . 89

20 An exact torus twist calculation and a constant orbit series 90
20.1 The length function and its integral . . . 90
20.2 A positive partition of unity . . . 91

21 Matching across cuffs: positions and directions 91

22 Cusps, McShane gaps, and the orthospectrum 92
22.1 Which geodesics enter a cusp? . . . 92
22.2 The family used in the gap argument . . . 92
22.3 An exact truncated cusp-arc formula . . . 93
22.4 The precise relation to Basmajian's identity . . . 94

23 Simple multicurve coefficients of a twisted trace formula 95
23.1 Broken geodesics and the effect of reconnection . . . 95
23.2 The finite coefficient distributions . . . 95
23.3 Length triangularity and the first atom . . . 96
23.4 An exponential coefficient bound from bracelet positivity . . . 98
23.5 Convergent heat and Laplace transforms . . . 99
23.6 All iterates of a simple curve . . . 101
23.7 A figure-eight family and an exact cancellation . . . 102
23.8 Metric variation and the twist derivative . . . 103
23.9 The logarithmic product and primitive extraction . . . 103
23.10 Scope and further questions . . . 104

24 Quantitative counting and coefficient cancellation 106
24.1 Effective counting and the meaning of an explicit saving . . . 106
24.2 A constructive lattice benchmark for all multicurves . . . 106
24.3 Primitivity, connectivity, and the torus benchmark . . . 107
24.4 A fixed-surface obstruction to uniform first-atom separation . . . 108
24.5 A constructive uniform isolation estimate . . . 109
24.6 A finite coefficient that selects one support . . . 111

24.7 An explicit criterion for passing from smooth to sharp counts . . . 112
24.8 The remaining analytic step . . . 113

25 Global coefficient filters and a length-valued Euler product 114
25.1 An explicit filter for the entire one-crossing sector . . . 114
25.2 Why a fixed linear filter cannot finish the job . . . 115
25.3 Retaining lengths: an Euler product with a unit diagonal . . . 116
25.4 Absolute convergence and spectral input . . . 117
25.5 What these constructions settle, and what remains . . . 117

V Mapping-class dynamics and mixing 118

26 When counting errors imply mixing 118
26.1 A direct converse in the compact Selberg setting . . . 118
26.2 A quantitative converse in homogeneous dynamics . . . 119
26.3 A converse for a finite Markov model . . . 119
26.4 Which scalar count is being estimated? . . . 121
26.5 A qualitative converse from full geodesic counting . . . 121
26.6 Localized counts and correlations . . . 121
26.7 Nonescape and the counting time parameter . . . 122

27 Filling-pair coefficients and a weighted mixing converse 124
27.1 Two independent character variables retain the pair . . . 124
27.2 Filling pairs as actual coordinates for the flow . . . 124
27.3 The correct quotient and the two different counting scales . . . 125
27.4 The existing square-tiled counting theorem and the new target . . . 126
27.5 What refining the coefficient mesh actually proves . . . 126
27.6 An exact weighted bisector kernel . . . 127
27.7 Removing the time average with a controlled loss . . . 128
27.8 A derivative endpoint weight preserves the full exponent . . . 130
27.9 Comparison with the usual bisector count . . . 131
27.10 A completely explicit arithmetic two-endpoint calculation . . . 132

28 Genus two: compact polyhedra and a marked-torus model 134
28.1 Six branch points and the compact polyhedron . . . 134
28.2 An explicit formula for every quadratic differential . . . 135
28.3 The nonsquare stratum with two double zeros . . . 135
28.4 The pillowcase cover and two marked torus points . . . 136
28.5 A homogeneous finite cover . . . 137
28.6 Haar measure and the period lattice . . . 138
28.7 Fourier decomposition and temperedness . . . 139
28.8 A global correlation bound with an explicit constant . . . 140
28.9 An exact finite lattice sum with two endpoint weights . . . 141
28.10 What the marked-torus calculation settles . . . 142

29 From simple-curve coefficients to controlled observables 143
29.1 The norm-ball obstruction and the correct error scale . . . 143
29.2 The Fourier and spectral meaning of the obstruction . . . 145
29.3 A coefficient observable from primitive pole-to-pole arcs . . . 146
29.4 Uniform rotation norms, including a sharp hyperbolic cutoff . . . 147
29.5 Control of the terminal derivative . . . 148
29.6 What the bridge retains . . . 149
29.7 Affine displacement sums: an exact isometric model . . . 149
29.8 Two marks: the rank-one divergence and its removal . . . 150
29.9 An exact Gaussian two-mark correlation . . . 152
29.10 Sharp boundaries and the rotation norm . . . 152

VI Sharp counting and arithmetic 153

30 Diagnosing the sharp remainder without fitting its main term 153
30.1 The normalized all-point discrepancy and the logarithmic loss . . . 153
30.2 A rational-polygon comparison . . . 154
30.3 An exact test independent of the leading coefficient . . . 155
30.4 A quantitative row-crossing formula . . . 156
30.5 Model flatness profiles and what they do not determine . . . 157
30.6 An exact row diagnostic and its precision requirements . . . 158
30.7 A trace-based numerical experiment . . . 159
30.8 A recent qualification of irrational flatness . . . 160

31 Arithmetic cancellation and the geometric remainder 160
31.1 A transfer estimate with both inputs explicit . . . 160
31.2 An unconditional sublinear consequence of the Dini condition . . . 161
31.3 Fixed zero-free half-planes and the recent quasi-RH claims . . . 162
31.4 Logarithmic geometric savings . . . 162
31.5 Power-saving geometric discrepancies and established benchmarks . . . 163
31.6 The remaining geometric obstruction . . . 164

VII Open problems and conjectural directions 165

32 Questions suggested by the calculations 165

A A trace-coordinate check of the WP factor 167

B Conventions at a glance 168

References 169

List of Figures

1 The geometric and analytic connections 8
2 Euclidean, compact, ideal, and hyperideal regimes 21
3 The cotangent Hessian and its finite-element interpretation 29
4 A diagonal exchange and the Delaunay energy comparison 31
5 Coalescing stationary points and the Airy normal form 38
6 Dehn filling in peripheral phase space 74
7 Crossing resolution and the first coefficient atom 98
8 The genus-two and marked-torus double covers 137
9 A flat supporting cap and a primitive-row burst 157

Introduction: the connections and the results

The organizing question is how information passes between apparently distant descriptions of geometric objects. A polyhedron can be studied through its intrinsic metric, its dihedral angles, its volume, or the conformal geometry of its boundary. A surface can be studied through geodesics, periods, character varieties, or a mapping-class quotient. Each description makes a different operation accessible: integration, optimization, gluing, quantization, or computation. The purpose of this monograph is to make the passages between these descriptions explicit.

Three operations recur. *Variation* produces a geometric equation and a Hessian. *Elimination of internal variables* produces a boundary response; the same Schur complement occurs in electrical networks, constrained volume optimization, and Gaussian integration. *Degeneration* tests the uniformity of a comparison, producing Euclidean limits, Airy transitions, cusp phenomena, and changes in numerical conditioning. These operations explain why the subjects belong together. Each comparison still has its own geometric space, measure, and normalization.

The computational interpretation is particularly useful. The companion study on Duffin networks and ideal polyhedra [116] identifies a volume Hessian with a cotangent finite-element matrix. Incorporating that identity here connects polyhedral realization to sparse linear algebra, and connects both to the later calculations of peripheral response and semiclassical determinants. Figure 1 gives a map of these connections. The arrows denote specified operations; the detailed statements identify which are exact, asymptotic, or conditional.

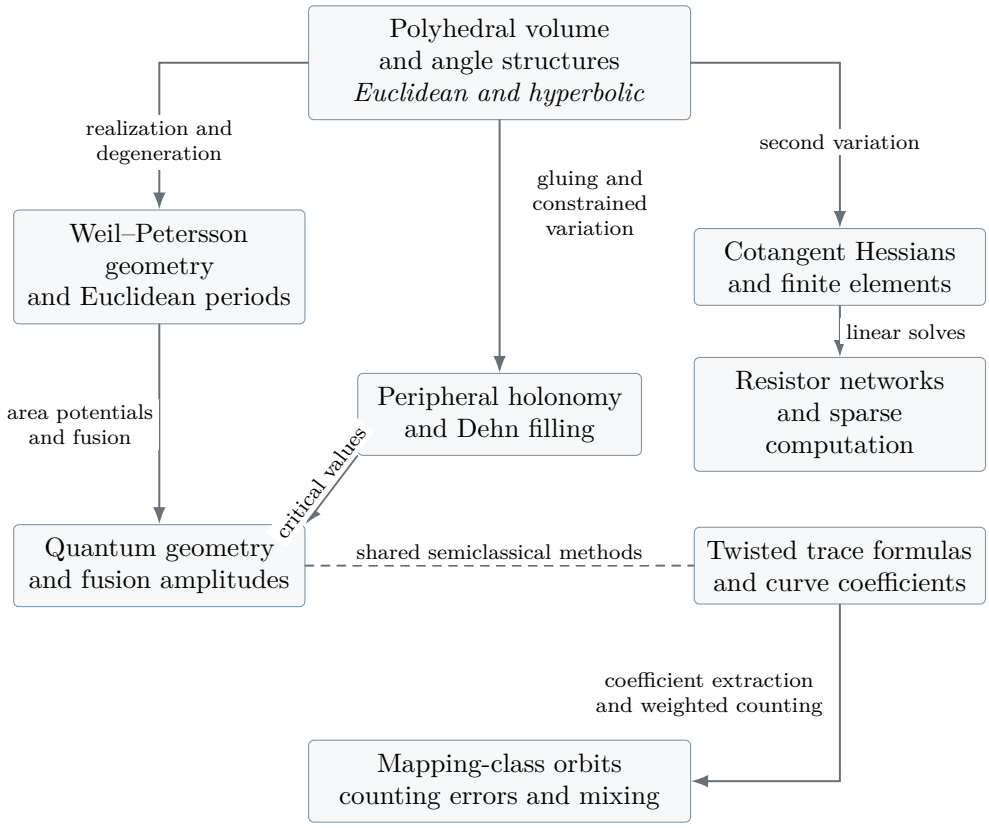


Figure 1: A guide to the constructions developed in the monograph. Solid arrows name operations whose domains and hypotheses are stated in the text; they do not assert equivalences of entire theories. The dashed link records the use of semiclassical analysis in two different settings: quantum three-manifold constructions and surface trace formulas. The counting-to-mixing implications retain their endpoint-weight and nonescape hypotheses.

The principal results

The main conclusions fall into six groups. Their significance lies both in the individual formulas and in the transfer of information between the subjects. We give their precise scopes here; the cited sections contain the hypotheses, proofs, and original sources of the inputs.

1. **Polyhedral geometry and moduli volumes.** The ideal-polyhedral cells give a normalized partition function for genus-zero WP volumes. For fixed admissible deficit weights, the final-chamber identity (8.9) gives an exact class polynomial, and (8.11) identifies $2^{-d}V_{\mathrm{CH}}(\mu)$ as the coefficient of A^d in its small-area volume expansion, with $d = n - 3$. This makes the Euclidean period geometry a computable limit of the hyperbolic moduli-volume theory. The chamber, labels, and orbifold conventions are part of the formulas.

2. **Quantum periods and a uniform degeneration.** The finite-b identity (12.6) identifies the Thurston–Deligne–Mostow form with the logarithmic curvature of a continued unit-volume Liouville correlation. The four-point residue calculation determines its fusion matrix and Hermitian area pairing. Proposition 15.1 gives a separate uniform Airy theorem, with its first correction, for the specified pole-cleared tetrahedral integral on the regular one-parameter family.

3. **Volume optimization, finite elements, and complexity.** Proposition 9.1 identifies the negative volume Hessian with twice the cotangent stiffness matrix in the disk's leading–trailing angle coordinates. Equations (9.9) and (9.11) give its boundary elimination and rooted-forest determinant. Proposition 9.2 supplies the compatible barrier and yields a quantitative path-following bound after initial centering. The distinction between a fast Laplacian solve and the full nonlinear realization problem makes the remaining feasibility, precision, and certification questions explicit.

4. **Dehn filling through a critical value.** Proposition 16.1 identifies the local filling volume with a holomorphic Legendre critical value as an equality of analytic germs. Thus every classical long-slope coefficient is fixed, including the Neumann–Zagier expansion. The inverse volume Hessian recovers the marked cusp quadratic form, and the Gaussian factor matches the peripheral factor in torsion surgery. The transfer to a full quantum invariant is conditional on the representation, cycle, noncancellation, and competing-sector estimates stated in Proposition 16.2.

5. **Length-sensitive reconstruction from trace data.** Proposition 25.2 assigns each multicurve an isolated unit first atom in a length-valued Euler product. A fixed linear character functional cannot preserve every simple primitive trace while annihilating every nonsimple trace; the explicit obstruction is (25.7). Keeping the length and twisting data therefore matters. The construction supplies exact coefficient objects and finite isolation bounds, while uniform stability and effective extraction remain additional analytic questions.

6. **Effective dynamics and the boundary of sharp counting.** Theorem 27.4 is a converse from a specified weighted two-endpoint count to mixing, with control of derivative endpoint weights. Theorem 28.1 gives an explicit global correlation bound for the marked-torus model of the nonsquare genus-two $(2, 2)$ stratum. The strict-convexity, inverse-flatness, and arithmetic transfer results then identify the separate geometric and Möbius inputs needed for sharp primitive counting. The proposed $O_X(L)$ remainder is a question to be decided, not a conclusion of these results.

The conjectural program

The larger quantum-volume problem asks whether controlled local tetrahedral asymptotics can be assembled into uniform asymptotics of three-manifold invariants, including Dehn filling. The local critical value and the Gaussian surgery factor specify two necessary matching conditions. Global gluing and contour estimates are the next steps. On the surface side, the coefficient constructions suggest a route from spectral data to effective connected-curve counting. The weighted mixing converse specifies the estimate that route would have to produce, including endpoint regularity and cusp control.

Two further questions join the computational and counting strands to this program. Can geometric angle bounds and verified linear solves turn variational existence into useful complexity and certification theorems beyond the planar model? Can the fine geometry of a torus norm ball provide the quantitative discrepancy modulus needed to decide the linear-error question? Section 32 states these targets alongside the other open problems, without assigning an answer where the evidence does not yet support one.

Structure and routes through the monograph

The exposition is divided into seven parts. The quantum introduction is written for geometers who do not assume familiarity with state sums, recoupling coefficients, conformal blocks, or stationary-phase terminology. Later calculations can be read after that introduction or reached directly through the result map above.

Part	Subject	Main passage
I	Polyhedra and moduli measures	Angle cells, WP volume, and Euclidean periods
II	Variational geometry and computation	Volume Hessians, finite elements, and complexity
III	Quantum geometry, degeneration, and surgery	Tetrahedral amplitudes, fusion, and peripheral response
IV	Curves, trace formulas, and quotients	Geodesic identities and length-valued coefficients
V	Mapping-class dynamics and mixing	Weighted orbit counts and the marked-torus model
VI	Sharp counting and arithmetic	Flat norm balls, diagnostics, and Möbius cancellation
VII	Open problems and conjectural directions	The estimates and geometric comparisons still required

A geometric route runs through Parts I–II and the classical Dehn-filling calculation in Part III. A quantum route begins with Section 10, then follows the period, fusion, Airy, and surgery comparisons. A surface and spectral route begins with the geodesic identities in Part IV and continues through Parts V–VI. For computation, Section 9 leads to the inverse-Hessian response in Section 16 and the counting diagnostics in Section 30. These routes share definitions and normalizations rather than requiring every calculation to be read in order.

Provenance and mathematical status. This research monograph grew from a mathematical discussion and includes both exposition and deductions from the cited literature. The supplied `igorrivin/ipoly` snapshot contributes the companion manuscript and its numerical scripts; Section 9 incorporates its core variational and computational arguments. Numerical checks are identified where they occur. A theorem proved for a normalized local integral, a specified stratum, or a conditional weighted count retains that scope when it is used elsewhere in the monograph.

Part I
Polyhedra and moduli measures

Polyhedral realization supplies coordinates; the measure comparison turns them into volume formulas. Fixed-deficit Euclidean shapes then reveal a period geometry at the flat limit.

1 The volume problem and its normalization

Let $\mathcal{M}_{g,n}(L_1,\dots,L_n)$ denote the moduli space of oriented hyperbolic surfaces of genus g with n labeled, unparametrized geodesic boundary components of prescribed lengths. A zero length denotes a cusp. Throughout, the curvature is -1, the parameters are fixed before taking a volume, and $2g-2+n>0$. Put

$$d = 3g-3+n, \qquad V_{g,n}(\mathbf{L}) = \int_{\mathcal{M}_{g,n}(\mathbf{L})} \frac{\omega_{\mathrm{WP}}^d}{d!}.$$

We use the Fenchel–Nielsen convention

$$\omega_{\mathrm{WP}} = \sum_{j=1}^{d} \mathrm{d}\ell_j \wedge \mathrm{d}\tau_j, \qquad \text{one full Dehn twist: } \tau_j \longmapsto \tau_j + \ell_j. \tag{1.1}$$

Volumes are orbifold, or stack, volumes: generic ineffective stabilizers are included. In particular,

$$V_{0,3} = 1, \qquad V_{0,4}(\mathbf{L}) = 2\pi^2 + \frac{1}{2}\sum_{i=1}^{4} L_i^2, \qquad V_{1,1}(L) = \frac{L^2+4\pi^2}{48}. \tag{1.2}$$

The effective torus volume is twice the last expression. This convention matters both in unfolding and in the torus–sphere correspondence; see Do–Norbury [55] and Section 6.

The original geometric proposal has three separate ingredients:

1. realize an intrinsic surface by a canonical convex polyhedron;
2. describe the allowed polyhedra by cells, ideally linear polytopes;
3. identify the pullback of $\omega_{\mathrm{WP}}^d/d!$ on each cell.

The realization and the cell decomposition identify the integration domain. They do not, by themselves, identify its measure. In the ideal genus-zero case the measure comparison is particularly favorable; in the hyperideal case a useful local density is still available, but it is generally no longer constant in the natural angle variables.

There are also two different cross-ratios in the four-point problem. Four ideal vertices on $\partial\mathbb{H}^3$ have an *ambient* cross-ratio. The intrinsic metric on the polyhedral boundary has a conformal uniformization, and its four punctures have a *conformal* cross-ratio. The realization theorem relates the two parameter spaces, but these two numerical coordinates are not generally equal. Keeping them distinct is essential to the first explicit integral below.

2 Ideal polyhedra, angle cells, and a partition function

2.1 The geometric parametrization

In the polyhedral cell discussion take $n \geq 4$; the three-ended sphere is the zero-dimensional pants base case.

Rivin's intrinsic realization theorem identifies a complete finite-area hyperbolic metric on an n-punctured sphere with the boundary metric of a convex ideal polyhedron, uniquely up to ambient isometry. Flat doubles of ideal polygons occur as limiting configurations [85]. For oriented moduli one retains the orientation of the boundary and uses orientation-preserving equivalences.

Fix a combinatorial polyhedron and write θ_e for its exterior dihedral angles. In the dual graph, the ideal conditions are

$$0 < \theta_e < \pi, \tag{2.1}$$

$$\sum_{e \in \partial f} \theta_e = 2\pi \qquad \text{for every dual face } f, \tag{2.2}$$

$$\sum_{e \in C} \theta_e > 2\pi \qquad \text{for every simple dual circuit not bounding a face.} \tag{2.3}$$

Together with the combinatorial hypotheses on a convex polyhedron, these are Rivin's characterization [86]. Thus the admissible angle assignments form an open convex polytope in an affine subspace. Generic polyhedra have triangular faces; their dual maps are trivalent and have $3n - 6$ edges. The n perimeter conditions leave dimension $2n - 6$, as required.

An edge of the dual graph has metric length θ_e. The resulting metric map has face perimeters 2π and no short nonfacial circuit. This is the genus-zero instance of a 2π-irreducible metric map. It gives the precise combinatorial object behind the proposed sum over polyhedral types.

2.2 A normalized cell integral

For a dual ribbon graph G with labeled faces, define its constrained Lebesgue measure by disintegrating $\prod_e \mathrm{d}\theta_e$ with respect to the face-perimeter map. Symbolically, at perimeters $\boldsymbol{\alpha}$,

$$\mathrm{d}\nu_G(\boldsymbol{\alpha}) = \left(\prod_{e \in E(G)} \mathrm{d}\theta_e \right) \left(\prod_{i=1}^{n} \delta \left(\sum_{e \in \partial f_i} \theta_e - \alpha_i \right) \right). \tag{2.4}$$

An edge is counted with its boundary multiplicity. This notation specifies a disintegrated measure, rather than Euclidean hypersurface measure with an unstated coarea factor.

The ideal cell description is known to give a volume calculation, beyond the possibility suggested by the realization theorem. Charbonnier–David–Eynard compare the angle measure with the WP form through lambda lengths [50, Section 3.2]. In Budd's normalization of metric-map volume,

$$\boxed{V_{0,n}(0, \ldots, 0) = 2^{n-2} \sum_{[G]} \frac{1}{|\operatorname{Aut}(G)|} \int_{\mathcal{P}_G} \mathrm{d}\nu_G(2\pi, \ldots, 2\pi).} \tag{2.5}$$

Here $\mathcal{P}_G$ imposes positivity and irreducibility, G ranges over oriented trivalent maps with labeled faces, and automorphisms preserve this data. Nontrivalent strata have zero volume. Budd gives this interpretation and the normalization explicitly [48, Section 1.3].

Thus there is a genuine partition function over polyhedral combinatorial types. Formula (2.5) is a finite cellwise integral at fixed n; it is not a sum over geodesics. The factors in it encode three choices that should never be suppressed: labels, orientation, and the normalization of the constrained measure. Section 3 checks them directly for $n = 4$.

2.3 The string-theoretic connection

Saadi–Zwiebach's closed-string construction uses equal-perimeter polyhedral diagrams and quadratic differentials [90]. Zwiebach's minimal-area formulation imposes a lower bound 2π

on nontrivial closed curves and uses uniqueness to control the covering of moduli space [98]. These give a concrete historical counterpart to the suggestion that such cells should be used for integration.

Several identifications nevertheless require care. A metric ribbon graph can be thickened into a flat surface using a Jenkins–Strebel differential. A hyperbolic polyhedron has an intrinsic hyperbolic boundary metric and also extrinsic dual data. These constructions may share the same combinatorial and length description without giving the same metric or the same conformal coordinate on moduli. Likewise, a string amplitude includes its specified integrand; a decomposition used for string integration does not automatically select the WP measure. Equation (2.5) uses an explicit measure comparison, not just the common vocabulary of polyhedra.

3 The ideal tetrahedron: an explicit volume integral

3.1 Lambda lengths and the ambient cross-ratio

Take the four labeled ideal vertices to be $0, 1, \infty, z$, where $z = x + iy \in \mathbb{C} \setminus \{0, 1\}$. For $y \neq 0$ their convex hull is a genuine ideal tetrahedron. The two half-planes describe its two orientation types.

For a decorated ideal triangulation, set $\lambda_e = e^{\delta_e/2}$, where δ_e is the signed distance between the chosen horocycles along the edge. An overall fixed multiplier in the definition of λ does not change its logarithmic differential. With cyclically ordered edge lengths a, b, c on each oriented triangle, Penner's formula in convention (1.1) is

$$\omega_{\mathrm{WP}} = -2 \sum_{\triangle} (\mathrm{d}\log a \wedge \mathrm{d}\log b + \mathrm{d}\log b \wedge \mathrm{d}\log c + \mathrm{d}\log c \wedge \mathrm{d}\log a). \tag{3.1}$$

This is a form on decorated Teichmüller space that descends after forgetting the horocycles [83]; its magnitude agrees with the cuspidal limit of Mondello's arc-coordinate formula [77].

The four decorations can be chosen so that

$$\lambda_{0\infty} = \lambda_{1\infty} = \lambda_{z\infty} = \lambda_{01} = 1, \qquad \lambda_{0z} = r = |z|, \quad \lambda_{1z} = s = |1 - z|. \tag{3.2}$$

For example, use equal Euclidean diameters for the three finite horospheres in the upper-half-space model and adjust their common scale together with the horosphere at infinity. The usual formula for the signed distance then gives (3.2).

Every face containing ∞ has at most one varying lambda length, so its contribution in (3.1) vanishes in this section of the decoration bundle. The remaining face has lengths $1, r, s$. Consequently the positive WP measure is

$$\boxed{\mathrm{d}V_{\mathrm{WP}} = 2\,|\mathrm{d}\log r \wedge \mathrm{d}\log s| = \frac{2|y|}{|z|^2|1 - z|^2}\ \mathrm{d}x\ \mathrm{d}y.} \tag{3.3}$$

The Jacobian calculation is elementary:

$$\mathrm{d}\log r = \frac{x\ \mathrm{d}x + y\ \mathrm{d}y}{r^2}, \qquad \mathrm{d}\log s = \frac{(x-1)\ \mathrm{d}x + y\ \mathrm{d}y}{s^2}.$$

Their wedge is $y\ \mathrm{d}x \wedge \mathrm{d}y/(r^2 s^2)$.

3.2 Integration in angle coordinates

For $y > 0$, let A, B, C be the interior angles of the Euclidean triangle with vertices $0, 1, z$. They satisfy

$$A > 0, \quad B > 0, \quad C > 0, \quad A + B + C = \pi.$$

They are also the three interior dihedral angles of the ideal tetrahedron, each repeated on the opposite edge. Direct differentiation gives

$$|\mathrm{d}A \wedge \mathrm{d}B| = \frac{y}{r^2 s^2}\, \mathrm{d}x\, \mathrm{d}y, \qquad \text{hence} \qquad \mathrm{d}V_{\mathrm{WP}} = 2\, \mathrm{d}A\, \mathrm{d}B.$$

One orientation sheet therefore contributes

$$\int_{\substack{A>0,\ B>0 \\ A+B<\pi}} 2\, \mathrm{d}A\, \mathrm{d}B = \pi^2. \tag{3.4}$$

Including the other sheet gives the promised direct calculation:

$$\boxed{V_{0,4} = 2\pi^2.} \tag{3.5}$$

The real configurations form the flat locus and contribute zero two-dimensional WP measure.

There is also a useful check on (2.5). In a tetrahedral angle cell, the four perimeter equations force opposite exterior angles to be equal, with three values a, b, c satisfying $a + b + c = 2\pi$. Disintegrating the six edge variables by the four perimeter equations gives $\frac{1}{2}\, \mathrm{d}a\, \mathrm{d}b$. Since $a = \pi - A$ and $b = \pi - B$, this is $\frac{1}{2}\, \mathrm{d}A\, \mathrm{d}B$. There are two oriented cells, each with trivial label-preserving automorphism group. The factor $2^{4-2} = 4$ in (2.5) converts their total map volume $\pi^2/2$ to $2\pi^2$.

3.3 What happens at the punctures of the parameter plane?

Finite volume requires integrability, not pointwise decay of a coordinate density. Near $z = 0$, write $z = \rho e^{i\vartheta}$. Then

$$\frac{2|y|}{|z|^2|1-z|^2} = \frac{2|\sin\vartheta|}{\rho}\,(1 + O(\rho)), \qquad \mathrm{d}x\, \mathrm{d}y = \rho\, \mathrm{d}\rho\, \mathrm{d}\vartheta.$$

The singularity is integrable. The same is true at 1. At infinity the density is $O(|z|^{-3})$, again integrable against Euclidean area. Thus the requested decay is best formulated as a condition on the measure, not on its density without its coordinate area element.

Remark 3.1 (The coordinate is not conformal). Formula (3.3) is the pullback to the ambient ideal-vertex coordinate. It is not a formula for WP measure in the intrinsic conformal cross-ratio. In particular, its vanishing on the real locus does not mean that the WP form degenerates at an ordinary interior point of moduli space. The polyhedral parametrization is not a smooth coordinate across that locus. The Delaunay literature also distinguishes the embedding coordinate from uniformization [50, Section 3.3.1].

Nor do two ordinary closed angle triangles describe the full topology of the compactification without additional data. As $y \downarrow 0$ with $0 < x < 1$, two of the triangle angles tend to zero and the third to π, independently of x. A whole family of flat quadrilaterals approaches one angle-simplex vertex. Retaining that modulus requires a refinement of the boundary, such as a real blow-up. This has no effect on the integral over the open cells.

4 Hyperideal polyhedra: an actual local integrand

4.1 From vertices to orthogeodesic lengths

A strictly hyperideal vertex lies outside the closed ball in the projective model, while every edge meets the ball. Its polar plane cuts off the vertex. After all such truncations, the original faces form a bordered hyperbolic surface; the truncation polygons give its geodesic boundaries. Ideal vertices can be included by a limiting procedure and give cusps.

For a hyperideal tetrahedron choose unit spacelike representatives $v_1, \dots, v_4$ in Minkowski space of signature $(3,1)$, with the appropriate common time orientation. Set

$$q_{ij} = -\langle v_i, v_j \rangle = \cosh a_{ij} > 1. \tag{4.1}$$

Here a_{ij} is the length of the original edge between the two polar truncation planes. It is an orthogeodesic of the intrinsic bordered surface. The six a_{ij} specify the vertex configuration up to ambient isometry on the admissible tetrahedral domain.

Let b_i^{jk} be the side of the truncation triangle at vertex i that lies in the original face ijk. The right-angled hexagon cosine rule gives

$$\cosh b_i^{jk} = \frac{\cosh a_{ij} \cosh a_{ik} + \cosh a_{jk}}{\sinh a_{ij} \sinh a_{ik}}, \qquad L_i = \sum_{\{j,k\} \subset \{1,2,3,4\} \setminus \{i\}} b_i^{jk}. \tag{4.2}$$

The four L_i are intrinsic boundary lengths. They are not the sums of the exterior dihedral angles around the hyperideal vertices.

These formulas also describe the domain concretely. The Gram matrix

$$G_{ii} = 1, \qquad G_{ij} = -q_{ij} \quad (i \neq j) \tag{4.3}$$

must have signature $(3,1)$. With all $q_{ij} > 1$, the strict tetrahedral condition is $\det G < 0$. For example, the 3×3 principal minors are negative and have exactly one negative eigenvalue; interlacing and the sign of the full determinant give the stated signature. The corresponding truncation triangles satisfy strict triangle inequalities.

It is possible to glue right-angled hexagons from a wider set of positive arc-length assignments and obtain a valid intrinsic bordered surface. Such arcs need not be the edges of its convex hyperideal realization. Thus the polyhedral domain must be imposed when the purpose is to integrate a canonical polyhedral cell.

4.2 The WP density on a fixed-boundary leaf

Mondello's Poisson formula for arc lengths has the boundary kernel

$$K_L(s) = \frac{\sinh(L/2 - s)}{\sinh(L/2)}, \qquad 0 < s < L. \tag{4.4}$$

For two different arcs e, f, the bracket in convention (1.1) is one half the sum of these kernels over their common boundary endpoints, with oriented boundary distances from the endpoint of e to that of f. Changing the order changes the sign, since $K_L(L - s) = -K_L(s)$. The factor $1/2$ results from evaluating the ordered wedge sum in [77, Theorem 4.1].

Restrict now to the two-dimensional leaf where $L_1, \dots, L_4$ are fixed. The arcs a_{12}, a_{13} share just the boundary at vertex 1, and the relevant distance is $b = b_1^{23}$, up to orientation. Inverting the two-dimensional Poisson tensor gives

$$\boxed{\mathrm{d}V_{\mathrm{WP}} = 2 \frac{\sinh(L_1/2)}{\sinh(L_1/2 - b_1^{23})} \left| \mathrm{d}a_{12} \wedge \mathrm{d}a_{13} \right|.} \tag{4.5}$$

Since the three truncation sides form a genuine triangle, $b_1^{23} < L_1/2$. Thus the denominator is positive, the bracket is nonzero, and these two arc lengths are local coordinates on this leaf.

Formula (4.5) is an actual expression for the volume integrand in terms of the hyperideal geometry. Its remaining variables are functions on the fixed-L leaf; they cannot be held independently fixed while a_{12}, a_{13} vary.

For an expression directly in the vertex Gram entries, put

$$D = q_{12}^2 + q_{13}^2 + q_{23}^2 + 2q_{12}q_{13}q_{23} - 1.$$

Equation (4.2) gives

$$\cosh b = \frac{q_{12}q_{13} + q_{23}}{\sqrt{q_{12}^2 - 1}\sqrt{q_{13}^2 - 1}}, \qquad \sinh b = \frac{\sqrt{D}}{\sqrt{q_{12}^2 - 1}\sqrt{q_{13}^2 - 1}}.$$

Substituting in (4.5) and using $\mathrm{d}q_{ij} = \sinh a_{ij}\ \mathrm{d}a_{ij}$ yields

$$\boxed{\mathrm{d}V_{\mathrm{WP}} = \frac{2\,|\mathrm{d}q_{12} \wedge \mathrm{d}q_{13}|}{q_{12}q_{13} + q_{23} - \coth(L_1/2)\sqrt{D}}.} \tag{4.6}$$

The positivity follows from the same truncation-triangle inequality.

The cusp limit requires renormalization. As a positive boundary shrinks, the raw distances to its truncation plane diverge; subtract the endpoint terms corresponding to a chosen limiting horocycle before comparing with lambda lengths. If $s/L \to \rho$, the kernel itself has the regular limit $K_L(s) \to 1 - 2\rho$. A fixed-perimeter horocycle gauge is different from the four-unit-lambda gauge in (3.2); changing that gauge changes coordinate Jacobians, while preserving the descended WP form. This is the cuspidal limiting procedure underlying [77, Theorem 4.5].

4.3 General cells and what remains to integrate

The same procedure applies to any arc system. On a fixed-boundary leaf, choose independent arc coordinates $x_1, \dots, x_{2d}$ and form the antisymmetric matrix

$$B_{ef} = \{x_e, x_f\}_{\mathrm{WP}}.$$

If it is nonsingular, then

$$\boxed{\mathrm{d}V_{\mathrm{WP}} = \frac{|\,\mathrm{d}x_1 \cdots \mathrm{d}x_{2d}|}{|\operatorname{Pf} B|} = \frac{|\,\mathrm{d}x_1 \cdots \mathrm{d}x_{2d}|}{\sqrt{\det B}}.} \tag{4.7}$$

This is simply the top exterior power of the inverse Poisson matrix. It turns a local arc-bracket formula into a local volume formula in every dimension.

Bao–Bonahon characterize hyperideal dihedral angles by linear inequalities [46]. Intrinsic realization results for complete genus-zero metrics with funnel ends supply the corresponding global geometric picture; one attaches a funnel to each positive boundary and then truncates the realization along its polar planes [91, 57]. The ideal and generalized realization statements have different historical scopes: Rivin's work supplies ideal existence and generalized uniqueness, while the broader existence statement is recorded in these later references.

For fixed L_i, however, the angle cell is cut out further by the nonlinear equations (4.2). The local density is therefore more complicated than in the ideal case. The known total

$$V_{0,4}(L_1, L_2, L_3, L_4) = 2\pi^2 + \tfrac{1}{2}\sum_i L_i^2$$

provides a stringent target for a direct evaluation of these cell integrals. No such full hyperideal evaluation is carried out in this monograph.

Budd's metric-map polynomials obey, for $g = 0, 1$,

$$V_{g,n}^{(2\pi)}(\alpha_1, \dots, \alpha_n) = 2^{2-2g-n} V_{g,n}(\sqrt{\alpha_1^2 - 4\pi^2}, \dots, \sqrt{\alpha_n^2 - 4\pi^2}). \tag{4.8}$$

This does not extend to $g \geq 2$. Nor does the displayed substitution describe the actual relation between hyperideal angle perimeters and truncation boundary lengths [48, Corollary 2 and Section 1.3]. It is useful evidence of a connection, with a separate geometric question behind it.

4.4 A further tree-coordinate result

There is now an additional concrete integration route for genus zero with a distinguished cusp. Budd–Meeusen–Zonneveld construct a spine bijection from a full-measure subset to explicitly decorated tree polytopes [49, Theorem 1]. Writing m for the total number of ends, their statement has the form

$$\mathcal{M}_{0,m}^{\circ}(0, \mathbf{L}) \longleftrightarrow \bigsqcup_{\mathrm{t}} \mathcal{A}_{\mathrm{t}}(\mathbf{L}), \qquad \mathrm{d}V_{\mathrm{WP}} = 2^{m-3}\, \mathrm{d}^{\,2m-6} u_{\mathrm{t}}. \tag{4.9}$$

Thus its volume is a finite sum of ordinary polytope volumes, with this specified constant density. Their construction uses spine and tree coordinates rather than the Bao–Bonahon hyperideal angle cells. It requires an origin cusp.

Zonneveld's sequel removes that restriction using decompositions into half-tight cylinders [97, Theorem 3]. For positive boundary lengths, with a selected pair satisfying $L_1 < L_2$, it gives a full-measure bijection to tree and double-tree cells and identifies their WP measures. The latter cells include a gluing length and a twist parameter. The resulting finite integrations compute the general genus-zero volume polynomials; equal-length volume values follow by continuity. These are further direct combinatorial integrations, but they do not identify the spine cells with the hyperideal dihedral-angle cells. The direct evaluation in those polyhedral coordinates remains a distinct calculation proposed here.

5 Boundary lengths, cone angles, and continuation

5.1 Fixed signature and the polynomial chamber

Allow a cone points with opening angles $\theta_1, \dots, \theta_a$ and b geodesic boundary components with lengths $L_1, \dots, L_b$, including zero lengths for cusps. Thus $n = a + b$ and $d = 3g - 3 + n$. The area is

$$\mathrm{Area}(X) = 2\pi(2g - 2 + n) - \sum_{i=1}^{a} \theta_i > 0. \tag{5.1}$$

There is no separate puncture parameter in the volume polynomial: a cusp is represented by $L = 0$, or equivalently by cone opening zero.

Use the logarithmic tautological class $\kappa_1 = \pi_*(\psi_{n+1}^2)$ on $\overline{\mathcal{M}}_{g,n}$. The fixed-boundary WP class, continued to cone parameters, is

$$\Omega_0 = 2\pi^2 \kappa_1 + \frac{1}{2} \sum_{j=1}^{b} L_j^2 \psi_j - \frac{1}{2} \sum_{i=1}^{a} \theta_i^2 \psi_i. \tag{5.2}$$

The labels in the two sums are understood to be disjoint. In its geometric validity chamber,

$$V_{g,n}(\mathbf{L}, i\boldsymbol{\theta}) = \frac{1}{d!}\int_{\overline{\mathcal{M}}_{g,n}} \Omega_0^d. \tag{5.3}$$

The boundary formula follows from the symplectic-reduction/intersection description of WP volumes [74]; imaginary lengths $L = i\theta$ account for the minus sign [55].

For mixed positive boundaries and cone points, a safe geometric regime is $0 \le \theta_i < \pi$. For pure cone/cusp surfaces, the uncorrected formula holds in the larger chamber where every pair of openings sums to less than 2π [45]. In particular, a single cone of arbitrary opening below 2π, with only cusps as the other ends, does not create a cone-collision wall.

For example, on $\overline{\mathcal{M}}_{0,4}$ one has $\int \kappa_1 = \int \psi_i = 1$, so

$$V_{0,4}(\mathbf{L}, i\boldsymbol{\theta}) = 2\pi^2 + \frac{1}{2}\sum L_j^2 - \frac{1}{2}\sum \theta_i^2 \tag{5.4}$$

whenever this chamber formula applies. On $\overline{\mathcal{M}}_{1,1}$, $\int \kappa_1 = \int \psi_1 = 1/24$, giving the torus normalization in (1.2) and its one-cone continuation.

5.2 Pure cone/cusp volumes beyond the first chamber

For $0 \le \theta_i < 2\pi$, put

$$w_i = 1 - \frac{\theta_i}{2\pi}, \qquad \Phi_S = \sum_{i\in S}\theta_i - 2\pi(|S| - 1), \qquad u_+ = \max(u, 0).$$

Assume $2g - 2 + \sum_i w_i > 0$. A collision of the marked points in S is allowed in the weighted compactification when $\sum_{i\in S} w_i \le 1$. The class pulled back to ordinary Deligne–Mumford space is

$$\boxed{\Omega_{\boldsymbol{\theta}} = 2\pi^2\kappa_1 - \frac{1}{2}\sum_i \theta_i^2\psi_i + \frac{1}{2}\sum_S (\Phi_S)_+^2\,\delta_{0:S}.} \tag{5.5}$$

The sum runs over stable rational-tail divisors, with $|S| \ge 2$. The geometric volume is $\int \Omega_{\boldsymbol{\theta}}^d/d!$. This is the wall-crossing class of Anagnostou–Mullane–Norbury [44, Definitions 2.12–2.13 and Lemma 2.14], also expressed in the cone-volume formulation of Eberhardt–Turiaci [56].

For completeness, the coefficient follows immediately from the weighted formula: its normalized correction is $(1 - \sum_{i\in S} w_i)^2\delta_{0:S}$, and multiplication by $2\pi^2$ gives $(\Phi_S)^2/2$. If r is the reduction to the relevant weighted compactification, then r is proper and birational. Therefore $r_*[\overline{\mathcal{M}}_{g,n}] = [\overline{\mathcal{M}}_{g,\mathbf{w}}]$, and the projection formula justifies integrating the top power of the pulled-back class on ordinary DM space.

In genus zero $\delta_{0:S} = \delta_{0:S^c}$. There is no positive-term double counting in (5.5): stability gives $\sum_i w_i > 2$, so both S and its complement cannot have weight at most one.

For four labeled cone/cusp points the formula becomes particularly simple:

$$\boxed{V_{0,4}(i\boldsymbol{\theta}) = 2\pi^2 - \frac{1}{2}\sum_{i=1}^4 \theta_i^2 + \frac{1}{2}\sum_{1\le i<j\le 4}(\theta_i + \theta_j - 2\pi)_+^2.} \tag{5.6}$$

For openings $(0, 0, 3\pi/2, 3\pi/2)$, the area is positive but the uncorrected polynomial gives $-\pi^2/4$. The single active correction contributes $\pi^2/2$, giving the actual positive volume $\pi^2/4$. This example pinpoints the failure of unrestricted polynomial continuation.

With a positive geodesic boundary and a cone opening greater than π, additional degenerations can occur when the cone approaches that boundary. Anagnostou–Norbury exhibit such a nonnodal degeneration [45, Lemma 2.10]. We consequently assert the full wall formula here for pure cones/cusps, and the preceding small-cone regime for arbitrary mixed positive boundaries. An unrestricted mixed-signature geometric theorem would require its own hypotheses and boundary analysis.

5.3 What continuation means for compact polyhedra

The boundary of a compact convex hyperbolic polyhedron has intrinsic cone opening θ_v equal to the sum of its face angles at v; its singular curvature is $2\pi - \theta_v$. Neither is an individual dihedral angle. Thus the substitution $L \mapsto i\theta$ concerns an intrinsic holonomy parameter. It is not a literal substitution of arbitrary hyperideal vertex coordinates by compact ones.

At a local algebraic level the continuation is natural. For example,

$$K_L(s) = \frac{\sinh(L/2 - s)}{\sinh(L/2)} \quad \longmapsto \quad \frac{\sin(\theta/2 - \alpha)}{\sin(\theta/2)} \quad \text{under } (L, s) = (i\theta, i\alpha). \tag{5.7}$$

Trace and Gram formulas have analogous analytic continuations. Recovering a real geometric volume also requires a real geometric domain, a valid coordinate chart, and the appropriate collision chamber. The polynomial example above shows why these cannot be inferred from the formal algebra.

6 Equal parameters and the torus–sphere correspondence

The familiar punctured correspondence extends geometrically to equal positive boundaries. Let $Q(c^4)$ be a four-holed sphere with all four boundary lengths c, and let $T(2c)$ be a one-holed torus with boundary length $2c$. The relevant construction is through a common orbifold:

$$T(2c) \xrightarrow{2:1} O(c; \pi, \pi, \pi) \xleftarrow{4:1} Q(c, c, c, c). \tag{6.1}$$

The middle surface is a sphere with one geodesic boundary and three angle-π orbifold points. The torus quotient uses its elliptic involution; the sphere quotient uses its Klein-four symmetry. This works for every $c > 0$ and limits to the cusped case [62, Section 3].

A matching simple curve on T covers an arc between two orbifold points twice; its counterpart on Q covers that arc four times. Thus

$$\ell_Q = 2\ell_T. \tag{6.2}$$

The construction gives a correspondence of simple-curve classes, not only an agreement of dimension or of volume polynomials.

WP forms scale by the degree of a local-isometric covering on a family of lifted metrics. This follows from naturality of the Goldman symplectic pairing: its surface integral is multiplied by the covering degree. On the common marked orbifold family,

$$\omega_T = 2\omega_O, \qquad \omega_Q = 4\omega_O, \qquad \text{so} \qquad \omega_Q = 2\omega_T.$$

Generically the equal-boundary sphere has its Klein-four isometry group. There are $4!/4 = 6$ inequivalent ways to label its boundaries. The torus has its generic elliptic involution, which acts ineffectively on Teichmüller space. Consequently

$$\boxed{V_{0,4}(c, c, c, c) = 12V_{1,1}^{\text{eff}}(2c) = 24V_{1,1}^{\text{stack}}(2c).} \tag{6.3}$$

Indeed, $24((2c)^2+4\pi^2)/48 = 2\pi^2+2c^2$, as required by (1.2). Exceptional extra automorphisms are accounted for by orbifold integration. Forgetting the ineffective torus involution is the source of a common factor-of-two discrepancy.

The cone analogue can be constructed directly. For $0 < \phi < \pi$, the elliptic quotient of a torus with one cone opening 2ϕ is a sphere with openings (ϕ, π, π, π). Its Klein-four cover branched over the three angle-π points has four openings equal to ϕ. Uniqueness of conical uniformization makes the conformal symmetries isometries. The same degree and labeling argument gives

$$V_{0,4}(i\phi, i\phi, i\phi, i\phi) = 2\pi^2 - 2\phi^2 = 24V_{1,1}(2i\phi). \tag{6.4}$$

The common quotient cone surface has area $\pi - \phi$, so $\phi < \pi$ is exactly the hyperbolic range. There are no collision corrections in this equal-cone sphere chamber.

This correspondence does not assert a direct unbranched degree-two map from the four-holed sphere to the one-holed torus. The common orbifold and its two different covering degrees are the precise geometric statement.

7 Euclidean shapes and Deligne–Mostow periods

7.1 The fixed-curvature leaf at the Euclidean end

Fix $n \geq 4$ and numbers

$$0 < \mu_i < 1, \qquad \sum_{i=1}^{n} \mu_i = 2. \tag{7.1}$$

The Euclidean cone opening at the ith marked point is $2\pi(1-\mu_i)$, so $2\pi\mu_i$ is its curvature deficit. A conformal sphere X with these marked points carries a unique such flat metric up to multiplication by a positive constant. In a coordinate with all marks finite it has the explicit form

$$g_{\mathrm{E},\mu} = C \prod_i |w - p_i|^{-2\mu_i} |\,\mathrm{d}w|^2. \tag{7.2}$$

The sum condition makes infinity a regular point. Near p_i, radial integration gives a cone of opening $2\pi(1-\mu_i)$. Thus the space of oriented, labeled Euclidean cone spheres with these deficits, up to similarity, has underlying complex manifold $\mathcal{M}_{0,n}$.

Alexandrov realization identifies these metrics with convex Euclidean polyhedral surfaces, allowing doubled planar polygons. This is the fixed-deficit space on which Thurston constructs his complex hyperbolic metric [95, 57]. The qualification "fixed deficits" is essential: all Euclidean polyhedra with n vertices, up to similarity, form a space of real dimension $3n-7$, whereas the fixed-deficit leaf has real dimension $2n-6$. The full space of compact hyperbolic n-vertex polyhedra has dimension $3n-6$; its fixed-cone-angle leaves again have dimension $2n-6$. The ideal locus has dimension $2n-6$. These dimension counts apply on the generic strata; changes of combinatorics and flattened realizations require the usual cell structure.

7.2 The same periods describe hypergeometric functions and developed edges

Define the multivalued differential

$$\eta_p(w) = \prod_i (w - p_i)^{-\mu_i}\, \mathrm{d}w. \tag{7.3}$$

Its branches differ by unit complex numbers. It therefore defines a twisted cohomology class for a rank-one unitary local system on the punctured sphere. The relevant twisted cohomology

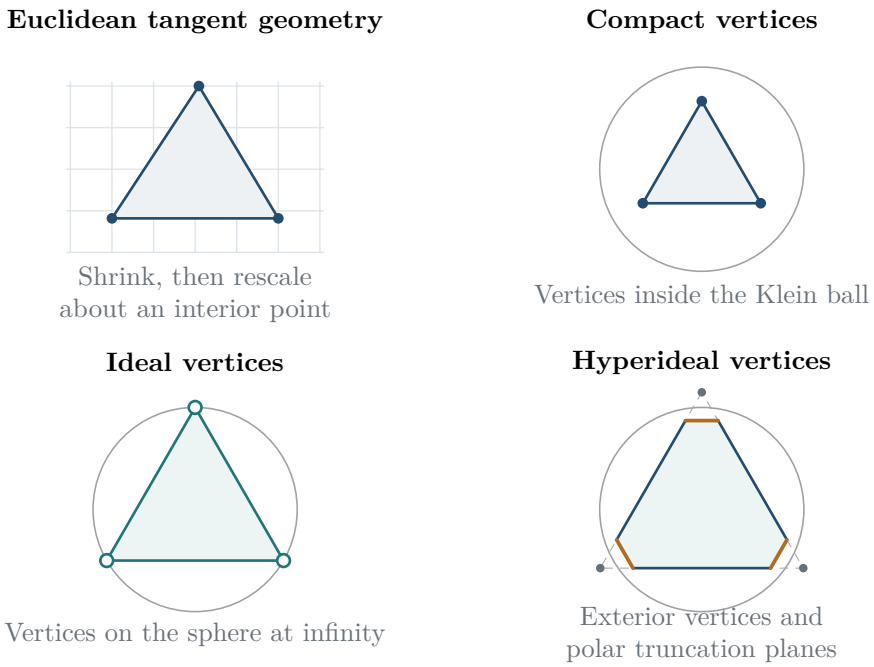


Figure 2: Local pictures for the geometric regimes. The three Klein-model drawings are planar sections illustrating the location of vertices; the ochre segments are polar truncations in the hyperideal case. The Euclidean picture represents a tangent-scale limit after shrinking and rescaling. These drawings do not assert a single global deformation through all four regimes with arbitrary prescribed data.

space has complex dimension $n-2$, and its Hermitian intersection form H has signature $(1, n-3)$, with the convention that the class of η_p has positive norm. In a local basis of twisted cycles or regularized relative arcs, the coordinates

$$Z_j(p) = \int_{\gamma_j} \eta_p \tag{7.4}$$

are the hypergeometric periods of Deligne–Mostow [53]. Projectivizing their positive line gives a local map to $\mathbb{CH}^{n-3}$. Analytic continuation produces projective unitary monodromy.

Thurston's Euclidean construction uses exactly the same coordinates. The local developing map is

$$F(w) = \int^{w} \eta_p.$$

Hence a developed edge vector is an integral of η_p along the corresponding arc. Cutting along a tree produces a Euclidean polygon whose edge vectors are periods, and whose area is the Hermitian form in these periods. Normalize H by this area pairing. Then, more intrinsically,

$$\mathcal{A}(p) = \frac{i}{2} \int_{\mathbb{P}^1} \eta_p \wedge \overline{\eta_p} = H(Z(p), Z(p)). \tag{7.5}$$

Multiplication of all periods by a nonzero complex number changes only the scale and rotation of the developed surface. Quotienting by this action is precisely the projectivization in the Deligne–Mostow construction.

Thus the connection is an equality of geometric structures: the hypergeometric periods are the developed edge vectors, and the Hermitian intersection pairing is the area pairing. Kojima proves that the two constructions are canonically isometric [64, Theorem 4.3]. Throughout this note the complex hyperbolic metric has *holomorphic sectional curvature* -1. Its Kähler form, in the preceding local period coordinates, is

$$\omega_{\mathrm{CH}} = -2i\,\partial\bar{\partial}\log\mathcal{A}. \tag{7.6}$$

A holomorphic change of local scale for η_p does not affect this form. A curvature--4 convention would divide the form by 4 and the volume in complex dimension d by 4^d.

If the weights are rational, say $\mu_i = m_i/q$, the normalized cyclic cover

$$y^q = \prod_i (w - p_i)^{m_i}$$

makes $\eta_p = \mathrm{d}w/y$ single-valued. The relevant twisted cohomology is an eigenspace in the ordinary cohomology of this cover. This gives a further useful interpretation in terms of periods of algebraic curves; it does not require the monodromy group to be a lattice.

7.3 Collision angles and the lattice conditions

Suppose two marks of total weight $\mu_i + \mu_j < 1$ collide, with local separation δ. Rescaling the integration variable in their vanishing period gives

$$Z_{ij}(\delta) = C\delta^{\,1-\mu_i-\mu_j}(1 + O(\delta)), \qquad C \neq 0. \tag{7.7}$$

The transverse complex hyperbolic cone angle is consequently $2\pi(1-\mu_i-\mu_j)$. Deligne–Mostow's condition INT is

$$(1 - \mu_i - \mu_j)^{-1} \in \mathbb{Z}_{>0} \quad \text{for every pair with } \mu_i + \mu_j < 1. \tag{7.8}$$

Under this condition the relevant completion has orbifold uniformization by complex hyperbolic space, and the monodromy is a lattice. If specified groups of equal-weight marks may be permuted, interchanging two such marks halves the transverse angle. Mostow's ΣINT condition accordingly permits half-integral reciprocals for pairs in the same permutation block, while retaining integrality for the other pairs [53, 79].

When $\mu_i + \mu_j = 1$, the collision is a cusp rather than a finite cone-angle divisor; logarithmic periods and parabolic monodromy enter. For general real weights satisfying (7.1), the local complex hyperbolic geometry and its finite volume still exist [95, Theorem 0.2]. One must not infer a lattice quotient for arbitrary weights. Conversely, INT and ΣINT are the classical sufficient uniformization conditions, not a claim to classify every exceptional discrete hypergeometric monodromy group. Our volumes retain all labels; a quotient by allowed equal-weight permutations changes the quotient and the volume factors.

7.4 Four equal deficits and elliptic periods

For $n = 4$ and $\mu_i = 1/2$, put the marks at $0, 1, z, \infty$. Then

$$\eta = \frac{\mathrm{d}w}{\sqrt{w(w-1)(w-z)}}, \qquad E_z:\ y^2 = w(w-1)(w-z). \tag{7.9}$$

The Euclidean cone sphere is the flat pillowcase quotient of this elliptic curve. The ratio of two periods gives a coordinate $\tau \in \mathbb{H}$. For $0 < z < 1$, one choice is

$$\tau(z) = i\frac{K(1-z)}{K(z)}, \qquad K(z) = \int_0^{\pi/2} \frac{\mathrm{d}u}{\sqrt{1 - z\sin^2 u}}. \tag{7.10}$$

Here z is the elliptic parameter, not its square root. The inverse map is the modular lambda function. Projective monodromy is $\mathrm{P}\Gamma(2) = \Gamma(2)/\{\pm I\}$, and the labeled moduli space is the thrice-punctured sphere with curvature--1 area 2π. Its effective sixfold permutation quotient is the modular orbifold [64, Example 2.1 and Section 5]. This is the flat endpoint of the equal-angle torus–sphere correspondence in Section 6.

8 From ideal polyhedra to the Euclidean limit of WP geometry

8.1 A prescribed intrinsic interpolation

For the weights in (7.1), define

$$\theta_i(t) = 2\pi t(1-\mu_i), \qquad a_i(t) = 1 - t(1-\mu_i), \qquad 0 \le t < 1. \tag{8.1}$$

Hold the marked conformal sphere $X \in \mathcal{M}_{0,n}$ fixed, and let $g_t(X)$ be its curvature--1 cone metric with these openings. Its area is

$$A_t = 2\pi\left(\sum_i a_i(t) - 2\right) = 2\pi(n-2)(1-t). \tag{8.2}$$

At $t = 0$ the metric is the complete cusped metric, realized by an ideal polyhedron, including doubled ideal polygons. For $0 < t < 1$ it is a positive-deficit hyperbolic cone metric, realized by a compact convex hyperbolic polyhedron, with flattened realizations included when necessary. The realization results are reviewed in [57]; the ideal and generalized settings are treated in [85, 91].

Normalize to unit area:

$$\widehat{g}_t = A_t^{-1} g_t, \qquad \mathrm{Area}(\widehat{g}_t) = 1, \qquad K(\widehat{g}_t) = -A_t. \tag{8.3}$$

The smooth conical uniformization theory of Mazzeo–Weiss gives the flat limit

$$\widehat{g}_t(X) \longrightarrow g_{\mathrm{E},\mu}(X) \quad (t \to 1), \tag{8.4}$$

with the Euclidean metric normalized to area one. The convergence is smooth away from the cone points, with controlled conic expansions; it is local over moduli away from collisions [69].

This constructs a family of complex fixed-angle leaves with two geometric endpoints: cusped hyperbolic surfaces/ideal polyhedra, and flat cone spheres/Euclidean shapes. It does not assign a single complex structure to all angle variables in the full real space of hyperbolic polyhedra. Nor is it generally radial scaling of a fixed Euclidean vertex configuration in the Klein ball. Such radial scaling makes all vertices ideal simultaneously only for a cospherical configuration. Every nondegenerate Euclidean tetrahedron is cospherical, but its radial path still need not preserve X or the angle path (8.1).

8.2 The normalized local WP form

The scale of the WP form can be seen directly from the cotangent pairing. If $g_t = \rho_t |\,\mathrm{d}w|^2$, its pairing on quadratic differentials is a fixed convention-dependent multiple of

$$Q_t(q,r) = \int_X \frac{q\overline{r}}{\rho_t}.$$

Since $\rho_t = A_t\widehat{\rho}_t$, the rescaled cometric A_tQ_t has a finite flat limit. A quadratic differential has at most a simple pole at a marked point, while $\widehat{\rho}_t \to \rho_{\mathrm{E}}$ with $\rho_{\mathrm{E}} = c_i(w)|w|^{-2\mu_i}$ for a smooth positive local factor c_i. The limiting integrand is $O(|w|^{2\mu_i-2})\,\mathrm{d}x\,\mathrm{d}y$, which is integrable because $\mu_i > 0$. Parameter derivatives add integrable logarithmic factors. Inverting the cometric gives a tangent metric and a Kähler form of order A_t.

Wang identifies the area-normalized flat WP pairing with the Thurston complex hyperbolic metric [96]. Combining that identification with the preceding limit, and using the WP normalization (1.2) and curvature -1 for the complex hyperbolic metric, gives

$$\boxed{A_t^{-1}\omega_{\mathrm{WP},t} \longrightarrow \frac{1}{2}\omega_{\mathrm{CH},\mu}.} \tag{8.5}$$

The factor is also checked independently by the class calculation below. A bare formula for Q_t without its Hermitian and Kähler conventions would not determine that factor. This local convergence is a deduction from cone-metric deformation and the pairing identification; it is not a theorem quoted verbatim from either source. Local convergence on a noncompact moduli space alone does not justify convergence of total volumes.

8.3 The class calculation and an exact final-chamber polynomial

Put $d = n-3$, $D = n-2$, and $\mu_I = \sum_{i\in I}\mu_i$. Let $\delta_I = \delta_{I^c}$ denote the divisor of the stable bipartition $I \mid I^c$. Define the real divisor class

$$B_\mu = \frac{1}{2}\left(-\sum_i \mu_i\psi_i + \sum_{\{I,I^c\}} \min(\mu_I, \mu_{I^c})\delta_I\right), \tag{8.6}$$

where the second sum counts each unordered stable bipartition once. The extension of the complex hyperbolic form to the ordinary Deligne–Mumford compactification satisfies

$$[\omega_{\mathrm{CH},\mu}] = 4\pi B_\mu, \qquad V_{\mathrm{CH}}(\mu) = \frac{(4\pi)^d}{d!}\int_{\overline{\mathcal{M}}_{0,n}} B_\mu^d. \tag{8.7}$$

This is the normalization of the intersection formula of Koziarz–Nguyen [65]. The class is a compactification class; the volume on the left is the finite volume of the open space of distinct marked points.

Let $\varepsilon = 1-t$, $r_i = 1-\mu_i$, and $r_I = \sum_{i\in I} r_i = |I| - \mu_I$. Then $a_i = \mu_i + \varepsilon r_i$ and $A = 2\pi D\varepsilon$. Write $\mathcal{L}_\mu$ for the stable subsets with $\mu_I < 1$, each light side counted once. For sufficiently small positive ε, these are precisely the light subsets in the adjacent hyperbolic chamber. Subsets with $\mu_I = 1$ are *not* light along this path. Define

$$\boxed{\Xi_\mu = \frac{1}{2D^2}\left(-\sum_i (1-\mu_i)^2\psi_i + \sum_{I\in\mathcal{L}_\mu}(|I|-\mu_I)^2\delta_I\right).} \tag{8.8}$$

Proposition 8.1 (A deduction from the two compactification formulas)**.** *In the final chamber of the path* (8.1)*,*

$$[\omega_{\mathrm{WP}}(A)] = \frac{A}{2}[\omega_{\mathrm{CH},\mu}] + A^2\Xi_\mu, \tag{8.9}$$

$$V_{\mathrm{WP}}(A) = \frac{A^d}{d!}\int_{\overline{\mathcal{M}}_{0,n}}\left(\frac{1}{2}[\omega_{\mathrm{CH},\mu}] + A\Xi_\mu\right)^d. \tag{8.10}$$

In particular,

$$\boxed{V_{\mathrm{WP}}(A) = 2^{-d}A^d V_{\mathrm{CH}}(\mu) + O(A^{d+1}).} \tag{8.11}$$

The volume in (8.10) *is exactly a polynomial in A with possible degrees $d, d+1, \dots, 2d$ within that chamber.*

Proof. In terms of deficit weights, the pulled-back WP class (5.5), divided by $2\pi^2$, is

$$\beta(a) = \kappa_1 - \sum_i (1-a_i)^2\psi_i + \sum_{a_I<1}(1-a_I)^2\delta_I.$$

The genus-zero divisor relations, or equivalently the flat universal-curve extension in [65, Section 8], give

$$\beta(\mu) = 0,$$
$$\sum_i r_i^2 \psi_i - \sum_{I \in \mathcal{L}_\mu} (1 - \mu_I) r_I \delta_I = D B_\mu.$$

One way to see the linear term is to use the corrected relative log-canonical divisor on the universal curve. At $\sum \mu_i = 2$ it is the pullback of B_μ; the perturbing divisor has fiber degree $\sum r_i = D$. Pushing forward its square therefore gives $2\varepsilon D B_\mu$ as the first-order term. The universal-curve argument starts with rational weights; the real-weight identities follow by continuity of the displayed classes. Balanced collision terms have zero coefficient at the endpoint. These are class identities, so no unproved convergence of integrals is involved.

Expanding the quadratic expression for β now gives

$$\beta(\mu + \varepsilon r) = 2\varepsilon D B_\mu + \varepsilon^2 \left(- \sum_i r_i^2 \psi_i + \sum_{I \in \mathcal{L}_\mu} r_I^2 \delta_I \right).$$

Multiplication by $2\pi^2$, followed by $A = 2\pi D\varepsilon$ and (8.7), proves (8.9). The WP intersection-volume theorem [44] proves (8.10); its leading term is (8.11). □

The final-chamber condition can be made explicit:

$$0 < \varepsilon < \min_{I \in \mathcal{L}_\mu} \frac{1 - \mu_I}{|I| - \mu_I}, \tag{8.12}$$

with no restriction from an empty light-subset set. Along the full path, a stable subset crosses its wall at $t_I = (|I| - 1)/(|I| - \mu_I)$ when $\mu_I < 1$. Thus the ideal-to-flat volume is piecewise polynomial, even though the final expansion is an exact polynomial.

8.4 Examples and nonuniformity at balanced weights

With four equal weights $\mu_i = 1/2$, all openings are πt, and there are no active collision corrections for $t < 1$. Thus

$$V_{\mathrm{WP}}(t) = 2\pi^2(1 - t^2), \qquad A = 4\pi(1 - t), \qquad V_{\mathrm{WP}}(A) = \pi A - \frac{A^2}{8}. \tag{8.13}$$

Here $\int B_\mu = 1/2$, $V_{\mathrm{CH}} = 2\pi$, and $\int \Xi_\mu = -1/8$, checking all constants directly.

For $\mu = (0.3, 0.4, 0.6, 0.7)$, one obtains $\int B_\mu = 0.3$ and $\int \Xi_\mu = 0.225$. The final-chamber volume is

$$V_{\mathrm{WP}}(A) = 0.6\pi A + 0.225 A^2, \qquad V_{\mathrm{CH}} = 1.2\pi.$$

At the flat endpoint the uncorrected cone polynomial would equal $-0.2\pi^2$; the two light-pair corrections, $0.18\pi^2$ and $0.02\pi^2$, cancel it. The wall corrections are indispensable to the existence of the flat limit.

The leading class B_μ and volume $V_{\mathrm{CH}}(\mu)$ vary continuously with the weights. The higher coefficients of the *final-chamber* expansion need not. For example, let

$$\mu(\delta) = \left(\tfrac{1}{2} - \delta, \tfrac{1}{2} - \delta, \tfrac{1}{2} + \delta, \tfrac{1}{2} + \delta \right), \qquad 0 < \delta < \tfrac{1}{2}.$$

The only light pair is $\{1, 2\}$, and

$$\int \Xi_{\mu(\delta)} = \frac{\delta}{2} \longrightarrow 0, \qquad \int \Xi_{(1/2)^4} = -\frac{1}{8}.$$

There is no contradiction: the final chamber has $\varepsilon < 2\delta/(1+2\delta)$ and collapses as $\delta \to 0$. The full piecewise WP volume remains continuous. The small-A expansion must therefore be taken at fixed weights; uniformity across balanced collision walls requires a separate analysis.

8.5 Why the complex hyperbolic volumes look different

Thurston proves finite volume and connects particular examples to counting triangulations [95]. A general explicit volume formula is obtained by McMullen's stratified Gauss–Bonnet calculation [70]. In the present labeled, curvature--1 convention it is

$$V_{\mathrm{CH}}(\mu) = \frac{(-4\pi)^d}{(d+1)!} \sum_{\substack{\mathcal{P} \text{ a partition of } \{1,\dots,n\} \\ |\mathcal{P}|\geq 3}} (-1)^{|\mathcal{P}|+1}(|\mathcal{P}|-3)! \prod_{B\in\mathcal{P}} (1-\mu_B)_+^{|B|-1}. \tag{8.14}$$

The blocks encode collisions. Their solid-angle factors weight Euler characteristics of the corresponding strata. Koziarz–Nguyen's intersection formula (8.7) computes the same quantity by a different route. Equation (8.11) recovers it from the normalized small-area limit $V_{\mathrm{CH}}(\mu) = 2^d \lim_{A\downarrow 0} A^{-d} V_{\mathrm{WP}}(A)$. Thus the Gauss–Bonnet, period, and intersection calculations refer to the same limiting geometry, with their normalization made explicit.

Part II
Variational geometry and computation

The volume Hessian becomes a discrete Dirichlet operator. Boundary elimination, Gaussian determinants, and barrier methods connect the geometry to computation.

9 Volume Hessians, finite elements, and computational complexity

The variational description of a polyhedron has a second interpretation: its infinitesimal deformation equation can be a discrete Dirichlet problem. This is an exact connection to finite elements and electrical networks, and it supplies the computational counterpart of the volume and symplectic constructions developed above. The companion study *Duffin's resistor networks, the volume of ideal polyhedra, and nearly-linear-time Laplacian solvers* [116] develops this connection and its numerical examples. We incorporate the essential identities here and connect them to boundary response, Gaussian integration, flips, and the Casson–Rivin program.

9.1 The cotangent form and the correct deformation coordinates

Let $\mathcal{T}$ triangulate a planar disk, with interior vertices I and boundary vertices B. For a piecewise affine function v on a geometric realization, integration over each triangle gives

$$\int_{|\mathcal{T}|} |\nabla v|^2 \, dA = \sum_{e=ij} w_e (v_i - v_j)^2 = v^T L v, \qquad w_e = \frac{1}{2}(\cot \alpha_e + \cot \beta_e). \tag{9.1}$$

Here α_e, β_e are the angles opposite e, and the second term is omitted on a boundary edge. To check the normalization, the off-diagonal stiffness entry on a triangle is $\int_T \nabla\phi_i \cdot \nabla\phi_j \, dA = -\frac{1}{2}\cot\alpha_k$ for its barycentric functions. Row sums vanish, giving the displayed edge form. This is the cotangent network of the finite element and electrical interpretations [117].

The energy is nonnegative on every geometric mesh, even if some edge weights are negative. The additional condition for nonnegative *interior* conductances is local Delaunayness:

$$w_e = \frac{\sin(\alpha_e + \beta_e)}{2 \sin\alpha_e \sin\beta_e} \geq 0 \quad \Longleftrightarrow \quad \alpha_e + \beta_e \leq \pi. \tag{9.2}$$

When the relevant weights are nonnegative, the grounded matrix L_{II} is a positive definite graph Laplacian, and solving a Dirichlet problem is solving Kirchhoff's equations. Boundary-to-boundary edges do not enter this Dirichlet system; positivity of their weights is an additional condition if the full matrix is used in a different boundary problem.

Now prescribe the sums $\Delta_e = \alpha_e + \beta_e \leq \pi$ opposite interior edges, and prescribe the single opposite angle on each boundary edge. An angle structure assigns positive corner angles, with sum π in every triangle and these prescribed edge sums. Assume this positive angle domain is nonempty. For planar realization the data also satisfy $\sum_{e \ni i}(\pi - \Delta_e) = 2\pi$ at each interior vertex; the left-hand side is its total angle in every such angle structure. If the intended planar boundary is strictly convex, impose also

$$0 < \kappa_b := \sum_{t \ni b} \alpha_{t,b} = \sum_{e \ni b} (\pi - \Delta_e) - \pi < \pi \qquad (b \in B).$$

Here Δ_e denotes the single prescribed opposite angle on a boundary edge. These boundary inequalities are needed for the convex polyhedral interpretation; the Hessian identities below already hold on the abstract positive angle domain. Positive feasibility alone does not impose convexity of the developed boundary. Orient the triangles consistently. If a corner c is opposite the directed edge $i \to j$, set

$$\alpha_c(u) = \alpha_c^0 + u_i - u_j, \qquad u|_B = 0. \tag{9.3}$$

These are leading–trailing coordinates, not logarithmic scale factors. Triangle sums telescope; increments opposite a shared edge cancel. For a triangulated disk they parametrize the entire affine space of solutions: the map is injective, and the connected triangle–edge incidence graph together with Euler's formula gives dimension $|I|$.

Let $\Lambda'(a) = -\log(2\sin a)$ be the Lobachevsky derivative and put $\mathcal{V}(\alpha) = \sum_c \Lambda(\alpha_c)$. At a geometric configuration this is the sum of the volumes of the ideal tetrahedra obtained by adjoining the vertex at infinity to the planar triangles. Rivin's variational principle uses this concave functional to impose geometric closure [118, 37].

Proposition 9.1 (The exact finite-element Hessian)**.** *On the positive angle domain in the coordinates* (9.3)*,*

$$\nabla_u \mathcal{V} = (\log H_i)_{i\in I}, \qquad -D_u^2 \mathcal{V} = 2L_{II}(\alpha), \tag{9.4}$$

where H_i is the dilation accumulated by developing the incident triangles once around i. In particular, at a geometric critical point the negative volume Hessian is twice the finite element stiffness matrix.

Proof. In a positively oriented triangle ijk, differentiation with respect to u_i gives $\log(\sin\alpha_{t,j}/\sin\alpha_{t,k})$. Summing over the star of i gives the logarithm of the product of side ratios, which is the dilation holonomy by the sine law. Since $\Lambda''(a) = -\cot a$, the two corners opposite an interior edge contribute $-(\cot\alpha_e + \cot\beta_e)(du_i - du_j)^2$ to $D^2\mathcal{V}$. This is -2 times the corresponding summand of (9.1). Boundary increments vanish. □

For a true interior edge with $0 < \Delta_e < \pi$, the weights satisfy

$$w_e = \frac{\sin\Delta_e}{2\sin\alpha_e \sin\beta_e} \geq \cot(\Delta_e/2) > 0. \tag{9.5}$$

Thus positivity holds throughout the angle domain, before the angles close to form a geometric mesh. At $\Delta_e = \pi$ the weight is zero; such a diagonal subdivides a flat face. The finite element interpretation as an actual planar mesh applies at realization, whereas the Laplacian identity and positive-network interpretation already apply to the intermediate angle structures. Newton's equation is precisely

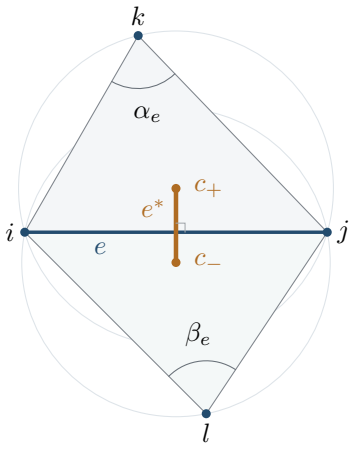
$$2L_{II}(\alpha(u))\,\delta u = \log H(u). \tag{9.6}$$

9.2 Boundary response, Schur complements, and a Gaussian factor

For the finite element network, eliminating the interior voltage gives

$$v_I = -L_{II}^{-1} L_{IB} v_B, \qquad \mathcal{N} = L_{BB} - L_{BI} L_{II}^{-1} L_{IB}. \tag{9.7}$$

The matrix $\mathcal{N}$ maps boundary voltages to boundary currents; it is the discrete Dirichlet-to-Neumann operator. It has the constant mode in its kernel. Grounding one boundary vertex, or working modulo constants, removes that mode.



One local weight, several roles

$$w_e = \frac{1}{2}(\cot \alpha_e + \cot \beta_e)$$
$$= \frac{|e^*|_\pm}{|e|}$$

$$\int |\nabla u|^2 = \sum_{e=ij} w_e (u_i - u_j)^2$$

$$-D_u^2 \mathcal{V} = 2L_{II}$$

The last identity uses the specified angle coordinates and fixed boundary values.

Figure 3: The cotangent weight for two adjacent Euclidean triangles. The dual segment joins their circumcenters and has signed length $|e^*|_\pm$; in the locally Delaunay configuration shown, its sign is positive. The finite-element stiffness matrix is the same cotangent matrix that appears as minus one half of the volume Hessian in the coordinates $\alpha_c(u) = \alpha_c^0 + u_i - u_j$, with fixed edge-angle sums and $u = 0$ on the boundary. This identity is specific to that coordinate and constraint system.

The same elimination appears in volume response. For a nondegenerate concave maximum, write

$$\mathcal{V}(x) = V_0 - \tfrac{1}{2} x^T K x + O(|x|^3), \qquad K > 0,$$

and let Px be the selected linear boundary data. Perturbation by a source j gives

$$\max_x \{\mathcal{V}(x) + j^T P x\} = V_0 + \tfrac{1}{2} j^T \chi j + O(|j|^3), \qquad \chi = P K^{-1} P^T,$$
$$\max_{Px=b} \mathcal{V}(x) = V_0 - \tfrac{1}{2} b^T \chi^{-1} b + O(|b|^3), \tag{9.8}$$

where the second formula assumes P has full row rank. The first is the response to a source; the second fixes its conjugate boundary variable. They are related by a Legendre transform. In the planar problem $K = 2L_{II}$; in the Dehn filling problem, the same algebra gives (16.7) and (16.8). This identifies an exact common operation without identifying their different spaces of boundary data.

More generally, split $x = (y, z)$ and eliminate y. The effective quadratic form is

$$K_{\text{eff}} = K_{zz} - K_{zy} K_{yy}^{-1} K_{yz}, \qquad \det K = \det K_{yy} \det K_{\text{eff}}. \tag{9.9}$$

Completing the square gives the exact real Gaussian identity

$$\int_{\mathbb{R}^{\dim y}} e^{-\frac{1}{2\hbar}(y,z)^T K (y,z)}\, dy = \frac{(2\pi\hbar)^{\dim y/2}}{\sqrt{\det K_{yy}}} e^{-z^T K_{\text{eff}} z/(2\hbar)}. \tag{9.10}$$

Thus elimination in the Dirichlet problem, reduction of the classical boundary action, and its leading Gaussian integration use the same Schur complement. An oscillatory or complex phase requires the specified contour and determinant branch, as in the later quantum calculations. The finite-dimensional identity does not supply those analytic data or identify an entire quantum invariant by itself.

There is also a combinatorial formula for the planar determinant. Let $\mathfrak{F}_B$ be the spanning forests of the weighted graph in which each component contains exactly one boundary vertex. Then

$$\boxed{\det(-D_u^2 \mathcal{V}) = 2^{|I|} \sum_{F \in \mathfrak{F}_B} \prod_{e \in F} w_e.} \tag{9.11}$$

To prove it, write $L_{II} = C_I \operatorname{diag}(w_e) C_I^T$, where C_I is the incidence matrix with boundary rows removed. Cauchy–Binet expands its determinant over $|I|$-edge subsets. An incidence minor is ± 1 exactly for a forest rooted at B, and zero otherwise. This rooted-forest formula expresses the local Gaussian determinant in geometric conductances. It is a deduction from the Hessian identity and finite-dimensional linear algebra.

9.3 Why the nonlinear iteration needs a barrier

A fast linear solve does not by itself bound the number of Newton steps. In particular, $-\mathcal{V}$ is not uniformly self-concordant: near a vanishing angle, its entropy-like behavior makes the normalized third derivative unbounded. The companion study supplies a more useful estimate involving the barrier

$$F(\alpha) = -\sum_c \log \sin \alpha_c. \tag{9.12}$$

The coefficient $\csc^2 \alpha_c$ occurs in both D^2F and $D^3\mathcal{V}$, but these are tensors of different order. The precise relation needed for optimization is the following inequality.

Proposition 9.2 (A barrier compatible with ideal volume)**.** *For every positive angle structure and tangent vector h whose coordinates sum to zero in each triangle,*

$$|D^3\mathcal{V}[h,h,h]| \le \sqrt{3}\,(-D^2\mathcal{V}[h,h])\sqrt{D^2F[h,h]}. \tag{9.13}$$

Consequently $F - t\mathcal{V}$ is self-concordant with the standard constant for every $t \ge 0$. The barrier F has parameter at most the number m of corner angles.

Proof. For one triangle set $s_i = \sin \alpha_i$ and choose the cyclic side directions $\psi = (0, \alpha_1 + \alpha_2, \pi + \alpha_2)$. They satisfy $\sum_i s_i e^{i\psi_i} = 0$. Every vector with $h_1 + h_2 + h_3 = 0$ can therefore be written $h_i = \lambda s_i \cos \varphi_i$, where $\varphi_i = \psi_i - \tau$. The cotangent energy identity and a trigonometric expansion give

$$A := -D^2\mathcal{V}[h,h] = \lambda^2 s_1 s_2 s_3, \quad B := D^2F[h,h] = \lambda^2 \sum_i \cos^2 \varphi_i,$$

$$D^3\mathcal{V}[h,h,h] = \lambda^3 s_1 s_2 s_3 \cos(\varphi_1 + \varphi_2 + \varphi_3).$$

For the last identity, use $\cos^3 x = (3\cos x + \cos 3x)/4$ and the vanishing side-vector sum. Expanding the cosine of the sum shows $|\cos(\varphi_1 + \varphi_2 + \varphi_3)| \le \sum_i |\cos \varphi_i| \le \sqrt{3 \sum_i \cos^2 \varphi_i}$. This proves the inequality on one triangle, and summation proves it globally since $\sum A_T \sqrt{B_T} \le (\sum A_T)\sqrt{\sum B_T}$. Each $-\log \sin \alpha_c$ is self-concordant, and, with $A = -tD^2\mathcal{V}[h,h]$ and $B = D^2F[h,h]$,

$$|D^3(F - t\mathcal{V})[h,h,h]| \le 2B^{3/2} + \sqrt{3} A \sqrt{B} \le 2(A+B)^{3/2}.$$

Finally Cauchy–Schwarz gives $|DF[h]| \le \sqrt{m}\sqrt{D^2F[h,h]}$, the required barrier-parameter bound. Restriction to additional affine angle constraints preserves these inequalities. □

For the central point $\alpha(t)$ minimizing $F - t\mathcal{V}$, the volume gap is at most m/t. Indeed, concavity of sine gives $DF(\alpha)[\beta - \alpha] \le m$ for every feasible β; combine this with stationarity and concavity of $\mathcal{V}$. Increasing t by a factor $1 + c/\sqrt{m}$, with fixed sufficiently small c, gives a bounded Newton decrement: the Hessian of $F - t\mathcal{V}$ dominates D^2F and the barrier gradient has local norm at most $\sqrt{m}$. Standard self-concordant Newton analysis therefore gives, starting from a centered point at $t_0 > 0$,

$$O\left(\sqrt{m} \log \max\left\{e, \frac{m}{t_0 \varepsilon}\right\}\right) \quad \text{Newton steps to volume gap at most } \varepsilon. \tag{9.14}$$

See [119, Chapter 11] for the underlying path-following analysis. This is a bound for the barrier method, not for unrestricted plain Newton iteration.

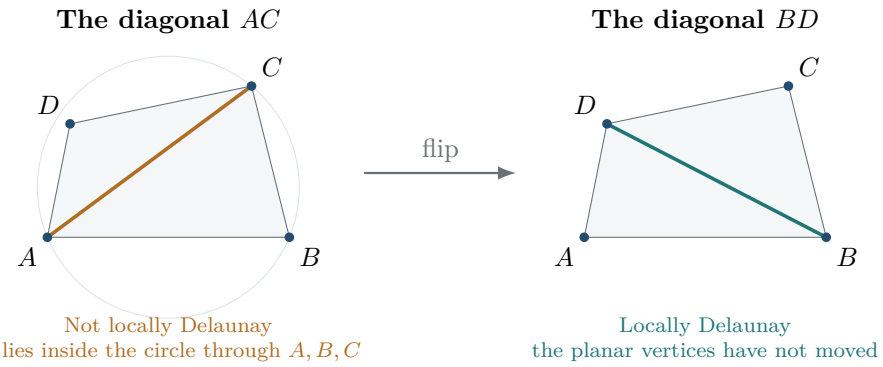
In the disk coordinates the Hessian of $F - t\mathcal{V}$ is again a grounded planar Laplacian, now with interior conductances

$$2tw_e + \csc^2 \alpha_e + \csc^2 \beta_e.$$

Nearly-linear Laplacian solvers approximate a linear solve in its energy norm [120]. With controlled inexact-Newton tolerances, their combination with (9.14) gives the usual $\widetilde{O}(n^{3/2})$ arithmetic-work scale after centering, up to accuracy factors, when $m = O(n)$. It does not give a nearly-linear bound for the whole nonlinear realization problem. Finding a feasible point, centering it, controlling arithmetic precision, and certifying the final geometric error are separate costs. The numerical timings in the companion are experiments on selected inputs, not bounds for these additional tasks.

9.4 Flips connect the local matrix to mapping-class geometry

A diagonal exchange has two different roles. Among triangulations of the convex hull of a fixed planar vertex set, with fixed nodal values, Rippa's theorem says that Delaunay triangulation minimizes the piecewise-linear Dirichlet energy; in matrix terms $L(\mathcal{T}_D) \preccurlyeq L(\mathcal{T})$ [121]. This is an ordering of quadratic forms for the same geometric data. It does not assert monotonicity of hyperbolic volume along an arbitrary sequence of topological flips.



For the same nodal values u: $D_{\mathcal{T}_{BD}}(u) \leq D_{\mathcal{T}_{AC}}(u)$.

Figure 4: A diagonal flip on a fixed Euclidean quadrilateral, shown with its actual vertex positions. Rippa's energy comparison selects the Delaunay diagonal for fixed nodal values. The same combinatorial move also occurs in flip graphs used in surface topology; the displayed energy comparison concerns the fixed Euclidean realization.

For a connected orientable finite-type marked surface admitting an ideal triangulation, with at least one marked point on every boundary curve, the graph of ideal triangulations and diagonal exchanges is instead a coarse geometric model of its mapping class group. With isotopies fixing the marked points, the action is proper and cocompact, and an orbit map is a quasi-isometry [122, Section 2.1, Lemma 2.3]. The pants graph is a different model, quasi-isometric to Teichmüller space with its WP metric, and Brock relates WP distance to convex-core volume [123]. His mapping-torus theorem relates WP translation distance and hyperbolic volume for closed pseudo-Anosov surface bundles [124]. These comparisons have constants depending on the surface topology; the flip graph and the pants graph should not be identified.

There is a direct volume estimate when a flip sequence gives a layered ideal triangulation of a hyperbolic mapping torus with N tetrahedra: straightening gives $\mathrm{Vol}(M) \le N v_3$, where v_3 is the volume of the regular ideal tetrahedron. This does not require the given triangulation to have positive geometric shapes; signed straightened volumes still have absolute value at most v_3. It connects a combinatorial decomposition length to three-dimensional volume. Comparing it further to WP geometry requires a theorem for the same surface and boundary conventions, rather than transferring the closed-surface theorem silently to a punctured surface.

9.5 The three-manifold extension and the remaining questions

The Casson–Rivin program uses the same ideal-tetrahedron volume and the same angle-sum constraints [37]. Its local Hessian is still a cotangent quadratic form. For leading–trailing edge variables y_e, put $p_\alpha = y_e + y_{e'}$ for the pair of opposite tetrahedron edges carrying α, and cyclically for β, γ. Then

$$-D^2\mathcal{V} = \sum_t \left[\cot\alpha_t (dp_\gamma - dp_\beta)^2 + \cot\beta_t (dp_\alpha - dp_\gamma)^2 + \cot\gamma_t (dp_\beta - dp_\alpha)^2\right]. \tag{9.15}$$

This displays the Euclidean cusp-link triangle inside the second variation. But a square now involves sums of opposite-edge variables, and the assembled matrix need not have nonpositive off-diagonal entries or be diagonally dominant. Additional peripheral directions and removal of linear dependencies must also be included in a full coordinate system. The planar positive-Laplacian solver conclusion therefore does not transfer automatically to arbitrary three-manifold triangulations. The trianglewise barrier estimate does transfer, because it survives affine restriction.

An interior critical point has the geometric interpretation supplied by the Casson–Rivin theorem. A boundary maximizer need not realize the chosen triangulation by positive ideal tetrahedra. Optimization on a given angle polytope therefore does not solve the problem of finding a suitable geometric triangulation or yield a polynomial-time geometrization algorithm.

The finite element comparison also does not identify its real symmetric Hessian with the alternating WP form. The exact links established here are the volume Hessian, its inverse response, its Schur complement, and its Gaussian determinant. A continuum limit relating appropriately renormalized polyhedral functionals to a WP potential would require geometric convergence, a specified normalization, and control of second variations. This is a natural question, not a theorem asserted here. Likewise, in three dimensions it remains useful to ask for an effective bound on the smallest angle at an interior maximum in terms of the number of tetrahedra and a quantitative feasible margin. Such a bound would connect variational existence to numerical certification and complexity. These questions describe how the computational strand can extend the monograph's variational and semiclassical program.

Part III
Quantum geometry, degeneration, and surgery

An introduction to the quantum constructions leads to exact period comparisons, a controlled tetrahedral degeneration, and the classical and Gaussian data of Dehn filling.

10 A geometric introduction to the quantum constructions

The next part uses several constructions that share the word *quantum* but have different definitions. This section introduces the objects, the elementary asymptotic mechanism, and the distinctions needed to read the calculations. The prerequisites are representation theory at the level of tensor products, elementary complex analysis, and the geometry already discussed. No physical interpretation is needed for the finite sums or special-function identities.

The organizing operation is cutting and gluing. A decomposition gives intermediate labels; changing the decomposition gives a change-of-basis matrix or integral kernel. The quantum constructions make these operations exact. In suitable large-parameter limits their equations recover geometric gluing, and their leading exponents recover geometric actions. The simplest complete example below is a finite sum for the figure-eight knot whose growth gives the volume of two regular ideal tetrahedra.

10.1 Six labels: recoupling and tetrahedral weights

Let V_j be the irreducible $\mathfrak{sl}_2$-representation of dimension $2j+1$, where j is a nonnegative half-integer. Consider

$$\mathrm{Hom}_{\mathfrak{sl}_2}(V_d, V_a \otimes V_b \otimes V_c).$$

There are two natural bases. One first decomposes $V_a \otimes V_b$ and records an intermediate representation V_s; the other first decomposes $V_b \otimes V_c$ and records V_t. Thus

$$B_s^{(ab)c} = \sum_t F_{st}^{abc;d} B_t^{a(bc)}. \tag{10.1}$$

The coefficients depend on six labels a, b, c, d, s, t. With conventional dimension and sign factors, they are the $6j$*-symbols*. The four admissible triples

$$(a,b,s), \quad (s,c,d), \quad (b,c,t), \quad (a,t,d)$$

can be assigned to the four faces of a tetrahedron. At this point the tetrahedron is a way to organize the labels. Its interpretation as a metric tetrahedron requires an additional asymptotic statement.

A quantum group deforms this tensor-product calculus. For $U_{q_{\mathrm{alg}}}(\mathfrak{sl}_2)$, ordinary integers and factorials in the explicit formulas are replaced by their quantum analogues, for example

$$[n]_{q_{\mathrm{alg}}} = \frac{q_{\mathrm{alg}}^n - q_{\mathrm{alg}}^{-n}}{q_{\mathrm{alg}} - q_{\mathrm{alg}}^{-1}}, \qquad [n]_{q_{\mathrm{alg}}} \longrightarrow n \quad (q_{\mathrm{alg}} \to 1). \tag{10.2}$$

The full definition also specifies tensor products and braidings; it is not merely a rule for replacing integers. At the usual compact root $q_{\mathrm{alg}} = e^{\pi i/r}$, the finite fusion theory retains the

labels $0, \frac{1}{2}, \dots, \frac{r-2}{2}$. A triple must satisfy the triangle inequalities, integral total spin, and $a+b+c \le r-2$. Taylor–Woodward give the representation-theoretic definition and these conventions explicitly [26].

Associating four tensor factors in the five possible ways gives the *pentagon identity*. In tetrahedral language it supplies the identity for replacing two tetrahedra by three in a triangulated ball. Together with orthogonality and normalization identities, this makes state sums independent of the triangulation. Schematically, a Turaev–Viro sum for a closed triangulated manifold has the form

$$Z_r(M) = C_r(T) \sum_{\text{admissible edge labels } \mathbf{j}} \prod_{e \in E(T)} d_{j_e} \prod_{\Delta \in T} \{6j\}_{\Delta,\mathbf{j}}. \tag{10.3}$$

Here T is a triangulation, d_j is a quantum dimension, and $C_r(T)$ is the prescribed normalization. This displayed formula explains the structure; the exact weights are part of the definition of the chosen theory [38, Sections 2.2–2.3]. A *state* in this sum is an admissible assignment of labels. The sum is over such assignments; a description as an integral over hyperbolic metrics, with a particular geometric measure, would be a further result.

10.2 The invariants and their volume limits

The main objects used below can be distinguished by their inputs and outputs.

Object	Input	Output and construction
Colored Jones J_N	Knot, color N, variable q_J	A Laurent polynomial, from braid or skein calculus.
Kashaev invariant	Knot and integer N	The specialization $J_N(K; e^{2\pi i/N})$ in the reduced convention.
Reshetikhin–Turaev (RT)	Closed oriented three-manifold and root/level	A scalar, from weighted colored-link evaluations of a surgery presentation.
Turaev–Viro (TV)	Closed three-manifold and root/level	A scalar state sum, built from tetrahedral $6j$-weights.
Continuous b-$6j$	Six continuous representation parameters and b	A change-of-basis kernel for a noncompact representation theory.
Liouville correlation	Marked conformal surface, insertion parameters, and b	A function of the marked points, obtained by averaging field insertions.

The first four objects are topological invariants with their prescribed normalizations. The last two supply analytic kernels and functions that enter the comparisons in this monograph. In particular, a local $6j$-weight is not already a three-manifold invariant.

The knot construction begins with a braid whose closure is the knot. A crossing acts by a braiding matrix on the tensor product of the strand representations; the Yang–Baxter equation gives the braid relation. A quantum trace, adjusted to the representation theory, closes the strands, and the prescribed normalization handles the moves that change the braid presentation [29, Section 2].

For the colored Jones polynomial, the color N denotes the N-dimensional representation at generic parameter. We use the *reduced* normalization $J_N(\text{unknot}; q_J) = 1$, and specialize only after forming that polynomial. Murakami–Murakami identify this specialization with Kashaev's invariant [80]. For a hyperbolic knot, the volume conjecture is

$$\lim_{N \to \infty} \frac{2\pi}{N} \log |J_N(K; e^{2\pi i/N})| = \operatorname{Vol}(S^3 \setminus K). \tag{10.4}$$

The right side is the ordinary three-dimensional hyperbolic volume, not a volume of moduli space. The assertion concerns a sequence in which both the color and the root change.

The closed-manifold version used in Section 16 is the Chen–Yang asymptotic at their root $\zeta_r = e^{2\pi i/r}$, with r odd:

$$\frac{4\pi}{r}\log|\mathrm{RT}_r(M;\zeta_r)| \longrightarrow \mathrm{Vol}(M), \qquad \frac{2\pi}{r}\log \mathrm{TV}_r(M;\zeta_r) \longrightarrow \mathrm{Vol}(M). \tag{10.5}$$

These statements are conjectural in general, with theorems for specified families; the hyperbolic fillings of the figure-eight used later are covered by Wong–Yang [40]. The doubling relation between matching RT and TV theories explains the factor two between their leading logarithms. Exact identities require consistent choices: generally $\mathrm{TV}_{\mathcal{C}}$ is the RT invariant of the categorical center, and the familiar modulus-square formula holds in the matching unitary modular setting [27]. One should not equate differently normalized SU(2) and SO(3) versions without conversion.

There are two root-of-unity subtleties. First, authors use parameters differing by a square: in a common convention $q_J = q_{\mathrm{alg}}^2$. Within Chen–Yang's conventions, the volume root $e^{2\pi i/r}$ differs from the standard compact root $e^{\pi i/r}$; the latter belongs to a different asymptotic problem [38]. Second, even where the numerical roots agree after changing notation, the colors may not belong to the same theory. At $q_{\mathrm{alg}} = e^{\pi i/N}$, the N-dimensional color has $[N]_{q_{\mathrm{alg}}} = 0$ and lies outside the retained simple labels of the level-$N-2$ compact fusion theory. Reducing J_N first and then specializing is essential. Thus the Kashaev limit should not be read as an allowed fixed-color observable in that finite unitary theory.

10.3 A complete example: the figure-eight knot

For the figure-eight knot 4_1, the root specialization is an especially simple positive sum:

$$\boxed{J_N(4_1;e^{2\pi i/N}) = \sum_{k=0}^{N-1} P_{N,k}, \qquad P_{N,k} = \prod_{j=1}^{k} 4\sin^2\frac{\pi j}{N}, \quad P_{N,0} = 1.} \tag{10.6}$$

Kashaev gives the product formula [28, eqs. (2.2), (2.5)]. The following elementary asymptotic proof is presented by Murakami and credited there to Ekholm [29, Section 3.2].

The ratio of consecutive products is

$$\frac{P_{N,k}}{P_{N,k-1}} = 4\sin^2\frac{\pi k}{N}.$$

It is less than one before $N/6$, greater than one between $N/6$ and $5N/6$, and less than one afterward. The global maximum is therefore attained at $m_N = \lfloor 5N/6 \rfloor$, allowing an adjacent tie. To compare with the initial value 1, use the elementary identity $P_{N,N-1} = N^2$. Positivity now gives

$$P_{N,m_N} \le J_N(4_1;e^{2\pi i/N}) \le N P_{N,m_N}.$$

The two logarithmic bounds differ by $\log N$, which disappears after division by N. The remaining Riemann sum has an integrable logarithmic singularity at zero, and yields

$$\lim_{N\to\infty}\frac{1}{N}\log J_N(4_1;e^{2\pi i/N}) = \int_0^{5/6}\log(4\sin^2\pi x)\,\mathrm{d}x. \tag{10.7}$$

The endpoint causes no difficulty: subtract $2\log x$ near zero; the remaining function is continuous and the logarithmic part is controlled by Stirling's formula.

Define the Lobachevsky function by

$$\Lambda(\theta) = -\int_0^{\theta}\log|2\sin t|\,\mathrm{d}t.$$

Its elementary symmetry and duplication identities give

$$\lim_{N\to\infty}\frac{2\pi}{N}\log J_N(4_1;e^{2\pi i/N}) = -4\Lambda(5\pi/6) = 4\Lambda(\pi/6)$$
$$= 6\Lambda(\pi/3) = 2v_3 = \mathrm{Vol}(S^3\setminus 4_1), \tag{10.8}$$

where $v_3 = 3\Lambda(\pi/3)$ is the volume of a regular ideal tetrahedron. The complete figure-eight complement consists of two such tetrahedra. Its volume is approximately 2.02988321.

This calculation already displays the mechanism: an exact quantum sum, a logarithmic classical phase, a dominant critical point, and a geometric critical value. It also shows why selecting the critical point matters. For $f(t) = \int_0^t \log(4\sin^2\pi x)\,\mathrm{d}x$, the stationary point $t = 1/6$ is a minimum, while $t = 5/6$ is the relevant maximum. Here positivity controls the entire sum. Complex-valued sums and integrals require a separate analysis of their contributing critical points and possible cancellation.

10.4 Why dilogarithms occur

An elementary convergent product explains the terminology. For $\varepsilon > 0$ and $|z| < 1$, set

$$(z;e^{-\varepsilon})_\infty = \prod_{n=0}^{\infty}(1 - ze^{-n\varepsilon}).$$

Expanding the logarithm and summing the geometric progression gives

$$\log(z;e^{-\varepsilon})_\infty = -\sum_{m=1}^{\infty}\frac{z^m}{m(1-e^{-m\varepsilon})}$$
$$= -\frac{\mathrm{Li}_2(z)}{\varepsilon} + \frac{1}{2}\log(1-z) + O(\varepsilon), \tag{10.9}$$

uniformly on compact subsets of $|z| < 1$. The logarithm of this "quantum dilogarithm" has the ordinary dilogarithm as its leading term. Faddeev–Kashaev's quantum pentagon identity deforms the classical five-term dilogarithm identity [30].

The geometry enters in two ways. For an oriented ideal tetrahedron with shape z in the upper half-plane,

$$\mathrm{Vol}(T_z) = D(z) = \Im\,\mathrm{Li}_2(z) + \arg(1-z)\log|z|. \tag{10.10}$$

The logarithmic correction matters: an arbitrary dilogarithm phase does not have volume as its imaginary part without the appropriate boundary terms. Also, $\mathrm{d}\,\mathrm{Li}_2(z) = -\log(1-z)\,\mathrm{d}\log z$. Consequently differentiation of a sum of dilogarithms produces logarithms of tetrahedral shapes, which can impose gluing equations.

Here is a particularly direct version in the convention of Dimofte–Garoufalidis [31, Section 5.1]. Their tetrahedral quantum function obeys

$$\psi_\hbar(Z+\hbar) = (1-e^{-Z})\psi_\hbar(Z), \qquad \log\psi_\hbar(Z) \sim \frac{1}{\hbar}\,\mathrm{Li}_2(e^{-Z}) + O(1). \tag{10.11}$$

Putting $\psi_\hbar \sim e^{S/\hbar}$ in the difference equation gives $e^{S'(Z)} = 1 - e^{-Z}$. With $z = e^Z$, the right side is the companion tetrahedral shape $1 - z^{-1}$. Thus a difference equation for a special function becomes an algebraic shape relation in the limit.

A *state integral* combines these tetrahedral functions with quadratic exponential factors determined by the gluing data and integrates over internal logarithmic variables. Differentiating its leading phase gives the logarithmic gluing equations; exponentiating them gives the familiar

products of shapes around edges. The boundary variables are retained as arguments of the integral. The precise quadratic term, logarithm branches, and contour are part of the construction. The convergent product in (10.9) explains the mechanism in a simple domain; one cannot substitute a unit-modulus root into that product and assume it converges. The continued or noncompact functions used later have their own definitions and contour prescriptions.

10.5 Stationary phase, loops, torsion, and degenerate saddles

The word *semiclassical* means an asymptotic regime in which the quantum parameter tends to zero while the intended geometric data remain controlled. Much of the calculation is ordinary asymptotic analysis. Consider

$$I(\hbar) = \int_\Gamma e^{S(x)/\hbar} a(x;\hbar)\ \mathrm{d}^m x, \qquad a(x;\hbar) = a_0(x) + \hbar a_1(x) + \cdots . \tag{10.12}$$

Suppose x_* is a nondegenerate critical point with $a_0(x_*) \neq 0$, the parameter tends to zero along a fixed ray, and Γ contains the corresponding steepest-descent cycle. Taylor expansion and Gaussian integration give its local contribution

$$I_{x_*}(\hbar) \sim e^{S(x_*)/\hbar} (2\pi\hbar)^{m/2} \frac{a_0(x_*)}{\sqrt{\det(-S''(x_*))}} (1 + c_1\hbar + \cdots). \tag{10.13}$$

The contour fixes the phase and square-root branch. For a real Laplace integral with a strict maximum this is the usual Gaussian formula; for a complex phase it is applied along suitable real m-dimensional contours in the complexified variables.

This formula gives a useful translation of the physics terminology. The *classical* or *tree-level* contribution is the critical value $S(x_*)/\hbar$. The *one-loop* contribution includes the Gaussian determinant and leading amplitude or measure factors. Higher-loop terms come from the higher Taylor derivatives and amplitude corrections. Feynman diagrams organize the Gaussian contractions needed to compute these coefficients; the terminology does not require an additional physical assumption. In logarithmic form one writes

$$\log I_{x_*} = \frac{S_0}{\hbar} + \nu \log \hbar + S_1 + S_2\hbar + \cdots ,$$

where ν depends on the integral's normalization and any zero modes (kernel directions of the linearized operator). A geometric expansion of S_0 in a shrinking area or an inverse Dehn-filling slope is a separate expansion. The cubic and quintic Neumann–Zagier coefficients are not, on that account, quantum higher-loop coefficients.

Torsion is the topological form taken by a determinant of linearized equations after gauge directions have been accounted for. For a flat holonomy representation ρ, the relevant cellular complex is $C^\bullet(M; \mathrm{Ad}\,\rho)$, with the holonomy acting on the Lie algebra by conjugation. Its H^0 records infinitesimal stabilizers and its H^1 records infinitesimal deformations. Reidemeister torsion compares the specified cellular bases with bases adapted to the differentials, using an alternating product of determinants. Nonzero cohomology requires cohomology bases as well; for a cusped three-manifold, peripheral curves enter this normalization. This explains the appearance of adjoint torsion in semiclassical amplitudes [32, 31].

An individual Hessian determinant is not the torsion by definition: measure factors, gauge directions, and basis conventions must be matched. In the surgery convention used below the relevant amplitude is proportional to $\mathrm{Tor}^{-1/2}$. Section 16 supplies the actual peripheral and solid-torus factors and an established figure-eight comparison.

There is also a global question. A *Lefschetz thimble* is a steepest-descent cycle attached to a critical point. Under the usual analytic hypotheses, the exact contour can be represented in

an appropriate relative homology group as

$$\Gamma = \sum_{\alpha} n_\alpha \mathcal{J}_\alpha.$$

The coefficients determine which saddles contribute. Their description can change across a *Stokes locus* as parameters vary; the exact integral can remain analytic while its decomposition into saddle contributions changes [33, Sections 2–3]. To recover the asymptotic of a whole invariant, one must control these coefficients, all competing contributions, and cancellation. The figure-eight positive sum supplied this global control by an elementary inequality. A local geometric saddle in a general state integral supplies only its local contribution until these additional questions are settled.

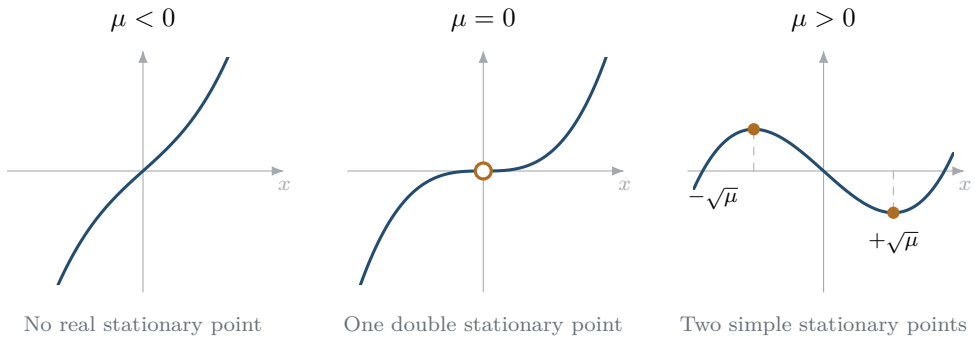
Finally, a degenerate Hessian changes the local model. At a simple fold, two saddles merge. After separating the nondegenerate transverse directions, the remaining phase has the form

$$S(x;t) = S_0 + \frac{x^3}{3} - tx + \text{higher terms}.$$

The substitution $x = \hbar^{1/3} y$, $t = \hbar^{2/3} s$ gives

$$\frac{S - S_0}{\hbar} = \frac{y^3}{3} - sy + \cdots .$$

The resulting contour integral is an Airy integral, rather than a sum of two separately uniform Gaussian approximations. This explains the exponent $2/3$ in the transition scale. For the regular family in Section 15, the real small parameter is $h = 2\pi b^2$, so the geometric deformation has scale $b^{4/3}$. When that deformation is proportional to boundary area A, the same scale follows from $W \asymp A^{3/2}$ and $W/h = O(1)$. The later theorem proves the necessary contour and error estimates for its specified family; the scaling argument alone does not.



$$\mathcal{S}_\mu(x) = \frac{x^3}{3} - \mu x, \qquad x = O(\hbar^{1/3}), \quad \mu = O(\hbar^{2/3}).$$

Figure 5: The cubic normal form for an Airy transition. The stationary equation is $x^2 = \mu$; two stationary points merge at $\mu = 0$. The scaling shown balances the cubic and linear terms in $\exp(i\mathcal{S}_\mu/\hbar)$. This is a normal-form illustration, not the full phase or integration contour of a quantum invariant; normalization, amplitudes, and contour tails require the separate estimates in the text.

10.6 Flat connections, Chern–Simons theory, and boundary states

The connection with hyperbolic geometry can be phrased using flat connections. If $\mathscr{A}$ is a connection one-form, its curvature is $F_{\mathscr{A}} = \mathrm{d}\mathscr{A} + \mathscr{A} \wedge \mathscr{A}$. In a chosen trivialization the Chern–Simons functional has the form

$$\mathrm{CS}(\mathscr{A}) = \frac{1}{4\pi} \int_M \mathrm{tr}\left(\mathscr{A} \wedge \mathrm{d}\mathscr{A} + \frac{2}{3} \mathscr{A} \wedge \mathscr{A} \wedge \mathscr{A} \right). \tag{10.14}$$

The trace convention fixes the periods and overall normalization. On a closed manifold its first variation is $\delta\,\mathrm{CS} = (2\pi)^{-1}\int_M \mathrm{tr}(\delta\mathscr{A}\wedge F_{\mathscr{A}})$. Thus the critical points are flat connections modulo gauge. The formal expression $\int e^{ik\,\mathrm{CS}(\mathscr{A})}\,D\mathscr{A}$ motivates the theory; exact algebraic constructions of quantum invariants do not depend on treating this symbol as an ordinary convergent integral [32].

For the compact group SU(2), the flat connections form the compact character-variety problem. Hyperbolic three-dimensional holonomy instead takes values in $\mathrm{PSL}_2(\mathbb{C})$, with a lift to $\mathrm{SL}_2(\mathbb{C})$ when chosen. It supplies a distinguished *complex* flat connection. The correctly normalized complex Chern–Simons action contains both the volume and the Chern–Simons invariant. Analytic continuation and the choice of integration cycle connect these settings; merely changing the name of the gauge group does not specify the cycle [34, 33].

For a manifold with boundary, variation of the functional has a boundary term. This is the source of the symplectic form on boundary flat connections and the need to specify boundary data. Abstractly, a topological quantum field theory assigns a vector space $\mathcal{H}(\Sigma)$ to an oriented surface, and a vector $\Psi_M \in \mathcal{H}(\partial M)$ to a three-manifold with boundary. Gluing is a pairing:

$$Z(M_1 \cup_\Sigma M_2) = \langle \Psi_{M_1}, \Psi_{M_2} \rangle,$$

with orientation reversal supplying the dual space and the theory's framing conventions understood. In finite root-of-unity theories this is finite-dimensional linear algebra. In the analytic models used for complex holonomy, a state is represented locally by a function or distribution of continuous boundary variables, and the pairing becomes an integral. Those analytic models require their own existence and convergence statements.

A *polarization* chooses which half of the conjugate boundary coordinates are arguments of the state. For a torus boundary, use the multiplier logarithms u, v of Section 16. The Neumann–Zagier normalization is

$$\Omega_\partial = \frac{1}{2}\,\mathrm{d}v \wedge \mathrm{d}u, \qquad \alpha_\partial = \frac{1}{2} v\;\mathrm{d}u. \tag{10.15}$$

The geometric deformations of the interior give a Lagrangian curve $v = v(u)$. A semiclassical boundary state associated with this branch has the form

$$\Psi_M(u;\hbar) \sim e^{\Phi(u)/\hbar}(a_0(u) + \hbar a_1(u) + \cdots), \qquad \Phi'(u) = \frac{1}{2}v(u). \tag{10.16}$$

The leading exponent is a primitive of the symplectic potential on the Lagrangian curve. This is the geometric content of the WKB (Wentzel–Kramers–Brillouin) ansatz [35].

This choice of half-boundary variables also explains the boundary correction to volume. With $F' = v$ and $F(0) = 0$, the note uses $\Phi = \Phi_0 + F/2$, with $\Im\Phi_0 = V_0$ the complete volume, while

$$V(u) - V_0 = \frac{1}{2}\Im F(u) - \frac{1}{2}\Re u\,\Im v(u). \tag{10.17}$$

The last term is essential. The imaginary part of an action in an arbitrary boundary polarization need not itself be the incomplete hyperbolic volume. Section 16.10 derives this identity and the corresponding filling transformation explicitly.

10.7 Quantizing peripheral holonomy and performing surgery

Set $P = v/2$, so that $\Omega_\partial = \mathrm{d}P \wedge \mathrm{d}u$. In the formal convention compatible with $e^{\Phi/\hbar}$, use operators

$$\widehat{u}\,\Psi = u\Psi, \qquad \widehat{P}\,\Psi = \hbar\partial_u\Psi, \qquad [\widehat{P}, \widehat{u}] = \hbar. \tag{10.18}$$

The usual physical factor i has been absorbed into the parameter $\hbar$; the formula defines the convention here. Acting on the leading exponential gives $\widehat{P}\Psi \sim \Phi'\Psi = P\Psi$. Thus differentiation becomes the classical longitude coordinate in the small-$\hbar$ limit.

In local eigenvalue coordinates $m = e^{u/2}$ and $l = e^{v/2}$, the corresponding operators are multiplication and translation:

$$\widehat{m}\Psi(u) = e^{u/2}\Psi(u), \qquad \widehat{l}\Psi(u) = \Psi(u+\hbar), \qquad \widehat{l}\widehat{m} = e^{\hbar/2}\widehat{m}\widehat{l}.$$

Fixed signs from an SL_2 lift can multiply m and l; they do not change this commutation relation. The factor $e^{\hbar/2}$ is the Weyl parameter in these coordinates. It should not be identified with every author's variable called q.

The classical boundary holonomies satisfy a polynomial relation $A(l,m) = 0$ on the appropriate character-variety component. To *quantize* such a relation is to seek a difference operator $\widehat{A}(\widehat{l},\widehat{m};\hbar)$ annihilating the boundary state. Substituting (10.16) gives, to leading order,

$$A(e^{\Phi'(u)}, e^{u/2}) = 0.$$

This recovers the classical curve because $\Phi' = v/2$. It does not uniquely determine the exact operator: ordering and lower-order terms remain. Identifying a particular operator with a recursion for a particular quantum invariant is a further mathematical statement [34, 35].

Dehn filling has a similarly concrete meaning. For a slope (p,q), choose $\rho,\sigma \in \mathbb{Z}$ with $p\sigma - q\rho = 1$ and put

$$a = pu + qv, \qquad b_{\mathrm{core}} = \rho u + \sigma v.$$

Then $\mathrm{d}b_{\mathrm{core}} \wedge \mathrm{d}a = \mathrm{d}v \wedge \mathrm{d}u$. For $q \neq 0$ the generating function

$$G(u,a) = \frac{pu^2}{4q} - \frac{ua}{2q} + \frac{\sigma a^2}{4q} \tag{10.19}$$

satisfies $\mathrm{d}G = \frac{1}{2}b_{\mathrm{core}}\,\mathrm{d}a - \frac{1}{2}v\,\mathrm{d}u$ on this relation. An integral transform with exponential kernel $e^{G/\hbar}$ therefore changes the boundary coordinate from u to a. Its stationary equation for $\Phi(u) + G(u,a)$ is exactly $a = pu + qv$. At the classical level, pairing with a solid-torus state imposes the geometric filling branch $a = 2\pi i$; the cone-angle family uses $a = i\theta$. The full surgery kernel includes amplitude and topological normalization factors in addition to this exponential.

In the closed-filling convention of this monograph, $\hbar_r = 4\pi i/r$. If the filled critical phase has imaginary part V, then

$$\left|e^{\Phi_{\mathrm{filled}}/\hbar_r}\right| = \exp\left(\frac{rV}{4\pi}\right). \tag{10.20}$$

This is the RT exponential in (10.5). The Gaussian change of boundary coordinate gives $(p + qv'(u))^{-1/2}$; the solid torus contributes further factors. The calculation in Section 16 makes each of these statements precise and states the hypotheses needed to pass from the geometric contribution to an exact invariant. Notice that the inverse level r^{-1} and the inverse slope length $|p+q\tau|^{-1}$ are independent small parameters; here $\tau = v'(0)$ is the complete cusp shape.

10.8 Liouville theory, conformal blocks, and continuous fusion

Take $0 < b < 1$. The other quantum construction in this monograph starts in two dimensions. In Liouville theory the variable is a conformal factor ϕ on a fixed Riemann surface with reference metric g; formally the new area element is $e^{2b\phi}\,\mathrm{d}A_g$. The field is distributional, so this exponential and its point insertions require renormalization. The Gaussian free field is

a generalized Gaussian function with inverse-Laplacian covariance. Gaussian multiplicative chaos constructs its exponential measure by smoothing the field, subtracting half the variance of the scaled field from the exponent, and taking a limit. In the subcritical range above this gives the rigorous random measures used here [66].

The basic observables are the renormalized exponential insertions $V_\alpha(p) \sim e^{2\alpha\phi(p)}$. A *correlation function* is the averaged product

$$C_b(\boldsymbol{\alpha};\boldsymbol{p}) = \left\langle \prod_i V_{\alpha_i}(p_i) \right\rangle .$$

It depends on the marked conformal surface. It is neither a three-manifold invariant nor, without a further operation, a measure on moduli space. The parameters governing conformal symmetry are

$$Q = b + b^{-1}, \qquad c_L = 1 + 6Q^2, \qquad \Delta(\alpha) = \alpha(Q-\alpha). \tag{10.21}$$

Here c_L is the *central charge* and Δ the *conformal weight*. Concretely the infinitesimal symmetry algebra is the Virasoro algebra

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c_L}{12}(m^3 - m)\delta_{m+n,0}.$$

A highest-weight vector has L_0-eigenvalue Δ and is killed by L_n for $n > 0$; its descendants are obtained with the negative generators. This representation theory determines the power series called conformal blocks [93].

On a four-marked sphere with marks $0, 1, z, \infty$, a pants decomposition first combines two of the insertions. Fixing the intermediate representation with label α_s gives a *chiral conformal block*: a holomorphic, generally multivalued function with a specified leading behavior, for example

$$\mathcal{F}^{(s)}_{\alpha_s}(z) = z^{\Delta(\alpha_s)-\Delta(\alpha_1)-\Delta(\alpha_3)}(1 + O(z)).$$

"Chiral" refers here to the holomorphic half. The full correlation pairs holomorphic and antiholomorphic blocks and integrates over the intermediate spectrum. For real external charges in its initial convergence domain this has the form

$$C_b(z,\bar z) = \int_0^\infty \rho(\alpha_1,\dots,\alpha_4;P)\, |\mathcal{F}^{(s)}_{Q/2+iP}(z)|^2 \, \mathrm{d}P,$$

with normalization and products of three-point structure constants included in ρ. The explicit three-point constants are the DOZZ constants, named for Dorn–Otto and Zamolodchikov–Zamolodchikov. The relevant three-point formula and four-point decomposition have rigorous constructions and proofs in their stated domains [66, 59]. Continuing the charges outside that domain can require additional residues when poles cross the spectral contour.

Changing the pants decomposition gives the *fusion kernel*:

$$\mathcal{F}^{(s)}_{\alpha_s}(z) = \int_{\mathcal{C}} F_{\alpha_s\alpha_t}\, \mathcal{F}^{(t)}_{\alpha_t}(z)\, \mathrm{d}\alpha_t. \tag{10.22}$$

The measure is either displayed separately or incorporated in F. This is the continuous counterpart of (10.1): four external labels and two internal ones again give six labels. Rescaling the basis blocks changes the displayed kernel. Accordingly the fusion kernel is related to a b-6j symbol through explicit trivalent normalization and spectral factors, not by an unqualified equality of scalar functions [94, Section 4].

The noncompact representation theory behind this kernel is the *modular double*. It uses the compatible deformation parameters

$$q_{\mathrm{alg}} = e^{\pi i b^2}, \qquad \widetilde{q}_{\mathrm{alg}} = e^{\pi i/b^2}.$$

The principal representation labels are $\alpha = Q/2 + iP$; tensor products decompose by direct integrals instead of finite sums. The double refers to the simultaneous roles of b and b^{-1}. The corresponding b-6j is a continuous recoupling kernel. This is a different setting from the finite root-of-unity label set in Section 10.1.

Its explicit formulas use Barnes' double Gamma function Γ_b, a meromorphic analogue of the Gamma function with the two step sizes b and b^{-1}, and the *double sine*

$$S_b(x) = \frac{\Gamma_b(x)}{\Gamma_b(Q-x)}, \qquad \begin{aligned} S_b(x+b) &= 2\sin(\pi b x) S_b(x), \\ S_b(x+b^{-1}) &= 2\sin(\pi x/b) S_b(x), \\ S_b(x)S_b(Q-x) &= 1. \end{aligned} \tag{10.23}$$

These identities explain why sine factors and pole movements occur in the later calculations. The double sine is related to Faddeev's noncompact quantum dilogarithm by shifted arguments and explicit Gaussian factors; conventions matter [94, Appendix A].

The first shift relation also explains its classical limit. Put $L_b(x) = b^2 \log S_b(x/b)$ for real x in the fundamental strip. Then

$$L_b(x+b^2) - L_b(x) = b^2 \log(2\sin \pi x).$$

A smooth leading limit must therefore have derivative $\log(2\sin \pi x)$. Reflection fixes the constant, giving $L(x) = -\Lambda(\pi x)/\pi$ for $0 < x < 1$. This is a formal derivation of the leading term; Lemma 15.2 supplies the uniform estimate used in the proof. Thus the same Lobachevsky function that appeared in the elementary figure-eight sum also appears in the continuous b-6j analysis.

10.9 Why a screening residue meets flat cone geometry

The particular Liouville specialization in this monograph can be understood through a one-variable integral. Separate the constant part c_0 of the field from its nonconstant part. After the insertions have shifted the Gaussian field, let $Z > 0$ be the remaining random area factor and put

$$s = \frac{\sum_i \alpha_i - Q}{b}.$$

With the overall normalization adopted in (12.2), the integral over the constant mode is

$$2\int_{\mathbb{R}} e^{2bsc_0 - \Lambda_L Z e^{2bc_0}}\, \mathrm{d}c_0 = \frac{1}{b}\Lambda_L^{-s}\Gamma(s)Z^{-s}, \qquad \Re s > 0. \tag{10.24}$$

This is just the substitution $t = \Lambda_L Z e^{2bc_0}$. The positive number Λ_L is the cosmological constant, or coefficient penalizing total area. The formula explains both the Gamma factor in the correlation and the usefulness of removing that factor in the unit-volume normalization.

Now take the deficit weights of the flat cone metric:

$$0 < \mu_i < 1, \quad \sum_i \mu_i = 2, \qquad \alpha_i = \frac{\mu_i}{2b}.$$

Then $\sum_i \alpha_i = 1/b = Q - b$, and therefore $s = -1$. The raw integral (10.24) does not converge there. Its meromorphic continuation has a Gamma pole, whereas the continued unit-volume expression contains $Z^{-s} = Z$: its *first moment*. The normalized random measure has the prescribed reference measure as its expectation. The reference factors cancel at these weights, leaving

$$\mathbb{E}[Z] = \mathcal{A}_\mu(\boldsymbol{p}) = \int_{\mathbb{C}} \prod_i |w - p_i|^{-2\mu_i}\, \mathrm{d}^2 w. \tag{10.25}$$

This is precisely the area of the flat cone metric that underlies the Thurston–Deligne–Mostow construction. The charge convention in the detailed probabilistic formulas is $\beta_i = 2\alpha_i$. Section 12 checks the continuation and integrability conditions and proves the exact finite-b identity.

The remaining factors in the correlation are pairwise powers of $|p_i - p_j|$, whose logarithms are pluriharmonic away from collisions. Consequently the operation that produces the geometric form is

$$-2i\,\partial\bar{\partial}\log C_b^{\mathrm{uv}} = \omega_{\mathrm{CH},\mu}, \qquad \mathrm{d}V_{\mathrm{CH},\mu} = \frac{\omega_{\mathrm{CH},\mu}^{n-3}}{(n-3)!}. \tag{10.26}$$

The correlation supplies a logarithmic potential; its logarithmic curvature supplies the form; the top exterior power supplies the moduli measure. These are three different operations. This is why the exact comparison does not identify an unmodified quantum amplitude with a WP or complex hyperbolic volume density.

The name *one screening* has an equally concrete explanation. The insertion V_b has conformal weight $\Delta(b) = b(Q-b) = 1$, so integrating it gives a conformally covariant insertion. At $\sum_i \alpha_i = Q - b$, one such insertion balances the background charge. Its chiral contour integrals are

$$\int_\gamma \prod_i (w - p_i)^{-\mu_i}\,\mathrm{d}w,$$

the Deligne–Mostow periods. For four points the relevant twisted period space is two-dimensional, and changing contour bases gives the finite hypergeometric fusion matrix calculated in Section 13. Generic Liouville fusion is still the integral operator (10.22); the 2×2 matrix belongs to this specific residue sector. Its auxiliary tetrahedron also need not be the convex realization of the original cone sphere. Section 14 checks that geometric comparison separately.

10.10 Parameters, limits, and a route through the calculations

The notation records several independent limits. The following dictionary prevents a formal resemblance of symbols from identifying different asymptotic questions.

Parameter	Meaning and controlled limit
N	Color and root order in the Kashaev specialization; both change in (10.4).
r	Root-order parameter in the finite three-manifold invariants; $\hbar_r = 4\pi i/r$ in the closed-filling phase here.
b	Continuous Liouville/modular-double parameter; the regular-family calculation uses $h = 2\pi b^2$.
α_i	Representation or insertion labels; the flat-weight Liouville specialization uses $\alpha_i = \mu_i/(2b)$.
A	Geometric boundary area tending to zero in the Euclidean degeneration; it is not a quantum parameter.
(p, q) and θ	Dehn-filling slope and core cone angle; $pu + qv = i\theta$, with $\theta = 2\pi$ for smooth filling.

The last charge prescription is the pure heavy-charge scaling: the labels grow like b^{-1} while the deficit weights stay fixed. The exact length-kernel dictionary has additional finite-b shifts, recorded in (12.22). Fixed labels with $q \to 1$, labels growing with the order of a root of unity, and the continuous $b \to 0$ limit are different problems. The classical large-spin $6j$ asymptotic involves Euclidean tetrahedra [89]; particular compact root regimes involve spherical tetrahedra [26]; the hyperbolic regimes used below have their own label and contour

hypotheses. Geometric edge lengths or angles are assigned only after specifying the relevant regime.

The quantum part of this monograph can now be read along three routes. For the boundary-moduli comparison, begin with Section 12, then follow the four-point residue and its normalization in Sections 13 and 14. For the degeneration of an individual tetrahedron, use Section 11, the regular family in Section 14, and the uniform Airy theorem in Section 15. For three-manifold volume and surgery, Section 16 is largely independent of the Liouville calculations; its symplectic explanation is in Section 16.10.

An exact identity of kernels, a local saddle expansion, and a theorem about a global invariant answer different questions. The later sections state which one is obtained at each step. This distinction permits the geometric calculations to be used fully without silently assuming an integration cycle, a measure on moduli, or a dominance theorem that has not been established.

11 Polyhedral degeneration and quantum volume asymptotics

11.1 Three different geometric quantities

The connection to the volume conjecture becomes precise only after separating three quantities:

$$A = \mathrm{Area}(\partial P), \qquad W(P) = \mathrm{Vol}_{\mathbb{H}^3}(P), \qquad V_{\mathrm{WP}} = \int_{\mathcal{M}} \frac{\omega_{\mathrm{WP}}^d}{d!}.$$

The first is a surface area, the second an enclosed three-dimensional volume, and the third a measure of a space of shapes. The ordinary knot volume conjecture concerns three-dimensional volume. For a hyperbolic knot, in the reduced colored-Jones normalization $J_N(\mathrm{unknot}; q) = 1$, its familiar form is

$$\lim_{N\to\infty} \frac{2\pi}{N} \log \left| J_N(K; e^{2\pi i/N}) \right| = \mathrm{Vol}(S^3 \setminus K). \tag{11.1}$$

This remains a conjecture in general; see the colored-Jones formulation in [80] and the tetrahedral results in [52].

Suppose a shrinking family P_A, after lengths are multiplied by $A^{-1/2}$, converges to a *nondegenerate* Euclidean polyhedron P_{E} of surface area one. Ordinary volume scaling gives

$$W(P_A) = A^{3/2} W_{\mathrm{E}}(P_{\mathrm{E}}) + o(A^{3/2}). \tag{11.2}$$

In contrast, the corresponding moduli volume satisfies

$$V_{\mathrm{WP}}(A) = 2^{-d} A^d V_{\mathrm{CH}}(\mu) + O(A^{d+1}).$$

The different powers reflect different dimensions. Flattened Euclidean limits and changes of polyhedral combinatorics require additional care; they are not covered by the nondegeneracy hypothesis in (11.2).

11.2 An explicit small-polyhedron expansion

A controlled model gives more than the leading power. Let Q be a bounded nondegenerate convex Euclidean polyhedron, and place $P_r = rQ$ in the curvature--1 Klein ball. Its Euclidean planar faces are hyperbolic geodesic faces. The Klein volume density is

$$\mathrm{d}\,\mathrm{Vol}_{\mathbb{H}^3} = (1 - |x|^2)^{-2}\ \mathrm{d}^3x.$$

Consequently, whenever $r \sup_Q |y| < 1$,

$$\begin{aligned} W(rQ) &= r^3 \int_Q (1 - r^2|y|^2)^{-2} \, \mathrm{d}^3y \\ &= \sum_{k=0}^{\infty} (k+1) r^{3+2k} \int_Q |y|^{2k} \, \mathrm{d}^3y \\ &= r^3 W_{\mathrm{E}}(Q) + 2r^5 \int_Q |y|^2 \, \mathrm{d}^3y + 3r^7 \int_Q |y|^4 \, \mathrm{d}^3y + \cdots . \end{aligned} \tag{11.3}$$

This convergent expansion is elementary and exact for the specified homothetic Klein path. It is not generally the fixed-conformal angle path (8.1). Nor does it imply a Taylor expansion of this form at ideal, flattening, or otherwise uncontrolled degenerations. Unlike (8.10), it is an infinite series: the finite polynomial for the moduli volume comes from a quadratic cohomology class and a top intersection.

11.3 Schläfli, doubling, and degeneration/regeneration

For a compact hyperbolic polyhedron with interior dihedral angles α_e and edge lengths ℓ_e, Schläfli's formula is

$$\mathrm{d}W = -\frac{1}{2} \sum_e \ell_e \, \mathrm{d}\alpha_e. \tag{11.4}$$

For the Klein family above, write

$$\ell_e(r) = rL_e + O(r^3), \qquad \alpha_e(r) = \alpha_e^{\mathrm{E}} + r^2 b_e + O(r^4).$$

The leading term of (11.3) and (11.4) imply

$$\boxed{\sum_e L_e b_e = -3W_{\mathrm{E}}(Q).} \tag{11.5}$$

Indeed, $\mathrm{d}W/\,\mathrm{d}r = 3r^2 W_{\mathrm{E}} + O(r^4)$, while the right side of (11.4), divided by $\mathrm{d}r$, is $-r^2 \sum_e L_e b_e + O(r^4)$. Thus the first curvature correction to the Euclidean dihedral angles already determines the Euclidean enclosed volume. The sign reverses if exterior dihedral angles are used. The same identity holds with $\ell_e = A^{1/2} L_e + O(A^{3/2})$ and $\alpha_e = \alpha_e^{\mathrm{E}} + A b_e + O(A^2)$ for a smooth unit-area Euclidean limiting family.

Doubling a convex polyhedron across its faces produces a cone three-manifold with singular graph given by its edges. The cone angle along an edge is $2\alpha_e$, and its volume is $2W(P)$. A shrinking polyhedron therefore supplies a direct example of degeneration to a Euclidean cone three-manifold after rescaling. This is the geometric setting suggested by Hodgson's thesis, *Degeneration and regeneration of geometric structures on 3-manifolds* [61]. Later precise regeneration results include Porti–Weiss [84]; their assumptions on cone angles and the limiting structure must be checked before using them for a particular doubled polyhedron. No unrestricted regeneration theorem is being asserted here. The edge cone angles $2\alpha_e$ are also distinct from the intrinsic vertex openings θ_i of the boundary surface.

Rivin's *Volumes of degenerating polyhedra—on a conjecture of J. W. Milnor* is directly relevant [87]. It studies continuous extension of volume toward degenerate angle data, the Euclidean vanishing boundary, and quantitative control through Schläfli's formula and polar geometry. Its estimates allow logarithmically growing volume derivatives at degenerations; continuity does not imply a smooth power series at every boundary point. The controlled small-polyhedron expansion above gives a particularly explicit local calculation on the Euclidean side of this broader degeneration picture.

11.4 The established tetrahedral quantum connection

Quantum $6j$-symbols provide a concrete bridge to three-dimensional geometry. Costantino obtains hyperbolic-volume exponential growth for appropriate root-of-unity label regimes associated to truncated tetrahedra [52]. Chen–Murakami give, for a fixed hyperbolic tetrahedron with at least one ideal or ultraideal vertex and their prescribed admissible label sequences, an asymptotic of the form

$$|\{6j\}_r| \sim C\, r^{-3/2}(-\det G)^{-1/4} \exp\left(\frac{r}{2\pi} W(T)\right), \tag{11.6}$$

where G is the dihedral-angle Gram matrix and the prefactor and labels depend on the stated conventions [51]. In that regime the exponential growth rate is the tetrahedron's hyperbolic volume. Different $6j$ normalizations and doubled tetrahedral blocks change factors; the formula is not meant to identify all such conventions.

At the Euclidean end, Roberts proves the Ponzano–Regge asymptotic for classical $6j$-symbols: Euclidean tetrahedral edge lengths and dihedral angles enter an oscillatory Regge phase, and Euclidean volume enters its amplitude [89]. This is an established appearance of Euclidean tetrahedral geometry in quantum asymptotics. It does not automatically identify the classical large-label limit with a chosen root-of-unity hyperbolic limit.

In particular, the stated Chen–Murakami theorem does not cover an arbitrary family of compact tetrahedra shrinking to Euclidean ones. At a Euclidean tetrahedron the angle Gram determinant vanishes, so the displayed fixed-shape prefactor itself warns of nonuniformity. A continuous classical volume limit, even with good quantitative estimates, does not by itself control the quantum Hessian, the prefactor, or the error term in a joint limit.

11.5 A proposed joint scale and a model integral

Suppose a suitable quantum object, in a specified geometric branch and choice of variables, has an exponential contribution $\exp(W(P_A)/\hbar)$. Equation (11.2) then singles out the dimensionless ratio

$$\lambda = \frac{A^{3/2}}{\hbar}, \qquad A \asymp \hbar^{2/3}. \tag{11.7}$$

This is the regime in which geometric degeneration competes with the semiclassical limit. At this stage it is a scaling prediction; Section 15 proves a uniform Airy statement for an explicitly normalized regular b-$6j$ family. A different polarization may involve a Regge or Legendre-transformed action, so the quantum object and phase must be specified first.

There is a useful classical model making the potential role of the WP–Thurston limit explicit. Let $K \subset \mathcal{M}_{0,n}$ be a fixed compact subset with controlled nondegenerate Euclidean realizations, and define

$$\mathcal{Z}_K(A,\hbar) = \int_K \exp\left(\frac{W(P_A(X))}{\hbar}\right) \frac{\omega_{\mathrm{WP},A}^d}{d!}. \tag{11.8}$$

This is a constructed geometric integral, not an identification with a colored-Jones or other known quantum invariant. Under the corresponding uniform geometric convergence on K, take $A \to 0$, $\hbar \to 0$, and $A^{3/2}/\hbar \to \lambda_0 \in (0,\infty)$. The local limits then give

$$2^d A^{-d} \mathcal{Z}_K(A,\hbar) \longrightarrow \int_K e^{\lambda_0 W_{\mathrm{E}}(P_{\mathrm{E}}(X))} \frac{\omega_{\mathrm{CH},\mu}^d}{d!}. \tag{11.9}$$

Extending this formula to the full noncompact space requires domination and control of collision and flattened strata. Identifying its measure and phase with those of an actual quantum state integral is a further mathematical problem.

This separates the two potentially compatible roles: enclosed three-dimensional volume supplies an action, while the WP form supplies a measure on boundary shapes. Their Euclidean limits are respectively Euclidean polyhedral volume and the Thurston–Deligne–Mostow measure. Establishing that both occur together in a particular quantum construction would provide the substantive connection to the volume conjecture.

12 An exact quantum comparison and the boundary measure

The preceding question has a concrete partial answer. A special *unit-volume Liouville correlation* gives the Deligne–Mostow Kähler potential exactly, at finite quantum parameter. Independently, the boundary Reidemeister-torsion measure reduces to the Goldman measure, whose rescaled flat limit is the Deligne–Mostow measure. Neither statement identifies an arbitrary tetrahedral quantum amplitude with the model integral (11.8). The distinction between a correlation, a Kähler potential, and a volume density is essential.

12.1 A first-moment identity at the flat weights

Fix the weights (7.1), all marked points finite in the chosen coordinate, and a Liouville parameter $0 < b < 1$. Set

$$\gamma = 2b, \qquad Q = b + b^{-1}, \qquad \beta_i = \frac{\mu_i}{b}, \qquad s = \frac{\sum_i \beta_i - 2Q}{\gamma} = -1. \tag{12.1}$$

Here β_i are the charge conventions of Kupiainen–Rhodes–Vargas; in the physics convention $V_{\alpha_i} = e^{2\alpha_i \phi}$, one has $\beta_i = 2\alpha_i$. Write $\Lambda_L > 0$ for the cosmological constant, to distinguish it from both the deficit weights and the geometric area.

The exact correlation formula of [66, eqs. (2.11)–(2.15)] is

$$\begin{aligned} C_b(\boldsymbol{\beta}; \boldsymbol{p}) &= \frac{1}{b} \Lambda_L^{-s} \Gamma(s) P_b(\boldsymbol{p})\, \mathbb{E}[Z_{\boldsymbol{\beta}}^{-s}], \\ P_b(\boldsymbol{p}) &= \prod_{i<j} |p_i - p_j|^{-\beta_i \beta_j}, \\ Z_{\boldsymbol{\beta}} &= \int_{\mathbb{C}} \prod_i \left(\frac{|w|_+}{|w - p_i|} \right)^{\gamma \beta_i} M_\gamma(\mathrm{d}^2 w), \qquad |w|_+ = \max(1, |w|), \end{aligned} \tag{12.2}$$

where M_γ is the normalized Gaussian multiplicative-chaos measure in their reference metric. The only probabilistic property needed for the following calculation is its first moment,

$$\mathbb{E}[M_\gamma(\mathrm{d}^2 w)] = |w|_+^{-4}\, \mathrm{d}^2 w. \tag{12.3}$$

The standard unit-volume normalization removes the zero-mode Gamma factor:

$$C_b^{\mathrm{uv}} := \frac{\Lambda_L^s}{\Gamma(s)} C_b = \frac{1}{b} P_b\, \mathbb{E}[Z_{\boldsymbol{\beta}}^{-s}]. \tag{12.4}$$

The right side defines the continuation at $s = -1$.

Proposition 12.1 (An exact Liouville–area identity)**.** *For every $0 < b < 1$ and the fixed weights (7.1),*

$$\boxed{C_b^{\mathrm{uv}}(\boldsymbol{p}) = \frac{1}{b} \prod_{i<j} |p_i - p_j|^{-\mu_i \mu_j / b^2} \mathcal{A}_\mu(\boldsymbol{p}), \qquad \mathcal{A}_\mu = \int_{\mathbb{C}} \prod_i |w - p_i|^{-2\mu_i}\, \mathrm{d}^2 w.} \tag{12.5}$$

Consequently, on the configuration space away from collisions,

$$\boxed{-2i\,\partial\bar\partial\log C_b^{\mathrm{uv}} = \omega_{\mathrm{CH},\mu}, \qquad \mathrm{d}V_{\mathrm{CH},\mu} = \frac{1}{d!}\left(-2i\,\partial\bar\partial\log C_b^{\mathrm{uv}}\right)^d, \quad d = n-3.} \tag{12.6}$$

The form descends to $\mathcal{M}_{0,n}$ with its fixed weights and the holomorphic-sectional-curvature--1 convention.

Proof. At (12.1), the exponent $-s$ is 1. Moreover, $\gamma\beta_i = 2\mu_i$ and $\gamma\sum_i \beta_i = 4$. Applying (12.3) therefore cancels the reference-metric factors exactly:

$$\mathbb{E}[Z_\beta] = \int_{\mathbb{C}} \prod_i \left(\frac{|w|_+}{|w-p_i|}\right)^{2\mu_i} |w|_+^{-4}\,\mathrm{d}^2 w = \mathcal{A}_\mu(\boldsymbol{p}).$$

This integral is finite because each $\mu_i < 1$ and, when all marks are finite, the integrand is $O(|w|^{-4})$ at infinity. It is exactly the area in (7.5). Each $\log|p_i - p_j|$ is pluriharmonic off the diagonals. Taking $-2i\partial\bar\partial\log$ consequently removes P_b/b, leaving (7.6). A change of the Möbius normalization changes the potential by a pluriharmonic function, so its curvature descends. □

This is a deduction from the established correlation formula and the area interpretation of the period pairing; it does not use the main DOZZ structure-constant theorem. It is also an exact finite-b identity, rather than an interchange of a semiclassical limit with two moduli derivatives. The raw expression $C_b^{\mathrm{uv}}\,\mathrm{d}^{2d}p$ is *not* the moduli-space volume density: the density is the top exterior power of its logarithmic curvature.

12.2 Analytic continuation and the equivalent residue

The original unconstrained zero-mode integral is not convergent at $s = -1$. The unit-volume function, however, is positive and finite there. Indeed, the extended Seiberg bounds of [66, eq. (2.14)] require

$$-s < \min\left\{\frac{4}{\gamma^2}, \min_i \frac{2}{\gamma}(Q-\beta_i)\right\}, \qquad \beta_i < Q.$$

At our parameters these inequalities reduce to

$$1 < b^{-2}, \qquad 1 < 1 + \frac{1-\mu_i}{b^2}, \qquad \beta_i < Q, \tag{12.7}$$

and are strict. In particular, the positive moment has some integrability beyond order one. Theorem 6.1 of [66] gives holomorphic continuation in the charges of the unit-volume n-point correlation to a complex neighborhood of this real domain. Thus the pole in (12.2) is an actual meromorphic pole, not just a formal screening calculation.

Since $\mathrm{Res}_{s=-1}\,\Gamma(s) = -1$,

$$\operatorname*{Res}_{s=-1} C_b = -\frac{\Lambda_L}{b} P_b \mathcal{A}_\mu. \tag{12.8}$$

This residue means the coefficient of $(s+1)^{-1}$ with the other charge coordinates fixed. With the transverse physics coordinate $t = \sum_i \alpha_i - Q + b = b(s+1)$, it becomes $\mathrm{Res}_{t=0}\,C_b = -\Lambda_L P_b \mathcal{A}_\mu$. Consequently the logarithmic curvature of the absolute residue also equals (12.6).

In Coulomb-gas terminology this is the one-screening resonance $\sum_i \alpha_i = Q - b$; see also [93, Section 3.4]. Its chiral contour integrals are

$$\int_\gamma \prod_i (w-p_i)^{-\mu_i}\,\mathrm{d}w, \tag{12.9}$$

precisely the Deligne–Mostow periods (7.4). Their Hermitian pairing is the area. Section 13 gives an independent finite period calculation of the four-point residue and fusion matrix, and states the additional continuation and tail-control input for the spectral-contour derivation. A pointwise identification with an unnormalized modular-double $6j$ function would still require its normalization and contour to be specified. Section 14 fixes these normalizations for the induced finite operator and examines the geometric meaning of that specialization.

12.3 Four points: the exact quantum potential and its measure

Put the marks at $(0, 1, z, \infty)$, with the standard primary-field normalization at infinity. Then (12.5) reads

$$C_b^{\mathrm{uv}}(z) = \frac{1}{b}|z|^{-\mu_1\mu_3/b^2}|1-z|^{-\mu_2\mu_3/b^2}\mathcal{A}_\mu(z), \tag{12.10}$$
$$\mathcal{A}_\mu(z) = \int_{\mathbb{C}} |w|^{-2\mu_1}|w-1|^{-2\mu_2}|w-z|^{-2\mu_3}\,\mathrm{d}^2w.$$

The condition at infinity is now $\mu_1+\mu_2+\mu_3 = 2-\mu_4 > 1$. The ordinary field normalization at infinity cancels the corresponding area scaling; there is no additional moduli-dependent factor.

For equal weights $\mu_i = 1/2$, the area is

$$\mathcal{A}(z) = \int_{\mathbb{C}} \frac{\mathrm{d}^2w}{|w(w-1)(w-z)|}. \tag{12.11}$$

Let Ω_1, Ω_2 be oriented lattice periods of $\mathrm{d}w/y$ on the elliptic curve (7.9), and put $\tau = \Omega_2/\Omega_1 = u + iv \in \mathbb{H}$. The pillowcase has half the area of its double cover, so

$$\mathcal{A}(z) = \frac{1}{2}\operatorname{Im}(\overline{\Omega_1}\Omega_2) = \frac{1}{2}|\Omega_1|^2 v.$$

Locally Ω_1 is nonzero and holomorphic. Hence

$$\boxed{-2i\partial\bar\partial \log C_b^{\mathrm{uv}} = \frac{\mathrm{d}u\wedge\mathrm{d}v}{v^2} = \frac{|\tau'(z)|^2}{(\operatorname{Im}\tau(z))^2}\,\mathrm{d}(\operatorname{Re}z)\wedge\mathrm{d}(\operatorname{Im}z).} \tag{12.12}$$

The elliptic-integral expression for τ is (7.10). Integrating this form over the labeled moduli space gives 2π, as in Section 7. This is the flat equal-deficit Deligne–Mostow area; the cusped hyperbolic WP area is the different quantity $V_{0,4}(0) = 2\pi^2$.

12.4 An actual boundary torsion measure

There is also a measure comparison on the representation side. For a good, boundary-regular $\mathrm{SL}_2(\mathbb{C})$ representation of an n-holed sphere, Heusener–Porti prove

$$\Omega_S = \pm\frac{\omega_G^d}{d!}\wedge\nu_1\wedge\cdots\wedge\nu_n, \qquad d = n-3, \tag{12.13}$$

where Ω_S is the holomorphic Reidemeister-torsion volume form, ω_G is the Goldman form on the fixed-conjugacy leaf, and the ν_i are the peripheral volume forms [60, Theorem 3]. With invariant pairing $B(U,V) = -\operatorname{tr}(UV)$ their peripheral formula is $\nu_i = \pm\sqrt{2}\,\mathrm{d}\operatorname{tr}\rho(\partial_i)$ [60, Proposition 2].

One must disintegrate by the peripheral forms *before* fixing the conjugacy classes. Simply pulling the unrestricted top-degree form back to a fixed-trace leaf gives zero. After this

disintegration, and on the relevant real component, the torsion density is the Goldman Liouville density, with the chosen normalization of B.

Mondello identifies the cone WP and Goldman Poisson structures up to a specified constant for openings in $(0, 2\pi)$ [78, Theorems 1.1–1.2]. We calibrate B so that the symplectic form on our real leaf is ω_{WP}. The holonomy statement is local in general; it is globally injective on marked cone metrics when every opening is at most π. Combining (12.13) with (8.5) gives, locally on a controlled family,

$$\boxed{\left(\frac{2}{A}\right)^{d} \mathrm{d}V_{\mathrm{torsion,relative}} \longrightarrow \mathrm{d}V_{\mathrm{CH},\mu}.} \tag{12.14}$$

Apply the torsion theorem for $A > 0$ and then take the rescaled limit. At the Euclidean endpoint the affine holonomies become reducible, and their ordinary characters discard translational shape data. The limiting twisted periods or a blowup retaining their scale are indispensable. In particular, an unqualified ordinary Goldman construction for the Euclidean isometry group would miss the needed nondegenerate pairing.

12.5 The four-point character surface, including a correction

The preceding measure limit can be calculated directly. Give the four cone points the common opening $\theta = \pi - A/4$, so their total hyperbolic area is A. Choose equal peripheral traces $a = -2\cos(\theta/2) = -2\sin(A/8)$. The Fricke relation of Appendix A becomes

$$\Phi = x^2 + y^2 + z^2 + xyz - 2a^2(x + y + z) + 4a^2 + a^4 - 4 = 0. \tag{12.15}$$

Blow up the limiting character by

$$x = -2 - AX, \qquad y = -2 - AY, \qquad z = -2 - AZ. \tag{12.16}$$

Direct substitution gives

$$\begin{aligned}\frac{\Phi}{A^2} &= X^2 + Y^2 + Z^2 - 2XY - 2XZ - 2YZ + 1 \\ &\quad + A\left(-XYZ + \frac{X + Y + Z}{8}\right) - \frac{A^2}{768} + O(A^3).\end{aligned} \tag{12.17}$$

On a chart with $u > 0$, $v > 0$, parametrize the limiting quadric by

$$X = \frac{1}{2v}, \qquad Y = \frac{u^2 + v^2}{2v}, \qquad Z_0 = \frac{(u-1)^2 + v^2}{2v}. \tag{12.18}$$

These are squared side lengths of a Euclidean triangle of area $1/4$; the quadric equation is Heron's identity. Since

$$X + Y - Z_0 = \frac{u}{v}, \qquad \mathrm{d}X \wedge \mathrm{d}Y = \frac{u}{2v^3}\,\mathrm{d}u \wedge \mathrm{d}v,$$

the calibrated density (A.2) yields $A^{-1}\omega_{\mathrm{WP}} \to \frac{1}{2}\,\mathrm{d}u \wedge \mathrm{d}v/v^2$, with the positive orientation on this chart.

In fact, the implicit-function expansion is

$$Z = Z_0 + AZ_1 + O(A^2), \qquad Z_1 = \frac{(1-u)[v^2 - u(1-u)]}{16v^2}.$$

Using $\Phi_z = 2A(X+Y-Z) + A^2(XY-1/8) + O(A^4)$ gives

$$\boxed{\frac{\omega_{\mathrm{WP}}}{A} = \frac{1}{2v^2}\left[1 - \frac{A(1+u^2+v^2)}{16v} + O(A^2)\right] \mathrm{d}u \wedge \mathrm{d}v.} \tag{12.19}$$

This is an explicit first correction in a fixed *blown-up character chart*. Here $\tau = u + iv$ is the limiting flat period coordinate, not the exact period ratio or conformal cross-ratio of the finite-A surface. The error is local on compact subsets of the chart. In particular, the displayed correction is not termwise integrable at $v = \infty$; it cannot be integrated naively to recover the global polynomial of (8.10).

12.6 What a normalized quantum kernel would have to produce

The phase-space measure of a unitary change of polarization is more constrained than a bare scalar $6j$ prefactor. In one degree of freedom, let $S(q, Q)$ generate a canonical relation, $p = \partial_q S$, away from its caustic. A branchwise normalized semiclassical kernel has the leading form

$$K_\hbar(q, Q) \sim \frac{|\partial_q \partial_Q S|^{1/2}}{\sqrt{2\pi\hbar}} \exp\left(\frac{iS}{\hbar} + \frac{i\pi m}{4}\right). \tag{12.20}$$

The amplitude follows by applying unitarity to the leading phase difference $(q - q')p$ and changing variables from Q to p. Thus

$$2\pi\hbar |K_\hbar|^2 \, \mathrm{d}q \, \mathrm{d}Q \sim |\partial_q \partial_Q S| \, \mathrm{d}q \, \mathrm{d}Q = |\, \mathrm{d}q \wedge \mathrm{d}p|. \tag{12.21}$$

This is a branchwise assertion. Several stationary branches produce interference; a pointwise square of an oscillatory cosine need not converge without averaging.

Teschner–Vartanov's Fenchel–Nielsen transition kernel contains the factors $\sqrt{M(\alpha_s)M(\alpha_t)}$ multiplying the b-$6j$ symbol [94, eq. (5.14)]. Consequently a proposed measure comparison must include the spectral measure, label Jacobians, and polarization. If a normalized kernel quantizes our cone WP leaf, its classical measure is constrained by (12.21), and its rescaled flat measure by (12.14). This conditional argument does not by itself establish the elliptic continuation of that kernel.

12.7 The auxiliary tetrahedron and the order of limits

There is a concrete reason not to identify the auxiliary tetrahedron in a fusion kernel with the convex realization of our cone sphere without further work. In the conventions of [94, Sections 3–5], the parameters satisfy

$$\alpha_k = \frac{Q}{2} + \frac{i\ell_k}{4\pi b}, \qquad A_k = e^{-2\pi i b \alpha_k + \pi i} = e^{\ell_k/2 - \pi i b^2}. \tag{12.22}$$

Here A_k is the paper's exponentiated tetrahedral parameter, not our geometric area A. Continuing a boundary length by $\ell_i = i\theta_i$ gives an auxiliary dihedral parameter $\eta_i = \theta_i/2 - \pi b^2$, modulo the chosen branch. Thus at the equal-angle Euclidean endpoint, and as $b \to 0$, the four external auxiliary dihedrals tend to $\pi/2$. The internal s, t edges are opposite. After reordering faces, the limiting angle Gram matrix would be

$$G = \begin{pmatrix} 1 & -\cos\eta_s & 0 & 0 \\ -\cos\eta_s & 1 & 0 & 0 \\ 0 & 0 & 1 & -\cos\eta_t \\ 0 & 0 & -\cos\eta_t & 1 \end{pmatrix}, \qquad \det G = \sin^2\eta_s \sin^2\eta_t. \tag{12.23}$$

For interior angles $\eta_s, \eta_t \in (0, \pi)$ this is nonsingular. A nondegenerate Euclidean tetrahedron has a rank-three angle Gram matrix. Hence this real slice cannot converge to the generic Euclidean tetrahedron realizing our flat cone metric. An additional degeneration, a different real slice, or a different geometric identification is needed. This is an obstruction to the proposed identification, not to the existence of Liouville theory with cone charges.

There is also a finite-b shift in the charge convention. At a fixed classical opening $\theta_i = 2\pi(1 - \mu_i)$, the exact dictionary (12.22) gives

$$\alpha_i = \frac{\mu_i}{2b} + \frac{b}{2},$$

whereas (12.1) uses $\alpha_i = \mu_i/(2b)$. The latter corresponds in that dictionary to the shifted opening $\theta_i + 2\pi b^2$. For unshifted flat openings the zero-mode parameter would be $s = n/2 - 1$, not -1. Thus the exact correlation identity is a specified charge specialization; it is not an identification with the finite-b kernel at unshifted classical cone angles. Indeed, treating the shifted openings as hyperbolic cone data would give the negative Gauss–Bonnet area $-2\pi n b^2$.

There is a second, independent distinction. Along a heavy-charge path with deficit weights $a_i(A)$, set $\alpha_i(A) = a_i(A)/(2b)$ and $A = 2\pi(\sum_i a_i(A) - 2)$. Then exactly

$$\boxed{s = \frac{A}{4\pi b^2} - 1.} \tag{12.24}$$

The finite-b flat endpoint of this pure-heavy charge path is $s = -1$, the resonance used above. The original zero-mode convergence region $s > 0$ requires $A > 4\pi b^2$. In contrast, under the proposed joint scaling $A \asymp \hbar^{2/3}$ with $\hbar \asymp b^2$, one has $A/b^2 \to \infty$ and hence $s \to +\infty$. The exact residue identity therefore establishes a quantum origin for the Deligne–Mostow curvature without proving the uniform quantum degeneration proposed in Section 11.

13 The four-point residue: fusion, area, and an explicit density

The one-screening calculation can be completed at the level of four-point residues. Two intermediate Virasoro channels contribute, their change of basis is an explicit hypergeometric connection matrix, and the invariant Hermitian pairing is exactly the flat area. This gives both a density and an evaluation of its total volume for arbitrary four deficit weights. At balanced weights, the two ordinary blocks coalesce and must be replaced by a regularized logarithmic basis.

13.1 The two spectral pinches and their coefficients

Put the fields with charges $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ at $0, 1, z, \infty$, respectively. Throughout this section

$$0 < b < 1, \qquad Q = b + b^{-1}, \qquad \alpha_i = \frac{\mu_i}{2b}, \qquad 0 < \mu_i < 1, \qquad \sum_i \mu_i = 2. \tag{13.1}$$

To take a transverse residue, vary the charges near this locus and write

$$\varepsilon = \sum_i \alpha_i - Q + b, \qquad a_s = \alpha_1 + \alpha_3. \tag{13.2}$$

The s-channel pairs the marks $0, z$; its spectral integrand contains

$$C(\alpha_1, \alpha_3, Q - \alpha) C(\alpha, \alpha_2, \alpha_4) \mathcal{F}_\alpha(z) \mathcal{F}_\alpha(\bar{z}).$$

For real charges the last product is the absolute square. For continuation in complex charges, both chiral factors are continued analytically, rather than conjugating the variable charges. We use the DOZZ/KRV normalization fixed in Section 12.

The bootstrap identity is established in an initial convergence domain by [59, Theorem 1.1]. Its continuation requires the discrete terms created by crossing poles, as the paragraph following that theorem also explains. The relevant contour mechanism and pole families are described in [93, Section 7.2, eqs. (87)–(90)].

Analytic input for the spectral derivation. The infinite integral requires a separate continuation statement. In the present charge convention and ordering, the initial real bootstrap domain has $\alpha_i < Q/2$, $\alpha_1 + \alpha_3 > Q/2$, and $\alpha_2 + \alpha_4 > Q/2$. The one-screening locus $\sum_i \alpha_i = Q - b$ lies outside that domain. Meromorphic dependence of each integrand factor alone does not justify continuing the integral. The spectral derivation below uses the following sufficient input, at fixed $0 < b < 1$.

Continue from the initial domain along a specified path of complex charges, retaining every residue of a pole crossed by the contour. Locally, the deformed contours are required to coincide with $Q/2 + i\mathbb{R}$ outside a fixed compact set. Near a generic target point, separate the finitely many pinch contributions, including their reflected partners, and write the remaining term using a locally fixed contour Γ_{reg} avoiding all poles. Its integrand $\mathcal{I}_{\mathrm{reg}}(\alpha; \varepsilon, \eta, z)$ must be holomorphic in the charge coordinates through $\varepsilon = 0$; here η denotes the remaining charge coordinates. Any additional discrete terms left in this remainder must also be holomorphic there. Require normal absolute convergence: on every compact set K of these complex charge parameters and of noncolliding z in the chosen branch chart, there is $G_K \in L^1(\mathbb{R})$ such that on the common contour ends

$$\sup_{(\varepsilon,\eta,z)\in K} |\mathcal{I}_{\mathrm{reg}}(Q/2 + iP; \varepsilon, \eta, z)| \le G_K(P). \tag{13.3}$$

The same normal-convergence requirement is imposed on compact continuation neighborhoods along the path, after their crossed-pole terms are separated. If another contour prescription moves the ends, the connecting pieces at infinity must additionally have integrals tending to zero uniformly on those parameter sets.

Under these conditions, the truncated regular integrals are holomorphic in ε and converge locally uniformly as the spectral cutoff tends to infinity. Their limit is therefore holomorphic through $\varepsilon = 0$, and Cauchy's residue integral in that parameter commutes with removal of the cutoff. The regular term contributes no residue; only the finite pinch terms do.

This is an additional analytic input for the contour proof. No estimate (13.3) is proved here, and the cited bootstrap theorem in its initial domain is not being used as such an estimate at the one-screening locus. The residue identity below also has an independent first-moment and period derivation. The curvature identity and finite fusion formulas use that derivation. Balanced limits are taken in the finite period expressions; no uniform spectral estimate through further pole coalescences or as $b \to 0$ is asserted.

The local pinch calculation. Two opposite pole families are

$$\begin{aligned} &\alpha = a_s + mb + n/b, \\ &\alpha = Q - \alpha_2 - \alpha_4 - m'b - n'/b = a_s + b - \varepsilon - m'b - n'/b, \end{aligned} \tag{13.4}$$

with nonnegative integer indices. At $\varepsilon = 0$ they collide only if

$$(m + m' - 1)b^2 + (n + n') = 0.$$

Since $0 < b^2 < 1$, this forces $m + m' = 1$ and $n = n' = 0$. Thus the two intermediate charges are

$$\boxed{\alpha_s^{(0)} = a_s, \qquad \alpha_s^{(1)} = a_s + b.} \tag{13.5}$$

Their reflected partners $Q - \alpha_s^{(j)}$ are included in the usual unfolded contour. We first work at generic weights and then take the regularized limits at coincident poles.

The coefficients require only elementary special-function identities. Define

$$\mathfrak{g}(x) = \frac{\Gamma(x)}{\Gamma(1-x)}. \tag{13.6}$$

For a three-point function with $\Sigma = \alpha_1 + \alpha_2 + \alpha_3$, the zero- and one-screening residues, with respect to the displayed sum coordinate, are

$$\begin{aligned} \operatorname*{Res}_{\Sigma=Q} C &= 1, \\ \operatorname*{Res}_{\Sigma=Q-b} C &= -\frac{\pi\Lambda_L}{\prod_{j=1}^{3} \mathfrak{g}(2b\alpha_j)}. \end{aligned} \tag{13.7}$$

For completeness these follow directly from DOZZ. Its denominator is $\Upsilon_b(\Sigma-Q)\prod_j \Upsilon_b(\Sigma-2\alpha_j)$; use

$$\Upsilon_b(Q-x) = \Upsilon_b(x), \qquad \Upsilon_b(x+b) = \mathfrak{g}(bx)b^{1-2bx}\Upsilon_b(x).$$

At $\Sigma = Q$ reflection cancels the other factors. At $\Sigma = Q - b = 1/b$, the product of shifted ratios is $b/\prod_j \mathfrak{g}(2b\alpha_j)$. Combining it with the DOZZ prefactor and $\mathfrak{g}(b^2)\mathfrak{g}(-b^2) = -b^{-4}$ proves the second identity. These are the conventions of [93, eqs. (9), (11), (13)].

Set

$$\delta_0 = 1 - \mu_1 - \mu_3, \qquad \delta_1 = 1 - \mu_2 - \mu_3, \qquad \delta_\infty = 1 - \mu_3 - \mu_4, \tag{13.8}$$

and, initially assuming $\delta_0 \neq 0$, define

$$h_0 = \frac{\pi\mathfrak{g}(\delta_0)}{\mathfrak{g}(\mu_2)\mathfrak{g}(\mu_4)}, \qquad h_1 = \frac{\pi\mathfrak{g}(-\delta_0)}{\mathfrak{g}(\mu_1)\mathfrak{g}(\mu_3)}. \tag{13.9}$$

At $\alpha = a_s$, the left three-point function has its zero-screening pole and the right one its one-screening pole, with coefficient $-\Lambda_L h_0$. At $\alpha = a_s + b$ their roles reverse, with coefficient $-\Lambda_L h_1$. In the local coordinate $x = \alpha - \alpha_s^{(j)}$ the singular product is

$$\frac{\Lambda_L h_j}{x(x+\varepsilon)}.$$

The contour separating $x = -\varepsilon$ and $x = 0$ gives the coefficient $-\Lambda_L h_j/\varepsilon$. In the unfolded $1/(4\pi i)$ spectral normalization the reflection partner supplies the other half of the local $1/(2\pi i)$ residue. Under the continuation and tail-control input above, the remaining contour integral is holomorphic in ε and contributes no residue. The contour calculation therefore gives

$$\boxed{\operatorname*{Res}_{\varepsilon=0} C_4(z) = -\Lambda_L \left(h_0 |\mathcal{F}_0(z)|^2 + h_1 |\mathcal{F}_1(z)|^2 \right).} \tag{13.10}$$

The identity itself also follows without continuing the infinite spectral integral: combine the first-moment residue (12.8), in the transverse coordinate $\varepsilon = b(s+1)$, with the independently proved period pairing (13.19) and the block factorization (13.12). This fixes the normalization and sign independently of the conditional contour derivation. Conventions for the overall normalization of correlations differ in the literature; the bootstrap normalization is calibrated by (13.7).

13.2 Why the screening integrals are the Virasoro blocks

Let $F(a,b;c;z) = {}_2F_1(a,b;c;z)$ and take branches positive for $0 < z < 1$, continued first to $\mathbb{C} \setminus ((-\infty, 0] \cup [1, \infty))$. The normalized s-channel periods are

$$\begin{aligned} f_0(z) &= F(1-\mu_4, \mu_3; \mu_1+\mu_3; z), \\ f_1(z) &= z^{\delta_0} F(\mu_2, 1-\mu_1; 1+\delta_0; z). \end{aligned} \tag{13.11}$$

The full normalized chiral blocks are

$$\mathcal{F}_j(z) = E_b(z) f_j(z), \qquad E_b(z) = z^{-\mu_1\mu_3/(2b^2)} (1-z)^{-\mu_2\mu_3/(2b^2)}. \tag{13.12}$$

Here the conformal weight is $\Delta(\alpha) = \alpha(Q-\alpha)$ and the central charge is $c_L = 1 + 6Q^2$.

To verify the block identification, the screening field has weight $\Delta(b) = b(Q-b) = 1$. Integration over a closed twisted cycle therefore commutes with the Virasoro action: its Ward commutator is a total derivative. A screening outside a circle separating $0, z$ from $1, \infty$ gives intermediate charge a_s, while a screening inside gives $a_s + b$. The corresponding Euler arc integrals are

$$\begin{aligned} S_0 &= \int_1^\infty w^{-\mu_1} (w-1)^{-\mu_2} (w-z)^{-\mu_3} \, \mathrm{d}w = B(1-\mu_2, 1-\mu_4) f_0(z), \\ S_1 &= \int_0^z w^{-\mu_1} (1-w)^{-\mu_2} (z-w)^{-\mu_3} \, \mathrm{d}w = B(1-\mu_1, 1-\mu_3) f_1(z). \end{aligned} \tag{13.13}$$

The positive arc branches are phase-normalized versions of the twisted periods; closed Pochhammer contours justify the Ward-identity argument, with the same normalized functions. Dividing out the beta constants sets the leading coefficient to one. The leading exponents agree because

$$\Delta(a_s) - \Delta(\alpha_1) - \Delta(\alpha_3) = -2\alpha_1\alpha_3, \qquad \Delta(a_s + b) - \Delta(a_s) = \delta_0.$$

The screened intertwiner construction and uniqueness of the normalized Ward solution identify the functions at every level; see [93, Section 18.1, eq. (224)]. Here $0 < a_s, a_s + b < Q$, so these internal weights are positive and avoid the degenerate Verma weights at $c_L > 25$. No external degenerate field is assumed.

The two blocks solve a second-order hypergeometric equation after removal of the common factor E_b. This reduction is a consequence of the specified one-screening charge locus, not a general second-order equation for arbitrary four-point Liouville blocks.

13.3 The finite fusion matrix

For $\delta_1 \neq 0$, the normalized t-channel periods, pairing $z, 1$, are

$$\begin{aligned} g_0(z) &= F(1-\mu_4, \mu_3; \mu_2+\mu_3; 1-z), \\ g_1(z) &= (1-z)^{\delta_1} F(1-\mu_2, \mu_1; 1+\delta_1; 1-z). \end{aligned} \tag{13.14}$$

Their internal charges are $\alpha_2 + \alpha_3$ and $\alpha_2 + \alpha_3 + b$, and their full chiral blocks are $E_b g_j$. Thus the same matrix relates periods and normalized screened Virasoro blocks:

$$\begin{pmatrix} f_0 \\ f_1 \end{pmatrix} = \mathsf{M} \begin{pmatrix} g_0 \\ g_1 \end{pmatrix}, \tag{13.15}$$

where

$$\boxed{\mathsf{M} = \begin{pmatrix} \dfrac{\Gamma(1-\delta_0)\Gamma(\delta_1)}{\Gamma(\mu_1)\Gamma(1-\mu_2)} & \dfrac{\Gamma(1-\delta_0)\Gamma(-\delta_1)}{\Gamma(\mu_3)\Gamma(1-\mu_4)} \\ \dfrac{\Gamma(1+\delta_0)\Gamma(\delta_1)}{\Gamma(1-\mu_3)\Gamma(\mu_4)} & \dfrac{\Gamma(1+\delta_0)\Gamma(-\delta_1)}{\Gamma(1-\mu_1)\Gamma(\mu_2)} \end{pmatrix}.} \tag{13.16}$$

This is the Gauss connection formula, obtained by substituting $a = 1-\mu_4$, $b = \mu_3$, $c = \mu_1+\mu_3$ in [54, eqs. 15.10.21–22]. The Wronskians give

$$\det \mathsf{M} = -\frac{\delta_0}{\delta_1}. \tag{13.17}$$

The matrix is independent of the Liouville parameter b when the deficits μ_i are fixed. All b-dependence in these blocks is then in the common factor E_b and in the overall correlation normalization.

There is a simpler expression in actual arc coordinates. For real $0 < z < 1$, put

$$T_0 = \int_{-\infty}^{0} (-w)^{-\mu_1}(1-w)^{-\mu_2}(z-w)^{-\mu_3}\,\mathrm{d}w = B(1-\mu_1, 1-\mu_4)g_0,$$
$$T_1 = \int_{z}^{1} w^{-\mu_1}(1-w)^{-\mu_2}(w-z)^{-\mu_3}\,\mathrm{d}w = B(1-\mu_2, 1-\mu_3)g_1.$$

Gamma reflection reduces (13.16) to

$$\boxed{\begin{pmatrix} S_0 \\ S_1 \end{pmatrix} = \frac{1}{\sin(\pi\delta_1)} \begin{pmatrix} \sin(\pi\mu_1) & -\sin(\pi\mu_3) \\ \sin(\pi\mu_4) & -\sin(\pi\mu_2) \end{pmatrix} \begin{pmatrix} T_0 \\ T_1 \end{pmatrix}.} \tag{13.18}$$

These arc pairs cease to be bases at the corresponding balanced weights. The divergent coefficients in this formula cannot be specialized there without first changing basis.

13.4 The Hermitian pairing and its direct verification

We now prove the area pairing directly from periods:

$$\boxed{\mathcal{A}_\mu(z) = h_0|f_0(z)|^2 + h_1|f_1(z)|^2.} \tag{13.19}$$

Together with the first-moment residue, this supplies an independent derivation of (13.10). The complex beta integral is

$$\int_{\mathbb{C}} |w|^{-2p}|1-w|^{-2q}\,\mathrm{d}^2w = \pi\frac{\Gamma(1-p)\Gamma(1-q)\Gamma(p+q-1)}{\Gamma(p)\Gamma(q)\Gamma(2-p-q)}, \tag{13.20}$$

for $0 < p, q < 1$ and $p+q > 1$. Introduce positive Laplace parameters for the two powers, perform the Gaussian integral in w, and then use their sum and ratio as variables. The remaining integrals are an ordinary Gamma integral and $B(1-q, 1-p)$, proving the formula.

The area is a constant Hermitian form in the period basis. For $\delta_0 \neq 0$, local monodromy has distinct eigenvalues $1, e^{2\pi i\delta_0}$, so the form is diagonal in (f_0, f_1). If $\delta_0 > 0$, sending $z \to 0$ in the area integral and using (13.20) gives h_0. If $\delta_0 < 0$, rescaling $w = zu$ gives the leading coefficient h_1. Continuation in the weights gives (13.9) throughout the nonresonant range.

The signs of h_0, h_1 are opposite. The pairing has signature $(1,1)$; the positive value of $\mathcal{A}_\mu$ on the period line is compatible with that indefinite signature. It is not a positive probability decomposition into two channels. In arc coordinates it becomes

$$\mathcal{A}_\mu = \frac{\sin(\pi\mu_2)\sin(\pi\mu_4)|S_0|^2 - \sin(\pi\mu_1)\sin(\pi\mu_3)|S_1|^2}{\sin(\pi\delta_0)}. \tag{13.21}$$

Define $H_s = \mathrm{diag}(h_0, h_1)$ and

$$H_t = \mathrm{diag}\left(\frac{\pi\mathfrak{g}(\delta_1)}{\mathfrak{g}(\mu_1)\mathfrak{g}(\mu_4)}, \frac{\pi\mathfrak{g}(-\delta_1)}{\mathfrak{g}(\mu_2)\mathfrak{g}(\mu_3)}\right).$$

The connection preserves the area exactly:

$$\mathsf{M}^* H_s \mathsf{M} = H_t. \tag{13.22}$$

This follows either from the two area decompositions or directly from Gamma recurrence and reflection. On the specified real branch $\mathsf{M}^* = \mathsf{M}^T$.

13.5 A closed formula for the four-point density and total volume

The area expression gives an explicit density in the conformal cross-ratio. For a holomorphic column $\boldsymbol{f} = (f_0, f_1)^T$ and constant Hermitian matrix H, differentiation gives

$$\partial_z \partial_{\bar{z}} \log(\boldsymbol{f}^* H \boldsymbol{f}) = \frac{\det H \, |f_0 f_1' - f_0' f_1|^2}{(\boldsymbol{f}^* H \boldsymbol{f})^2}. \tag{13.23}$$

The hypergeometric equation yields

$$W(f_0, f_1) = \delta_0 z^{-(\mu_1+\mu_3)}(1-z)^{-(\mu_2+\mu_3)}, \qquad -h_0 h_1 \delta_0^2 = \frac{\pi^2}{\prod_i \mathfrak{g}(\mu_i)}. \tag{13.24}$$

Using $\mathrm{d}z \wedge \mathrm{d}\bar{z} = -2i \, \mathrm{d}x \wedge \mathrm{d}y$ and the curvature--1 convention therefore gives

$$\boxed{\omega_{\mathrm{DM},\mu} = \frac{4\pi^2}{\prod_{i=1}^4 \mathfrak{g}(\mu_i)} \frac{|z|^{-2(\mu_1+\mu_3)}|1-z|^{-2(\mu_2+\mu_3)}}{\mathcal{A}_\mu(z)^2} \, \mathrm{d}x \wedge \mathrm{d}y, \qquad z = x + iy.} \tag{13.25}$$

The area in the denominator is explicitly the hypergeometric combination (13.19). Although that diagonal basis is singular at balanced weights, the density extends through them by its logarithmic limit.

Its integral can also be evaluated. At $z = 0$, if $\delta_0 > 0$ then $\mathcal{A}_\mu \to h_0$, while if $\delta_0 < 0$ then $\mathcal{A}_\mu \sim h_1 |z|^{2\delta_0}$. In both cases the metric density has the form const. $|z|^{-2(1-|\delta_0|)}$, giving cone angle $2\pi|\delta_0|$. At $\delta_0 = 0$ the area has logarithmic growth and the end is a cusp. The same argument at $1, \infty$ gives the other two angles $2\pi|\delta_1|$ and $2\pi|\delta_\infty|$. This agrees with the three-ended cone-sphere completion described in [64, Section 5].

Gauss–Bonnet, retaining the labels, gives the integral

$$\boxed{V_{\mathrm{DM}}(\mu) = 2\pi(1 - |\delta_0| - |\delta_1| - |\delta_\infty|) = 4\pi \min_{1 \le i \le 4} \{\mu_i, 1-\mu_i\}.} \tag{13.26}$$

For the last equality, order the weights as $x_1 \le x_2 \le x_3 \le x_4$ and put $r = 1 - x_4$. The sum constraint forces $x_2, x_3 \ge r$. The three absolute bipartition differences have sum

$$\sum_{j=1}^{3} |r - x_j| = 1 - r - x_1 + |r - x_1| = 1 - 2\min(r, x_1).$$

This proves the stated minimum formula. Equal weights give $V_{\mathrm{DM}} = 2\pi$; for example $\mu = (1/4, 1/2, 1/2, 3/4)$ gives $V_{\mathrm{DM}} = \pi$.

The integral here is of the smooth curvature on the open moduli space, with its cone/cusp completion. If one instead extends $\partial\bar{\partial} \log C_b^{\mathrm{uv}}$ as a distribution to the compactified sphere, the explicit pairwise powers have delta masses at collisions. Their logarithms are pluriharmonic only off those collisions.

13.6 Four equal weights: the logarithmic fusion calculation

When all $\mu_i = 1/2$, both δ_0 and δ_1 vanish. The functions f_0, f_1 coincide. On the representation side, $a_s + b = Q - a_s$, so the two intermediate charges have the same Virasoro weight. The second solution is a regularized logarithmic, or parameter-derivative, block; it is not a second independent ordinary normalized highest-weight block at the same weight.

Use the independent periods

$$f(z) = F(1/2, 1/2; 1; z) = \frac{2}{\pi} K(z), \qquad g(z) = F(1/2, 1/2; 1; 1-z) = \frac{2}{\pi} K(1-z). \tag{13.27}$$

Here K uses the parameter convention of (7.10). Put $L = \log 16$ and define

$$\ell_0 = Lf - \pi g, \qquad \ell_1 = Lg - \pi f. \tag{13.28}$$

The logarithmic expansion [54, Sections 15.10 and 19.12] gives

$$\ell_0(z) = f(z) \log z + 4 \sum_{m \geq 1} \left(\frac{(1/2)_m}{m!} \right)^2 (H_{2m} - H_m) z^m,$$

where H_m is a harmonic number. Thus the holomorphic part has zero constant term. The analogous statement for ℓ_1 uses $\log(1-z)$. Solving (13.28) gives the finite connection matrix

$$\boxed{\begin{pmatrix} f \\ \ell_0 \end{pmatrix} = \frac{1}{\pi} \begin{pmatrix} L & -1 \\ L^2 - \pi^2 & -L \end{pmatrix} \begin{pmatrix} g \\ \ell_1 \end{pmatrix}.} \tag{13.29}$$

It has determinant -1 and its square is the identity.

The limiting area pairing is

$$\begin{aligned} \mathcal{A}(z) &= 2\pi^2 \operatorname{Re}(f \overline{g}) = 8 \operatorname{Re}(\overline{K(z)} K(1-z)), \\ &= 2\pi L |f|^2 - 2\pi \operatorname{Re}(f \overline{\ell_0}). \end{aligned} \tag{13.30}$$

The Hermitian matrices in the two displayed bases are

$$H_{fg} = \pi^2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad H_{\log} = \begin{pmatrix} 2\pi L & -\pi \\ -\pi & 0 \end{pmatrix}. \tag{13.31}$$

The connection (13.29) preserves $H_{\log}$ exactly.

One can see the cancellation rather than just substitute into a singular formula. Take $\mu_1 = \mu_3 = (1-\delta)/2$ and $\mu_2 = \mu_4 = (1+\delta)/2$. Then

$$h_0 = \frac{\pi}{\delta} + \pi L + O(\delta), \qquad h_1 = -\frac{\pi}{\delta} + \pi L + O(\delta), \qquad f_1 - f_0 = \delta \ell_0 + O(\delta^2).$$

The pole terms cancel in (13.19), leaving (13.30). This explains precisely how the regularized residue sector survives the coincidence of the ordinary blocks.

For counterclockwise loops about 0 and 1, respectively, the period column $(f, g)^T$ has monodromies

$$\begin{pmatrix} 1 & 0 \\ -2i & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -2i \\ 0 & 1 \end{pmatrix}.$$

In the oriented torus-period basis $\Omega_1 = 4K(z)$, $\Omega_2 = 4iK(1-z)$, these become

$$N_0 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \qquad N_1 = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}. \tag{13.32}$$

They give the expected projective $\Gamma(2)$ monodromy. The channel exchange $z \leftrightarrow 1-z$ swaps f, g and acts on $\tau = \Omega_2/\Omega_1$ by $\tau \mapsto -1/\tau$. For oriented integral cycles this is a symplectic swap up to a common phase of the developing differential.

Finally,

$$\boxed{\omega_{\mathrm{DM}} = \frac{4\pi^2}{|z(1-z)|^2 \mathcal{A}(z)^2}\, \mathrm{d}x \wedge \mathrm{d}y, \qquad \mathcal{A}(z) = 8\,\mathrm{Re}(\overline{K(z)}K(1-z)).} \tag{13.33}$$

Indeed $W(f,g) = -1/[\pi z(1-z)]$, or equivalently $\tau' = -i\pi/[4z(1-z)K(z)^2]$. Near $z = 0$,

$$\mathcal{A}(z) = 2\pi \log \frac{16}{|z|} + O\left(|z| \log \frac{1}{|z|}\right), \qquad \omega_{\mathrm{DM}} \sim \frac{\mathrm{d}x \wedge \mathrm{d}y}{|z|^2 \log^2(16/|z|)}.$$

This gives the integrable cusp density directly in the cross-ratio coordinate.

13.7 What the calculation establishes

The first-moment residue, the screened Virasoro blocks, and the independently computed Hermitian period pairing agree with all constants fixed. Under (13.3) and the stated contour-continuation conditions, the infinite spectral representation gives the same answer through its two pinches. The finite fusion transformation on the two-dimensional one-screening residue sector is the period connection matrix (13.16), with (13.29) at the fully balanced point. Its independent derivation requires no infinite-contour interchange. Matrix entries are not being identified with unspecified pointwise values of a bare b-6j function.

This completes the finite four-point residue calculation. Section 14 matches its quantum-group normalization explicitly, strengthens the Gram-matrix obstruction, and computes a classical cubic saddle for the separate compact-tetrahedron degeneration.

14 The normalized quantum-group comparison and the compact limit

The further comparison has two distinct outcomes. The finite screening operator can be matched explicitly to a specified b-6j normalization: its double-Gamma factors reduce to ordinary Gamma factors and its analytic quantum-group matrix is a sine matrix. Here the operator is fixed by the independent period calculation in Section 13; its interpretation through a continued infinite spectral representation retains the analytic input specified there. Its auxiliary angle Gram matrix approaches rank two, whereas a nondegenerate shrinking compact tetrahedron approaches rank three. A direct calculation for the regular compact family gives a cubic classical saddle and fixes the proposed joint scale. The exact integration coordinate requires care: a previously stated pole accumulation at the saddle came from an incorrect shift. The corrected dictionary below and Section 15 give a uniform quantum Airy formula after removal of an explicit external scalar pole factor.

14.1 Fixing the Teschner–Vartanov normalization

Write R^{an} for the analytic b-6j symbol and R for the symmetric symbol in [94]. Their equations (4.5)–(4.6) give the standard Virasoro fusion kernel as $F_{st} = (N_s/N_t)m(t)R^{\mathrm{an}}_{st}$, where

$$N(z,y,x) = \frac{\Gamma_b(2Q-2z)\Gamma_b(2y)\Gamma_b(2x)}{\Gamma_b(2Q-x-y-z)\Gamma_b(Q-x-y+z)\Gamma_b(x+z-y)\Gamma_b(y+z-x)}. \tag{14.1}$$

Here N_s and N_t are products of the two trivalent normalizations in each channel. On the principal contour, $m(a) = |S_b(2a)|^2$ is the Plancherel density. The trivalent factor distinguishing the symmetric symbol is

$$D(z,y,x) = \big[S_b(2Q-x-y-z)S_b(Q-x-y+z)S_b(x+z-y)S_b(y+z-x)\big]^{-1/2}.$$

If D_s, D_t are its channel products, then $R = (D_s/D_t)R^{\mathrm{an}}$. Equivalently, equations (4.10)–(4.11) of that paper read

$$G_{st} = \sqrt{\frac{C_s}{C_t}}\,F_{st} = \sqrt{m(s)m(t)}\,R_{st}, \tag{14.2}$$

where C_s, C_t are the corresponding products of DOZZ three-point functions. These identities specify the scalar factors that were left open in Section 13.

For our ordering of fields, use TV's external labels

$$(p,q,r,d_{\mathrm{out}}) = (\alpha_1,\alpha_3,\alpha_2,Q-\alpha_4). \tag{14.3}$$

Reflecting the infinity charge preserves its Virasoro weight. At the one-screening resonance,

$$d_{\mathrm{out}} = p+q+r+b, \qquad s_j = p+q+jb, \qquad t_k = q+r+kb, \qquad j,k\in\{0,1\}.$$

Consequently each channel contains one zero-screening and one one-screening trivalent vertex. Explicitly,

$$N_{s,j} = N(s_j,q,p)N(d_{\mathrm{out}},r,s_j), \qquad N_{t,k} = N(t_k,r,q)N(d_{\mathrm{out}},t_k,p).$$

The reflection convention matters for the analytic symbol, and is retained throughout the following formulas. Initially the weights and auxiliary factors are generic, with $\delta_0\delta_1 \neq 0$; balanced limits are taken afterward.

14.2 The double-Gamma factors and the finite sine matrix

The shift identity

$$\frac{\Gamma_b(x+b)}{\Gamma_b(x)} = \frac{\sqrt{2\pi}\,b^{bx-1/2}}{\Gamma(bx)}$$

immediately gives

$$\begin{aligned} N(x+y,y,x) &= \Gamma_b(Q)^{-1},\\ N(x+y+b,y,x) &= \frac{\Gamma(1+b^2)\Gamma(2bx)\Gamma(2by)\Gamma(2-2b(x+y))}{(2\pi)^2 b^{1+b^2}\Gamma_b(Q)}. \end{aligned} \tag{14.4}$$

Thus

$$N_{s,j} = \mathcal{K}_b n_{s,j}, \qquad N_{t,k} = \mathcal{K}_b n_{t,k}, \qquad \mathcal{K}_b = \frac{\Gamma(1+b^2)}{(2\pi)^2 b^{1+b^2}\Gamma_b(Q)^2},$$

with

$$\begin{aligned} n_{s,0} &= \Gamma(1-\delta_0)\Gamma(\mu_2)\Gamma(\mu_4), & n_{s,1} &= \Gamma(1+\delta_0)\Gamma(\mu_1)\Gamma(\mu_3),\\ n_{t,0} &= \Gamma(1-\delta_1)\Gamma(\mu_1)\Gamma(\mu_4), & n_{t,1} &= \Gamma(1+\delta_1)\Gamma(\mu_2)\Gamma(\mu_3). \end{aligned} \tag{14.5}$$

All b-dependence in these normalization products is in the same scalar $\mathcal{K}_b$.

Rescale the screened channel vectors by $e_{s,j} = \mathcal{F}_{s,j}/N_{s,j}$ and $e_{t,k} = \mathcal{F}_{t,k}/N_{t,k}$. Their matrix is $\mathsf{B}_{jk} = (n_{t,k}/n_{s,j})\mathsf{M}_{jk}$, where M is (13.16). Gamma reflection gives the explicit answer

$$\boxed{\mathsf{B} = \frac{1}{\sin(\pi\delta_1)} \begin{pmatrix} \sin(\pi\mu_2) & -\sin(\pi\mu_4) \\ \sin(\pi\mu_3) & -\sin(\pi\mu_1) \end{pmatrix}.} \tag{14.6}$$

For example its upper-left entry is

$$\frac{n_{t,0}}{n_{s,0}}\mathsf{M}_{00} = \frac{\Gamma(1-\delta_1)\Gamma(\delta_1)}{\Gamma(\mu_2)\Gamma(1-\mu_2)} = \frac{\sin(\pi\mu_2)}{\sin(\pi\delta_1)}.$$

This is the induced finite action of $m(t)R^{\mathrm{an}}$ in the specified convention. It is obtained from the exact normalization identity and the screened invariant space; it is not a pointwise substitution into a singular Barnes integral. Its determinant is $-\sin(\pi\delta_0)/\sin(\pi\delta_1)$.

For reference, put $\sigma_i = \sin(\pi\mu_i)$ and $d_\nu = \sin(\pi\delta_\nu)$. After removing the common $E_b/\mathcal{K}_b$ from these vectors, their Hermitian matrices are

$$H_s^{\mathrm{an}} = \frac{\pi^4}{d_0}\,\mathrm{diag}\left(\frac{1}{\sigma_2\sigma_4}, -\frac{1}{\sigma_1\sigma_3}\right), \qquad H_t^{\mathrm{an}} = \frac{\pi^4}{d_1}\,\mathrm{diag}\left(\frac{1}{\sigma_1\sigma_4}, -\frac{1}{\sigma_2\sigma_3}\right). \tag{14.7}$$

They follow from $H_s^{\mathrm{an}} = \mathrm{diag}(h_j n_{s,j}^2)$, and satisfy $\mathsf{B}^* H_s^{\mathrm{an}} \mathsf{B} = H_t^{\mathrm{an}}$.

14.3 Symmetric normalization and the residue prescription

The square roots in (14.2) also have a concrete interpretation at a pinch. Write $\alpha_s = s_j + \varepsilon\xi$, with fixed $\xi \neq 0, -1$. The product of three-point factors has the form

$$C_s \sim \frac{\Lambda_* h_{s,j}}{\varepsilon^2 \xi(\xi+1)}.$$

The common Λ_* includes the infinity reflection factor and cancels in channel ratios. With square-root branches fixed by continuation, the DOZZ-normalized vector $\mathcal{G}_s = \sqrt{C_s}\mathcal{F}_s$ obeys

$$\lim_{\varepsilon\to 0} \varepsilon \sqrt{\frac{\xi(\xi+1)}{\Lambda_*}}\, \mathcal{G}_s(s_j + \varepsilon\xi; \varepsilon) = \sqrt{h_{s,j}}\mathcal{F}_{s,j}. \tag{14.8}$$

Hence its finite analytic matrix is

$$\mathsf{O} = \mathrm{diag}(\sqrt{h_s})\mathsf{M}\,\mathrm{diag}(\sqrt{h_t})^{-1}, \qquad \mathsf{O}^T\mathsf{O} = I.$$

This is complex orthogonality: negative Hermitian coefficients give factors of i in analytic square roots.

Using absolute square roots instead gives a real matrix with an indefinite pairing:

$$\boxed{\mathsf{U} = \frac{\mathrm{sgn}(d_1)}{\sqrt{|d_0 d_1|}} \begin{pmatrix} \sqrt{\sigma_1\sigma_2} & -\sqrt{\sigma_3\sigma_4} \\ \sqrt{\sigma_3\sigma_4} & -\sqrt{\sigma_1\sigma_2} \end{pmatrix}.} \tag{14.9}$$

Indeed,

$$\sigma_1\sigma_2 - \sigma_3\sigma_4 = d_0 d_1, \qquad \mathsf{U}^T J_s \mathsf{U} = J_t.$$

$$J_s = \mathrm{sgn}(d_0)\,\mathrm{diag}(1,-1), \quad J_t = \mathrm{sgn}(d_1)\,\mathrm{diag}(1,-1).$$

Equations (14.6) and (14.9) are fully normalized finite operator comparisons. The logarithmic basis of Section 13 supplies the balanced limits; one cannot set $d_0 = d_1 = 0$ in the displayed matrices.

There is a separate reason to retain the operator prescription. The analytic continuation of the Plancherel factor is

$$\begin{aligned} m(a) &= S_b(2a)S_b(2Q-2a) \\ &= -4\sin\bigl(\pi b(2a-Q)\bigr)\sin\bigl(\pi(2a-Q)/b\bigr). \end{aligned} \tag{14.10}$$

The two S_b shift equations prove the second equality. At the screened labels this becomes

$$m(s_j) = 4\sin(\pi[\delta_0-(2j-1)b^2])\sin(\pi\delta_0/b^2), \tag{14.11}$$

with the analogous formula for t_k, δ_1. These factors oscillate for small real b and vanish at balanced weights. The trivalent symmetric normalizations also encounter $S_b(Q+jb)=0$. Thus dividing a finite matrix by pointwise Plancherel values does not define a bare-symbol limit.

14.4 The actual screened Gram matrix has rank two at the limit

The tetrahedral dictionary of [94, Section 3] is

$$A_e = \exp(\pi i - 2\pi i b\alpha_e) = e^{i\eta_e}.$$

The vertex triples are $(1,2,s),(3,4,s),(1,4,t),(2,3,t)$; compare the original convention in [81]. Here A_e is an exponentiated angle, distinct from surface area. Put

$$\vartheta_i = \pi\mu_i, \qquad h = 2\pi b^2, \qquad \sum_i \vartheta_i = 2\pi.$$

With the reflected labels (14.3), a face ordering gives the formal angle Gram matrix

$$\mathbf{G}_{jk}(h) = \begin{pmatrix} 1 & \cos(\vartheta_2+\vartheta_3+kh) & \cos\vartheta_2 & \cos(\vartheta_4-h) \\ \cos(\vartheta_2+\vartheta_3+kh) & 1 & \cos\vartheta_3 & \cos\vartheta_1 \\ \cos\vartheta_2 & \cos\vartheta_3 & 1 & \cos(\vartheta_1+\vartheta_3+jh) \\ \cos(\vartheta_4-h) & \cos\vartheta_1 & \cos(\vartheta_1+\vartheta_3+jh) & 1 \end{pmatrix}. \tag{14.12}$$

Calling this a formal Gram matrix does not assert that the angle data are realized by a convex tetrahedron.

At $h=0$ it is the Euclidean planar Gram matrix

$$(\mathbf{G}_0)_{ab} = \cos(\phi_a-\phi_b), \qquad (\phi_1,\phi_2,\phi_3,\phi_4) = (0,\vartheta_2+\vartheta_3,\vartheta_2,\vartheta_1+\vartheta_2+\vartheta_3). \tag{14.13}$$

Since $0<\vartheta_2<\pi$, these unit vectors span a plane. Therefore

$$\boxed{\operatorname{rank}\mathbf{G}_0 = 2.} \tag{14.14}$$

For equal half-weights it is

$$\mathbf{G}_0 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

By contrast, the outward normals of a nondegenerate Euclidean tetrahedron span three dimensions. This difference holds for every four-weight choice on the screening locus, and remains under varying such weights because rank at most two is closed.

There is already a finite-b obstruction. If $\mathcal{C}_{ab}$ denotes the cofactor matrix, then

$$\mathcal{C}_{1+j,1+j} = 0, \qquad \mathcal{C}_{4-k,4-k} = 0. \tag{14.15}$$

Each identity is the determinant of three planar unit vectors. At a finite vertex of a compact hyperbolic tetrahedron, the three incident face normals lie in the positive-definite orthogonal complement of a timelike vertex vector and are independent. Their Gram determinant is positive. Thus (14.15) excludes compact tetrahedra already before taking the limit.

For completeness, the determinant and orientation checks are also explicit. Each $\mathbf{G}_{jk}$ is a planar rank-two Gram matrix perturbed in a single symmetric off-diagonal entry. The entry, its increment ϵ, and a positive factor q are

(j,k)	entry	ϵ	q
$(0,0)$	14	$\cos(\vartheta_4 - h) - \cos\vartheta_4$	$\sin^2\vartheta_3$
$(0,1)$	13	$\cos\vartheta_2 - \cos(\vartheta_2 + h)$	$\sin^2\vartheta_1$
$(1,0)$	24	$\cos\vartheta_1 - \cos(\vartheta_1 + h)$	$\sin^2\vartheta_2$
$(1,1)$	23	$\cos\vartheta_3 - \cos(\vartheta_3 + h)$	$\sin^2(\vartheta_4 - h)$

In every row, $\det\mathbf{G} = -q\epsilon^2$ and the cofactor at that off-diagonal entry is $-q\epsilon$. For fixed interior weights and small positive h, $\epsilon > 0$. Hence the matrices have Lorentz signature but fail the positive off-diagonal cofactor condition for an ordinary convex hyperbolic tetrahedron; see [47, Theorem 2.1]. The zero principal cofactors therefore do not justify calling these configurations tetrahedra with two ideal vertices.

14.5 The screened classical saddle also degenerates

The dilogarithmic potential gives an independent check. At $b = 0$, write $X = A_1, Y = A_2, Z = A_3$. The reflected screening labels satisfy

$$A_4 = XYZ, \qquad A_s = -XY, \qquad A_t = -YZ.$$

Set $P = X^2Y^2 = e^{2\pi i\delta_0}$ and $R = Y^2Z^2 = e^{2\pi i\delta_1}$. Substitution into [94, eq. (3.1)] cancels two pairs of dilogarithms and leaves

$$U(u) = \mathrm{Li}_2(u) + \mathrm{Li}_2(PRu) - \mathrm{Li}_2(Pu) - \mathrm{Li}_2(Ru). \tag{14.16}$$

The exponentiated saddle equation $uU_u = -2\pi i$ becomes

$$(1 - Pu)(1 - Ru) = (1 - u)(1 - PRu), \qquad (1 - P)(1 - R)u = 0.$$

For nonbalanced weights there is no nonzero regular saddle. Before cancellation, the equation contains the factor $(1 - (XYZ)^2u)^2$; its double root is at the very branch factors that cancel in (14.16). This is not a regular double critical point of the reduced phase. If $P = 1$ or $R = 1$, the reduced potential itself vanishes on a consistent local branch. Thus the finite screening operator survives in a singular pole/branch collision regime, rather than in the ordinary compact Euclidean saddle regime.

14.6 A genuine compact family: the regular tetrahedron

For comparison, prescribe the actual six interior dihedral angles η_e of a small compact tetrahedron and set

$$\alpha_e = \frac{\pi - \eta_e}{2\pi b}. \tag{14.17}$$

These quantum labels encode edge dihedrals. The μ_i of the flat boundary metric encode intrinsic vertex deficits, so the two kinds of data must be related geometrically before they can be identified.

For a regular tetrahedron of edge length ℓ, let β be a face angle and η an interior dihedral. Hyperbolic trigonometry gives

$$\cos\beta = \frac{\cosh\ell}{1+\cosh\ell}, \qquad \cos\eta = \frac{\cosh\ell}{1+2\cosh\ell}, \qquad A = 4\pi - 12\beta. \tag{14.18}$$

At the Euclidean limit $\eta_{\mathrm{E}} = \arccos(1/3)$. Its angle Gram has

$$\det G = (1+\cos\eta)^3(1-3\cos\eta),$$

and rank three at η_{E}. All six limiting heavy edge weights are

$$2b\alpha_e = 1 - \eta_{\mathrm{E}}/\pi \simeq 0.608173.$$

In the reflected screening convention, $2bd_{\mathrm{out}} \to 2 - \mu_4 > 1$, whereas this geometric prescription gives the same number below one for all six labels. Moreover the screened internal charges are sums of pairs of external charges. The rank obstruction is independent of reflected representatives; these explicit prescriptions already differ at leading order $1/b$, beyond an order-b normalization shift.

Expansion of (14.18) gives

$$\begin{aligned} A &= \sqrt{3}\,\ell^2 - \frac{\sqrt{3}}{8}\ell^4 + O(\ell^6), \\ t := \eta_{\mathrm{E}} - \eta &= \frac{\ell^2}{12\sqrt{2}} - \frac{23\ell^4}{1152\sqrt{2}} + O(\ell^6) = \frac{A}{12\sqrt{6}} + O(A^2). \end{aligned} \tag{14.19}$$

Using $\mathrm{d}W = -3\ell\,\mathrm{d}\eta$ and $W(0) = 0$,

$$\boxed{W(A) = \frac{A^{3/2}}{6\sqrt{2}\,3^{3/4}}\left(1 - \frac{A}{10\sqrt{3}} + O(A^2)\right).} \tag{14.20}$$

In edge length this is $W = \ell^3/(6\sqrt{2}) - 23\ell^5/(480\sqrt{2}) + O(\ell^7)$.

14.7 The regular classical fold and its scale

Put all six exponentiated angles equal to $a = e^{i\eta}$. The TV potential becomes

$$U(u,\eta) = \mathrm{Li}_2(u) + 3\,\mathrm{Li}_2(a^4u) - 4\,\mathrm{Li}_2(-a^3u).$$

Exponentiating the saddle equation gives $(1+a^3u)^4 = (1-u)(1-a^4u)^3$. After removing $u(a+1)^2$, the quadratic is

$$P_\eta(u) = a^8(a^2-2a+3)u^2 - 3a^4(a-1)^2u + (3a^2-2a+1). \tag{14.21}$$

Its discriminant is

$$-a^8(a+1)^2(3a^2-2a+3), \qquad 3a^2-2a+3 = 2a(3\cos\eta - 1).$$

Thus the saddles merge simply at the Euclidean angle. The double root and its real logarithmic lift are

$$a_{\mathrm{E}} = \frac{1+2\sqrt{2}\,i}{3}, \qquad u_{\mathrm{E}} = a_{\mathrm{E}}^{-5} = \frac{241+22\sqrt{2}\,i}{243}, \qquad v_{\mathrm{E}} = 2\pi - 5\eta_{\mathrm{E}}.$$

None of the dilogarithm arguments is 1 at this point.

For $u = e^{iv}$ the action, including a term independent of v, is $S(v,\eta) = U(e^{iv},\eta) - 2\pi v + \text{const}$. Choose the local branches satisfying the saddle equation. Direct differentiation at the fold yields

$$S_v = S_{vv} = 0, \qquad S_{vvv} = 54i, \qquad S_{v\eta} = 18\sqrt{2}\,i.$$

For example,

$$S_{v\eta} = -12\left(\frac{a^3u}{1+a^3u} + \frac{a^4u}{1-a^4u}\right).$$

Writing $x = v - v_{\mathrm{E}}$ gives the concrete normal form

$$\boxed{S(v_{\mathrm{E}}+x, \eta_{\mathrm{E}}-t) - S(v_{\mathrm{E}}, \eta_{\mathrm{E}}-t) = i(9x^3 - 18\sqrt{2}\,tx) + O(x^4 + tx^2 + t^2|x|).} \tag{14.22}$$

Use the saddle exponential preceding [94, eq. (3.11)], namely $e^{S/(ih)}$ with $h = 2\pi b^2$. Rescaling $x = (h/27)^{1/3}y$ gives $y^3/3 - \zeta y$ as the leading exponent, where

$$\boxed{\zeta = \frac{6\sqrt{2}\,t}{h^{2/3}} = \frac{A}{2\sqrt{3}\,h^{2/3}} + O\left(\frac{A^2}{h^{2/3}}\right).} \tag{14.23}$$

This proves the cubic normal form of the *classical* phase and derives the scale $A \asymp b^{4/3}$ with its normalization. A smooth-amplitude integral on a specified Airy contour has the corresponding uniform asymptotic; see [54, Section 36.12]. Section 15 checks those hypotheses for the exact regular-family integral, with an explicit scalar regularization.

With TV's choice $\Delta - 5\pi^2/3$ for the term independent of v, the volume formula of [81], in TV's doubled potential convention, gives $\operatorname{Im} S_\pm = \pm 2W$. Thus the two saddle exponentials contain $e^{\pm 2W/h}$. If the volume exponent is written as $W/\hbar_{\mathrm{vol}}$, its parameter is

$$\hbar_{\mathrm{vol}} = \pi b^2 = h/2. \tag{14.24}$$

The leading cubic saddle values agree: $(2/3)\zeta^{3/2} \sim 2W/h$.

14.8 Correction of the integration-coordinate dictionary

An earlier version of this monograph placed a dense pole lattice at the classical fold. That conclusion was incorrect: the logarithmic coordinate of the dilogarithm potential was matched to the wrong shifted integration variable. The exact integral fixes the dictionary without an asymptotic choice of branch.

With all six charges $\alpha = (\pi - \eta)/(2\pi b)$, write the original integration variable in [94, eq. (2.27)] as w. Its integrand is

$$I_b(w,\eta) = S_b(w - 3\alpha)^4 S_b(4\alpha - w)^3 S_b(2Q - w). \tag{14.25}$$

Set $a = e^{i\eta}$. Directly,

$$e^{2\pi i b(w-3\alpha)} = -a^3 e^{2\pi i b w}.$$

Thus the variable u in the classical potential is

$$u = e^{2\pi i b w}, \qquad v = 2\pi b w. \tag{14.26}$$

The $Q/2$-shift used previously is incompatible with this identification. A second check is the exact shift quotient

$$\begin{aligned}\frac{I_b(w+b,\eta)}{I_b(w,\eta)} &= \frac{[2\sin\pi b(w-3\alpha)]^4}{[2\sin\pi b(4\alpha - w - b)]^3[2\sin\pi b(2Q - w - b)]} \\ &\longrightarrow \frac{(1+a^3u)^4}{(1-a^4u)^3(1-u)}.\end{aligned} \tag{14.27}$$

The limit reproduces the saddle polynomial (14.21).

The lift of the double root lying in the separating contour strip is

$$v_* = 4\pi - 5\eta_{\mathrm{E}}, \qquad w_* = \frac{v_*}{2\pi b}.$$

It differs by 2π from the logarithmic lift used in the classical calculation; all its local derivatives remain the same. At this lift,

$$\begin{aligned} w_* - 3\alpha_{\mathrm{E}} &= \frac{\pi - 2\eta_{\mathrm{E}}}{2\pi b}, \\ 4\alpha_{\mathrm{E}} - w_* &= \frac{\eta_{\mathrm{E}}}{2\pi b}, \\ 2Q - w_* &= \frac{5\eta_{\mathrm{E}}}{2\pi b} + 2b. \end{aligned} \tag{14.28}$$

All three arguments lie in $0 < \operatorname{Re} z < Q$ for sufficiently small b, a fixed scaled distance from its boundary. There is therefore no integration-pole accumulation at the fold. The external trivalent factors do have parameter poles; the next section removes them exactly and proves a uniform transition formula.

15 A uniform quantum Airy transition for the regular family

This section proves a statement for an explicitly specified meromorphic Barnes integral on the one-parameter regular slice. It does not assert a transition theorem for arbitrary six angle parameters, a three-manifold invariant, or an identification with the cone-metric measure. No novelty claim is made.

15.1 The exact normalization and the result

Take $b > 0$, $Q = b + b^{-1}$, and

$$\alpha = \frac{\pi - \eta}{2\pi b}, \quad \eta_{\mathrm{E}} = \arccos(1/3), \quad h = 2\pi b^2, \quad \lambda = h^{1/3}.$$

For η close to η_{E} and sufficiently small b, choose an upward vertical contour C satisfying $Q < \operatorname{Re} w < 4\alpha$. This lies in the usual separating interval $3\alpha < \operatorname{Re} w < \min(4\alpha, 2Q)$ and puts all three integrand arguments in their common fundamental strip. Define

$$\mathcal{R}_b(\eta) = \frac{S_b(3\alpha - Q)^2}{S_b(\alpha)^6} \int_C I_b(w, \eta)\, \frac{\mathrm{d}w}{i}. \tag{15.1}$$

This is the equal-label symmetric formula (2.27) of [94], with the measure convention specified as $\mathrm{d}w/i$. If that formula's printed $\mathrm{d}w$ is read literally, its value is $i\mathcal{R}_b$. Keeping this constant phase explicit avoids imposing an unstated Barnes-measure convention.

The functional equation $S_b(x + b^{-1}) = 2\sin(\pi x/b) S_b(x)$ gives the exact identity

$$\frac{S_b(3\alpha - Q)^2}{S_b(\alpha)^6} = \frac{1}{4\sin^2((3\eta - \pi)/(2b^2))} \frac{S_b(3\alpha - b)^2}{S_b(\alpha)^6}. \tag{15.2}$$

Accordingly define

$$\widehat{\mathcal{R}}_b(\eta) = 4\sin^2\!\left(\frac{3\eta - \pi}{2b^2}\right) \mathcal{R}_b(\eta) = \frac{S_b(3\alpha - b)^2}{S_b(\alpha)^6} \int_C I_b(w, \eta)\, \frac{\mathrm{d}w}{i}. \tag{15.3}$$

The right-hand side defines the continuation across the apparent singularities in the product on the left.

Proposition 15.1 (Uniform regular-family transition)**.** *For every fixed $Z > 0$, put*

$$\eta = \eta_{\mathrm{E}} - \frac{h^{2/3}}{6\sqrt{2}}\,\zeta, \qquad |\zeta| \le Z.$$

As $b \downarrow 0$, the integral (15.3) satisfies, uniformly for these real ζ,

$$\boxed{\widehat{\mathcal{R}}_b(\eta) = \frac{3}{4}\,\frac{h^{1/3}}{b}\left\{\mathrm{Ai}(\zeta) - 3\sqrt{2}\,h^{1/3}\,\mathrm{Ai}'(\zeta) + O_Z(h^{2/3})\right\}.} \tag{15.4}$$

The error inside braces is additive. The physical compact hyperbolic side has $\zeta \ge 0$; negative ζ describes the analytic continuation across the Euclidean transition.

In particular the leading prefactor is $\frac{3}{4}(2\pi)^{1/3}b^{-1/3}$. The correction can equivalently be written by replacing the Airy argument with $\zeta - 3\sqrt{2}\,h^{1/3}$, at an error $O_Z(h^{2/3})$. The proof below controls the exact quantum integrand, the whole contour, its orientation, and the first correction.

15.2 A uniform double-sine estimate from its real integral

Let Λ be the analytic primitive on $0 < \operatorname{Re} z < \pi$ specified by

$$\Lambda'(z) = -\log(2\sin z), \qquad \Lambda(\pi/2) = 0,$$

with the logarithm real on $(0,\pi)$. On that real interval it is the usual Lobachevsky function.

Lemma 15.2 (Strip estimate)**.** *For z in a compact subset of $0 < \operatorname{Re} z < 1$ and real κ in a bounded set,*

$$S_b(z/b + \kappa b) = \exp\!\left(-\frac{\Lambda(\pi z)}{\pi b^2}\right)(2\sin \pi z)^{\kappa - 1/2}(1 + O(b^2)). \tag{15.5}$$

The logarithmic error and its derivatives are uniform on smaller compact subsets. For $\varepsilon \le \operatorname{Re} z \le 1 - \varepsilon$, the real part of the logarithmic error is $O_\varepsilon(b^2)$ uniformly for all $\operatorname{Im} z$.

Proof. Subtract the two double-Gamma integrals defining $\log S_b(x) = \log \Gamma_b(x) - \log \Gamma_b(Q - x)$; see [94, Appendix A]. At the centered argument, the substitution of the integration variable by $2bs$ gives

$$\log S_b(z/b + b/2) = -\frac{\Lambda(\pi z)}{\pi b^2} + D_b(z), \tag{15.6}$$

where

$$D_b(z) = \frac{1}{2}\int_0^\infty \frac{\sinh((1-2z)s)}{s\sinh s}\left(\operatorname{csch}(b^2 s) - \frac{1}{b^2 s}\right)\mathrm{d}s. \tag{15.7}$$

The leading primitive is fixed by its derivative $\log(2\sin \pi z)$ and its value zero at $z = 1/2$. For $u > 0$, $|\operatorname{csch} u - u^{-1}| \le Cu$. Near $s = 0$ the numerator $\sinh((1-2z)s)$ is $O(s)$, and at infinity the strip condition supplies exponential decay. This proves $D_b = O(b^2)$ on compact subsets. Holomorphy and Cauchy's estimate give the derivative statement.

For the global real-part bound use

$$\operatorname{Re}\sinh((1-2z)s) = \sinh((1 - 2\operatorname{Re} z)s)\cos(2\operatorname{Im} z\, s).$$

The same integrable bound is then independent of $\operatorname{Im} z$. Finally replace z in (15.6) by $z + (\kappa - \frac{1}{2})b^2$ and Taylor expand the leading primitive. Its second and higher derivatives are bounded on closed substrips, including as $|\operatorname{Im} z| \to \infty$. This gives (15.5) and the corresponding global modulus estimate. □

15.3 The smooth quantum phase and amplitude

Use the corrected variable $v = 2\pi b w$, and set

$$\begin{aligned} \theta_1 &= \tfrac{1}{2}(v - 3\pi + 3\eta), \quad \theta_2 = \tfrac{1}{2}(4\pi - 4\eta - v), \\ \theta_3 &= \tfrac{1}{2}(4\pi - v), \qquad\qquad s = \tfrac{1}{2}(\pi - \eta). \end{aligned} \tag{15.8}$$

The three integrand arguments are $\theta_1/(\pi b)$, $\theta_2/(\pi b)$, and $\theta_3/(\pi b) + 2b$. The cleared prefactor uses $3s/(\pi b) - b$ and $s/(\pi b)$. All five scaled arguments are in a closed substrip near the fold. Lemma 15.2 therefore gives

$$\frac{S_b(3\alpha - b)^2}{S_b(\alpha)^6} I_b(w, \eta) = \mathcal{A}(v, \eta)\, e^{G(v,\eta)/h}(1 + O(h)) \tag{15.9}$$

on compact subsets, where

$$\begin{aligned} G(v, \eta) &= -2\{4\Lambda(\theta_1) + 3\Lambda(\theta_2) + \Lambda(\theta_3)\} + 12\Lambda(s) - 4\Lambda(3s), \\ \mathcal{A}(v, \eta) &= (2\sin\theta_1)^{-2}(2\sin\theta_2)^{-3/2}(2\sin\theta_3)^{3/2} \left(\frac{\sin s}{\sin 3s}\right)^3. \end{aligned} \tag{15.10}$$

The powers are continued from their positive real values. For modulus estimates on the full vertical line, the same formula holds with a bounded multiplicative error, uniformly in its imaginary coordinate; this follows from the real-part estimate in the lemma.

At $v_* = 4\pi - 5\eta_{\mathrm{E}}$ the exact values are

	θ_1	θ_2	θ_3
$\sin\theta_j$	$1/3$	$1/\sqrt{3}$	$1/(9\sqrt{3})$
$\cot\theta_j$	$2\sqrt{2}$	$\sqrt{2}$	$-11\sqrt{2}$

$$\frac{\sin s}{\sin 3s} = 3.$$

Consequently

$$\mathcal{A}_* = \frac{9}{4}, \qquad (\log \mathcal{A})_v\big|_* = 7\sqrt{2}. \tag{15.11}$$

Differentiation of (15.10) gives

$$\begin{gathered} G_* = G_v|_* = G_{vv}|_* = G_\eta|_* = 0, \\ G_{vvv}|_* = 54, \qquad G_{v\eta}|_* = 18\sqrt{2}, \qquad G_{\eta\eta}|_* = 36\sqrt{2}, \\ G_{vvvv}|_* = -648\sqrt{2}, \qquad G_{vv\eta}|_* = -18. \end{gathered} \tag{15.12}$$

The value $G_* = 0$ can be checked without a numerical evaluation of a Lobachevsky identity. For real v, η near the fold, the identity $\operatorname{Im}\operatorname{Li}_2(e^{i\vartheta}) = 2\Lambda(\vartheta/2)$, together with periodicity and oddness, identifies the entire term $-2\{4\Lambda(\theta_1) + 3\Lambda(\theta_2) + \Lambda(\theta_3)\}$ with $\operatorname{Im} U(e^{iv}, \eta)$. The four trivalent terms in the doubled Murakami–Yano correction have imaginary part

$$-6\operatorname{Im}\operatorname{Li}_2(-e^{i\eta}) - 2\operatorname{Im}\operatorname{Li}_2(-e^{-3i\eta}) = 12\Lambda(s) - 4\Lambda(3s).$$

The remaining logarithm products are real. Thus G at the two classical saddles is $\pm 2W$ by [81, Theorem 2]; at their Euclidean collision $W \to 0$. This fixes the constant in (15.12).

15.4 The full contour and its tail estimate

The poles of the first double-sine factor are at $w = 3\alpha - mb - n/b$; the other two factors have their poles at $w = 4\alpha + mb + n/b$ and $w = 2Q + mb + n/b$. Thus the separating interval remains nonempty for $0 < \alpha < 2Q/3$. Decreasing the equal real labels from $\alpha = Q/2$ to

the compact values preserves this interval. Along this path $\alpha > Q/4$, so one may choose $\max(3\alpha, Q) < \operatorname{Re} w < 4\alpha \le 2Q$ and keep all three arguments in the common strip throughout. The integral has no pinch on this path. The singularities in the external trivalent factor have already been removed by (15.2).

On the vertical line $v = v_* + iy$ at $\eta = \eta_{\mathrm{E}}$, put $r = \tanh(y/2)$. For $y > 0$, direct differentiation gives

$$-\frac{\mathrm{d}}{\mathrm{d}y} \operatorname{Re} G(v_* + iy, \eta_{\mathrm{E}}) = \Psi(r), \tag{15.13}$$

where

$$\Psi(r) = 4 \arctan(2\sqrt{2}r) + 3 \arctan(\sqrt{2}r) - \arctan(11\sqrt{2}r).$$

Here $\Psi(0) = 0$ and

$$\Psi'(r) = \frac{864\sqrt{2}\, r^2(3 + 11r^2)}{(1 + 8r^2)(1 + 2r^2)(1 + 242r^2)} > 0 \qquad (r > 0). \tag{15.14}$$

It follows that the fold is the unique maximum of the real part of G on this vertical line. In addition,

$$\operatorname{Re} G(v_* + iy, \eta_{\mathrm{E}}) = -27\sqrt{2}\, y^4 + O(y^6), \qquad \Psi(\tanh(y/2)) \longrightarrow 2\pi \quad (y \to +\infty).$$

Complex conjugation gives the corresponding estimates for negative y.

These bounds persist in the form needed below under small changes of η and of the real contour coordinate. On bounded annuli in y this follows by continuity. For large $|y|$, the identity $4\theta_1 + 3\theta_2 + \theta_3 = 2\pi$ on the real line shows directly that the limiting decay slope is 2π, uniformly for nearby real parameters. The amplitude in (15.10) is $O(e^{-|y|})$ there. The global modulus bound from Lemma 15.2 transfers these estimates to the exact integrand, so all tails outside a fixed fold neighborhood are exponentially small in h^{-1}.

15.5 Airy localization and the first correction

Proof of Proposition 15.1. Write $t = \eta_{\mathrm{E}} - \eta$ and $\lambda = h^{1/3}$. Choose a fixed $c > 0$ and move the contour to

$$v = v_* + c\lambda + iy.$$

For all sufficiently small b this deformation stays within the common pole-free strip. The tail bounds just proved make the horizontal closing segments vanish, so no residues or undetermined cycle coefficients enter.

At this displaced real point, $G_{vv} = 54c\lambda + O_Z(\lambda^2) > 0$ and $G_{vvvv} = -648\sqrt{2} + O_Z(\lambda) < 0$. Taylor expansion in iy, with real derivatives at the real point, gives constants $c_1, c_2, \delta > 0$ such that

$$\begin{aligned} &\operatorname{Re}\{G(v_* + c\lambda + iy, \eta) - G(v_* + c\lambda, \eta)\} \\ &\qquad \le -c_1 \lambda y^2 - c_2 y^4, \qquad |y| \le \delta. \end{aligned} \tag{15.15}$$

The constants can be taken uniform for $|\zeta| \le Z$. Outside this neighborhood use the previously established fixed tail gap and linear decay.

Now set

$$x = v - v_* = \frac{\lambda}{3} Y, \qquad t = \frac{\lambda^2}{6\sqrt{2}} \zeta.$$

The rescaled contour is upward with $\operatorname{Re} Y = 3c$. The derivative table gives, through weighted order four,

$$G(v_* + x, \eta_{\mathrm{E}} - t) = 9x^3 - 18\sqrt{2}\,tx + 18\sqrt{2}\,t^2 - 27\sqrt{2}\,x^4 + 9tx^2 + \text{higher weighted terms},$$
$$\frac{\mathcal{A}(v_* + x, \eta_{\mathrm{E}} - t)}{9/4} = 1 + 7\sqrt{2}\,x + O(x^2 + |t|). \tag{15.16}$$

The exact quantum error in (15.9) is $O(h) = O(\lambda^3)$ and hence does not affect the first correction. Consequently, the rescaled integrand is

$$\frac{9}{4}\, e^{Y^3/3 - \zeta Y} \left\{ 1 + \lambda P(Y, \zeta) + O_Z(\lambda^2) \right\}, \tag{15.17}$$

where the pointwise remainder is locally uniform and

$$P(Y, \zeta) = \frac{7\sqrt{2}}{3} Y - \frac{\sqrt{2}}{3} Y^4 + \frac{\zeta}{6\sqrt{2}} Y^2 + \frac{\sqrt{2}}{4} \zeta^2. \tag{15.18}$$

For completeness, the integrated remainder is also $O_Z(\lambda^2)$. On $|Y| \le \lambda^{-1/12}$, analytic Taylor remainders in (15.16), followed by exponentiation, are bounded by λ^2 times a fixed polynomial in $|Y|$, against a Gaussian on the vertical line. Equation (15.15) supplies a uniform Gaussian majorant for the exact integrand. The part with $\lambda^{-1/12} < |Y|$ inside the fixed fold neighborhood is smaller than every power of λ; the global contour tail is exponentially small in h^{-1}. This justifies integration of the expansion and its first correction.

With the chosen orientation,

$$\frac{1}{2\pi i} \int_{\operatorname{Re} Y = 3c} e^{Y^3/3 - \zeta Y}\, \mathrm{d}Y = \mathrm{Ai}(\zeta);$$

this is the vertical-contour form of the Airy representation in [54, Section 9.5]. Inserting Y^k yields $(-1)^k \mathrm{Ai}^{(k)}(\zeta)$. Using

$$\mathrm{Ai}''(\zeta) = \zeta\, \mathrm{Ai}(\zeta), \qquad \mathrm{Ai}^{(4)}(\zeta) = 2\, \mathrm{Ai}'(\zeta) + \zeta^2\, \mathrm{Ai}(\zeta)$$

in (15.18), all terms proportional to $\zeta^2\,\mathrm{Ai}$ cancel. The result is $-3\sqrt{2}\ \mathrm{Ai}'(\zeta)$. Finally, $\mathrm{d}w/i = \lambda\, \mathrm{d}Y/(6\pi i b)$. Multiplication by $9/4$ gives the coefficient $3\lambda/(4b)$ in (15.4). □

15.6 A numerical check of the exact integral

The following check evaluates the convergent real integral for $\log S_b$, combines all factors before exponentiating, and integrates the exact cleared integrand on the displaced vertical contour. It uses no cubic replacement for that integrand. Write

$$Q_b(\zeta) = \frac{4b}{3h^{1/3}} \widehat{\mathcal{R}}_b(\eta), \qquad Q_b^{(1)}(\zeta) = \mathrm{Ai}(\zeta) - 3\sqrt{2}\, h^{1/3}\, \mathrm{Ai}'(\zeta).$$

b	ζ	Numerical Q_b	$Q_b^{(1)}$
0.0125	0	0.459087	0.464164
0.00625	0	0.422572	0.423779
0.003125	0	0.398047	0.398339
0.0125	1	0.214333	0.202400
0.00625	1	0.182316	0.177567
0.003125	1	0.163805	0.161924

For the three $\zeta = 1$ rows, the differences divided by $h^{2/3}$ are approximately 1.208, 1.211, and 1.209. The quadrature is an independent numerical check of the normalization and first correction, not a certified error bound; the proof of the proposition supplies the uniform estimate.

15.7 Interpretation and scope

The correction to the earlier pole argument is substantive. The exact quantum integrand is smooth at the correct fold; the remaining real-b resonances are an explicit scalar factor outside the integral. For example, at $\eta = \eta_{\mathrm{E}}$ the original normalization has poles along

$$b^2 = \frac{3\eta_{\mathrm{E}} - \pi}{2\pi m}, \qquad m = 1, 2, \dots.$$

The numerator in (15.3) is nonzero for sufficiently small b in a bounded physical transition window, because $\mathrm{Ai}(\zeta) > 0$ for $\zeta \geq 0$. Thus the poles are not canceled by the integral. The cleared quantity has the uniform asymptotic in the proposition even through these resonances.

In geometric terms,

$$\zeta = \frac{A}{2\sqrt{3}\, h^{2/3}} + O\left(\frac{A^2}{h^{2/3}}\right), \qquad A \asymp b^{4/3}.$$

The decaying Airy branch selected by the contour agrees with the decaying volume exponential: its classical saddle exponent is $-2W/h = -W/(\pi b^2)$. The proposition is uniform for bounded ζ; it does not, by itself, justify taking $\zeta \to \infty$ simultaneously with $b \to 0$ or keeping a nonzero tetrahedral volume fixed. Those matching limits require additional estimates.

This proves a quantum transition for one specified regular family, beyond the classical cubic calculation. Extension to arbitrary nondegenerate Euclidean tetrahedral shapes, comparison of normalizations used in a global state integral, and gluing into a topological invariant remain separate problems. In particular no new case of the colored-Jones volume conjecture, and no identification of the integrated quantum measure with the cone-metric measure, follows from this proposition alone.

16 Dehn filling, variational response, and quantum surgery

There is a second route from volume optimization to a quantum-volume comparison: perturb a complete cusped structure by prescribing peripheral holonomy. This route has a particularly explicit answer. The critical-value response of the real concave optimization recovers the Euclidean cusp quadratic form. After passage to a holomorphic boundary potential, the filled volume is exactly the imaginary part of a Legendre critical value. Its Taylor coefficients reproduce the Neumann–Zagier expansion, and its Hessian is the derivative entering the torsion surgery formula. These are deductions from the established variational, deformation, and torsion identities cited below; no novelty claim is made.

16.1 Peripheral conventions and the two small parameters

Let M be an oriented, finite-volume, one-cusped hyperbolic three-manifold. Fix an oriented meridian–longitude basis (μ, λ) and the geometric deformation branch through the complete structure. We use logarithms of the $\mathrm{PSL}_2(\mathbb{C})$ multipliers, equivalently of the eigenvalue ratios, so that

$$u = H(\mu), \qquad v = H(\lambda), \qquad v(u) = \tau u + c_3 u^3 + c_5 u^5 + O(u^7), \qquad B = \Im\tau > 0. \tag{16.1}$$

Thus u is twice a local logarithm of an SL_2 eigenvalue. In particular, the smooth filling equation is $pu + qv = 2\pi i$, with the branch chosen continuously from the cusp. Using eigenvalue logarithms instead changes this right-hand side and the nonlinear coefficients. The oddness and the holomorphic potential are those of Neumann–Zagier [36, Sections 4–6].

For a primitive slope (p, q) put

$$A = p + q\tau, \qquad Q(p,q) = \frac{|A|^2}{B}, \qquad pu + qv(u) = i\theta. \tag{16.2}$$

Here θ is the cone angle, with $\theta = 0$ the complete cusp and $\theta = 2\pi$ a smooth hyperbolic filling when the geometric branch extends there. The quantity Q is the squared length of the slope on a Euclidean cusp cross section divided by its area. For sufficiently long slopes the whole interval $0 \le \theta \le 2\pi$ stays in the local holonomy neighborhood. The holomorphic formulas below also apply to any local branch without requiring a separate construction of a cone path.

Two expansions must be kept distinct: $|A|^{-1} \to 0$ is a *classical long-slope expansion*, whereas the quantum level $r \to \infty$ is a semiclassical expansion. The coefficients c_3, c_5 belong to the first. They are not by themselves quantum two-loop or higher-loop coefficients.

16.2 What the linear term does to real volume optimization

Suppose a chosen ideal triangulation has a positive complete geometric angle structure. Use linear coordinates x on its angle space, with tetrahedral and edge-angle constraints already imposed, and center them at the complete maximum $x = 0$. Write

$$\mathscr{V}(x) = \sum_{\text{tetrahedra}} \mathrm{Vol}(\Delta_x), \qquad V_0 = \mathscr{V}(0), \qquad K = -D^2\mathscr{V}(0) > 0.$$

Peripheral angular holonomies are affine in the angles and vanish at 0 in the selected branches. Thus there is a real linear map P such that

$$Px = \begin{pmatrix} \Im u \\ \Im v \end{pmatrix}, \qquad L = (p \;\; q)P, \qquad Lx = \theta.$$

An interior critical point of $\mathscr{V}$ on this one-constraint slice gives the filling structure. The missing real peripheral equation follows from stationarity, together with the real edge equations; see [37, Proposition 5.4 and Theorem 6.1].

To see the peripheral part directly, on the edge-glued deformation space the volume differential is

$$\mathrm{d}V = \frac{1}{2}(\Re v \; \mathrm{d}\Im u - \Re u \; \mathrm{d}\Im v). \tag{16.3}$$

Requiring this covector to be proportional to $p\,\mathrm{d}\Im u + q\,\mathrm{d}\Im v$ gives $p\Re u + q\Re v = 0$. Hence the single angular condition, together with constrained criticality, gives the full complex equation (16.2).

The proposed linear perturbation can therefore be written as

$$\mathscr{B}(j) = \max_x\{\mathscr{V}(x) + jLx\}, \qquad \theta = \mathscr{B}'(j), \qquad V_{p,q}(\theta) = \mathscr{B}(j) - j\theta. \tag{16.4}$$

All statements here are local near the unique interior complete maximum. The envelope theorem gives

$$V'_{p,q}(\theta) = -j.$$

The Schlaefli formula gives $V'_{p,q}(\theta) = -\ell_{\mathrm{core}}/2$, so the coefficient of the added real linear term is

$$\boxed{j = \frac{\ell_{\mathrm{core}}}{2}.} \tag{16.5}$$

The cone angle is its conjugate variable. In particular, the critical value $\mathscr{B}$ of the biased objective and the physical volume $V_{p,q}$ must be distinguished by the Legendre subtraction in (16.4).

16.3 The cusp shape in the inverse volume Hessian

Set

$$\chi = LK^{-1}L^T.$$

Differentiating the stationary equation for the perturbed objective gives

$$\left.\frac{\mathrm{d}x_j}{\mathrm{d}j}\right|_{j=0} = K^{-1}L^T, \qquad \mathscr{B}''(0) = \chi.$$

Consequently

$$V_{p,q}(\theta) = V_0 - \frac{\theta^2}{2\chi} + O(\theta^4). \tag{16.6}$$

Evenness follows from the signed-cone continuation $u \mapsto -u$. Comparison with the geometric calculation below yields

$$\boxed{LK^{-1}L^T = \frac{2|p+q\tau|^2}{\Im\tau} = 2Q(p,q).} \tag{16.7}$$

Equivalently, writing $a = \Re\tau$,

$$\boxed{PK^{-1}P^T = \frac{2}{B}\begin{pmatrix} 1 & a \\ a & |\tau|^2 \end{pmatrix}, \qquad \det(PK^{-1}P^T) = 4.} \tag{16.8}$$

Thus the peripheral response of the angle-volume Hessian determines the cusp shape. Although K and P depend on the triangulation, this two-by-two matrix depends only on the marked cusp. The formula assumes the nondegenerate positive angle maximum above; it is not a claim that an arbitrary chosen ideal triangulation has one.

16.4 An exact holomorphic critical-value identity

Define

$$F(u) = \int_0^u v(t)\,\mathrm{d}t = \frac{\tau u^2}{2} + \frac{c_3 u^4}{4} + \frac{c_5 u^6}{6} + O(u^8). \tag{16.9}$$

With Neumann–Zagier's convention $\Phi'_{\mathrm{NZ}} = 2v$, this is $F = \Phi_{\mathrm{NZ}}/2$, after fixing the additive constant. Assume $q \neq 0$ and set

$$\boxed{\mathcal{S}_{p,q}(u;\theta) = F(u) + \frac{p}{2q}u^2 - \frac{i\theta}{q}u.} \tag{16.10}$$

The quadratic term implements the change of peripheral polarization; the final term is linear in the boundary coordinate. Its critical equation is exactly $pu + qv = i\theta$. Let u_θ denote the small critical solution, and write $\mathcal{S}^*(\theta) = \mathcal{S}_{p,q}(u_\theta;\theta)$.

Proposition 16.1 (Local filling critical value)**.** *On the geometric branch through the complete cusp,*

$$\boxed{V_{p,q}(\theta) - V_0 = \frac{1}{2}\Im\mathcal{S}^*(\theta).} \tag{16.11}$$

Moreover,

$$\frac{\mathrm{d}\mathcal{S}^*}{\mathrm{d}\theta} = -\frac{iu_\theta}{q}, \qquad \frac{\mathrm{d}^2\mathcal{S}^*}{\mathrm{d}\theta^2} = \frac{1}{q(p+qv'(u_\theta))}. \tag{16.12}$$

Proof. Choose integers ρ, σ with $p\sigma - q\rho = 1$. The complementary peripheral holonomy is

$$w = \rho u + \sigma v = \frac{i\sigma\theta - u}{q}.$$

In these orientation conventions the physical core complex length has positive real part when represented by $-w$, and therefore $\ell_{\text{core}} = \Re(u/q)$ near the positive cone branch. The envelope theorem gives the first identity in (16.12). Its half imaginary part is $-\ell_{\text{core}}/2$, exactly the Schlaefli derivative of volume. Both sides of (16.11) vanish at $\theta = 0$. Differentiating $pu_\theta + qv(u_\theta) = i\theta$ gives the second identity in (16.12).

There is also an algebraic proof valid directly on the local holonomy branch. Neumann–Zagier's volume formula is

$$\begin{aligned} V(u) - V_0 &= \frac{1}{4}\Im(u\bar{v}) + \Im\left(\frac{1}{2}F(u) - \frac{1}{4}uv\right) \\ &= \frac{1}{2}\Im F(u) - \frac{1}{2}\Re u\,\Im v. \end{aligned} \tag{16.13}$$

Substituting $\Im v = (\theta - p\Im u)/q$ yields (16.11). This proof avoids assuming an independently constructed cone path. □

This is an exact identity of analytic germs, so it compares all orders of the two classical expansions. For $q = 0$, exchange the oriented basis to $(\mu', \lambda') = (\lambda, -\mu)$; the new slope has $q' = -p \neq 0$. The real angle optimization itself needs no such basis change. Primitive $q = 0$ slopes have $p = \pm 1$ and do not occur in a long-slope sequence in a fixed basis.

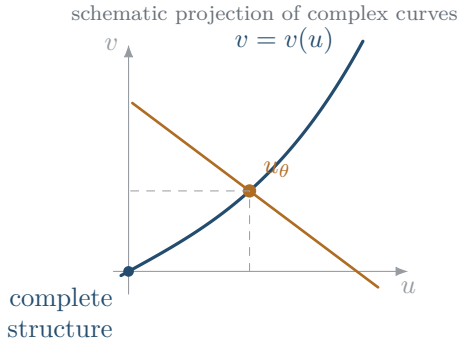


Figure 6: Local Dehn filling as an intersection in the complex peripheral holonomy space. The drawing records the intersection condition; it is not a real slice of the complex deformation curve. Here u, v are logarithms of multipliers, $\theta = 0$ is the complete structure, and $\theta = 2\pi$ gives smooth filling when the local geometric branch extends to that value. The Hessian response formula assumes the nondegenerate positive angle maximum specified in the text.

16.5 The Neumann–Zagier coefficients through sixth order

Series inversion of (16.2) gives

$$u_\theta = \frac{i\theta}{A} + \frac{iqc_3\theta^3}{A^4} + i\theta^5\left(\frac{3q^2c_3^2}{A^7} - \frac{qc_5}{A^6}\right) + \cdots. \tag{16.14}$$

Substitution in the action, or integration of the envelope derivative, gives

$$\mathcal{S}^*(\theta) = \frac{\theta^2}{2qA} + \frac{c_3\theta^4}{4A^4} + \frac{\theta^6}{6}\left(\frac{3qc_3^2}{A^7} - \frac{c_5}{A^6}\right) + \cdots. \tag{16.15}$$

Taking half its imaginary part yields the volume loss

$$V_0 - V_{p,q}(\theta) = \frac{B\theta^2}{4|A|^2} - \frac{\theta^4}{8}\Im\frac{c_3}{A^4} - \frac{\theta^6}{12}\Im\left(\frac{3qc_3^2}{A^7} - \frac{c_5}{A^6}\right) + O(|A|^{-8}). \tag{16.16}$$

Here the displayed remainder is a long-slope estimate for bounded θ. Since q/A is bounded, the three displayed terms have orders $|A|^{-2}, |A|^{-4}, |A|^{-6}$. For a fixed slope and $\theta \to 0$, the corresponding Taylor remainder is $O(\theta^8)$.

For the smooth hyperbolic filling $M_{p/q}$ this becomes

$$\boxed{V_0 - \mathrm{Vol}(M_{p/q}) = \frac{\pi^2 B}{|A|^2} - 2\pi^4\Im\frac{c_3}{A^4} - \frac{16\pi^6}{3}\Im\left(\frac{3qc_3^2}{A^7} - \frac{c_5}{A^6}\right) + O(|A|^{-8}).} \tag{16.17}$$

These are the Neumann–Zagier coefficients, with the multiplier normalization fixed explicitly [36, Section 6]. In particular the leading loss is $\pi^2/Q(p,q)$, and the quartic term detects the first nonlinear deformation coefficient c_3.

16.6 Computing the quartic coefficient from angle-volume derivatives

The comparison supplies a more detailed computational identity. Keep K, L, χ as above, and write

$$T_3 = D^3\mathscr{V}(0), \qquad T_4 = D^4\mathscr{V}(0), \qquad e = \frac{K^{-1}L^T}{\chi}, \qquad R = K^{-1} - \frac{K^{-1}L^T L K^{-1}}{\chi}.$$

Let $t = T_3(e, e, \cdot)$ be the resulting column covector. Solving the constrained stationary equation to second order gives

$$x_\theta = \theta e + \frac{\theta^2}{2}Rt + O(\theta^3).$$

Indeed $Le = 1$, $LR = 0$, and RK is the identity on $\ker L$. Inserting this expression into the Taylor expansion of $\mathscr{V}$ gives

$$V_{p,q}(\theta) = V_0 - \frac{\theta^2}{2\chi} + \frac{\theta^3}{6}T_3[e,e,e] + \frac{\theta^4}{24}\Big(T_4[e,e,e,e] + 3t^T Rt\Big) + O(\theta^5). \tag{16.18}$$

The cross term with the unknown cubic displacement vanishes because that displacement is in $\ker L$ and $Ke = L^T/\chi$. The remaining quadratic-displacement terms combine to $t^T Rt/8$. Evenness of the geometric branch then removes the odd coefficients and improves the remainder to $O(\theta^6)$. Comparison with (16.16) yields

$$\boxed{T_3[e,e,e] = 0, \qquad T_4[e,e,e,e] + 3t^T Rt = 3\Im\frac{c_3}{(p+q\tau)^4}.} \tag{16.19}$$

Thus c_3 can be accessed from derivatives of the angle-volume functional by varying the slope. The correction $3t^T Rt$ accounts for readjustment of all other angles. Keeping only the raw fourth derivative would miss that response.

16.7 The closed quantum-volume statement and its local consequence

A precise closed-manifold analogue is the Chen–Yang volume conjecture. At the specified root $\zeta_r = e^{2\pi i/r}$, with r odd, its absolute-growth statements are

$$\frac{4\pi}{r}\log|\mathrm{RT}_r(N;\zeta_r)| \longrightarrow \mathrm{Vol}(N), \qquad \frac{2\pi}{r}\log \mathrm{TV}_r(N;\zeta_r) \longrightarrow \mathrm{Vol}(N). \tag{16.20}$$

The root and normalization matter; these are the Chen–Yang versions for hyperbolic manifolds [38]. We use absolute growth to avoid mixing the different Chern–Simons sign conventions in the literature.

For a boundary state whose geometric semiclassical phase satisfies

$$\Phi(u) = \Phi_0 + \frac{1}{2}F(u), \qquad \Im\Phi_0 = V_0,$$

the filling phase is

$$\mathcal{T}_\theta(u) = \Phi_0 + \frac{1}{2}\mathcal{S}_{p,q}(u;\theta). \tag{16.21}$$

Possible real framing constants do not alter its imaginary part. With $\hbar_r = 4\pi i/r$, Proposition 16.1 immediately gives

$$\left|e^{\mathcal{T}_\theta(u_\theta)/\hbar_r}\right| = \exp\left(\frac{r}{4\pi}V_{p,q}(\theta)\right). \tag{16.22}$$

In particular, the deformation of the geometric exponential is

$$\log\left|\frac{e^{\mathcal{T}_\theta(u_\theta)/\hbar_r}}{e^{\Phi_0/\hbar_r}}\right| = -\frac{r}{4\pi}(V_0 - V_{p,q}(\theta)). \tag{16.23}$$

Every classical coefficient in (16.16) therefore appears with precisely the required quantum exponential coefficient. At $\theta = 2\pi$, Yoshida's complex-volume formula has this normalization; see [40, equation (5.9)]. Quadratic and linear filling kernels also appear explicitly in [39, equations (12)–(14)]; their boundary coordinates and level conventions must be converted before numerical coefficients are compared.

Proposition 16.2 (Conditional local quantum transfer)**.** *Suppose an exact surgery representation for a specified quantum invariant has, near the geometric branch, phase $\mathcal{T}_\theta$ and a uniform classical-symbol expansion of its amplitude. Suppose the geometric critical point is nondegenerate, its steepest-descent cycle occurs with nonzero coefficient in the specified integration cycle, the net leading stationary-phase coefficient of that geometric sector is uniformly bounded away from zero, and the remaining integral or sum is exponentially smaller, uniformly on a chosen compact deformation interval. Then stationary phase gives the volume exponent* (16.22) *for the full invariant, with a uniform expansion on that interval. Without the last global condition, the same statement holds for the specified local geometric contribution.*

Proof. At the geometric critical point,

$$\mathcal{T}_\theta''(u_\theta) = \frac{p + qv'(u_\theta)}{2q} \neq 0.$$

Uniform stationary phase on the stipulated cycle gives a nonzero amplitude times $e^{\mathcal{T}_\theta(u_\theta)/\hbar_r}$, with powers of r^{-1} and a Gaussian power of r according to the normalization of the kernel. Such powers contribute only $O(\log r/r)$ to normalized logarithmic growth. The exponential gap permits the addition of the remaining terms without changing the leading growth. The critical-value proposition identifies the exponent with the geometric volume. □

For a general knot this proposition identifies the needed analytic input; it does not establish that input from a complete-point Kashaev asymptotic alone. A value at $u = 0$ does not provide a uniform boundary-state expansion near 0, and a local critical point does not by itself specify its coefficient in the global surgery sum. The regular tetrahedral Airy formula in Section 15 concerns a different integral and does not supply that missing manifold-level information.

16.8 The Gaussian factor and the torsion surgery formula

The derivative of the filling map is

$$D_* := p + qv'(u_\theta), \qquad \frac{\mathrm{d}u_\theta}{\mathrm{d}\theta} = \frac{i}{D_*}, \qquad \mathcal{S}_{uu}(u_\theta;\theta) = \frac{D_*}{q}. \tag{16.24}$$

Thus the same quantity controls the classical susceptibility and the Gaussian surgery integral. With the usual square-root normalization of the boundary Fourier kernel, its contribution has dependence $D_*^{-1/2}$, up to the common normalization and phase choices. A full amplitude also contains the boundary-state torsion and a solid-torus factor.

More precisely, at $\theta = 2\pi$ the representation descends to the smooth filled manifold. Suppose its adjoint representation is acyclic and its restriction is peripheral-regular. The change-of-peripheral-curve and surgery formulas for adjoint Reidemeister torsion then give

$$\begin{aligned} \mathrm{Tor}_{p\mu+q\lambda}(M_u) &= (p + qv'(u))\,\mathrm{Tor}_\mu(M_u), \\ \mathrm{Tor}(M_{p/q}) &= \frac{\mathrm{Tor}_{p\mu+q\lambda}(M_u)}{4\sinh^2(w/2)}, \qquad w = \rho u + \sigma v. \end{aligned} \tag{16.25}$$

These are understood up to the standard torsion signs and consistent choices of square root. They are the identities used in [40, equations (5.16) and (5.23)]. Consequently the expected full geometric amplitude is proportional to

$$\boxed{\frac{2\sinh(w/2)}{\sqrt{(p + qv'(u))\,\mathrm{Tor}_\mu(M_u)}} = \frac{1}{\sqrt{\mathrm{Tor}(M_{p/q})}}.} \tag{16.26}$$

The identity shows what must accompany the boundary Hessian. Matching the Gaussian factor alone would leave out both the cusp torsion and the solid-torus amplitude. This one-loop comparison uses more information than the classical volume exponent. The closed-manifold torsion in this formula is evaluated at smooth filling; it is not being assigned to a generic cone holonomy. At the cusp endpoint $\theta = 0$ the solid-torus factor vanishes, so a uniform quantum statement including that endpoint requires a separate normalization and analysis. The compact parameter interval in Proposition 16.2 is subject to its nonvanishing-amplitude hypothesis.

16.9 A figure-eight benchmark and the scope of the conclusion

For the figure-eight knot complement, in the standard peripheral basis Neumann–Zagier compute

$$v(u) = 2\sqrt{3}\,i\left(u + \frac{u^3}{3} + \frac{23u^5}{180} + O(u^7)\right). \tag{16.27}$$

Thus $\tau = 2\sqrt{3}\,i$, $c_3 = 2i/\sqrt{3}$, and $c_5 = 23i/(30\sqrt{3})$. The first two volume-loss terms are

$$\begin{aligned} V_0 - \mathrm{Vol}(M_{p/q}) = {} & \frac{2\sqrt{3}\,\pi^2}{p^2 + 12q^2} \\ & - \frac{4\pi^4}{\sqrt{3}}\frac{p^4 - 72p^2q^2 + 144q^4}{(p^2 + 12q^2)^4} + O\big((p^2 + q^2)^{-3}\big). \end{aligned} \tag{16.28}$$

This agrees directly with the expansion following equation (68) in [36]. The angle-Hessian response is correspondingly

$$\chi = \frac{p^2 + 12q^2}{\sqrt{3}}.$$

An independent calculation starts with the two regular ideal tetrahedra and their Lobachevsky volume. In the Neumann–Zagier presentation their shapes z_1, z_2 satisfy

$$z_1 z_2 (z_1 - 1)(z_2 - 1) = 1, \qquad u = \log(z_2(1 - z_1)), \qquad v = 2\log(z_1(1 - z_1)).$$

Use $\alpha_i = \arg z_i$, $\beta_i = -\arg(1 - z_i)$, and $\gamma_i = \pi - \alpha_i - \beta_i$. Choose the independent variations $\xi = (\delta\alpha_1, \delta\beta_1, \delta\alpha_2)$. The edge equation then makes the six angle variations

$$E\xi = (\xi_1, \xi_2, -\xi_1 - \xi_2, \xi_3, \xi_1 - \xi_2 + \xi_3, -\xi_1 + \xi_2 - 2\xi_3).$$

At the complete structure all six angles are $\pi/3$. Since $\Lambda''(\alpha) = -\cot\alpha$, direct differentiation gives

$$K = \frac{1}{\sqrt{3}} \begin{pmatrix} 4 & -1 & 3 \\ -1 & 4 & -3 \\ 3 & -3 & 6 \end{pmatrix}, \qquad P = \begin{pmatrix} 0 & -1 & 1 \\ 2 & -2 & 0 \end{pmatrix}, \qquad PK^{-1}P^T = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 \\ 0 & 12 \end{pmatrix}. \quad (16.29)$$

This recovers the cusp shape directly from the optimizer. Put $D_0 = p^2 + 12q^2$. The unit angular-response vector is

$$e = \frac{(6q - p, -6q - p, p - 6q)^T}{2D_0}.$$

Using

$$\Lambda'''(\pi/3) = \frac{4}{3}, \qquad \Lambda''''(\pi/3) = -\frac{8}{3\sqrt{3}},$$

the six entries of Ee occur in opposite pairs, so $T_3[e, e, e] = 0$. Direct contraction of the fourth derivative and the constrained inverse-Hessian correction gives

$$T_4[e, e, e, e] + 3t^T Rt = \frac{2\sqrt{3}\,(p^4 - 72p^2q^2 + 144q^4)}{D_0^4}. \quad (16.30)$$

Multiplying by $\theta^4/24$ and setting $\theta = 2\pi$ gives the quartic filled-volume term in (16.28) with the opposite sign, as required for volume rather than volume loss. Thus the real optimization and the holomorphic series agree in a calculation made directly from the tetrahedral volume derivatives.

For this example there is also a theorem about the actual quantum invariants: Wong–Yang prove the RT asymptotic expansion and the RT/TV volume conjectures for every hyperbolic rational Dehn filling of the figure-eight knot [40, Theorems 1.1 and 1.2]. Their leading amplitude is the inverse square root of adjoint torsion, with their specified overall constant. Thus the local consistency calculation above is contained in a stronger result in that example. The present discussion reorganizes the comparison around volume optimization and its derivatives; it does not produce a new case of the quantum-volume conjecture.

For a general knot, the unconditional results here are the classical critical-value identity and its coefficient dictionary. The quantum conclusion for a specified local geometric contribution follows under the stated stationary-phase hypotheses. Passing to the full invariant requires control of its exact surgery representation and all competing contributions. One should first keep the filling slope fixed while taking $r \to \infty$, and then study the long-slope coefficients. A simultaneous limit in which the slope grows with r requires uniform estimates not supplied by separate fixed-slope expansions.

16.10 The symplectic pairing behind the calculation

The pairing in Neumann–Zagier is the structure behind the preceding formulas. It identifies the peripheral variables as conjugate coordinates, makes the geometric deformation locus Lagrangian, and fixes the canonical transformation for Dehn filling. We record the normalization explicitly, since interchanging meridian and longitude or using eigenvalue rather than multiplier logarithms changes the displayed signs or factors.

16.10.1 From the tetrahedra to the boundary intersection form

For n ideal tetrahedra introduce independent ambient coordinates

$$Z_j = \log z_j, \qquad W_j = \log(1 - z_j), \qquad \Omega_{\mathrm{tet}} = \sum_{j=1}^{n} \mathrm{d}Z_j \wedge \mathrm{d}W_j.$$

The actual shapes satisfy $e^{Z_j} + e^{W_j} = 1$. Their product is a complex Lagrangian of dimension n in this $2n$-dimensional phase space. The logarithmic edge equations and peripheral holonomies are affine linear functions of $X = (Z_1, \dots, Z_n, W_1, \dots, W_n)^T$.

Let C_λ, C_μ be the longitude and meridian coefficient rows, let R_E be the edge-equation matrix, and set

$$J_{2n} = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}, \qquad U = \begin{pmatrix} C_\lambda \\ C_\mu \\ R_E \end{pmatrix}.$$

For a triangulated manifold with h torus cusps, the original matrix identity is

$$\boxed{U J_{2n} U^T = 2 \begin{pmatrix} 0 & I_h & 0 \\ -I_h & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.} \tag{16.31}$$

This is [36, Theorem 2.2, equation (21)], in their longitude–meridian row order. They normalize the coefficient-row pairing as $\langle a, b \rangle = \frac{1}{2} a J_{2n} b^T$. Equivalently, the unnormalized coefficient pairing of peripheral curves is twice their oriented intersection number; edge rows pair to zero with one another and with all peripheral rows. The topological interpretation is proved in their Section 3. A geometric construction of symplectic bases is given in [42].

The vanishing edge-edge pairings say that the linear constraints Poisson-commute. Since $\operatorname{rank} R_E = n - h$, their coisotropic reduction has complex dimension $2h$. Corollary 2.4 of [36] identifies the reduced form as

$$\boxed{\Omega_\partial = \frac{1}{2} \sum_{i=1}^{h} \mathrm{d}v_i \wedge \mathrm{d}u_i, \qquad \alpha_\partial = \frac{1}{2} \sum_{i=1}^{h} v_i \, \mathrm{d}u_i, \qquad \mathrm{d}\alpha_\partial = \Omega_\partial.} \tag{16.32}$$

Locally near the regular geometric point, imposing the edge equations on the tetrahedral Lagrangian and reducing gives the geometric deformation Lagrangian $\mathcal{L}_M$ in boundary holonomy space. For a modern formulation of the reduction, see [41]. The symplectic form lives on the ambient boundary phase space; its pullback to $\mathcal{L}_M$ is zero.

16.10.2 The potential and the volume primitive

Writing $v = v(u)$ on the geometric branch, Lagrangianity gives

$$\sum_i \mathrm{d}v_i \wedge \mathrm{d}u_i = 0 \quad \Longleftrightarrow \quad \frac{\partial v_i}{\partial u_j} = \frac{\partial v_j}{\partial u_i}. \tag{16.33}$$

Thus $\sum_i v_i\,\mathrm{d}u_i = \mathrm{d}F$ locally, and $\alpha_\partial|_{\mathcal{L}_M} = \mathrm{d}(F/2)$. This is the reciprocity and potential statement of [36, Theorem 3]. For one cusp the existence of a primitive is automatic once $v = v(u)$; the substantial content is the triangulation-level pairing and, for several cusps, the symmetry of the mixed derivatives. In the original normalization, $\Phi_{\mathrm{NZ}} = 2F$, whereas the complex-volume phase used above is $\Phi = \Phi_0 + F/2$.

For one cusp write

$$u = x + iy, \qquad v = r + is.$$

The real volume one-form is an exact modification of the imaginary part of the symplectic primitive:

$$\beta = \frac{1}{2}(r\,\mathrm{d}y - x\,\mathrm{d}s) = \Im\alpha_\partial - \frac{1}{2}\,\mathrm{d}(xs), \qquad \mathrm{d}\beta = \Im\Omega_\partial. \tag{16.34}$$

It is therefore closed on $\mathcal{L}_M$. Integrating it gives exactly (16.13). The factor $1/2$ can also be seen from the leading–trailing angle variations: angular holonomy varies by twice the boundary intersection number, while volume varies by the real holonomy [37, Lemma 4.4 and Proposition 5.4]. The same pairing thus supplies the real optimization differential and the holomorphic generating function.

16.10.3 Filling as a canonical transformation

Choose $\rho, \sigma \in \mathbb{Z}$ with $p\sigma - q\rho = 1$ and use the filling slope as the new meridian:

$$a = pu + qv, \qquad b = \rho u + \sigma v, \qquad \frac{1}{2}\,\mathrm{d}b \wedge \mathrm{d}a = \frac{1}{2}\,\mathrm{d}v \wedge \mathrm{d}u. \tag{16.35}$$

For $q \neq 0$, the type-I generating function is

$$\boxed{G(u,a) = \frac{p}{4q}u^2 - \frac{ua}{2q} + \frac{\sigma}{4q}a^2.} \tag{16.36}$$

Indeed, $v = (a - pu)/q$ and $b = (\sigma a - u)/q$ give the exact identity

$$\mathrm{d}G = \frac{1}{2}b\,\mathrm{d}a - \frac{1}{2}v\,\mathrm{d}u. \tag{16.37}$$

Thus

$$\widetilde{\Phi}(a) = \mathrm{crit}_u(\Phi(u) + G(u,a)) \quad \text{satisfies} \quad \widetilde{\Phi}'(a) = \frac{b}{2}.$$

The stationary equation is $a = pu + qv$. Setting $a = i\theta$ gives

$$\widetilde{\Phi}(i\theta) = \Phi_0 + \frac{1}{2}\mathcal{S}^*_{p,q}(\theta) - \frac{\sigma\theta^2}{4q}. \tag{16.38}$$

The last term is real, hence invisible to the volume. At $\theta = 2\pi$ it is $-\pi^2\sigma/q$, the framing term in the Yoshida normalization used in [40, equation (5.9)]. The quadratic and linear terms of the surgery action are therefore fixed by a symplectic change of coordinates.

The complementary core holonomy b is conjugate to the filling holonomy a. Along $a = i\theta$,

$$\mathrm{d}\Im\widetilde{\Phi} = \Im\left(\frac{1}{2}b\,i\,\mathrm{d}\theta\right) = \frac{1}{2}\Re b\,\mathrm{d}\theta = -\frac{1}{2}\ell_{\mathrm{core}}\,\mathrm{d}\theta.$$

This gives Schlaefli and the real source $j = \ell_{\mathrm{core}}/2$ in the canonical coordinates.

On the canonical relation, $G = (ab - uv)/4$, so

$$\boxed{\widetilde{\Phi} - \frac{ab}{4} = \Phi - \frac{uv}{4}.} \tag{16.39}$$

Equivalently $\widetilde{\Phi}_{\mathrm{NZ}} - ab = \Phi_{\mathrm{NZ}} - uv$, where $\widetilde{\Phi}_{\mathrm{NZ}} = 4(\widetilde{\Phi} - \Phi_0)$. The equality is local with consistently fixed logarithmic branches and additive constants. It explains why the combination of potential and peripheral product occurs in the complex-volume formula. The differential of the boxed expression is the invariant one-form $\frac{1}{4}(v\ \mathrm{d}u - u\ \mathrm{d}v)$.

16.10.4 Covariance of the filling coefficients

The powers of $p + q\tau$ in (16.17) also have a canonical-coordinate interpretation. Write the same geometric Lagrangian in the new coordinates as

$$b(a) = \widetilde{\tau}a + \widetilde{c}_3 a^3 + \widetilde{c}_5 a^5 + O(a^7).$$

Using the inverse series of $a = pu + qv(u)$ gives, with $A = p + q\tau$,

$$\widetilde{\tau} = \frac{\rho + \sigma\tau}{p + q\tau}, \qquad \Im\widetilde{\tau} = \frac{\Im\tau}{|p + q\tau|^2},$$
$$\widetilde{c}_3 = \frac{c_3}{A^4}, \qquad \widetilde{c}_5 = \frac{c_5}{A^6} - \frac{3qc_3^2}{A^7}. \tag{16.40}$$

The cancellation in the cubic coefficient is

$$\sigma A - q(\rho + \sigma\tau) = 1.$$

It is exactly the symplectic determinant condition. Let $\widetilde{F}(0) = 0$ and $\widetilde{F}' = b$. Since $\Re a = 0$ on the cone locus, the real boundary correction in the new coordinates vanishes. Therefore

$$V_{p,q}(\theta) - V_0 = \frac{1}{2}\Im\widetilde{F}(i\theta) = -\frac{\theta^2}{4}\Im\widetilde{\tau} + \frac{\theta^4}{8}\Im\widetilde{c}_3 - \frac{\theta^6}{12}\Im\widetilde{c}_5 + \cdots. \tag{16.41}$$

This is the previously computed volume expansion in filling coordinates. The cubic coefficient has a weight--4 covariance law under a peripheral basis change; this alone does not assert that it is a modular form as a function of cusp shape. The pairing fixes the transformation laws, while the actual values of $c_3, c_5, \ldots$ describe nonlinear geometry of $\mathcal{L}_M$.

16.10.5 An exact Hessian and Monge–Ampere identity

There is a useful strengthening of (16.8). Eliminate internal angles at fixed peripheral angular holonomies, writing

$$\overline{V}(y, s) = \max_{P\xi = (y,s)} \mathscr{V}(\xi).$$

All statements are local in the positive geometric neighborhood, where the maximum is interior and nondegenerate. Let $K(\xi) = -D^2\mathscr{V}(\xi) > 0$ at the maximizing point. The boundary intersection pairing gives $\operatorname{rank} P = 2$, so the constrained implicit-function theorem supplies a unique nearby interior maximizer for every sufficiently small (y, s). The Schur-complement identity for the linear constraints is

$$-D^2\overline{V} = (PK(\xi)^{-1}P^T)^{-1}. \tag{16.42}$$

Set $T = v'(u) = A_0 + iB_0$, with $B_0 > 0$ near the complete point. The volume differential and Cauchy–Riemann equations give

$$\overline{V}_y = \frac{r}{2}, \qquad \overline{V}_s = -\frac{x}{2}, \qquad \mathrm{d}s = B_0\,\mathrm{d}x + A_0\,\mathrm{d}y, \qquad \mathrm{d}r = A_0\,\mathrm{d}x - B_0\,\mathrm{d}y.$$

Solving these two linear differential equations yields

$$\boxed{-D^2\overline{V} = \frac{1}{2B_0}\begin{pmatrix} |T|^2 & -A_0 \\ -A_0 & 1 \end{pmatrix}, \qquad \det(-D^2\overline{V}) = \frac{1}{4}.} \tag{16.43}$$

Equivalently,

$$PK(\xi)^{-1}P^T = \frac{2}{\Im v'(u)}\begin{pmatrix} 1 & \Re v'(u) \\ \Re v'(u) & |v'(u)|^2 \end{pmatrix}. \tag{16.44}$$

The determinant-4 identity holds throughout this local branch. At $u = 0$ it reduces to the cusp formula already derived. For a filling path it gives the exact response

$$V''_{p,q}(\theta) = -\frac{\Im v'(u_\theta)}{2|p + qv'(u_\theta)|^2}, \tag{16.45}$$

in agreement with (16.12).

For a geometric interpretation of the positive Hessian metric $g = -D^2\overline{V}$, multiplication by i in the u-coordinate induces, in coordinates (y, s),

$$J_T = \frac{1}{B_0}\begin{pmatrix} -A_0 & 1 \\ -|T|^2 & A_0 \end{pmatrix}, \qquad J_T^2 = -I.$$

Then

$$\omega_{\mathrm{ang}} = -\frac{1}{2}\,\mathrm{d}y \wedge \mathrm{d}s, \qquad g(X, Y) = \omega_{\mathrm{ang}}(X, J_T Y), \qquad g = \frac{B_0}{2}(\mathrm{d}x^2 + \mathrm{d}y^2). \tag{16.46}$$

This real area form comes from the angular polarization and the positive complex tangent. It is not the pullback of Ω_∂, which vanishes on the geometric Lagrangian. Nor has it been identified with the earlier Thurston–Deligne–Mostow complex hyperbolic metric.

For completeness, with h cusps let $T = (\partial v_i/\partial u_j) = A_0 + iB_0$. It is symmetric, and $B_0 > 0$ near the complete point. In coordinates $(y, s) = (\Im u, \Im v)$ the same calculation gives

$$-D^2\overline{V} = \frac{1}{2}\begin{pmatrix} B_0 + A_0B_0^{-1}A_0 & -A_0B_0^{-1} \\ -B_0^{-1}A_0 & B_0^{-1} \end{pmatrix}, \qquad \det(-D^2\overline{V}) = 4^{-h}. \tag{16.47}$$

At the complete structure T is diagonal, with the individual cusp shapes on the diagonal. Coupling between distinct cusps appears in higher derivatives of the even Neumann–Zagier potential, for example through terms $u_i^2 u_j^2$.

16.10.6 The quantum Gaussian as a half-density transformation

Let a geometric boundary state have leading form $A_0(u)e^{\Phi(u)/\hbar}$, with $\Phi' = v/2$. A quantization of the canonical transformation uses the exponential kernel $e^{G(u,a)/\hbar}$ and its mixed-derivative normalization. The relevant derivatives are

$$-G_{ua} = \frac{1}{2q}, \qquad (\Phi + G)_{uu} = \frac{p + qv'(u)}{2q}.$$

The kernel normalization divided by the stationary Gaussian therefore has the dependence

$$\boxed{\sqrt{\frac{-G_{ua}}{(\Phi + G)_{uu}}} = \frac{1}{\sqrt{p + qv'(u)}} = \sqrt{\frac{\mathrm{d}u}{\mathrm{d}a}}.} \tag{16.48}$$

Universal powers of $\hbar$, phases, and square-root branches are fixed by the chosen integral convention. The displayed dependence is the transformation law for a half-density along $\mathcal{L}_M$. It is exactly the law obeyed by the inverse square root of peripheral torsion in (16.25). The extra $2\sinh(b/2)$ factor comes from the solid torus and is additional input. Explicit quantum filling kernels appear in [39, equations (12)–(14)].

The symplectic structure fixes the conjugate variables, the canonical generating function, and this Gaussian Jacobian. It does not select the integration cycle or prove that the geometric contribution survives cancellation in the exact quantum invariant. The analytic hypotheses in Proposition 16.2 remain necessary for that conclusion.

Part IV
Curves, trace formulas, and quotients

Cutting and gluing geodesics produces identities and twist integrals. Twisted trace formulas retain the multicurve coefficients and their lengths.

17 Unfolding and the problem of finding constant series

17.1 What a volume computation determines

Let Γ be the mapping class group and let γ be a fixed curve or multicurve class. For an appropriate nonnegative or absolutely integrable function, the basic unfolding identity is

$$\int_{\Gamma\backslash\mathcal{T}} \sum_{\delta\in\Gamma\gamma} F(\ell_\delta(X))\, \mathrm{d}V_{\mathrm{WP}}(X) = \int_{\mathrm{Stab}(\gamma)\backslash\mathcal{T}} F(\ell_\gamma(X))\, \mathrm{d}V_{\mathrm{WP}}(X). \tag{17.1}$$

Multiplicity conventions belong in the stabilizer. This identity is just integration over coset translates of a fundamental domain; it does not require the curves in the orbit to be simple.

Simplicity becomes important when evaluating the right side. If γ is a disjoint simple multicurve, cut along it and use Fenchel–Nielsen coordinates. Each twist period integrates to ℓ_j, so the resulting integral has factors $\prod_j \ell_j\, \mathrm{d}\ell_j$ times WP volumes of the components, with the relevant symmetry factors. This is the geometric integration mechanism in Mirzakhani's volume recursion [73, Theorem 8.1].

Suppose a series $S(X)$ is independently known to be constant, say C. Then

$$C = \frac{1}{V_{g,n}} \int_{\mathcal{M}_{g,n}} S(X)\, \mathrm{d}V_{\mathrm{WP}}(X), \tag{17.2}$$

and unfolding can calculate its value. If constancy has not been proved, the same quotient computes only its average. This is the precise logical distinction behind the proposal to recover an identity from a volume.

On the punctured torus, for an unoriented essential simple curve orbit, the stack-normalized formula is

$$\int_{\mathcal{M}_{1,1}} \sum_{\alpha \text{ simple}} f(\ell_\alpha)\, \mathrm{d}V_{\mathrm{WP}} = \frac{1}{2}\int_0^\infty x f(x)\, \mathrm{d}x. \tag{17.3}$$

For $f(x) = (1+e^x)^{-1}$, expand in e^{-x} to obtain

$$\int_0^\infty \frac{x\, \mathrm{d}x}{1+e^x} = \sum_{m\ge 1} \frac{(-1)^{m-1}}{m^2} = \frac{\pi^2}{12}.$$

The integral of the series is $\pi^2/24$; division by $V_{1,1} = \pi^2/12$ gives $C = 1/2$. Its pointwise constancy is McShane's theorem [71]:

$$\sum_{\alpha \text{ simple, unoriented}} \frac{1}{1+e^{\ell_\alpha(X)}} = \frac{1}{2}. \tag{17.4}$$

The tetrahedral volume calculation and the covering argument provide an independent route to the volume used in this normalization check.

17.2 Operations that preserve constancy

There is indeed a large closure of any known collection of identities. Finite linear combinations preserve constancy. So do absolutely convergent parameter integrals and locally uniformly convergent differentiations of identities on an identified parameter family. Products of identities give constant double series when Tonelli's theorem or absolute convergence applies. These are useful mechanisms for constructing new kernels.

Their geometric interpretation needs to be retained. A product usually indexes pairs of curves that may intersect; it is not automatically a disjoint-multicurve sum. A parameter derivative also differentiates lengths if the metric varies with that parameter. In particular, differentiating a constant series with respect to the metric gives the one-form identity

$$\sum_{\gamma} F'(\ell_\gamma)\, \mathrm{d}\ell_\gamma = 0, \tag{17.5}$$

assuming sufficient local uniform control. It does not generally give another scalar identity by differentiating only the displayed weight.

Two possible tests for a proposed constant series follow. One can try to show that (17.5) vanishes on a spanning set of tangent directions. Alternatively, if $S \in L^2$ and its mean is $\overline{S}$, then

$$\int_{\mathcal{M}} (S - \overline{S})^2 \, \mathrm{d}V_{\mathrm{WP}} = 0 \tag{17.6}$$

proves constancy almost everywhere, and everywhere if S is continuous. The second-moment calculation involves pairs of curves, including their intersection patterns. It can be substantially harder than the first unfolding integral.

Thus a useful search for further identities should ask for more than formal closure: a natural index set, a manageable positive weight, a geometric partition or telescoping mechanism, and a convergence proof. The constructions in Sections 18 and 20 illustrate these requirements.

17.3 Other established constant series

There are also constant series with substantially different kernels. For example, write

$$\mathcal{L}(x) = \mathrm{Li}_2(x) + \tfrac{1}{2} \log x \log(1 - x)$$

for the Rogers dilogarithm on $(0, 1)$, extended continuously to its endpoints, and define

$$\begin{aligned} D(r) = {} & \mathcal{L}\big(\operatorname{sech}^2(r/2)\big) \\ & + 2\mathcal{L}\left(\frac{1 + e^{-r}}{2}\right) - 2\mathcal{L}\left(\frac{1 - e^{-r}}{2}\right). \end{aligned}$$

Hu–Tan's punctured-torus identity is

$$\sum_{\alpha \text{ simple, unoriented}} D(\ell_\alpha) = \frac{\pi^2}{2}. \tag{17.7}$$

The same paper treats a positive torus boundary [62, Theorems 1–2].

For a closed surface of genus $g \geq 2$, Luo–Tan give explicit dilogarithmic weights $f(P), g(T)$ on embedded geometric pairs of pants and one-holed tori such that

$$\sum_{P} f(P) + \sum_{T} g(T) = 8\pi^2(g - 1) = \mathrm{Vol}(T^1 X). \tag{17.8}$$

Their construction partitions unit tangent directions [68]. It illustrates another source of constant series: the invariant measure being partitioned need not live only on a horocycle or a geodesic boundary. The precise, longer subsurface weights are given in that reference.

18 Figure-eight curves and simple lifts

In this section surfaces have cusps or geodesic boundary and no cone singularities.

18.1 The pants trace relation

Let γ be a closed geodesic with exactly one transverse self-intersection, not a proper iterate. Its two-lobed neighborhood is a pair of pants. After tightening the cuffs, allow coincident cuff images in the ambient surface; this occurs on a one-holed torus. If the three cuff lengths are a, b, c, write

$$t_a = 2\cosh(a/2), \quad t_b = 2\cosh(b/2), \quad t_c = 2\cosh(c/2).$$

A cusp has positive trace 2. For the figure-eight that wraps the a and b cuffs, the trace identity is

$$\boxed{t_\gamma = t_a t_b + t_c.} \tag{18.1}$$

This is Rivin's pants relation [88, Theorem 1.3]. It also follows directly from $\operatorname{tr}(AB^{-1}) = \operatorname{tr} A \operatorname{tr} B - \operatorname{tr}(AB)$ with the consistent signs of pants holonomies; rewriting in positive traces produces the plus sign in (18.1).

For pants with lengths $(a, b, 0)$, the two figures that use the cusp as one lobe have traces

$$u = 2t_a + t_b, \qquad v = t_a + 2t_b, \qquad t_a = \frac{2u - v}{3}, \quad t_b = \frac{2v - u}{3}. \tag{18.2}$$

Consequently any weight assigned to this pair of cuffs can be rewritten as a weight in the two associated figure-eight lengths. This gives a direct reindexing of a pants identity. It produces a paired-curve weight; it does not automatically produce a universal function of one figure-eight length.

18.2 What the local double cover accomplishes

Use pants generators a, b, with third cuff $c = (ab)^{-1}$ and figure $\gamma = ab^{-1}$. Consider the connected double cover given by

$$\chi(a) = \chi(b) = 1 \in \mathbb{Z}/2\mathbb{Z}.$$

Then $\chi(c) = \chi(\gamma) = 0$. The a, b cuffs each lift with degree two, while c has two degree-one lifts. The upstairs surface is a four-holed sphere with cuff lengths

$$(2a, 2b, c, c). \tag{18.3}$$

The figure-eight has two closed degree-one lifts, each of which is simple. One way to see this is to lift its two-loop spine: the lifted graph has two vertices, and either lift uses one a edge and one b edge to form a simple cycle.

The two lifted cycles meet at both vertices of that graph. Geometrically, the two simple lifts intersect each other twice, once over each lift of the original crossing. Thus two components of the lifted domain do not mean a disjoint multicurve upstairs.

There is a further global issue: this particular character of the pants group need not extend to the whole surface. A lobe that is zero in $H_1(S; \mathbb{Z}/2\mathbb{Z})$—for example, a separating lobe on a closed surface— cannot be assigned value one by an ambient character. Rivin proves a global substitute: a one-double-point geodesic has a simple lift in a fourfold cover [88, Theorem 4.2].

Even when a suitable cover exists, applying an entire simple-curve identity upstairs and projecting downstairs introduces other projected topological types, covering degrees, and multiplicities. Desingularization supplies an important geometric tool, but selecting just the desired downstairs curves requires additional accounting.

18.3 The two types on a punctured torus

On a once-punctured torus, a figure-eight has a unique essential simple curve α disjoint from it. Its pants neighborhood has one peripheral cuff and two parallel essential cuffs, which both tighten to α. Cutting along α gives pants of lengths $(a, a, 0)$, where $a = \ell_\alpha$ and $t = 2\cosh(a/2) > 2$.

There are exactly three unoriented figure-eight classes in these cut pants:

Lobes	Number	Positive trace
One essential cuff and the cusp	2	$3t$
Both essential cuff copies	1	$t^2 + 2$

The first two are exchanged by the hyperelliptic involution; call their family $\mathcal{N}$. The third is hyperelliptic; call its family $\mathcal{H}$. Equal lengths in the first row do not identify the two free-homotopy classes. In torus homology the first family has a primitive class up to sign and the second twice a primitive class, although its geodesics are not proper iterates.

For a word check, glue pants generators by $b = sa^{-1}s^{-1}$. The first two figures can be represented by a^2b and ab^2, while the third is $ab^{-1} = asas^{-1}$. The last word is not a square. The common disjoint class α is recoverable from the figure's pants neighborhood, so grouping all figure-eights by α has no duplication.

Lei–Zhang correct the index set in Rivin's earlier torus identity: its stated formula accounts for $\mathcal{N}$, while $\mathcal{H}$ requires separate treatment [67]. The separation is important here because the next construction is intended to include *both* families.

18.4 A positive single-length kernel for all figure-eights

Define, for $t > 2$,

$$h(t) = 1 - \sqrt{1 - \frac{4}{t^2}}. \tag{18.4}$$

If $t = 2\cosh(\ell/2)$, then $h(t) = 2/(1 + e^\ell)$. Hence McShane's identity is $\sum_\alpha h(t_\alpha) = 1$. The classification immediately gives the two separate identities

$$\sum_{\gamma \in \mathcal{N}} \left(1 - \sqrt{1 - \frac{36}{t_\gamma^2}}\right) = 2, \tag{18.5}$$

$$\sum_{\gamma \in \mathcal{H}} \left(1 - \sqrt{1 - \frac{4}{t_\gamma - 2}}\right) = 1. \tag{18.6}$$

To use the same kernel on every one-crossing geodesic, let

$$\beta(x) = \frac{x^2}{9} + 2, \qquad F(x) = \sum_{n=0}^{\infty} (-1)^n 2^{-n-1} h\left(\frac{\beta^{\circ n}(x)}{3}\right), \qquad x > 6. \tag{18.7}$$

Here $\beta^{\circ n}$ means iteration, with $\beta^{\circ 0}(x) = x$.

Proposition 18.1. *On a complete finite-area once-punctured hyperbolic torus,*

$$\boxed{\sum_{\substack{\gamma \text{ unoriented, not iterated} \\ i(\gamma,\gamma)=1}} F\big(2\cosh(\ell_\gamma/2)\big) = 1.} \tag{18.8}$$

The kernel is positive and satisfies

$$\frac{1}{4}h(x/3) \le F(x) \le \frac{1}{2}h(x/3), \qquad F(x) \sim \frac{9}{x^2} \quad (x \longrightarrow \infty). \tag{18.9}$$

Proof. For $x > 6$, $\beta(x) - x = (x-3)(x-6)/9 > 0$. Since h is positive and decreasing, the absolute values of the summands in (18.7) decrease, and the series is absolutely convergent, bounded termwise by 2^{-n-1}. Its alternating-series bounds place F between the first term and the first term minus the second. The second is at most half the first, proving (18.9)'s inequalities.

Shifting the series index yields

$$2F(x) + F(\beta(x)) = h(x/3).$$

For $x = 3t$ this reads

$$2F(3t) + F(t^2 + 2) = h(t). \tag{18.10}$$

The three figure-eights associated with a simple α have exactly the traces on the left. Sum over α and use McShane's identity. Positivity justifies the regrouping and proves convergence of the full geodesic sum. Finally $h(x/3) = 18x^{-2} + O(x^{-4})$ and the second and subsequent terms contribute $O(x^{-4})$, giving the stated asymptotic. □

This is an elementary construction from the corrected classification, not a claim of a new identity in the literature. It is restricted to the cusped torus. Cone continuation requires checking existence of the indexed geodesics as well as continuing the weight; the cone obstruction discussed by Lei–Zhang is a reason to keep that distinction explicit [67].

19 Cutting into pants and varying twists

19.1 What data a cut curve retains

Fix a pants decomposition with cuff lengths $a_1, \dots, a_N$, where $N = 3g - 3 + n$. Cutting a simple multicurve gives disjoint essential arcs in each pair of pants. Cutting a general immersed curve, even a filling one, need not give disjoint arcs: self-intersections may remain inside a pair of pants. Nor are the actual segments generally orthogonal to the cuffs.

Boundary intersection counts alone lose the relative homotopy classes of the arcs, their cyclic joining pattern, and winding information. For simple multicurves this information is organized by Dehn–Thurston coordinates. For a general immersed curve one must retain a fuller itinerary. Once this topological data is fixed, varying the metric twist is a well-defined and tractable problem.

There are two complementary exact descriptions. One uses arc lengths as generating functions in endpoint positions. The other uses products of holonomy matrices. Both keep the information that intersection counts alone discard.

19.2 An exact local distance function

Take two ultraparallel geodesic lifts of cuffs in $\mathbb{H}^2$, at common perpendicular distance d. Let u, v be signed distances on the two lines from the perpendicular feet, with compatible orientations. Then

$$\boxed{D_d(u,v) = \operatorname{arcosh}\bigl(\cosh d \cosh u \cosh v - \sinh u \sinh v\bigr).} \tag{19.1}$$

For example, in the hyperboloid model use

$$p(u) = (\cosh u, \sinh u, 0), \qquad q(v) = (\cosh d \cosh v, \sinh v, \sinh d \cosh v).$$

Then $\cosh D = -\langle p, q \rangle$. Reversing one endpoint coordinate changes the sign of the mixed term.

For the ordinary seam between cuffs a, b of pants with third cuff c,

$$\cosh d = \frac{\cosh(c/2) + \cosh(a/2)\cosh(b/2)}{\sinh(a/2)\sinh(b/2)}. \tag{19.2}$$

For a returning or winding arc, choose the boundary lifts appropriate to its relative homotopy class and use their common perpendicular instead. There is no single seam length that describes all relative arc classes.

Given a fixed itinerary, cuff gluing imposes affine relations on the endpoint variables involving the twists τ_i and integer winding data. The length of its closed geodesic representative is obtained by minimizing the sum of the corresponding functions (19.1) over the crossing positions. The first-variation equations impose smooth tangent matching at every crossing. This provides an exact formulation even when the resulting minimization is not elementary to evaluate.

19.3 Holonomy and first variation

In a basis adapted to a cuff axis, a twist displacement is represented by

$$T_i(\tau_i) = \operatorname{diag}(e^{\tau_i/2}, e^{-\tau_i/2}).$$

For a fixed curve word γ, insert these matrices into the fixed pants holonomies. The resulting product W_γ satisfies

$$2\cosh(\ell_\gamma/2) = \left|\operatorname{tr} W_\gamma(\mathbf{a}, T_1(\tau_1), \dots, T_N(\tau_N))\right|. \tag{19.3}$$

With the cuff lengths fixed, its trace is a finite Laurent polynomial in $e^{\tau_i/2}$. Its coefficients need not all be positive in an arbitrary choice of lifts and coordinates.

The twist derivative has the cosine formula

$$\frac{\partial \ell_\gamma}{\partial \tau_i} = \sum_{p \in \gamma \cap \alpha_i} \cos \vartheta_p, \qquad \left|\frac{\partial \ell_\gamma}{\partial \tau_i}\right| \le i(\gamma, \alpha_i). \tag{19.4}$$

Angles and twist orientation are chosen consistently. Length is strictly convex along an individual nontrivial twist when the curve crosses the twisting cuff; the argument also applies to nonsimple geodesics [63]. We do not need a stronger claim about arbitrary lines with several twist parameters varying in different signs.

These ideas overlap with several parts of Mirzakhani's work, but serve different roles there. Dehn–Thurston estimates compare combinatorial and geometric length for simple curves [75, Section 3]. The work on general curve orbits studies asymptotically piecewise-linear length, including filling curves and length sublevel volumes [76, Sections 4 and 8]. In the basic volume recursion, the cutting curves are simple and the kernel depends on their lengths, so the twist integral itself is just the period a_i.

19.4 Unfolding a twist lattice

Let D_i be the Dehn twist about the ith cuff. With a consistent sign convention,

$$\ell_{D_1^{k_1}\cdots D_N^{k_N}\gamma}(\mathbf{a}, \boldsymbol{\tau}) = \ell_\gamma(\mathbf{a}, \boldsymbol{\tau} + \mathbf{k}\mathbf{a}).$$

For a nonnegative function F, Tonelli's theorem gives

$$\boxed{\int_{\prod_i [0,a_i)} \sum_{\mathbf{k}\in\mathbb{Z}^N} F\big(\ell_{D_1^{k_1}\cdots D_N^{k_N}\gamma}(\mathbf{a}, \boldsymbol{\tau})\big)\, \mathrm{d}\boldsymbol{\tau} = \int_{\mathbb{R}^N} F\big(\ell_\gamma(\mathbf{a}, \mathbf{u})\big)\, \mathrm{d}\mathbf{u}.} \tag{19.5}$$

For signed functions, impose absolute integrability. If the curve has a nontrivial stabilizer in this twist lattice, remove the redundant directions or account for their multiplicity; in particular, twisting about a disjoint cuff should not produce infinitely many copies of the same term.

A fixed curve's length is not periodic in a twist coordinate. The sum over its twist orbit is periodic, which is exactly what makes this unfolding work. The accompanying WP measure is $\prod_i \mathrm{d}a_i\ \mathrm{d}\tau_i$. However, the positive cuff orthant times one twist box is the quotient by this particular twist subgroup, not a fundamental domain for the full mapping class group. A further full-moduli integration needs the appropriate orbit unfolding.

For example, with weight $(2\cosh(\ell_\gamma/2))^{-s}$ and $x_i = e^{u_i/2}$, the right side becomes the Mellin-type integral

$$2^N \int_{(0,\infty)^N} |\operatorname{tr} W_\gamma(\mathbf{a},\mathbf{x})|^{-s} \prod_i \frac{\mathrm{d}x_i}{x_i}, \tag{19.6}$$

whenever it converges. This is an explicit analytic object attached to the curve itinerary and the fixed pants lengths.

20 An exact torus twist calculation and a constant orbit series

20.1 The length function and its integral

Let X be a one-holed torus with boundary length $b \geq 0$. Cut along a simple curve α of length $a > 0$, and choose a curve β crossing α once. Normalize the twist origin so that this crossing curve has minimum length at $\tau = 0$. Set

$$A^2 = \frac{\cosh a + \cosh(b/2)}{\cosh a - 1}. \tag{20.1}$$

The exact formula is

$$\boxed{\cosh\bigl(\ell_\beta(\tau)/2\bigr) = A\cosh(\tau/2).} \tag{20.2}$$

It follows from the right-angled-hexagon description of the cut pants; see also [43, Proposition 3.1]. For $b = 0$, $A = \coth(a/2)$. The minimum length is $r_0 = 2\operatorname{arcosh} A$.

For any nonnegative measurable function f, change variables separately on the positive and negative twist axes to obtain

$$\int_{\mathbb{R}} f(\ell_\beta(u))\ \mathrm{d}u = 2\int_{r_0}^{\infty} f(r)\frac{\sinh(r/2)}{\sqrt{\cosh^2(r/2) - A^2}}\ \mathrm{d}r. \tag{20.3}$$

Indeed, differentiating $\cosh(r/2) = A\cosh(u/2)$ gives $\mathrm{d}u/\,\mathrm{d}r = \sinh(r/2)/\sqrt{\cosh^2(r/2) - A^2}$ for $u > 0$. The two signs of u explain the factor 2.

Several useful checks are immediate:

$$\int_{\mathbb{R}} \operatorname{sech}^2\bigl(\ell_\beta(u)/2\bigr)\ \mathrm{d}u = \frac{4}{A^2}, \tag{20.4}$$

$$\operatorname{meas}\{u : \ell_\beta(u) \leq R\} = 4\operatorname{arcosh}\left(\frac{\cosh(R/2)}{A}\right), \quad R \geq r_0, \tag{20.5}$$

$$\int_{\mathbb{R}} \cosh^{-s}\bigl(\ell_\beta(u)/2\bigr)\ \mathrm{d}u = \frac{2\sqrt{\pi}\,\Gamma(s/2)}{A^s\Gamma((s+1)/2)}, \quad s > 0. \tag{20.6}$$

The sublevel measure is zero for $R < r_0$. These are exact twist integrals, not asymptotic estimates.

20.2 A positive partition of unity

Define

$$H_a(u) = \frac{\sinh(a/2)}{\cosh u + \cosh(a/2)} = \frac{1}{2}\left[\tanh\left(\frac{u + a/2}{2}\right) - \tanh\left(\frac{u - a/2}{2}\right)\right]. \tag{20.7}$$

The second expression telescopes under translation by a, so

$$\sum_{n\in\mathbb{Z}} H_a(\tau + na) = 1. \tag{20.8}$$

All terms are positive and the convergence is exponential on compact twist intervals.

Since (20.2) implies $\cosh r + 1 = A^2(\cosh u + 1)$, eliminate u from (20.7) and set

$$K_{a,b}(r) = \frac{A^2 \sinh(a/2)}{\cosh r + 1 + A^2(\cosh(a/2) - 1)}. \tag{20.9}$$

For $r_n = \ell_{D_\alpha^n \beta}(X)$ we obtain the pointwise identity

$$\boxed{\sum_{n\in\mathbb{Z}} K_{a,b}(r_n) = 1.} \tag{20.10}$$

This is a constant series proved by telescoping, rather than inferred from its average. Its kernel depends on the distinguished cuff length a and the boundary length b, and its index set is one twist orbit. It is not an unrestricted universal sum over every curve on the surface.

There is a simple source of many such examples. If ρ is an even probability density on $\mathbb{R}$, then

$$H_{a,\rho}(u) = \int_{u-a/2}^{u+a/2} \rho(v)\, \mathrm{d}v \quad \text{satisfies} \quad \sum_{n\in\mathbb{Z}} H_{a,\rho}(u + na) = 1.$$

The translated intervals partition the line. Evenness makes $H_{a,\rho}$ a function of $|u|$, hence of the length r through (20.2). The choice $\rho(v) = \frac{1}{4}\operatorname{sech}^2(v/2)$ gives (20.7). This realizes the proposed "zoo" of identities in a concrete way, with the source of constancy and the parameter dependence both visible.

21 Matching across cuffs: positions and directions

The pants viewpoint suggests that orthogeodesic pieces should glue when their boundary data match. There are two levels of matching.

First consider a cuff of length a, with endpoint configurations $x_1, \dots, x_m$ and $y_1, \dots, y_m$ on the two sides. Choose compatible coordinates in $\mathbb{R}/a\mathbb{Z}$, reversing one induced boundary orientation as needed. For a prescribed cyclic matching σ, a twist τ must satisfy

$$y_{\sigma(j)} - x_j \equiv \tau \pmod{a} \qquad (1 \le j \le m). \tag{21.1}$$

One endpoint determines τ modulo a; the remaining $m-1$ equations require the cyclic spacings to agree. Integer winding data must also be retained if the curve is marked, rather than only considered modulo a twist.

Second, a general geodesic segment has a tangential velocity component p at an endpoint. If S is the sum of segment lengths, the first variation at a glued endpoint has the form

$$\frac{\partial S}{\partial x} = p_{\text{in}} - p_{\text{out}}. \tag{21.2}$$

Stationarity imposes tangent matching as well as position matching. For orthogonal segments the tangential components vanish, so smooth gluing is automatic once the endpoints match. A generic closed geodesic does not satisfy this orthogonality condition along an arbitrarily chosen pants decomposition.

Equivalently, boundary data live in a phase space (s, p): endpoint position and tangential component. The length of an arc is a generating function, with

$$\mathrm{d}L = -p_s\ \mathrm{d}s + p_t\ \mathrm{d}t, \qquad \text{and cuff gluing acts by } (s,p) \longmapsto (s+\tau, p).$$

A closed geodesic is a periodic compatible chain of these relations. The orthospectrum records the special $p = 0$ representatives. It is excellent input for gap identities and local geometry, but endpoint positions and directions are needed to reconstruct arbitrary glued curves.

22 Cusps, McShane gaps, and the orthospectrum

22.1 Which geodesics enter a cusp?

The statements that simple closed geodesics avoid a sufficiently deep cusp and that simple cusp-ended geodesics enter the cusp are compatible. The first concerns a returning excursion; the second concerns an actual end.

Here is an elementary version of the distinction. Normalize a cusp stabilizer in the upper half-plane to $z \mapsto z + 1$. A geodesic with two finite ideal endpoints $u < v$ is a semicircle of radius $(v - u)/2$. If $v - u > 1$, its endpoints interlace those of its translate:

$$u < u + 1 < v < v + 1.$$

The two geodesics intersect, so the projected geodesic is not simple. Therefore a simple geodesic with two finite endpoints cannot reach height greater than $1/2$. A geodesic reaching farther into this lifted cusp while remaining simple must have ∞ as an endpoint, and is vertical. It is perpendicular to every horocycle based at that endpoint.

Choose Y large enough that the horocycle $H = \{y = Y\}/\langle z \mapsto z + 1\rangle$ bounds an embedded sufficiently deep cusp. Simple closed geodesics never cross H. Complete simple geodesics with a cusp end do cross it, normally, and continue indefinitely toward that end.

22.2 The family used in the gap argument

At a point $x \in H$, take the normal geodesic directed away from the cusp. A lift is

$$\widetilde{\gamma}_x(t) = x + iYe^{-t}. \tag{22.1}$$

Its backward end tends to the cusp. Its forward continuation enters the rest of the surface. The relevant question is whether that whole continuation remains simple; the short portion inside the cusp is automatically simple.

McShane's gap argument partitions the horocycle using the behavior of these cusp-issued geodesics. Almost every direction eventually develops the relevant intersection pattern. The exceptional simple directions have measure zero and include cusp-to-cusp arcs and rays spiraling toward closed geodesics or minimal laminations [71, 72]. In some conventions a gap contains an isolated simple cusp-to-cusp direction; its interior is therefore not literally made entirely of nonsimple directions. The measure statement is unaffected.

For a cusp pants contribution whose two other cuffs are interior, with lengths a, b (allowing their identification on the torus), the combined normalized gap is

$$\frac{2}{1+e^{(a+b)/2}}. \tag{22.2}$$

On the punctured torus the two cuffs are copies of one simple curve α, so $a = b = \ell_\alpha$. The gap contributions sum to one, which is precisely (17.4). The simple closed curve occurring in the formula labels the pants and their gap. It is not the geodesic crossing the deep horocycle.

22.3 An exact truncated cusp-arc formula

Fix a sufficiently small cusp horocycle of length $H > 0$; here H now denotes its length rather than the horocycle as a set. In pants with other cuffs a, b, let η be the simple cusp-to-cusp geodesic arc separating those cuffs. Truncate its two ends at the horocycle and let d_η be the resulting length. With the convention $\lambda_\eta = e^{d_\eta/2}$,

$$\boxed{e^{d_\eta/2} = \frac{t_a + t_b}{H}, \qquad t_a = 2\cosh(a/2), \quad t_b = 2\cosh(b/2).} \tag{22.3}$$

The formula also holds for signed truncated lengths before the horocycle has been made small enough.

Proof. Normalize the cusp parabolic and horocycle to

$$P = \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix}, \qquad y = Y, \qquad H = q/Y.$$

Choose cuff holonomy lifts A, B with $AB = P$, $\operatorname{tr} A = t_a$ and $\operatorname{tr} B = -t_b$, and write c_A for the lower-left entry of A. Then

$$\operatorname{tr}(A^{-1}P) = t_a - qc_A = -t_b, \qquad qc_A = t_a + t_b.$$

The image under A of the horizontal horocycle has Euclidean diameter $1/(c_A^2 Y)$ and is based at $A(\infty)$. The vertical ideal geodesic from ∞ to this point represents the separating return arc. Its signed truncated distance satisfies $e^{d_\eta/2} = |c_A|Y$. Substituting $H = q/Y$ proves the formula. □

On the once-punctured torus, essential unoriented simple cusp-to-cusp arc classes correspond to essential unoriented simple closed-curve classes: each is the unique such class disjoint from the other. Cutting along the closed curve gives $(a, a, 0)$ pants, hence

$$e^{d_\alpha/2} = \frac{2t_\alpha}{H} = \frac{4\cosh(\ell_\alpha/2)}{H}. \tag{22.4}$$

As a check, an ideal triangulation with lambda lengths u, v, w has horocycle length $H = 2(u^2+v^2+w^2)/(uvw)$; its trace associated with the curve disjoint from edge u is $Hu/2$, in agreement with (22.4); see [92, Section 12].

Substitute $t_\alpha = (H/2)e^{d_\alpha/2}$ into $\sum_\alpha h(t_\alpha) = 1$. This gives a literal orthospectral form of McShane's identity:

$$\boxed{\sum_{\eta \text{ simple, unoriented cusp-to-cusp}} \left(1 - \sqrt{1 - \frac{16e^{-d_\eta}}{H^2}}\right) = 1.} \tag{22.5}$$

The index set retains the topological arc labels. Scaling the horocycle length by k changes each truncated length to $d_\eta - 2\log k$; therefore e^{-d_η}/H^2 and every summand are decoration-independent.

22.4 The precise relation to Basmajian's identity

For a compact hyperbolic surface with nonempty geodesic boundary, Basmajian's boundary-length identity is, with oriented orthogeodesics,

$$\ell(\partial X) = \sum_{\eta \text{ oriented}} 2 \log \coth\left(\frac{\ell_\eta}{2}\right). \tag{22.6}$$

If each unoriented orthogeodesic is counted only once, the coefficient is 4 instead of 2. This spectrum includes nonsimple arcs. The gap comes from the shadow on a boundary lift of another boundary lift, and hyperbolic trigonometry expresses its length in terms of their common perpendicular.

Equation (22.5) uses the *simple, truncated cuspidal orthospectrum.* Generic normal rays in the McShane gaps are not the canonical returning orthogeodesics, and adding all nonsimple cusp arcs would change the identity. Thus the orthospectrum interpretation is exact after the index set, truncation, and weights have been specified.

Parlier's coherent-marking framework gives a systematic bridge between these boundary and cusp identities [82]. The common mechanism is a measure decomposition of boundary initial data, with orthogeodesic geometry computing gaps. What differs is the chosen stopping or marking rule and the resulting index set. This is the useful sense in which the pants, McShane, and Basmajian pictures fit together.

23 Simple multicurve coefficients of a twisted trace formula

This section develops a spectral version of the crossing-resolution idea. We work throughout with a marked, closed, oriented hyperbolic surface X of genus $g \geq 2$ and curvature -1. All closed curves are unoriented, unless conjugacy classes are explicitly being counted with orientation. A multicurve may have parallel components; its length and number of components count those multiplicities. The empty multicurve is included. Cusps and boundary require additional trace terms and are not included in the assertions below.

The inputs are the twisted Selberg formula, the simple-multicurve basis of the character ring, and Dylan Thurston's sign-coherence theorem for the bracelet basis. We prove the intervening length and coefficient estimates explicitly. The resulting construction is a deduction from these established inputs, with no novelty claim.

23.1 Broken geodesics and the effect of reconnection

Colin de Verdière's construction approximates loop space by broken geodesics and applies stationary phase to a complex heat parametrix [19]. For an equal time subdivision the finite-dimensional energy is

$$E_N(x_0, \dots, x_{N-1}) = N \sum_{i=0}^{N-1} d(x_i, x_{i+1})^2, \qquad x_N = x_0,$$

with consecutive points within the injectivity radius. The critical equations match incoming and outgoing velocities; the critical value for a constant-speed closed geodesic is ℓ_γ^2. On a hyperbolic surface every closed geodesic, including a self-intersecting one, has positive transverse length Hessian

$$\delta^2 L_\gamma(f, f) = \int_0^{\ell_\gamma} ((f')^2 + f^2)\, \mathrm{d}s.$$

Consequently simplicity cannot be selected by retaining only positive Hessians in the ordinary loop space.

Reconnection enlarges that space. In a small embedded disk of radius r about a transverse crossing of angle $\theta \in (0, \pi)$, the two original diameters have total length $4r$. The two noncrossing pairings have lengths

$$\begin{aligned} L_0 &= 4\,\mathrm{arsinh}\big(\sinh r \sin(\theta/2)\big), \\ L_\infty &= 4\,\mathrm{arsinh}\big(\sinh r \cos(\theta/2)\big). \end{aligned} \tag{23.1}$$

Both are strictly less than $4r$. In particular,

$$4r - L_0 = 4r(1 - \sin(\theta/2)) + O(r^3).$$

The descent is first order in the surgery radius. Tightening the resolved components to geodesics, and discarding contractible components, cannot increase total length. This is a change of connectivity, not a negative direction of the original Hessian. An unconstrained global minimization would select the empty cycle; we instead keep the different multicurves as algebraic observables.

23.2 The finite coefficient distributions

Let $\mathscr{C}(S)$ be the set of essential simple multicurves on the underlying marked surface. For $\rho : \pi_1(S) \to \mathrm{SU}(2)$ set

$$T_C(\rho) = \prod_{\alpha \subset C} \operatorname{tr} \rho(\alpha), \qquad T_\varnothing = 1. \tag{23.2}$$

The phrase “unnegated trace basis” refers to this sign convention, not to positivity of the function on $\mathrm{SU}(2)$. The functions T_C are linearly independent on the unitary character variety and form a basis of the complex character ring after complexification; see Charles–Marché [20]. The usual Kauffman convention uses $-\operatorname{tr}$ per component, so differs by a basis sign. The relation

$$\operatorname{tr} A \operatorname{tr} B = \operatorname{tr}(AB) + \operatorname{tr}(AB^{-1})$$

resolves crossings. For every nontrivial closed curve γ there is a unique finite expansion

$$\operatorname{tr}\rho(\gamma) = \sum_{C \in \mathscr{C}(S)} a_C(\gamma) T_C(\rho), \qquad a_C(\gamma) \in \mathbb{Z}. \tag{23.3}$$

These integers depend only on topology, not on X.

Let Δ_ρ be the rank-two flat unitary bundle Laplacian, and write $\lambda_j(\rho) = 1/4 + r_j(\rho)^2$. For an even $\varphi \in C_c^\infty(\mathbb{R})$ put

$$h(r) = \int_{\mathbb{R}} \varphi(L) e^{irL} \, \mathrm{d}L,$$

and define

$$\mathscr{T}_\varphi^X(\rho) = \sum_j h(r_j(\rho)) - \frac{2\operatorname{Area}(X)}{4\pi} \int_{\mathbb{R}} r h(r) \tanh(\pi r) \, \mathrm{d}r.$$

With the two orientations grouped, the twisted Selberg formula is

$$\mathscr{T}_\varphi^X(\rho) = \sum_{\delta \in \mathcal{P}_{\mathrm{un}}} \sum_{k \geq 1} \frac{\ell_\delta}{\sinh(k\ell_\delta/2)} \varphi(k\ell_\delta) \operatorname{tr}\rho(\delta^k). \tag{23.4}$$

Here $\mathcal{P}_{\mathrm{un}}$ consists of unoriented primitive closed geodesics. Naud’s statement [21, Theorem 2.2] sums over oriented primitive conjugacy classes with denominator $2\sinh$; grouping inverses gives (23.4). The compact length support makes the geometric sum finite.

Definition 23.1 (Coefficient distribution)**.** For $C \in \mathscr{C}(S)$ define the locally finite signed measure on $(0, \infty)$

$$\boxed{\sigma_C^X = \sum_{\delta \in \mathcal{P}_{\mathrm{un}}} \sum_{k \geq 1} \frac{\ell_\delta a_C(\delta^k)}{\sinh(k\ell_\delta/2)} \, \delta_{k\ell_\delta}.} \tag{23.5}$$

Thus

$$[T_C]\mathscr{T}_\varphi^X = \langle \sigma_C^X, \varphi \rangle.$$

The notation δ_L denotes the unit point mass at L.

Local finiteness follows from the finiteness of the closed-geodesic length set below any bound. The identity term in the trace formula is a constant function of ρ, hence affects only the empty multicurve coefficient. The input is the *labelled family* $\rho \mapsto \operatorname{Spec}\Delta_\rho$ on the marked character variety. A scalar spectrum alone does not retain these coefficients.

23.3 Length triangularity and the first atom

Lemma 23.2 (Length ordering of resolutions)**.** *If δ is primitive and $k \geq 1$, then*

$$a_C(\delta^k) \neq 0 \quad \Longrightarrow \quad \ell_X(C) \leq k\ell_X(\delta).$$

Equality can occur only when δ is simple and $C = k\delta$, the multicurve consisting of k parallel copies. Its coefficient in that case is 1. If δ is nonsimple, the inequality is strict.

Proof. For a primitive nonsimple geodesic, completely resolve its crossings. Each state has a representative obtained by reconnecting the original arcs in small disks. Formula (23.1) gives a strict total saving; tightening and removing contractible loops cannot increase length. At a multiple intersection, resolve the endpoints in a small convex disk by a noncrossing pairing. The original pairing uses intersecting diameters, so at least one new chord is shorter than a diameter. The same strict conclusion holds. Coefficient cancellation cannot introduce a state not produced by a resolution.

The product version is also useful: every multicurve in the expansion of $T_{C_1}\cdots T_{C_j}$ has length at most $\sum_i \ell_X(C_i)$. Superimpose representatives, separate parallel components, resolve crossings, and tighten.

For powers use the Chebyshev polynomials

$$P_0(x)=2,\quad P_1(x)=x,\quad P_{k+1}(x)=xP_k(x)-P_{k-1}(x),\qquad \operatorname{tr}(A^k)=P_k(\operatorname{tr}A).$$

If δ is simple, $P_k(T_\delta)$ contains only parallel copies $m\delta$ with $m\le k$, and its leading term is T_δ^k. If δ is nonsimple, every basis term in $\operatorname{tr}\rho(\delta)$ has length strictly below ℓ_δ. Apply the product inequality to each monomial in P_k: every nonconstant output has length strictly below $k\ell_\delta$; the constant term has length zero. □

Proposition 23.3 (First atom for a simple support)**.** *For a simple primitive curve α and $m\ge 1$,*

$$\boxed{\sigma^X_{m\alpha}=\frac{\ell_\alpha}{\sinh(m\ell_\alpha/2)}\delta_{m\ell_\alpha}+\text{a measure supported strictly above } m\ell_\alpha.} \tag{23.6}$$

For each fixed C the gap is locally uniform in X, and uniform on compact subsets of marked Teichmüller space. If C has two or more nonparallel support components, every contributing orbit has length strictly greater than $\ell_X(C)$; this statement does not assert that the resulting signed measure is nonzero.

Proof. Lemma 23.2 gives the lower support bound. At $m\ell_\alpha$ the only contributing pair is $(\delta,k)=(\alpha,m)$, with coefficient 1. This gives the stated mass, whose numerator is the *primitive* length. Every other contributing pair has strictly greater length. Local finiteness gives a gap at each X.

For uniformity on a compact set of marked metrics, use uniform bilipschitz comparison with one reference metric. The target lengths are bounded, and only finitely many orbit labels can have lengths within one unit above them somewhere on the compact set. The strictly positive differences for those finitely many labels have a positive minimum; all remaining labels have a gap at least one. The argument also applies to nonparallel C, without requiring a nonzero first mass after cancellation. □

For $m=1$ the first atom determines both the simple length and its normalization. Choose a fixed smooth cutoff χ, equal to one near ℓ_α and supported away from every other atom in a small neighborhood of X. Then

$$\boxed{\ell_X(\alpha)=\frac{\langle\sigma^X_\alpha,L\chi(L)\rangle}{\langle\sigma^X_\alpha,\chi(L)\rangle},\qquad \langle\sigma^X_\alpha,\chi\rangle=\frac{\ell_\alpha}{\sinh(\ell_\alpha/2)}.} \tag{23.7}$$

This is an exact localized identity, with no limiting interchange.

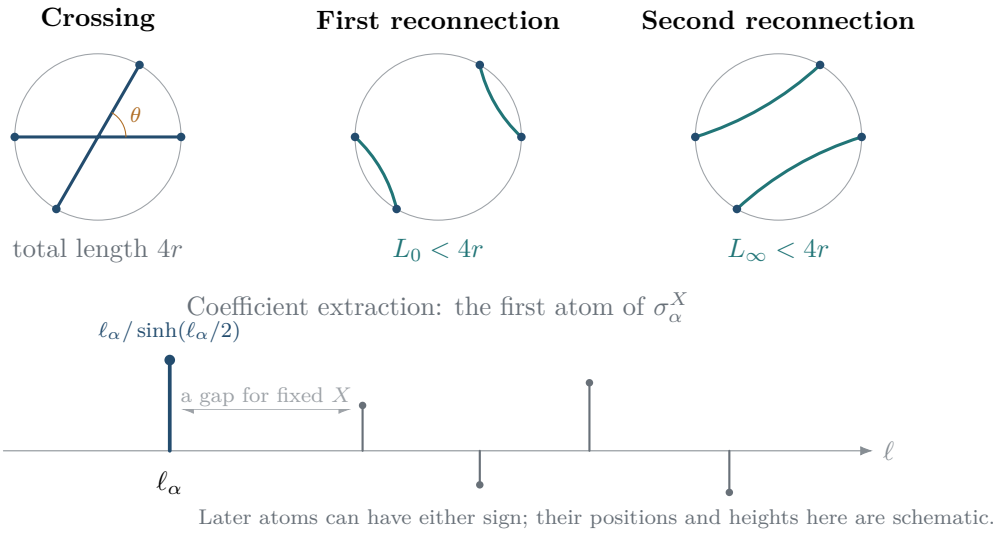


Figure 7: Reconnection and coefficient extraction are different operations. The upper pictures show the two noncrossing pairings in a hyperbolic metric disk; both shorten the original pair of diameters. Reconnection changes connectivity and is not a negative direction of the length Hessian within the original loop space. For a simple primitive support α, the lower schematic records the exact first atom of σ_α^X at ℓ_α, with the stated mass; all remaining atoms lie strictly above that length. Their depicted positions and weights are illustrative, not numerical data.

23.4 An exponential coefficient bound from bracelet positivity

The passage from compactly supported tests to heat kernels needs a growth estimate. A direct crossing count gives an exponential in ℓ_γ^2 and only a small-time conclusion. Bracelet positivity gives a stronger estimate.

At skein parameter $+1$, Dylan Thurston proves that every diagram has a sign-coherent expansion in the bracelet basis [22, Theorem 2]. A bracelet supported on a simple curve β with winding n is the Chebyshev element $P_n(z_\beta)$. Distinct bracelet support curves are disjoint. The canonical twisted Fuchsian evaluation associated to X is positive on bracelet basis elements:

$$B_D(X) = \prod_i 2\cosh(n_i \ell_X(\beta_i)/2) > 0.$$

For a one-component diagram representing γ, its evaluation has absolute value $2\cosh(\ell_X(\gamma)/2)$. Sign coherence therefore gives

$$\sum_D |b_D(\gamma)| B_D(X) = 2\cosh(\ell_X(\gamma)/2). \tag{23.8}$$

A spin structure identifies the $+1$ and -1 skein algebras by signs on diagrams. Changing to (23.2) also changes only basis signs. All the absolute estimates below are consequently valid in our unnegated trace basis. The positive Fuchsian evaluation is an auxiliary algebraic estimate; the Laplacians defining σ_C^X still use unitary twists.

Proposition 23.4 (Fixed-coefficient bound)**.** *Write $C = \sum_{j=1}^q r_j \beta_j$ with distinct support curves. For $r \geq 1$ and $\ell > 0$ put*

$$M_r(\ell) = \sup_{n \geq 1} \frac{|[x^r] P_n(x)|}{2\cosh(n\ell/2)}, \qquad K_{C,X} = \prod_{j=1}^q M_{r_j}(\ell_X(\beta_j)).$$

Set $K_{\varnothing,X} = 1$. These constants are finite, and

$$\boxed{|a_C(\gamma)| \leq K_{C,X}\, 2\cosh(\ell_X(\gamma)/2)} \tag{23.9}$$

for every nontrivial closed curve, including iterates.

Proof. When $n \geq r \geq 1$ have the same parity,

$$|[x^r]P_n(x)| = \frac{2n}{n+r}\binom{(n+r)/2}{r} \leq \frac{2n^r}{r!};$$

otherwise the coefficient is zero. Consequently

$$M_r(\ell) \leq \frac{2}{r!}\left(\frac{2r}{e\ell}\right)^r < \infty. \tag{23.10}$$

To contribute to C, a bracelet element must contain every β_j in its support. Additional support curves must contribute constant terms, which are either zero or ± 2. Their absolute constant coefficient divided by $2\cosh(n\ell/2)$ is at most one. It follows that

$$|[T_C]B_D| \leq K_{C,X} B_D(X).$$

Apply this bound to (23.8) and use the triangle inequality. Spin and basis signs do not affect the estimate. □

For a single simple α one can use

$$K_{\alpha,X} = \sup_{\substack{n\geq 1\\ n \text{ odd}}} \frac{n}{2\cosh(n\ell_\alpha/2)} \leq \frac{2}{e\ell_\alpha}.$$

The estimates are uniform for a fixed target C on compact subsets of Teichmüller space. Related positivity and Chebyshev estimates for analytic lamination expansions on punctured planar surfaces appear in Basok–Chelkak [23, Section 4.4, Theorems 4.7 and 4.9]. Their full analytic expansion theorem is not needed here.

Corollary 23.5 (Growth of the signed measure)**.** *For each fixed X, C there is a constant $B_{C,X}$ such that*

$$|\sigma_C^X|([0,R]) \leq B_{C,X}(1+R)e^R. \tag{23.11}$$

Proof. Put $\varepsilon = \mathrm{sys}(X) > 0$ and $L = k\ell_\delta$. Each orbital summand has absolute mass at most

$$\frac{\ell_\delta}{\sinh(L/2)} K_{C,X} 2\cosh(L/2) \leq 2K_{C,X}\coth(\varepsilon/2)L.$$

The number of primitive/iterate classes of length at most R is $O_X(e^R)$. One elementary proof conjugates each class so its axis passes a bounded distance from a fixed point in $\mathbb{H}^2$; the chosen element then moves that point by at most $R + O_X(1)$. Orbit points are uniformly separated, and hyperbolic disk area is $O(e^R)$. Summing the displayed mass bound proves the claim. No prime-geodesic asymptotic is required. □

23.5 Convergent heat and Laplace transforms

Define

$$G_C^X(t) = \int_0^\infty e^{-L^2/(4t)}\, \mathrm{d}\sigma_C^X(L), \qquad H_C^X(t) = \frac{e^{-t/4}}{\sqrt{4\pi t}} G_C^X(t), \tag{23.12}$$

$$D_C^X(s) = \int_0^\infty e^{-sL}\, \mathrm{d}\sigma_C^X(L). \tag{23.13}$$

By (23.11), G_C and H_C are absolutely convergent for every $t > 0$, and holomorphic for $\Re t > 0$ with the principal square root. The Laplace transform is absolutely convergent and holomorphic for $\Re s > 1$. Differentiation in these variables is permitted locally uniformly in the stated domains.

There is also a convergent expansion of the whole heat trace. For a finite character polynomial set

$$\|f\|_2 = \sum_C |[T_C]f|\, 2^{\#C}, \qquad c_X = \frac{\log(1+\sqrt{2})}{\operatorname{sys}(X)}.$$

The subscript 2 denotes the trace bound $|\operatorname{tr}\rho(\alpha)| \le 2$; this is a weighted coefficient norm, not an L^2 norm. The Chebyshev recurrence gives

$$\sum_r |[x^r]P_n(x)|2^r \le 2(1+\sqrt{2})^n.$$

A bracelet has at most $3g-3$ distinct support curves, so $\|B_D\|_2 \le 2^{3g-3}e^{c_X\ell_X(D)}$. In (23.8) the nonzero coefficients are integers, and $B_D(X) \ge e^{\ell_X(D)/2}$. Hence $\ell_X(D) \le \ell_X(\gamma) + 2\log 2$ for every occurring D, and $\sum_D |b_D| \le 2e^{\ell_X(\gamma)/2}$. Therefore

$$\|\operatorname{tr}_\gamma\|_2 \le C_X e^{(1/2+c_X)\ell_X(\gamma)}, \qquad C_X = 2^{3g-2}e^{2c_X\log 2}. \tag{23.14}$$

Gaussian decay dominates this exponential bound and the orbit count. It follows that

$$\boxed{\operatorname{Tr} e^{-t\Delta_\rho} - 2I_X(t) = \sum_{C\in\mathscr{C}(S)} H_C^X(t)T_C(\rho), \qquad t>0,} \tag{23.15}$$

where

$$I_X(t) = \frac{\operatorname{Area}(X)}{4\pi}\int_{\mathbb{R}} r\tanh(\pi r)e^{-t(r^2+1/4)}\,\mathrm{d}r, \qquad \sum_C |H_C^X(t)|2^{\#C} < \infty.$$

The series is uniform on the unitary character variety and locally uniform for $\Re t > 0$. It is obtained by taking the limit of the finite length-cutoff expansions; the preceding absolute estimate justifies the rearrangement.

This construction does not assert uniqueness among arbitrary infinite expansions of a continuous function. For example, $\|\operatorname{tr}\rho(\alpha^{2n+1})\|_\infty \le 2$, while its T_α coefficient is $(-1)^n(2n+1)$. Coefficient extraction is therefore not bounded for the uniform norm on unitary characters. Our coefficients are fixed by the finite spectral tests first; the estimates then justify their heat transform.

Proposition 23.6 (Simple-curve semiclassical term)**.** *Let α be simple and $\ell = \ell_X(\alpha)$. Choose $b > \ell$ so that every other orbital contribution to σ_α^X has length at least b. Then*

$$G_\alpha^X(t) = \frac{\ell}{\sinh(\ell/2)}e^{-\ell^2/(4t)} + O\big(e^{-b^2/(4t)}\big),$$

$$H_\alpha^X(t) = \frac{e^{-t/4}}{\sqrt{4\pi t}}\frac{\ell}{\sinh(\ell/2)}e^{-\ell^2/(4t)}\left[1 + O\Big(e^{-(b^2-\ell^2)/(4t)}\Big)\right], \tag{23.16}$$

$$D_\alpha^X(s) = \frac{\ell}{\sinh(\ell/2)}e^{-s\ell}\left[1 + O\big(e^{-s(b-\ell)}\big)\right] \quad (s\to+\infty). \tag{23.17}$$

In particular,

$$\boxed{\ell^2 = -\lim_{t\downarrow 0} 4t\log G_\alpha^X(t), \qquad \ell = -\lim_{s\to+\infty} s^{-1}\log D_\alpha^X(s).} \tag{23.18}$$

Both logarithms are real for sufficiently small t, respectively sufficiently large positive s.

Proof. Separate the first atom in (23.6). For its remaining measure τ, supported on $[b,\infty)$, fix $t_0>0$. For $0<t\le t_0$,

$$\int e^{-L^2/(4t)}\,\mathrm{d}|\tau|(L)\le e^{-b^2/(4t)+b^2/(4t_0)}\int e^{-L^2/(4t_0)}\,\mathrm{d}|\tau|(L).$$

The last integral is finite. The same argument with a fixed real $s_0>1$ gives the Laplace remainder. The leading masses are positive, so the leading terms eventually dominate the signed tails. □

After removing the explicit $e^{-t/4}$ factor, the correction to the single-geodesic term is exponentially small, rather than a new power series in t. Thus the constant-curvature heat coefficient has an exact isolated-orbit prefactor, followed by contributions at larger length actions. For $m\alpha$ replace the action by $(m\ell_\alpha)^2/(4t)$ and the mass by $\ell_\alpha/\sinh(m\ell_\alpha/2)$.

23.6 All iterates of a simple curve

Write $\sigma_C^{X,[\delta]}$ for the contribution of one primitive unoriented class δ and all of its iterates. The polynomial identity

$$P_k(x)=\sum_{j=0}^{\lfloor k/2\rfloor}(-1)^j\frac{k(k-j-1)!}{j!(k-2j)!}x^{k-2j},\qquad k\ge 1, \tag{23.19}$$

follows from the recurrence. If α is simple, the coefficient of $m\alpha$ vanishes unless $k=m+2j$, and for $m\ge 1$ is

$$b_{m+2j,m}=(-1)^j\frac{(m+2j)(m+j-1)!}{j!\,m!}. \tag{23.20}$$

In particular, with $\ell=\ell_\alpha$,

$$\boxed{\sigma_\alpha^{X,[\alpha]}=\ell\sum_{j\ge 0}\frac{(-1)^j(2j+1)}{\sinh((2j+1)\ell/2)}\delta_{(2j+1)\ell\cdot}} \tag{23.21}$$

The first masses have multipliers $1,-3,5,-7,\dots$ at $\ell,3\ell,5\ell,7\ell,\dots$. For 2α the multipliers are $1,-4,9,-16,\dots$ at $2\ell,4\ell,6\ell,8\ell,\dots$. These are explicit later atoms, not changes to the first-atom mass.

For $m\ge 1$ the coefficient generating function is

$$B_m(q):=\sum_{k\ge 1}b_{k,m}q^k=\frac{q^m(1-q^2)}{(1+q^2)^{m+1}}. \tag{23.22}$$

Indeed, $\sum_{k\ge1}P_k(x)q^k=(xq-2q^2)/(1-xq+q^2)$, and extracting x^m gives this formula. Expanding the hyperbolic denominator yields, for $\Re s>-1/2$,

$$\begin{aligned}D_{m\alpha}^{X,[\alpha]}(s)&=2\ell\sum_{n\ge0}B_m(q_n), & q_n&=e^{-(s+n+1/2)\ell},\\ &=2^{1-m}\ell\sum_{n\ge0}\operatorname{sech}^m(u_n)\tanh(u_n), & u_n&=(s+n+1/2)\ell.\end{aligned} \tag{23.23}$$

This larger convergence domain concerns the one fixed primitive family; the proved half-plane for the full D_C is $\Re s>1$.

The heat contribution of the same family is

$$H_\alpha^{X,[\alpha]}(t) = \frac{e^{-t/4}}{\sqrt{4\pi t}} \frac{\ell e^{-\ell^2/(4t)}}{\sinh(\ell/2)} \times \left[1 - 3\frac{\sinh(\ell/2)}{\sinh(3\ell/2)} e^{-2\ell^2/t} + 5\frac{\sinh(\ell/2)}{\sinh(5\ell/2)} e^{-6\ell^2/t} - \cdots\right]. \tag{23.24}$$

Other primitive classes may supply corrections before these particular iterate actions. Once ℓ_α is recovered, subtracting the $j \geq 1$ terms of (23.21) removes exactly the nonprimitive powers of the simple curve α; it leaves powers of nonsimple curves that resolve with an α coefficient.

23.7 A figure-eight family and an exact cancellation

Take an embedded geodesic pair of pants with three distinct essential cuffs a, b, c, and $c = (ab)^{-1}$. Set

$$x = T_a, \qquad y = T_b, \qquad z = T_c, \qquad \gamma = a^{-1}b, \qquad T_\gamma = xy - z.$$

These are functions of arbitrary unitary holonomy. The minus sign therefore differs from the positive hyperbolic length convention in (18.1). Powers are given exactly by $T_{\gamma^k} = P_k(xy - z)$. For example,

$$\begin{aligned} T_\gamma &= xy - z, \\ T_{\gamma^2} &= x^2y^2 - 2xyz + z^2 - 2, \\ T_{\gamma^3} &= x^3y^3 - 3x^2y^2z + 3xyz^2 - z^3 - 3xy + 3z. \end{aligned}$$

Only multicurves $C = ra + rb + mc$ occur. Put $d = r + m \geq 1$. Their coefficients vanish unless $k = d + 2j$, and in that case

$$\boxed{a_{ra+rb+mc}(\gamma^k) = (-1)^{j+m}\frac{k(k-j-1)!}{j!\,r!\,m!} = (-1)^m\binom{r+m}{r} b_{k,r+m}.} \tag{23.25}$$

This follows by substituting $xy - z$ into (23.19) and expanding the relevant monomial. Consequently

$$\sum_{k\geq 1} a_{ra+rb+mc}(\gamma^k)q^k = (-1)^m\binom{r+m}{r}\frac{q^{r+m}(1-q^2)}{(1+q^2)^{r+m+1}}.$$

For $r + m = 1$ the two target coefficients are opposite for every iterate. If $L = \ell_\gamma$, then

$$\boxed{\sigma_{a+b}^{X,[\gamma]} = -\sigma_c^{X,[\gamma]} = L\sum_{j\geq 0}\frac{(-1)^j(2j+1)}{\sinh((2j+1)L/2)}\delta_{(2j+1)L}.} \tag{23.26}$$

Thus $\sigma_c^X + \sigma_{a+b}^X$ cancels the entire chosen figure-eight family, including all its powers. This does not assert that the full sum vanishes: other primitive classes contribute to both coefficients. It gives an exact family-by-family cancellation which is preserved by either transform in (23.12)–(23.13).

The distinct-cuff assumption makes the monomials above distinct multicurve basis elements. If cuff images coincide, as in some one-holed-torus configurations of Section 18, the corresponding variables must first be identified and equal basis terms collected; the displayed coefficient indexing cannot be carried over unchanged.

23.8 Metric variation and the twist derivative

For a fixed topological C, the map $X \mapsto \sigma_C^X$ is smooth as a distribution-valued map. After applying any compactly supported test function, it is locally a finite sum of smooth length functions. It is not asserted to be real analytic for arbitrary smooth test functions.

For an orbital label (δ, k) write

$$L = k\ell_\delta, \qquad w = \ell_\delta \operatorname{csch}(L/2).$$

For a tangent vector v to Teichmüller space one has

$$D_v \sigma_C^X = \sum_{\delta,k} a_C(\delta^k) \left[(D_v w)\delta_L - w(D_v L)\delta'_L \right], \tag{23.27}$$

where $\langle \delta'_L, \varphi \rangle = -\varphi'(L)$ and

$$D_v L = k D_v \ell_\delta, \qquad D_v w = \operatorname{csch}(L/2) \left(1 - \frac{L}{2} \coth \frac{L}{2} \right) D_v \ell_\delta.$$

The coefficients a_C are constant under metric variation. This formula describes movement of the atoms as well as variation of their masses; omitting the δ'_L term would lose the length variation itself.

For a twist of displacement τ_β about a simple geodesic β, Wolpert's cosine formula gives

$$\frac{\partial \ell_\delta}{\partial \tau_\beta} = \sum_{p \in \delta \cap \beta} \cos \theta_p, \tag{23.28}$$

with the consistent right-twist angle convention and intersections counted with multiplicity [24]. Substitution in (23.27) gives an exact variation formula for every coefficient distribution. For a simple α, one can equivalently differentiate the localized ratio (23.7); it recovers (23.28) from two spectral coefficient moments, without an infinite limit.

The compact-set gap in Proposition 23.3 also permits differentiating the leading heat asymptotic locally. The leading logarithmic response is

$$D_v \log H_\alpha^X(t) = \left[-\frac{\ell_\alpha}{2t} + \frac{1}{\ell_\alpha} - \frac{1}{2} \coth \frac{\ell_\alpha}{2} \right] D_v \ell_\alpha + O(t^{-1} e^{-c/t}) \tag{23.29}$$

for some local $c > 0$. To justify the differentiated tail, metric length derivatives on a compact parameter neighborhood are bounded by constants times the length, while (23.9) is locally uniform. Gaussian decay absorbs the additional polynomial factors.

23.9 The logarithmic product and primitive extraction

A useful connection to the Selberg product keeps the orientation normalization explicit. Define, for $\Re v > 1$,

$$\mathcal{Z}_\rho^X(v) = \prod_{\delta \in \mathcal{P}_{\mathrm{un}}} \prod_{n \geq 0} \det \left(I - e^{-(v+n)\ell_\delta} \rho(\delta) \right). \tag{23.30}$$

This definition uses unoriented primitives; a product over both oriented conjugacy classes squares it. Expanding the logarithm gives

$$2 \, \frac{\partial}{\partial v} \log \mathcal{Z}_\rho^X(v) \Big|_{v = s + 1/2} = \sum_{\delta,k} \frac{\ell_\delta e^{-sk\ell_\delta}}{\sinh(k\ell_\delta/2)} \operatorname{tr} \rho(\delta^k). \tag{23.31}$$

With coefficients defined by finite length cutoffs, the T_C coefficient on the right is $D_C^X(s)$ for $\Re s > 1$. The coefficient estimate proves this assertion; it is not an appeal to continuity of coefficient extraction in the uniform norm. No meromorphic continuation of the individual D_C is established by this calculation.

Removing all iterates by a global Chebyshev substitution is not legitimate in the rank-two character ring. If $x = \operatorname{tr} A$, $y = \operatorname{tr} B$, $u = \operatorname{tr}(AB)$, and $v = \operatorname{tr}(AB^{-1})$, then $u + v = xy$, but

$$P_2(u) + P_2(v) - P_2(x)P_2(y) = 2\big(\operatorname{tr}[A, B] - 2\big), \tag{23.32}$$

which is generally nonzero. Thus sending every trace to its Chebyshev polynomial does not define a multiplicative operation on this character ring.

There is an exact repair if extra twisting data are supplied. For the original SU(2) representation ρ, work in the virtual representation ring generated by its symmetric powers. Let $\psi^k\rho$ denote its Adams operation, characterized by $\chi_{\psi^k\rho}(\gamma) = \chi_\rho(\gamma^k)$. For the standard rank-two representation,

$$\psi^k\rho = \operatorname{Sym}^k \rho - \operatorname{Sym}^{k-2} \rho \quad (k \geq 2), \qquad \psi^1\rho = \rho.$$

Trace formulas extend additively to virtual representations. All twists used here have inverse-invariant characters, so the unoriented normalization is preserved. Let Σ_ρ denote the character-valued geometric measure on the right of (23.4), and set

$$Q_\rho = \sinh(L/2)\Sigma_\rho, \qquad P_\rho = \sum_{\delta \in \mathcal{P}_{\mathrm{un}}} \ell_\delta \chi_\rho(\delta)\delta_{\ell_\delta}.$$

For $D_k\delta_L = \delta_{kL}$,

$$Q_\rho = \sum_{k\geq 1} D_k P_{\psi^k\rho}, \qquad \boxed{P_\rho = \sum_{k\geq 1} \mu(k) D_k Q_{\psi^k\rho}.}$$

These sums are locally finite, and the second identity follows from $\psi^j\psi^k = \psi^{jk}$ and divisor Möbius inversion. It requires spectra for higher symmetric-power twists, which are additional data beyond the original rank-two family. The explicit subtraction in (23.21) needs no such enlargement.

23.10 Scope and further questions

The construction now supplies signed coefficient distributions, an isolated first simple-curve atom, globally convergent heat transforms, a Laplace half-plane, explicit iterate families, a figure-eight cancellation, and metric derivatives. It has not produced a positive measure with the weights of simple-geodesic counting, nor a pointwise constant McShane series.

The scalar limitation is genuine: length-isospectral closed hyperbolic surfaces can have different simple length spectra [25]. The labelled holonomy family retains the information used by (23.7).

For a counting comparison, let $d = 6g - 6$ and normalize Thurston measure by the integral multicurve lattice. If

$$B(X) = \mu_{\mathrm{Th}}\{\lambda : \ell_X(\lambda) \leq 1\},$$

ordinary lattice counting gives $\#\{C \in \mathcal{ML}_{\mathbb{Z}} : \ell_X(C) \leq R\} \sim B(X)R^d$ [75, Proposition 3.1]. Its Laplace consequence is

$$\sum_{C\in\mathcal{ML}_{\mathbb{Z}}} e^{-t\ell_X(C)} \sim \Gamma(d+1)B(X)t^{-d} \qquad (t \downarrow 0).$$

This is a target for a measure comparison, not a consequence of the coefficient heat expansion. It counts all weighted multicurves; connected or fixed-orbit counts require additional information. The large-length limit here is also different from the minimum-selection limits in (23.18).

Two concrete further problems are to continue $D_C(s)$ beyond its proved half-plane, with control of the coefficient operation, and to find combinations of coefficients whose cancellations organize globally over pants. Formula (23.26) provides an exact local model for the second problem. An identity constant in X would require an additional normalization or geometric partition; the existence of these transforms alone does not supply that constancy. Section 24 strengthens the lattice benchmark and gives explicit orbital gaps, uniform recovery estimates, and a conditional counting criterion.

24 Quantitative counting and coefficient cancellation

The preceding construction raises a quantitative question: can explicit crossing resolutions give a counting remainder with a specified exponent and geometric constants? We first give an elementary counting benchmark. We then prove a systole-dependent gap and a uniform recovery estimate, and show on a fixed surface why their exponential separation scale is unavoidable for estimates of individual orbital terms. Finally, a finite support coefficient and a smoothing criterion isolate the additional estimate needed for connected-curve counting.

24.1 Effective counting and the meaning of an explicit saving

Continue with a closed hyperbolic surface X of genus $g \geq 2$, and put $d = 6g - 6$. For a fixed rational multicurve γ_0, write

$$N_{X,\gamma_0}(R) = \#\{\gamma \in \mathrm{Mod}(S)\gamma_0 : \ell_X(\gamma) \leq R\}.$$

Eskin–Mirzakhani–Mohammadi prove that there exists $\kappa(g) > 0$ such that

$$N_{X,\gamma_0}(R) = c_{X,\gamma_0} R^d + O_{X,\gamma_0}(R^{d-\kappa(g)}). \tag{24.1}$$

Their theorem also allows variable negative curvature; the hyperbolic case suffices here [14, Theorem 1.1]. The theorem states the existence of the exponent rather than a numerical value. Its proof uses exponential mixing of the Teichmüller flow. A specified exponent and certified constants for a given surface would therefore be additional quantitative information. This distinction does not diminish the proved power saving or establish that its constants are intrinsically noncomputable.

The coefficient distributions offer explicit local formulas, but their present estimates do not improve (24.1). The task is to control a growing collection of connected supports. The qualitative leading term, treated in [75, 17], and an effective remainder are separate questions. It is useful first to record how much quantitative control follows from ordinary lattice geometry alone.

24.2 A constructive lattice benchmark for all multicurves

Let $\mathcal{ML}_{\mathbb{Z}}(S)$ denote integral measured laminations, equivalently essential simple multicurves with positive integer weights. These include disconnected supports. Normalize Thurston measure by the actual integral-lamination lattice in each chart, and set

$$M_X(R) = \#\{C \in \mathcal{ML}_{\mathbb{Z}}(S) \setminus \{0\} : \ell_X(C) \leq R\}, \qquad B(X) = \mu_{\mathrm{Th}}\{\lambda : \ell_X(\lambda) \leq 1\}.$$

Proposition 24.1 (Total multicurve error)**.** *For the fixed surface X,*

$$\boxed{M_X(R) = B(X)R^d + O_X(R^{d-1}).} \tag{24.2}$$

The error can be bounded using finitely many lattice covolumes, fundamental-cell radii, convex-body inradii and volumes, and lower-dimensional face bounds.

Proof. Here is the quantitative lattice estimate needed for the argument. Let $\Lambda \subset \mathbb{R}^d$ have covolume v and a fundamental cell $F \subset B(0, \rho)$. Suppose the convex compact body K has volume V and contains $B(x_0, r)$, where $r > 0$. Put $a = \rho/r$. Then

$$\frac{V}{v}(R - a)_+^d \leq \#(\Lambda \cap RK) \leq \frac{V}{v}(R + a)^d. \tag{24.3}$$

Indeed, up to null sets the lattice cells tile Euclidean space, so

$$(RK) \ominus B_\rho \;\subset \bigcup_{p \in \Lambda \cap RK} (p + F) \;\subset RK + B_\rho.$$

Here $A \ominus B_\rho = \{x : x + B_\rho \subset A\}$. Writing $K_0 = K - x_0$, we have $B_r \subset K_0$ and

$$RK + B_\rho \subset Rx_0 + (R + a)K_0.$$

For $R \geq a$, also $Rx_0 + (R - a)K_0 \subset (RK) \ominus B_\rho$. Taking volumes proves (24.3); the lower bound is zero when $R < a$. Notice that the homotheties are centered at Rx_0, since the origin can be a boundary point of K. For $R \geq a$ this gives the explicit bound

$$\left| \#(\Lambda \cap RK) - \frac{V}{v} R^d \right| \leq d\, 2^{d-1} \frac{V}{v} \frac{\rho}{r}\, R^{d-1}. \tag{24.4}$$

The upper increment $(R + a)^d - R^d$ dominates the lower decrement.

To apply this estimate, use finitely many rational polyhedral chambers subordinate to train-track charts. Length is convex in each train-track cone, by the surgery argument in [14, Theorem 5.1]. The finite polyhedral handling of chart overlaps is described at the end of Section 8 of that paper. Refining into convex chambers Q_i, the sets $K_i = Q_i \cap \{\ell_X \leq 1\}$ are compact convex bodies. Apply (24.3) in every full-dimensional chamber with its own integral lattice. Their normalized volumes sum to $B(X)$. In particular, a coordinate parity condition must be absorbed into the lattice covolume, not silently discarded.

Shared faces can be assigned to one chamber. Alternatively, count closed chambers and bound the overcount. If a face has dimension $1 \leq k \leq d - 1$, its unit-length part lies in a ball of radius A_F, and its induced lattice has shortest nonzero vector of length δ_F, disjoint-ball packing gives

$$\#(\Lambda_F \cap RK_F) \leq \left(1 + \frac{2RA_F}{\delta_F}\right)^k.$$

Zero-dimensional faces contribute at most one point each. The finite sum of these bounds is $O_X(R^{d-1})$. Removing the zero multicurve changes the answer by one. □

This refines the elementary total-multicurve lattice asymptotic in [75, Proposition 3.1]; it is a benchmark for the new machinery. Formula (24.4) is explicit once bounds on its geometric inputs are certified. It does not supply numerical values for an unspecified surface. Fixed train-track representatives give linear upper bounds for length, and coordinate length estimates such as [15] help certify the required geometric bounds. Hyperbolic length itself need not be piecewise linear. For example, choose an interior rational weight vector in a chamber and scale it until a certified linear upper bound for length is less than one. A small coordinate ball then remains both in that chamber and below the upper bound, producing an inradius certificate. For an outer-radius certificate one needs a lower bound for length on the chamber's unit coordinate sphere. A finite filling collection of curves, together with collar inequalities and their piecewise-linear intersection functions, supplies such a bound. The constants can deteriorate when the surface degenerates; no uniformity over the entire moduli space is being asserted. On a fixed compact subset of Teichmüller space the same finite coordinate construction and uniform length comparisons give uniform versions of the bounds.

24.3 Primitivity, connectivity, and the torus benchmark

Write $C = \sum_j m_j \alpha_j$ with distinct simple supports. Call C primitive as an integral lamination if $\gcd_j m_j = 1$. Every nonzero integral multicurve is uniquely kC_0 with C_0 primitive. This

intrinsic definition avoids coordinate conventions: division by k is allowed precisely when the resulting lamination remains integral. Let $P_X(R)$ count these primitive laminations. Möbius inversion gives

$$P_X(R) = \sum_{k \le R/\operatorname{sys}(X)} \mu(k) M_X(R/k).$$

Since $d > 2$, the error series $\sum k^{-(d-1)}$ converges, while the truncated leading-term tail is $O_X(R)$. Therefore

$$P_X(R) = \frac{B(X)}{\zeta(d)} R^d + O_X(R^{d-1}). \tag{24.5}$$

A union $\alpha + \beta$ of distinct disjoint simple curves is primitive but disconnected. Thus (24.5) does not count connected simple curves in higher genus.

For a once-punctured torus every essential multicurve has one distinct support. The McShane–Rivin norm on $H_1(T;\mathbb{R})$ identifies simple curves with primitive lattice vectors modulo sign. If B_X is its full centrally symmetric unit ball, our unoriented convention gives

$$N_X(R) = \frac{\operatorname{area}(B_X)}{2\zeta(2)} R^2 + O_X(R \log R). \tag{24.6}$$

This is the norm-and-Möbius-inversion argument of [16, Theorem 4.1]; the factor $1/2$ identifies opposite homology vectors. The logarithm comes from $\sum_{k \le O_X(R)} k^{-1}$. This independent torus benchmark does not extend the closed-surface trace formula across a cusp.

24.4 A fixed-surface obstruction to uniform first-atom separation

The exponentially small remainder for one fixed simple curve does not remain small uniformly over curves whose length tends to infinity. There is an explicit obstruction on a fixed hyperbolic surface, with no degeneration of its metric.

Proposition 24.2 (An exponentially close pants contribution)**.** *On every closed hyperbolic surface X of genus at least two there are fixed distinct simple curves a, b, and simple curves c_n with $R_n = \ell_X(c_n) \to \infty$, such that a, b, c_n bound embedded geodesic pairs of pants. Let γ_n be their figure-eight curves with $T_{\gamma_n} = T_a T_b - T_{c_n}$, and set*

$$A = \cosh \frac{\ell_X(a)}{2} \cosh \frac{\ell_X(b)}{2}, \qquad L_n = \ell_X(\gamma_n).$$

Then

$$\begin{aligned} L_n - R_n &= 8Ae^{-R_n/2} + O(e^{-R_n}), \\ L_n^2 - R_n^2 &= 16AR_n e^{-R_n/2} + O(R_n e^{-R_n}). \end{aligned} \tag{24.7}$$

The coefficient of T_{c_n} in T_{γ_n} is -1.

Proof. Choose distinct disjoint nonseparating curves a, b whose complement after cutting is connected, of genus $g-2$ with four boundary components. In this cut surface let c bound a pair of pants with one copy of a and one copy of b. Choose an essential interior curve intersecting c. A Dehn twist about that curve fixes the boundary and gives infinitely many distinct curves c_n, each bounding a pair of pants with a, b. Straightening preserves this configuration. Finiteness of the closed-geodesic length set below a bound implies $R_n \to \infty$.

Rivin's pants identity [88, Theorem 1.3] gives

$$\cosh(L_n/2) = \cosh(R_n/2) + 2A.$$

Expanding its inverse for $R_n \to \infty$ gives (24.7). The coefficient assertion is the ordinary trace identity in the displayed pants variables. □

Normalize the c_n heat coefficient by its first-atom contribution. The single primitive orbital term supplied by γ_n becomes

$$-r_n \exp\left(-\frac{L_n^2 - R_n^2}{4t}\right), \qquad r_n = \frac{L_n}{R_n}\frac{\sinh(R_n/2)}{\sinh(L_n/2)} = 1 + O(e^{-R_n/2}). \tag{24.8}$$

For $t = R_n^{-p}$, with any fixed $p > 0$, this tends to -1. Thus the absolute-valued orbital remainder contains an order-one term at every polynomial time scale. Suppressing this one term below R_n^{-M} requires

$$t \le \left(\frac{4A}{M} + o(1)\right)\frac{R_n e^{-R_n/2}}{\log R_n}. \tag{24.9}$$

This is a statement about an individual orbital contribution and the absolute-valued remainder. It does not exclude cancellations with other curves, including other curves of exactly the same length, and does not by itself identify a nonzero net atom of $\sigma^X_{c_n}$.

Such cancellation is precisely what the coefficient approach can organize. By (23.26), $\sigma^X_{c_n} + \sigma^X_{a+b}$ cancels the entire chosen γ_n family, including all iterates, exactly. A useful counting construction should exploit these relations before taking absolute values. The proposition explains why the cancellation is more than a small correction to an already uniform first-atom estimate.

24.5 A constructive uniform isolation estimate

The same bracelet estimate that bounded the coefficients also provides an explicit lower bound on their orbital separation. Write $\varepsilon = \mathrm{sys}(X) > 0$, and put

$$c_\varepsilon = \min\{1, 2\cosh(3\varepsilon/2) - 2\cosh(\varepsilon/2)\} > 0. \tag{24.10}$$

Lemma 24.3 (A quantitative trace gap)**.** *Let α be simple, $x = \ell_X(\alpha)$, and let $\gamma = \delta^k$ be an orbital label contributing to σ^X_α, other than $(\alpha, 1)$. If $L = \ell_X(\gamma)$, then*

$$2\cosh(L/2) - 2\cosh(x/2) \ge c_\varepsilon. \tag{24.11}$$

Consequently, for $x \le R$,

$$L^2 - x^2 \ge \Delta_\varepsilon(R), \qquad \Delta_\varepsilon(R) = 4\,\mathrm{arcosh}^2\left(\cosh(R/2) + \frac{c_\varepsilon}{2}\right) - R^2 \sim 4c_\varepsilon R e^{-R/2}. \tag{24.12}$$

Proof. By (23.8), the positive Fuchsian value $2\cosh(L/2)$ is an integer positive sum of bracelet values. At least one contributing bracelet has an α coefficient. It therefore has an odd winding $n \ge 1$ around α; any additional support components must contribute constant Chebyshev terms and have even windings at least two.

If $n \ge 3$, that bracelet exceeds $2\cosh(x/2)$ by at least $2\cosh(3\varepsilon/2) - 2\cosh(\varepsilon/2)$. If $n = 1$ and additional support occurs, its value is at least $2\cosh(x/2)\,2\cosh\varepsilon$, giving a difference greater than two. The remaining case is the bracelet α itself. Unless it is the only term with coefficient of absolute value one, integrality gives an additional positive value of at least one. If it were that sole term, the Fuchsian lengths would satisfy $L = x$, contrary to the strict length-ordering lemma for the orbital labels under consideration. This proves (24.11).

Set $f(x) = 2\,\mathrm{arcosh}(\cosh(x/2) + c_\varepsilon/2)$. The trace gap gives $L \ge f(x) > x$. Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(x)^2 - x^2) = 2\left[f(x)\frac{\sinh(x/2)}{\sinh(f(x)/2)} - x\right] < 0,$$

because $y/\sinh(y/2)$ decreases for $y > 0$. This proves the uniform action bound; its asymptotic follows by expanding $f(R) - R = 2c_\varepsilon e^{-R/2} + O(e^{-R})$. □

Here and below "orbital gap" refers to the individual summands before cancellation at coincident lengths. This convention both makes the estimate stronger and avoids assuming that a later summand survives in the signed measure.

For explicit constants let $D = \operatorname{diam}(X)$ and define

$$\begin{aligned} Q_X &= \frac{e^{2D}}{(1-e^{-\varepsilon/2})^2}, \\ A_\varepsilon &= \frac{1+e^{-\varepsilon}}{(1-e^{-\varepsilon})^2} \geq \sup_{\substack{n\geq 1\\ n \text{ odd}}} n e^{-(n-1)\varepsilon/2}, \\ C_X &= \frac{Q_X A_\varepsilon \coth(\varepsilon/2) e^2}{(1-e^{-1})^2}\left(1+\frac{2}{\varepsilon}\right). \end{aligned} \tag{24.13}$$

For a simple α let b_α be the smaller of $\ell_\alpha + 1$ and its first later contributing orbital length. For $R \geq \varepsilon$ put

$$\eta_X(R) = \min_{\substack{\alpha \ msimple\\ \ell_\alpha \leq R}} (b_\alpha^2 - \ell_\alpha^2), \qquad \underline{\eta}_X(R) = \min\{2\varepsilon + 1, \Delta_\varepsilon(R)\}.$$

Only finitely many labels enter the first minimum, and Lemma 24.3 gives $\eta_X(R) \geq \underline{\eta}_X(R) > 0$.

Proposition 24.4 (Uniform recovery with explicit constants)**.** *For $0 < t \leq \varepsilon/4$,*

$$\boxed{\sup_{\substack{\alpha \ msimple\\ \ell_\alpha \leq R}} \left| \frac{\sinh(\ell_\alpha/2)}{\ell_\alpha} e^{\ell_\alpha^2/(4t)} G_\alpha^X(t) - 1 \right| \leq C_X \exp\left(R - \frac{\eta_X(R)}{4t}\right).} \tag{24.14}$$

The explicit $\underline{\eta}_X(R)$ may replace $\eta_X(R)$.

Proof. Orbit packing gives a bound $N_{\mathrm{all}}(U) \leq Q_X e^U$ for primitive/iterate labels of length at most U. Indeed, choose a representative whose axis comes within D of a fixed lift of a basepoint. It moves that point at most $U + 2D$. The orbit points have separation at least ε, and comparison of disjoint disks of radius $\varepsilon/2$ with the disk of radius $U + 2D + \varepsilon/2$ gives the stated constant.

Set $x = \ell_\alpha$ and $w_\alpha = x/\sinh(x/2)$. The sharper form of the fixed-coefficient bound is

$$\frac{K_{\alpha,X}}{w_\alpha} \leq \frac{A_\varepsilon}{2x}.$$

Thus the absolute mass of a length-L orbital term, divided by w_α, is at most $A_\varepsilon \coth(\varepsilon/2) L/x$. Sum the tail in unit shells starting at $b = b_\alpha$. Since $b \geq \varepsilon$ and $t \leq \varepsilon/4$, its normalized absolute value is at most

$$\frac{Q_X A_\varepsilon \coth(\varepsilon/2)}{x} e^{b+1-(b^2-x^2)/(4t)} \sum_{j\geq 0} (b+j+1) e^{-j} \leq \frac{Q_X A_\varepsilon \coth(\varepsilon/2)}{(1-e^{-1})^2} \frac{b+1}{x} e^{b+1-(b^2-x^2)/(4t)}.$$

Use $b \leq x + 1 \leq R + 1$ and $(b+1)/x \leq 1 + 2/\varepsilon$. This proves (24.14). □

For $R > 1$ and $M > 0$, a sufficient condition for relative error at most R^{-M} is

$$t \leq \min\left\{\frac{\varepsilon}{4}, \frac{\underline{\eta}_X(R)}{4(R + M\log R + \log\max\{1, C_X\})}\right\}. \tag{24.15}$$

Every displayed constant is determined by geometric data, and certified lower bounds for the systole and upper bounds for the diameter suffice for a conservative version. A sharper $\eta_X(R)$

can also be found from the finite cutoff $R+1$ when length comparisons and the metric are supplied effectively. The guaranteed time scale in (24.15) is exponential in $-R/2$. Together with Proposition 24.2, this identifies the exponential separation scale constructively.

This is a quantitative recovery theorem, not a counting asymptotic. Summing its normalized expression over $\ell_\alpha \le R$ already presupposes the target set and its lengths. A counting argument must construct an aggregate observable, perform the necessary signed cancellations, and estimate its large-length behavior without assuming that list.

24.6 A finite coefficient that selects one support

There is a useful way to state the remaining selection problem without an infinite-dimensional coefficient operation. For an integral multicurve $C = \sum_{i=1}^{r} m_i \alpha_i$, with distinct supports α_i, put $r(C) = r$ and $r(0) = 0$. Define

$$\mathcal{P}_X(s,z) = \sum_{C \in \mathcal{ML}_{\mathbb{Z}}(S)} z^{r(C)} e^{-s\ell_X(C)}, \qquad \Re s > 0. \tag{24.16}$$

The estimate (24.2) implies absolute convergence. Since $r(C) \le 3g-3$, this is a polynomial in z of degree at most $3g-3$, with holomorphic coefficients in s. Here $r(C)$ counts distinct geodesic supports, not the number of parallel copies used in the trace monomial T_C.

Its linear coefficient has an exact interpretation:

$$F_X(s) := [z]\mathcal{P}_X(s,z) = \sum_{\alpha \text{ simple}} \sum_{m \ge 1} e^{-sm\ell_\alpha} = \sum_{\alpha \text{ simple}} \frac{1}{e^{s\ell_\alpha} - 1}. \tag{24.17}$$

All curves in these sums are unoriented. Divisor inversion gives

$$\boxed{Z_X^{\mathrm{simple}}(s) := \sum_{\alpha \text{ simple}} e^{-s\ell_\alpha} = \sum_{k \ge 1} \mu(k) F_X(ks), \qquad \Re s > 0.} \tag{24.18}$$

There is no factor $1/k$. Indeed, after expanding the right side, the coefficient of $e^{-sn\ell_\alpha}$ is $\sum_{k \mid n} \mu(k)$, which is one for $n = 1$ and zero otherwise. The rearrangement is absolutely convergent: for $\sigma = \Re s > 0$ and $\varepsilon = \mathrm{sys}(X)$,

$$\sum_{\alpha,k,m \ge 1} e^{-\sigma k m \ell_\alpha} \le \frac{Z_X^{\mathrm{simple}}(\sigma)}{(1 - e^{-\sigma\varepsilon})^2} < \infty,$$

using $km \ge k + m - 1$.

Even the support-coefficient extraction is explicit and bounded. Set $q = 3g-2$ and $\omega = e^{2\pi i/q}$. Then

$$F_X(s) = \frac{1}{q} \sum_{j=0}^{q-1} \omega^{-j} \mathcal{P}_X(s, \omega^j). \tag{24.19}$$

Thus an error bounded by $E(s)$ at each of these finitely many values remains bounded by $E(s)$ after extracting the linear coefficient. This operation has none of the growing-degree conditioning of arbitrary trace-coefficient extraction.

For example, suppose a future argument establishes

$$\mathcal{P}_X(s, \omega^j) = s^{-d} A_X(\omega^j) + O_X(s^{-d+\eta}), \qquad 0 < s \le 1,$$

for every j, with a specified $0 < \eta < d-1$ and a polynomial A_X of degree at most $3g-3$. If $a_1 = [z]A_X(z)$, then

$$Z_X^{\mathrm{simple}}(s) = \frac{a_1}{\zeta(d)} s^{-d} + O_X(s^{-d+\eta}). \tag{24.20}$$

To check the error, split the sum in (24.18) at $k = 1/s$. The small-ks remainder is bounded by $s^{-d+\eta} \sum k^{-d+\eta}$; this converges because $d - \eta > 1$. The omitted leading tail and the large-ks tail are $O_X(s^{-1})$, the latter by exponential decay from the positive systole. Both are absorbed in the displayed error.

This is a conditional transfer of a remainder, not a proof of the hypothesized expansion. At $z = 1$ ordinary lattice counting already gives

$$\mathcal{P}_X(s, 1) = \Gamma(d + 1)B(X)s^{-d} + O_X(s^{-d+1}) \qquad (s \downarrow 0).$$

It does not determine the other values in (24.19). The support count is a discrete connectivity statistic, and a general weighted multicurve cannot be assigned that statistic by a continuous function on all of measured-lamination space. Knowing its value for each integral point does not establish its density in expanding regions.

Connectivity itself is computationally accessible: the Agol–Hass–Thurston interval-pairing algorithm counts components of a normal curve in time polynomial in the triangulation size and the logarithm of its normal weight [18, Theorem 12 and Corollary 13]. The missing quantitative input is an average estimate for this selector, or a cancellation formula that supplies such an estimate. Moreover, (24.16) is a target positive partition function. It has not been identified with an evaluation of the signed coefficient heat expansion. Defining a linear operation $T_C \mapsto z^{r(C)}$ on finite trace polynomials alone does not change an orbital length ℓ_γ into the resolved length $\ell_X(C)$, and does not supply that identification.

24.7 An explicit criterion for passing from smooth to sharp counts

Proposition 24.1 also bounds the number of connected simple curves in a shell: for $R \geq 2$ and $0 \leq \Delta \leq R/2$,

$$\#\{\alpha \text{ simple} : |\ell_X(\alpha) - R| \leq \Delta\} = O_X\big((\Delta + 1)R^{d-1}\big). \tag{24.21}$$

Subtract the total-multicurve estimates at the shell endpoints; one-sided limits handle endpoint atoms. The simple curves form a subset, so positivity gives the stated bound.

Proposition 24.5 (Conditional quantitative desmoothing)**.** *Let $N_X(R)$ count connected unoriented simple curves. For $0 < h \leq 1/2$, choose smooth $0 \leq f_h \leq 1$ equal to one on $[0, 1-h]$ and zero on $[1 + h, \infty)$. Suppose, with specified $\eta > 0$ and $m \geq 0$, a future argument proves uniformly in R, h that*

$$\sum_{\alpha \text{ simple}} f_h(\ell_X(\alpha)/R) = c_X R^d \, d \int_0^\infty f_h(u) u^{d-1} \, \mathrm{d}u + O_X(h^{-m} R^{d-\eta}). \tag{24.22}$$

Then

$$\boxed{N_X(R) = c_X R^d + O_X(R^{d-\delta}), \qquad \delta = \min\left\{1, \frac{\eta}{m+1}\right\}.} \tag{24.23}$$

Proof. The cutoff differs from the sharp indicator only in a shell of width hR. Equation (24.21) bounds this loss by $O_X(hR^d + R^{d-1})$. The normalized integral in (24.22) differs from one by $O(h)$. Hence the total error is

$$O_X(hR^d + R^{d-1} + h^{-m} R^{d-\eta}).$$

Choose $h = R^{-\delta}$ for sufficiently large R, with δ as in the statement. □

This criterion identifies the necessary analytic input and its explicit payoff. The present coefficient estimates do not establish (24.22). In particular, fixed-target first-atom asymptotics must acquire quantitative control of the sum over growing connected supports before they imply a counting remainder.

24.8 The remaining analytic step

The two limiting regimes must be distinguished. The first-atom theorem uses $t \downarrow 0$ for $G_\alpha(t)$, or a large positive Laplace parameter. A positive sum over simple curves detects growing lengths through $Z_X^{\mathrm{simple}}(s)$ as $s \downarrow 0$. With the denominator Gaussian $e^{-L^2/(4t)}$, lengths of order R correspond to t of order R^2. These are different asymptotic problems, and the latter positive measure has not been constructed by summing the present signed coefficient measures.

The support polynomial (24.16) specifies a useful aggregate target. Its finite Fourier extraction is stable, and its primitive inversion is exact. A positive-axis Laplace remainder such as (24.20) does not by itself justify an equally strong sharp counting error; quantitative desmoothing or suitable complex-parameter estimates are still required. Proposition 24.5 gives one sufficient condition with an explicit loss. The unresolved task is to construct and bound that smoothed connected-support observable using the coefficient relations, including their pants cancellations, before taking absolute values. No improved connected-simple-curve exponent is claimed here.

25 Global coefficient filters and a length-valued Euler product

Continue throughout with a closed hyperbolic surface of genus at least two; the pants cuffs below are essential curves, not specialized peripheral trace constants. There are two concrete ways to advance the preceding construction. First, an explicit bounded coefficient functional cancels every one-self-intersection primitive geodesic together with all its iterates, simultaneously over all pants. Second, a product in the length-convolution algebra gives a unit first atom for every weighted multicurve, including those with disconnected support. Neither result proves a new counting remainder. They also reveal why a fixed linear functional cannot remove every nonsimple primitive curve.

25.1 An explicit filter for the entire one-crossing sector

Write $n(C) = \#C$, counting parallel copies, and $n(\varnothing) = 0$. This differs from the number $r(C)$ of distinct supports in Section 24.6. We seek a functional

$$\Lambda(T_C) = w_{n(C)}$$

which preserves simple primitive curves and annihilates the complete iterate family of every primitive figure-eight.

For $0 \le \theta \le \pi$, put $v = 2\cos\theta$ and let u_+, u_- be the two roots of

$$u^2 - u = v.$$

The unordered pair of roots is unambiguous, including where the discriminant vanishes. Define

$$w_n = \frac{1}{\pi}\int_0^\pi \left(u_+(\theta)^n + u_-(\theta)^n\right)\,\mathrm{d}\theta, \qquad n \ge 0. \tag{25.1}$$

In particular $w_0 = 2$ and $w_1 = 1$. If $v \ge -1/4$, the roots are real and lie in $[-1, 2]$. If $v < -1/4$, they are complex conjugates of modulus $\sqrt{-v} \le \sqrt{2}$. Thus

$$|w_n| \le 2^{n+1}, \qquad |\Lambda(f)| \le 2\|f\|_2. \tag{25.2}$$

Here $\|\cdot\|_2$ is the weighted coefficient norm from Section 23, not a norm obtained by evaluating characters.

The weights have a closed generating function. The root sums $S_n(v) = u_+^n + u_-^n$ satisfy $S_0 = 2$, $S_1 = 1$, and $S_n = S_{n-1} + vS_{n-2}$, whence

$$\boxed{\sum_{n\ge0} w_n q^n = \frac{2-q}{\sqrt{(1-q)^2 - 4q^4}}, \qquad |q| < \frac{1}{2}.} \tag{25.3}$$

The square root has value one at zero. To derive the formula, average $(2-q)/(1-q-vq^2)$ over $v = 2\cos\theta$ and use the elementary integral of $(a - b\cos\theta)^{-1}$. The first weights are

$$2,\ 1,\ 1,\ 1,\ 5,\ 11,\ 19,\ 29,\ 53, \ldots.$$

They are nonnegative integers: the coefficients of $S_n(v)$ are nonnegative integers and the averages of v^{2j} and v^{2j+1} are $\binom{2j}{j}$ and zero, respectively.

Proposition 25.1 (A global one-crossing filter)**.** *For every simple primitive curve α, $\Lambda(T_\alpha) = 1$. For every primitive geodesic γ with exactly one transverse self-intersection and every $k \ge 1$,*

$$\boxed{\Lambda(T_{\gamma^k}) = 0.} \tag{25.4}$$

Proof. Use its pants relation $T_\gamma = xy - z$, where x, y, z are the unnegated cuff traces. Replacing every trace monomial of total degree n by u^n sends $P_k(xy - z)$ to $P_k(u^2 - u)$. This remains true if two cuff images coincide: the identification preserves total degree, including parallel multiplicities. Since $u_\pm^2 - u_\pm = 2\cos\theta$,

$$\Lambda(T_{\gamma^k}) = \frac{2}{\pi}\int_0^\pi P_k(2\cos\theta)\ \mathrm{d}\theta = \frac{4}{\pi}\int_0^\pi \cos(k\theta)\ \mathrm{d}\theta = 0.$$

The simple-curve assertion is $w_1 = 1$. □

This is a single global functional, so no choice of a pants decomposition and no sum of separately normalized pants projections is required. It is not evaluation at a flat representation having the same trace on every simple curve; no such assertion is used.

Equation (23.15) and (25.2) justify the convergent aggregate

$$\mathcal{H}_X^\Lambda(t) = \sum_C w_{n(C)} H_C^X(t), \qquad \Re t > 0. \tag{25.5}$$

It is holomorphic in that half-plane. Since $\Lambda(1) = 2$, the constant identity contribution must also be multiplied by two if one applies the functional before subtracting that contribution. Absolute convergence of the orbital weighted-norm estimate permits applying Λ before the orbital sum. Every one-crossing primitive and all its powers therefore disappear. Simple iterates remain with the explicit coefficients $\Lambda(P_k(T_\alpha))$; their first values are $1, -3, -2, 5, 11, -6, \ldots$. Primitives with at least two self-intersections also remain.

If the higher symmetric-power twisting spectra used for primitive extraction in Section 23 are supplied, let

$$\Pi_\rho^X = \sum_{\delta \in \mathcal{P}_{\mathrm{un}}} T_\delta(\rho)\delta_{\ell_\delta}.$$

This is the primitive measure from that section divided by its positive length variable. Then the filtered primitive measure satisfies

$$\Lambda(\Pi_\rho^X) = \sum_{\alpha \text{ simple}} \delta_{\ell_\alpha} + \mathcal{R}_X, \qquad \mathcal{R}_X = \sum_{\substack{\delta \in \mathcal{P}_{\mathrm{un}} \\ i(\delta,\delta) \geq 2}} \Lambda(T_\delta)\delta_{\ell_\delta}. \tag{25.6}$$

Here $i(\delta, \delta)$ is the minimal self-intersection number. The remainder is signed. Its Gaussian transforms converge by the same norm estimate; no smaller-order counting bound for it has been established.

25.2 Why a fixed linear filter cannot finish the job

There is an exact obstruction already in one pair of pants. Let a, b generate its fundamental group and put

$$\gamma_1 = a^2ba^{-1}b^{-1}, \qquad \gamma_2 = a^3b^2, \qquad \gamma_3 = a^2bab.$$

Then, as functions on the SL_2 character variety,

$$\boxed{T_a = T_{\gamma_1} + T_{\gamma_2} - T_{\gamma_3}.} \tag{25.7}$$

For a direct verification set $x = T_a$, $y = T_b$, $z = T_{ab}$ and $k = x^2 + y^2 + z^2 - xyz - 2$. Cayley–Hamilton gives

$$\begin{aligned} T_{\gamma_1} &= x(k-1), \\ T_{\gamma_2} &= (x^2-1)yz - xy^2 - x^3 + 3x, \\ T_{\gamma_3} &= xz^2 - yz - x. \end{aligned}$$

The indicated combination is x.

All three words are primitive in the geodesic sense of not being proper powers: their homology vectors are $(1,0)$, $(3,2)$, and $(3,2)$. They are nonperipheral cyclically reduced words in the pants group, so their geodesics are nonsimple. Embedded geodesic pants are convex; these intersections persist in the ambient surface. An ambient proper root would traverse the same geodesic image inside the pants and give a root in the pants group, which the homology vectors exclude. Here the word primitive does not assert membership in a free basis of the pants group.

Consequently, any linear functional L satisfying $L(T_a) = 1$ must satisfy

$$\max_{1\le j\le 3} |L(T_{\gamma_j})| \ge \frac{1}{3}. \tag{25.8}$$

In particular, no linear functional on this character algebra preserves every primitive simple trace and annihilates every primitive nonsimple trace. The assertion applies also to cutoff-dependent functionals once the same functional is applied to all four labels. It does not exclude functionals depending on each atom's length, nonlinear operations, or richer representation data. Our explicit filter has values $-5, 5, -1$ on these three words, in agreement with the identity.

25.3 Retaining lengths: an Euler product with a unit diagonal

To preserve the information discarded by a fixed character functional, work in the convolution algebra of locally finite length measures, with $\delta_u * \delta_v = \delta_{u+v}$. Its coefficients lie in the character algebra and its unit is $\mathbf{1} = T_\varnothing \delta_0$. Define

$$\mathbf{E}_X(\rho) = \prod_{\delta\in\mathcal{P}_{\mathrm{un}}} \left(\mathbf{1} - T_\delta(\rho)\delta_{\ell_\delta}\right)^{*-1}. \tag{25.9}$$

This is a product of scalar trace geometric series, not the determinant Selberg product. The positive systole makes every length cutoff finite. Equivalently,

$$\mathbf{E}_X(\rho) = \sum_{\mathfrak{m}} \left(\prod_{\delta} T_\delta(\rho)^{m_\delta}\right) \delta_{\sum_\delta m_\delta \ell_\delta} = \sum_C T_C(\rho)\nu_C^X,$$

where $\mathfrak{m}$ runs over finite multisets of primitive unoriented geodesics, including the empty multiset.

Proposition 25.2 (A unit first atom for every multicurve)**.** *For every integral weighted multicurve C, including the empty one,*

$$\boxed{\nu_C^X = \delta_{\ell_X(C)} + \tau_C^X, \qquad \operatorname{supp}\tau_C^X \subset (\ell_X(C), \infty).} \tag{25.10}$$

The first atom is isolated.

Proof. Resolve the crossings of the union of geodesics in an input multiset. Every resulting multicurve has total length at most the input length. If some primitive component is nonsimple, or two distinct components intersect, resolving a genuine transverse crossing and tightening gives strict shortening, as in the earlier length-ordering argument. Repeated components can be separated by small perturbations; a transverse crossing already present in a component or between distinct geodesics retains a strict shortening.

Equality is therefore possible only when all primitive components are simple and mutually disjoint, allowing repeated copies. Their product is then exactly T_C, and the unique multiset giving this C has coefficient one. The geometric series in (25.9) supplies no factorial. Local finiteness isolates the first length. The empty multiset gives the empty multicurve's unit atom at zero. □

25.4 Absolute convergence and spectral input

The fixed-coefficient bracelet estimate extends to multicomponent diagrams. With $t(u) = 2\cosh(u/2)$ and the same $K_{C,X}$ as before,

$$\left| [T_C] \prod_{\delta} T_\delta^{m_\delta} \right| \le K_{C,X} \prod_{\delta} t(\ell_\delta)^{m_\delta}.$$

Indeed, sign coherence and positive Fuchsian evaluation apply to the entire diagram, whose absolute evaluation is that product. It follows that

$$\int e^{-sL} \, \mathrm{d}|\nu_C^X|(L) \le K_{C,X} \prod_{\delta \in \mathcal{P}_{\mathrm{un}}} \left(1 - t(\ell_\delta) e^{-s\ell_\delta}\right)^{-1} < \infty \tag{25.11}$$

for real

$$s > s_X^* = \max\left\{ \frac{3}{2}, \frac{\log(2\cosh(\varepsilon/2))}{\varepsilon} \right\}, \qquad \varepsilon = \mathrm{sys}(X). \tag{25.12}$$

The second condition makes every geometric-series ratio less than one: $\log t(u)/u$ decreases in u. The first makes the sum of those ratios converge using the elementary $O_X(e^R)$ primitive-orbit bound. This also proves coefficient holomorphy for $\Re s > s_X^*$ and convergence of every coefficient Gaussian transform for $\Re t > 0$.

Recovering this product spectrally requires its higher twisting data to be kept straight. For a representation V, write $\Pi_V = \sum_\delta \chi_V(\delta)\delta_{\ell_\delta}$ and let $D_m \delta_L = \delta_{mL}$. The precise reconstruction formula is

$$\boxed{\log_* \mathbf{E}_X(\rho) = \sum_{m \ge 1} \frac{1}{m} D_m \Pi_{\rho^{\otimes m}}.} \tag{25.13}$$

The tensor-power character is T_δ^m. It must not be replaced by $P_m(T_\delta) = \operatorname{tr} \rho(\delta^m)$. Tensor powers decompose into finite sums of SU(2) symmetric powers, and the earlier Adams inversion recovers their primitive measures from the corresponding higher-twist spectra. Thus all those twisting spectra suffice for (25.13). A single rank-two primitive trace measure does not directly provide the required power sums at coincident lengths.

25.5 What these constructions settle, and what remains

The global filter gives an actual convergent aggregate with the entire one-crossing sector removed. Equation (25.7) proves that eliminating all nonsimple primitives by a fixed linear character functional is impossible. The length-valued product avoids that obstruction: its first-atom extraction is nonlinear and retains the source lengths.

In particular, the positive support partition function of Section 24.6 has the exact reconstruction

$$\mathcal{P}_X(s, z) = \sum_C z^{r(C)} \exp\left[-s \, \inf \operatorname{supp} \nu_C^X \right], \qquad \Re s > 0.$$

The unit first atoms make this identity unambiguous. It is a reconstruction statement, not a remainder estimate. Quantitative counting now requires either control of the residual $\mathcal{R}_X$ in (25.6), using cancellations among higher-intersection curves, or uniform control of the nonlinear diagonal extraction from $\mathbf{E}_X$. Neither follows from fixed-coefficient convergence alone.

Part V
Mapping-class dynamics and mixing

Weighted endpoint counts encode correlations. The general converse specifies the required estimates, while the marked-torus model supplies an explicit calculation on one genus-two stratum.

26 When counting errors imply mixing

A counting remainder can imply mixing when the counts determine an appropriate operator, or when they retain sufficiently uniform information about both endpoints of orbit segments. The following elementary converses make these hypotheses and their quantitative losses explicit. They do not identify the simple-curve counting function of Section 24 with such an operator.

26.1 A direct converse in the compact Selberg setting

There is a particularly close answer to the original trace-formula question. Let $X = \Gamma\backslash\mathbb{H}^2$ be a closed connected oriented surface of curvature -1. Let $\pi_X(L)$ count primitive hyperbolic conjugacy classes in Γ of length at most L. Thus the two orientations are counted separately: $[\gamma]$ and $[\gamma^{-1}]$ are distinct. This is the primitive-class convention in the Selberg product; see [7, Section 2.1].

Proposition 26.1 (A prime-geodesic remainder forces a mixing rate)**.** *Suppose that, for some* $0 < \delta < \frac{1}{2}$,

$$\pi_X(L) = \operatorname{Li}(e^L) + O(e^{(1-\delta)L}). \tag{26.1}$$

Then the first positive Laplace eigenvalue satisfies

$$\lambda_1(X) \geq \delta(1-\delta). \tag{26.2}$$

For every $0 < r < \delta$*, the unit-speed geodesic flow* g_t *on* T^1X *therefore satisfies*

$$|\operatorname{Corr}_t(f,h)| \leq C_{X,r} S_m(f) S_m(h) e^{-rt} \tag{26.3}$$

for some finite Sobolev order m*. Here* S_m *is an* L^2 *Sobolev norm and the correlation uses normalized Liouville measure.*

Proof. Put $\alpha = 1 - \delta > \frac{1}{2}$. For $\Re s > 1$, define

$$F(s) = \sum_{P \text{ primitive}} \ell(P) e^{-s\ell(P)} = \int_0^\infty t e^{-st}\, d\pi_X(t).$$

Above a fixed positive cutoff L_0, write $\pi_X(t) = \operatorname{Li}(e^t) + E(t)$ with $E(t) = O(e^{\alpha t})$. Since $d\operatorname{Li}(e^t) = e^t\, dt/t$, the main term contributes $e^{-(s-1)L_0}/(s-1)$. Integration by parts gives the error contribution, up to an entire cutoff term, as

$$\int_{L_0}^\infty (st-1) e^{-st} E(t)\, dt.$$

This integral is locally uniformly convergent for $\Re s > \alpha$. Consequently $F(s) - 1/(s-1)$ is holomorphic there.

The logarithmic derivative of the Selberg product is

$$\frac{Z_X'(s)}{Z_X(s)} = \sum_P \sum_{k\ge 1} \frac{\ell(P) e^{-sk\ell(P)}}{1 - e^{-k\ell(P)}}. \tag{26.4}$$

Subtracting $F(s)$ leaves a function holomorphic for $\Re s > \frac{1}{2}$: the terms with $k \ge 2$ have exponent at least $2\Re s > 1$, and the $k = 1$ denominator correction supplies an extra factor $e^{-\ell(P)}$. The positive systole and the counting bound justify local uniform convergence. Hence $Z_X'/Z_X - 1/(s-1)$ is holomorphic for $\Re s > \alpha$, excluding zeros of Z_X there except at 1.

Every Laplace eigenvalue $\lambda_j = s_j(1-s_j) < \frac{1}{4}$, with $\frac{1}{2} < s_j < 1$, gives a zero s_j of Z_X; see [8, Theorems 1.1 and 4.1]. Thus $s_j \le 1 - \delta$, which proves (26.2); equality is allowed. Set $s_* = \frac{1}{2}$ if there is no such positive eigenvalue, and otherwise let s_* be the largest such parameter. The matrix-coefficient estimate of [9, Theorem 1.1], specialized to a compact surface, gives Sobolev correlation decay at every rate below $1 - s_*$. Since $s_* \le 1 - \delta$, this proves (26.3). The time normalization is $a_t = \operatorname{diag}(e^{t/2}, e^{-t/2})$ in $\mathrm{PSL}_2(\mathbb{R})$. □

Two qualifications delimit this converse. First, (26.1) contains only the leading term: an error estimate *after subtracting* exceptional contributions $\operatorname{Li}(e^{s_j L})$ does not exclude those small eigenvalues. Second, the conclusion concerns the complete primitive spectrum. Applying it to simple curves or coefficient-filtered counts would require a corresponding spectral identity. If geometric geodesics are counted without orientation, the main term is $\frac{1}{2}\operatorname{Li}(e^L)$; doubling that count recovers the stated convention. The exponent is quantitative, while explicit numerical or family-uniform mixing constants require additional estimates.

26.2 A quantitative converse in homogeneous dynamics

There is an established arithmetic converse. In the S-arithmetic setting of Gorodnik–Nevo [10, Theorems 1.4 and 5.2], let $E_{S,m}(r, B)$ denote the L^2 discrepancy of normalized rational point counts in translated real balls of radius r and an S-adic set B. For a discretely embedded irreducible representation π in the mean-zero automorphic spectrum, their estimate is

$$\|\pi(\beta_B)\| \le 2[\Gamma_S : \Gamma_{S,m}] m_{G_\infty}(B(e,r))^{-1} E_{S,m}(r, B), \qquad 0 < r \le r_0(\pi). \tag{26.5}$$

Here the real factor is noncompact, the places in S are unramified, B is bounded, positive-measure, and bi-invariant under the specified hyperspecial compact subgroup, and β_B is its normalized Haar measure. The small-radius threshold depends on π. Thus sufficiently strong localized discrepancy controls averaging operators quantitatively. Their application to forms of SL_2 uses arithmetic counting to obtain a spectral gap. This statement is not a general implication from one scalar count to mixing of a flow.

26.3 A converse for a finite Markov model

Let A be an irreducible nonnegative integer $q \times q$ adjacency matrix, allowing parallel edges, and let $\lambda > 0$ be its Perron eigenvalue. The number F_n of periodic edge sequences fixed by the nth shift, with a distinguished starting position, is

$$F_n = \operatorname{tr}(A^n).$$

It includes repetitions and is not the number of primitive orbits. The associated identity

$$\zeta_A(z) = \exp\left(\sum_{n\ge 1} \frac{\operatorname{tr}(A^n)}{n} z^n\right) = \frac{1}{\det(I - zA)}$$

is the usual finite-state trace formula; see [11, Sections 1 and 4].

Proposition 26.2 (A trace-counting remainder forces a gap)**.** *Suppose that, for a specified* $\kappa > 0$*,*

$$\operatorname{tr}(A^n) = \lambda^n + O\big((\lambda e^{-\kappa})^n\big). \tag{26.6}$$

Then every eigenvalue of A *other than its simple Perron eigenvalue satisfies*

$$|\lambda_j| \le \lambda e^{-\kappa}. \tag{26.7}$$

Proof. Put $\rho = \lambda e^{-\kappa}$. The power series

$$E(z) = \sum_{n \ge 1} \big(\operatorname{tr}(A^n) - \lambda^n\big) z^n$$

is holomorphic for $|z| < \rho^{-1}$. On its initial convergence disk,

$$E(z) = \sum_{j=2}^{q} \frac{z\lambda_j}{1 - z\lambda_j},$$

where algebraic multiplicities are included. A nonzero λ_j produces a pole at $z = \lambda_j^{-1}$. Distinct eigenvalues produce different poles, and repeated copies of one eigenvalue have the same nonzero residue and add. No pole can therefore lie inside $|z| < \rho^{-1}$. Zero eigenvalues already satisfy the conclusion. □

To convert the gap into a mixing statement, choose positive Perron vectors r, l normalized by $\sum_i l_i r_i = 1$, and set

$$P_{ij} = \frac{A_{ij} r_j}{\lambda r_i}, \qquad \pi_i = l_i r_i, \qquad \Pi = \mathbf{1}\pi^T.$$

Then P is stochastic, π is stationary, and P is similar to A/λ. For every $0 < \kappa' < \kappa$,

$$\|P^n - \Pi\|_\infty \le C_{\kappa', A} e^{-\kappa' n}. \tag{26.8}$$

At the endpoint κ a polynomial factor can arise from Jordan blocks. In particular, the stationary Markov chain satisfies

$$|\mathbb{E}_\pi[f(X_0) g(X_n)] - \mathbb{E}_\pi f \, \mathbb{E}_\pi g| \le C_{\kappa', A} e^{-\kappa' n} \|f\|_\infty \|g\|_\infty.$$

Fixed-width cylinder observables have the same decay after the corresponding time offset. Approximation by cylinder functions gives mixing of the shift, and quantitative approximation gives a rate for Hölder observables with the additional regularity loss.

The matrix constant can itself be bounded explicitly. If $r_0 = e^{-\kappa} < r < 1$, the resolvent integral around $|z| = r$ gives

$$\|P^n - \Pi\|_\infty \le \frac{r \, q! (1 + r)^{q-1}}{(1 - r)(r - r_0)^{q-1}} \, r^n. \tag{26.9}$$

Indeed, on this circle the determinant has modulus at least $(1 - r)(r - r_0)^{q-1}$, while the row-sum norm of its adjugate is at most $q!(1 + r)^{q-1}$. Thus a numerical counting saving produces a numerical mixing estimate in this model.

26.4 Which scalar count is being estimated?

An exact count of unrestricted paths need not imply mixing. For example, with $m \geq 2$,

$$A = \begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix}, \qquad \mathbf{1}^T A^n \mathbf{1} = 2m^n.$$

This count has no error at all. Its normalized transition chain alternates between two states; the correlation of $(1, -1)$ with itself is $(-1)^n$. The closed-path count detects what the total path count misses:

$$\operatorname{tr}(A^n) = m^n + (-m)^n.$$

Thus a scalar sequence can suffice when it is a complete trace encoding of a suitable operator. The decisive issue is what the chosen count retains.

Continuous time introduces a separate return-time issue. The full shift on m symbols has exactly m^n fixed points for its nth iterate and is mixing for its Bernoulli measure. Suspend it with constant roof 1:

$$Y = (\Sigma \times [0, 1])/((x, 1) \sim (\sigma x, 0)).$$

For the suspension flow ϕ^t, the mean-zero function $F([x, u]) = e^{2\pi i u}$ is well defined and satisfies

$$F \circ \phi^t = e^{2\pi i t} F, \qquad \int \overline{F}\, F \circ \phi^t \, \mathrm{d}\nu = e^{2\pi i t}.$$

The flow is therefore not mixing. Its physical periods lie in a lattice, so this is not an example of an unoscillating continuous-time prime-orbit asymptotic. It shows why a discrete-time counting converse does not automatically pass to a suspension. Ruling out arithmetic return times is necessary; quantitative flow estimates also require control of approximate eigenfunctions or high-frequency return-time operators. The role of these additional geometric conditions is discussed in [13].

26.5 A qualitative converse from full geodesic counting

There is also a simple geometric deduction. In the negatively curved setting of Babillot [12, Theorem 1], quasi-product measures for the geodesic flow mix if the closed length spectrum is not contained in a discrete subgroup of $\mathbb{R}$. Assume a finite such measure, normalized to probability, and suppose the full primitive closed-orbit count is finite at each cutoff and satisfies

$$\pi(T) \sim C\, \frac{e^{hT}}{T}, \qquad C, h > 0,$$

as $T \longrightarrow \infty$ through all real values. If all periods belonged to $a\mathbb{Z}$, $a > 0$, then

$$\pi(na + a/4) = \pi(na + 3a/4),$$

whereas the ratio of their proposed main terms tends to $e^{ha/2} > 1$. This contradiction proves nonarithmeticity and hence qualitative mixing under the stated structural hypotheses. No rate follows from this argument. For a polynomial simple-curve main term CT^d, the corresponding fixed-shift ratio tends to one, so this argument does not transfer to simple-curve counting.

26.6 Localized counts and correlations

A useful general converse applies once geometric counts are related to joint transition masses. Let (Y, d, μ) be a compact metric probability space, with a measure-preserving flow ϕ^t. Suppose that for every $0 < h \leq 1$ there is a measurable partition $\mathcal{P}_h = \{P_1, \ldots, P_{N(h)}\}$ with

$$\operatorname{diam}(P_i) \leq h, \qquad N(h) \leq B h^{-D}.$$

Zero-measure cells may be discarded. Define $p_{ij}(t,h) = \mu(P_i \cap \phi^{-t} P_j)$.

Proposition 26.3 (A quantitative localized converse)**.** *Assume that geometric weighted counts give, uniformly in i, j, h, t,*

$$|p_{ij}(t,h) - \mu(P_i)\mu(P_j)| \le A h^{-m} e^{-\kappa t}, \qquad m \ge 0, \quad \kappa > 0. \tag{26.10}$$

For $0 < \beta \le 1$ and bounded β-Hölder functions f, g,

$$\left| \int f\, g \circ \phi^t \, \mathrm{d}\mu - \int f \, \mathrm{d}\mu \int g \, \mathrm{d}\mu \right| \le C \|f\|_{C^\beta} \|g\|_{C^\beta} e^{-\eta t}, \qquad \eta = \frac{\kappa \beta}{m + 2D + \beta}. \tag{26.11}$$

Proof. Replace f, g by their cellwise averages f_h, g_h. Their uniform approximation errors are at most $[f]_\beta h^\beta$ and $[g]_\beta h^\beta$. Invariance bounds the resulting correlation error by $C\|f\|_{C^\beta}\|g\|_{C^\beta} h^\beta$. For the step functions the centered correlation is

$$\sum_{i,j} f_i g_j (p_{ij}(t,h) - \mu(P_i)\mu(P_j)).$$

There are at most $B^2 h^{-2D}$ terms. Equation (26.10) therefore bounds the original correlation by

$$C\|f\|_{C^\beta}\|g\|_{C^\beta}(h^\beta + h^{-(m+2D)} e^{-\kappa t}).$$

Choose $h = e^{-\kappa t/(m+2D+\beta)}$. □

If the right side of (26.10) contains the additional factor $\mu(P_i)\mu(P_j)$, its sum costs one, and the exponent improves to $\kappa\beta/(m+\beta)$. The normalization of a localized count therefore materially affects the resulting rate.

The substantive hypothesis is the geometric identification of counts with p_{ij}. In an application one needs a formula of the form

$$p_{ij}(t,h) = \frac{\text{weighted segment count from } P_i \text{ to } P_j}{\text{normalizing growth}} + \text{controlled error}.$$

Jacobian weights, multiplicities, boundary effects, and time-window errors must be controlled uniformly at both endpoints and at shrinking resolution h. If the comparison error satisfies the same bound as the normalized counting error, their constants add and the proposition applies. The proposition does not supply this comparison for an arbitrary raw counting problem.

26.7 Nonescape and the counting time parameter

There is an explicit version with a noncompact tail. Suppose an exhaustion K_H satisfies $\mu(Y \setminus K_H) \le C e^{-aH}$, with $a > 0$. Assume that K_H has at most $B e^{bH} h^{-D}$ cells of diameter h, and that the entrywise bound is $A e^{cH} h^{-m} e^{-\kappa t}$, where $b, c, m \ge 0$. For bounded globally β-Hölder observables the preceding argument, extending the step approximations by zero off K_H, gives

$$|\operatorname{Corr}_t(f,g)| \le C\|f\|_{C^\beta}\|g\|_{C^\beta} \left(h^\beta + e^{-aH} + e^{(2b+c)H} h^{-(m+2D)} e^{-\kappa t} \right).$$

The tail error uses invariance and the L^1 approximation bound; no uniform approximation on the discarded region is required. Taking $h = e^{-\eta t/\beta}$ and $H = \eta t/a$ yields

$$\boxed{\eta = \frac{\kappa}{1 + (2b+c)/a + (m+2D)/\beta}.} \tag{26.12}$$

These are declared sufficient hypotheses, not estimates established here for Teichmüller flow.

Finally, under the counting-time conversion $R = e^t$, a relative power saving becomes an exponential one:

$$\frac{O(R^{d-\delta})}{R^d} = O(e^{-\delta t}).$$

This explains why the exponents in effective counting and mixing are related. The conversion alone gives no transition-mass identity. A converse for our curve-counting program would require the estimates to retain independent endpoint or sector weights, with controlled dependence on localization and on excursions into the noncompact part of the parameter space. Alternatively, an appropriate trace or transfer-operator representation could provide a spectral converse. Neither representation follows merely from a good remainder for the single unweighted simple-length count.

27 Filling-pair coefficients and a weighted mixing converse

The preceding converse criteria suggest retaining two independent sets of curve data. This section carries that proposal through three concrete steps. First, a tensor product of the coefficient construction recovers marked filling pairs, hence square-tiled differentials. Second, an exactly weighted count of Teichmüller orbit segments unfolds into a time integral of correlations. Third, an elementary de-averaging argument converts an assumed counting remainder into a mixing rate. An arithmetic torus calculation illustrates the additional cancellation needed to prove such a remainder. The identities and conditional implications below do not constitute a new proof of Teichmüller mixing.

27.1 Two independent character variables retain the pair

Keep a closed marked hyperbolic surface X as an auxiliary device, and let ρ, σ be independent rank-two character variables. The length-valued product of Section 25 can be tensored with itself, retaining two separate length coordinates:

$$\mathbf{E}_X(\rho) \boxtimes \mathbf{E}_X(\sigma) = \sum_{C,D} T_C(\rho) T_D(\sigma)\, (\nu_C^X \boxtimes \nu_D^X).$$

Here $\boxtimes$ is the external product of locally finite discrete measures, and both C and D range over integral weighted multicurves. The two variables are essential: multiplication in one character algebra would resolve crossings between C and D, whereas their ordered pair is precisely the information required here.

Proposition 27.1 (The first corner of a pair coefficient)**.** *The coefficient of $T_C(\rho)T_D(\sigma)$ is supported in*

$$[\ell_X(C), \infty) \times [\ell_X(D), \infty),$$

and has an isolated atom of mass one at $(\ell_X(C), \ell_X(D))$. Its remaining support misses this corner.

Proof. The one-variable unit-diagonal identity (25.10) gives

$$\nu_C^X = \delta_{\ell_X(C)} + \tau_C^X, \qquad \operatorname{supp} \tau_C^X \subset (\ell_X(C), \infty),$$

and similarly for D. Expanding the external product proves the claim; local finiteness isolates the corner. Each finite rectangle of source lengths involves only finitely many source configurations, so the argument also applies before taking any transform. □

After this extraction, the marked labels retain $i(C, D)$ and the topological filling condition. Thus one can restrict to filling pairs and apply weights to their associated flat differentials. The auxiliary lengths ℓ_X provide the triangular extraction; the area of the resulting differential will be $i(C, D)$, a different quantity. In particular, a bound on area is not a bound on both auxiliary hyperbolic lengths on the marked space. Passing to an area cutoff and a mapping-class quotient is an additional operation, whose quantitative cost is not supplied by fixed-coefficient convergence.

27.2 Filling pairs as actual coordinates for the flow

For definiteness let S be closed, oriented, and of genus $g \geq 2$. A pair (A, B) of nonzero measured foliations is *filling* if

$$i(A, C) + i(B, C) > 0 \qquad \text{for every nonzero measured foliation } C.$$

The condition is stronger than $i(A,B) > 0$. The realization theorem of Gardiner–Masur, also proved by Wentworth using harmonic maps, identifies filling pairs with marked nonzero holomorphic quadratic differentials [1, Theorem 1.6 and Corollary 1.7]; see also [2, Theorem 1.6] for the equivariant homeomorphism and its integral interpretation. Use the convention that $A = \mathcal{V}(q)$ is the vertical foliation with transverse measure $|dx|$, and $B = \mathcal{H}(q)$ is the horizontal foliation with transverse measure $|dy|$, in coordinates $q = dz^2$. Write $q(A,B)$ for the resulting differential. Then

$$\operatorname{Area}(q(A,B)) = i(A,B), \tag{27.1}$$

$$g_t q(A,B) = q(e^t A, e^{-t} B), \qquad g_t = \operatorname{diag}(e^t, e^{-t}). \tag{27.2}$$

In particular

$$\widehat{q}(A,B) := q\left(\frac{A}{\sqrt{i(A,B)}}, \frac{B}{\sqrt{i(A,B)}}\right)$$

has unit area. This is a genuine parametrization of points of the flow space; the two foliations are the stable and unstable data of a single differential, rather than two unrelated curves.

For an integral pair $A = \sum_i a_i \alpha_i$, $B = \sum_j b_j \beta_j$ in minimal position, there is a concrete realization. Put a rectangle of horizontal width a_i and vertical height b_j at each intersection of α_i with β_j, and glue the sides in the cyclic orders supplied by the embedded curves. The filling condition makes the complementary regions disks. Minimal position excludes bigons. A complementary $2k$-gon yields total angle $k\pi$, hence a zero of order $k-2$; $k = 2$ gives a regular point. Integral side lengths subdivide the rectangles into unit squares, with edge identifications by translations or half-turns. The total number of squares is

$$\sum_{i,j} a_i b_j i(\alpha_i, \beta_j) = i(A,B).$$

Thus the discrete objects suggested by the pair coefficients are precisely square-tiled half-translation surfaces. Translation surfaces form the subcollection in which the half-translation structure has trivial linear holonomy. Formula (27.2) is then visible by stretching and contracting the rectangles.

27.3 The correct quotient and the two different counting scales

Let $\Gamma = \operatorname{Mod}(S)$. Equivariance of the pair parametrization identifies the stabilizer of the measured pair with the finite automorphism group of its differential:

$$\operatorname{Stab}_\Gamma(A,B) = \operatorname{Stab}_\Gamma(A) \cap \operatorname{Stab}_\Gamma(B).$$

Finiteness also follows by acting on the finite filling graph $A \cup B$ and then on its complementary disks. It matters here that one fixes the measured foliations, including their scales. Stabilizers of projective foliations can contain nontrivial flow translations.

One simultaneous Γ-orbit of a filling pair represents just one point in moduli space and has constant intersection number. To obtain a growing family, prescribe the two curve types separately and take the diagonal quotient:

$$(\Gamma\alpha \times \Gamma\beta)_{\text{fill}}/\Gamma.$$

If $\Gamma_\alpha = \operatorname{Stab}_\Gamma(\alpha)$ and $\Gamma_\beta = \operatorname{Stab}_\Gamma(\beta)$, its classes are the relevant double cosets in $\Gamma_\alpha \backslash \Gamma / \Gamma_\beta$, represented by $(\alpha, \phi\beta)$. For an orbifold volume comparison the natural weight of such a class is

$$\frac{1}{|\Gamma_\alpha \cap \phi \Gamma_\beta \phi^{-1}|}.$$

This is the same stabilizer factor appearing in the square-tiled count of [2, Section 1 and Section 3].

Fix a connected component $\mathcal{Q}$ of a stratum and put $h = \dim_{\mathbb{C}} \mathcal{Q}$. The real period-coordinate dimension of its cone is $2h$, and the unit-area locus has real dimension $2h-1$. If D_r multiplies every period by r, then area is multiplied by r^2. With the appropriate period lattice and its covolume fixed, Lebesgue measure ν scales by

$$\nu(D_r E) = r^{2h}\nu(E).$$

Consequently an *area* cutoff L has natural count exponent h, whereas a common *linear* scale $R = \sqrt{L}$ has exponent $2h$. These conventions must not be interchanged when converting powers into rates in flow time.

For the closed principal quadratic stratum $h = 6g-6$. For a quadratic stratum with n specified singularities, including marked regular points, $h = 2g-2+n$ when the differential is not a square; the corresponding square locus has $h = 2g-1+n$. Lower-dimensional strata require their own period-coordinate lattice and Masur–Veech measure. The ambient product of two full-dimensional Thurston measures cannot serve as their volume form. Likewise, counting all square-tiled surfaces in a stratum does not by itself approximate an arbitrary affine invariant measure on a smaller locus.

27.4 The existing square-tiled counting theorem and the new target

The passage from pairs to square tilings is established mathematics, not a new consequence of the coefficient construction. Arana-Herrera's effective theorem counts square-tiled surfaces in a fixed connected component with prescribed horizontal and vertical multicurve types, with a power-saving error in the area cutoff [3, Theorem 1.1]. Its proof explicitly passes through Eskin–Mirzakhani–Mohammadi orbit counting and effective horosphere equidistribution, using exponential Teichmüller mixing [3, Section 3.2]. Using that estimate to recover mixing would therefore not provide an independent proof of mixing.

There is a useful distinction already visible in that paper. When one foliation type is fixed and the other is unrestricted, standard lattice counting replaces the mapping-class-orbit estimate [3, Theorem 3.9]. This provides a tractable elementary benchmark. It does not control arbitrary initial and terminal observables at long times: their weights must still be tracked through the reciprocal scaling in (27.2).

27.5 What refining the coefficient mesh actually proves

There are two independent limiting operations: approximation of a phase-space integral by rational points, and factorization of that integral under a long flow. The first can be effective without implying the second. Here is a precise local calculation. Put

$$A_t = \operatorname{diag}(e^t I_d, e^{-t} I_d), \qquad H_t(x) = F(x)G(A_t x),$$

where $F \in C_c^1(\mathbb{R}^{2d})$ has support in a fixed compact set K and $G \in C_b^1(\mathbb{R}^{2d})$. For $R \geq 1$ define

$$Q_R(t) = R^{-2d} \sum_{k \in \mathbb{Z}^{2d}} H_t(k/R), \qquad I(t) = \int_{\mathbb{R}^{2d}} H_t(x)\, dx.$$

Then, for $t \geq 0$,

$$|Q_R(t) - I(t)| \leq C_K R^{-1} e^t \|F\|_{C^1} \|G\|_{C^1}. \tag{27.3}$$

Indeed, partition space into cubes of side R^{-1}, compare the value at one corner with the integral on its cube, and use

$$\|\nabla H_t\|_\infty \le \|\nabla F\|_\infty \|G\|_\infty + e^t \|F\|_\infty \|\nabla G\|_\infty.$$

Only $O_K(R^{2d})$ cubes contribute. A fixed smooth density can be absorbed in F. To model unit-area data by an ambient lattice, one can first use a smooth radial cutoff in a compact area shell bounded away from area zero; radial normalization is then smooth. This local argument does not supply uniform estimates across arbitrarily many period-coordinate charts, their boundaries, or the cusp.

At the diagonal scale $R = e^t$, the displayed error has no decay. Taking $R = e^{\sigma t}$, with $\sigma > 1$, gives a quadrature error $O(e^{-(\sigma-1)t})$. This removes the resolution obstruction in this particular local model. It does not evaluate the remaining integral $I(t)$. For Schwartz data, Poisson summation makes the distinction exact:

$$Q_R(t) = \sum_{n \in \mathbb{Z}^{2d}} \widehat{H}_t(Rn) = I(t) + \sum_{n \neq 0} \widehat{H}_t(Rn). \tag{27.4}$$

The zero Poisson frequency is the two-time integral itself, not the product of its one-time means. Suppressing the nonzero Poisson frequencies therefore proves accurate quadrature, not mixing. On the quotient, one must still analyze the sum over changes of marking that returns $A_t K$ to a chosen collection of charts.

Even exact sampling identities do not force mixing. On $\mathbb{R}/\mathbb{Z}$, with $g_t(x) = x + t$ and $F(x) = G(x) = \cos(2\pi x)$,

$$\int_0^1 F(x) G(g_t x)\, dx = \tfrac{1}{2} \cos(2\pi t).$$

For every integer $R \ge 3$, the corresponding equal-mesh average is exactly this integral, by finite Fourier orthogonality. Both means vanish, while the correlation does not decay. This example isolates the logical issue; it is not a model of the full geometry of a stratum.

27.6 An exact weighted bisector kernel

There is a concrete discrete count whose remainder does imply mixing. The necessary weight can be specified exactly, and the passage from a cumulative count to a correlation has an elementary quantitative proof. This construction uses the full unit cotangent bundle, with its Masur–Veech measure supported on the principal stratum; it is not an automatic statement for an arbitrary lower-dimensional stratum or affine invariant measure.

Let S be closed of genus $g \ge 2$, put $h = 6g - 6$, and replace the mapping class group by a torsion-free finite-index subgroup Γ. This removes stabilizer factors. Write

$$Y = \Gamma \backslash Q^1 \mathcal{T}(S), \qquad B = \Gamma \backslash \mathcal{T}(S),$$

and normalize the invariant measure μ on Y to have mass one. Its projection m to B also has mass one. Use the same notation for the lifted measures, and disintegrate

$$\mathrm{d}\mu(q) = \mathrm{d}s_x(q)\ \mathrm{d}m(x), \qquad q \in S(x),$$

where $S(x) = \pi^{-1}(x)$ and s_x is a probability measure. The Teichmüller polar-coordinate Jacobian $J_x(q,t) > 0$ is defined almost everywhere by

$$\mathrm{d}m(\pi(g_t q)) = J_x(q,t)\ \mathrm{d}s_x(q)\ \mathrm{d}t. \tag{27.5}$$

This is the polar-coordinate formula of [4, Section 2.4, equation (2.3)]. We use the conventional Teichmüller time $g_t = \mathrm{diag}(e^t, e^{-t})$.

Let $D \subset \mathcal{T}(S)$ be a measurable fundamental domain. For $x, y \in D$ and $\gamma \in \Gamma$, let

$$\ell_\gamma = d_{\mathcal{T}}(x, \gamma y), \qquad q_\gamma = q_s(x, \gamma y) \in S(x)$$

be the length and initial unit quadratic differential of the unique Teichmüller segment. Its terminal differential, transported to $S(y)$, is $\gamma^{-1} g_{\ell_\gamma} q_\gamma$. For bounded observables F, G on Y, understood through their Γ-invariant lifts, define

$$K_T^{F,G}(x,y) = \sum_{\substack{\gamma \in \Gamma \\ 1 \le \ell_\gamma \le T}} \frac{e^{h\ell_\gamma}}{J_x(q_\gamma, \ell_\gamma)} \, F(q_\gamma) \, G\big(\gamma^{-1} g_{\ell_\gamma} q_\gamma\big). \tag{27.6}$$

The lower cutoff avoids the polar singularity at zero. The sum is finite for fixed x, y, T. Define it for $m \times m$-almost every (x, y) where the polar data are regular for all contributing γ. Such a set has full measure, since Γ is countable; values on its complement can be assigned arbitrarily. This convention makes no assertion about an exceptional prescribed pair of basepoints.

Proposition 27.2 (Exact unfolding of the weighted discrete count)**.** *For bounded F, G and $T \ge 1$,*

$$\begin{aligned} A_{F,G}(T) &:= \int_D \int_D K_T^{F,G}(x,y) \, \mathrm{d}m(y) \, \mathrm{d}m(x) \\ &= \int_1^T e^{ht} \left(\int_Y F(q) G(g_t q) \, \mathrm{d}\mu(q) \right) \mathrm{d}t. \end{aligned} \tag{27.7}$$

The same identity holds for $F, G \in L^2(\mu)$ whenever the almost-everywhere kernel is interpreted through the corresponding absolutely integrable unfolded expression.

Proof. First use nonnegative functions, so Tonelli's theorem applies. The y-integral over D and the sum over γ unfold to an integral over $\mathcal{T}(S)$. Substitute $z = \pi(g_t q)$ with $q \in S(x)$. The factor $J_x(q,t)$ from (27.5) cancels its reciprocal in (27.6), giving

$$\int_D \int_1^T \int_{S(x)} e^{ht} F(q) G(g_t q) \, \mathrm{d}s_x(q) \, \mathrm{d}t \, \mathrm{d}m(x).$$

Disintegration gives (27.7). The integral of the absolute value is finite for bounded functions, and for L^2 functions it is at most

$$\|F\|_2 \|G\|_2 \int_1^T e^{ht} \, \mathrm{d}t$$

by invariance and Cauchy–Schwarz. Thus the signed or complex-valued identity follows as well. □

The count is genuinely discrete before the basepoint integrals are taken. It records both the initial and the terminal differential, and its geometric weight is explicit. The identity does not prove its asymptotic: that is the new counting theorem one would need.

27.7 Removing the time average with a controlled loss

The next lemma applies to any probability-preserving flow. Write $U_t G = G \circ g_t$ for its strongly continuous unitary action on $L^2(\mu)$, and let V be the infinitesimal generator, so that $\partial_t U_t G = U_t V G$ on its domain. Put

$$C_{F,G}(t) = \int_Y F \, (G \circ g_t) \, \mathrm{d}\mu, \qquad M_{F,G} = \mu(F)\mu(G).$$

There is no complex conjugation in this notation; the usual Hermitian correlation is obtained by replacing F by $\overline{F}$.

Theorem 27.3 (A cumulative counting converse)**.** *Suppose, for some $\delta > 0$, that*

$$\left| \int_1^T e^{ht} C_{F,G}(t)\, \mathrm{d}t - M_{F,G} \frac{e^{hT} - e^h}{h} \right| \le A(F,G) e^{(h-\delta)T}, \qquad T \ge 1. \tag{27.8}$$

Let $k \ge 1$, $F \in L^2(\mu)$, and $G \in \mathrm{Dom}(V^k)$. Then, for $t \ge 2$,

$$|C_{F,G}(t) - M_{F,G}| \le C_{h,\delta,k} (A(F,G) + \|F\|_2 \|V^k G\|_2) e^{-\delta k t/(k+1)}. \tag{27.9}$$

Consequently a norm-controlled remainder for the discrete kernel (27.6) gives a norm-controlled mixing estimate. For one flow derivative the exponent is $\delta/2$. With arbitrarily many controlled flow derivatives one obtains every exponent $\eta < \delta$, at the cost of a higher regularity norm. No endpoint rate δ follows from this single-test estimate alone; Theorem 27.4 below gives a stronger conclusion for generator-differentiated endpoint weights.

Proof. Let $E(T)$ be the expression inside absolute values in (27.8), and set $R(T) = e^{-hT} E(T)$. Then $|R(T)| \le A(F,G) e^{-\delta T}$ and, distributionally (in fact continuously here),

$$R'(T) + hR(T) = C_{F,G}(T) - M_{F,G}. \tag{27.10}$$

Choose a smooth, compactly supported kernel ρ on $(-1,1)$ with

$$\int \rho(u)\, \mathrm{d}u = 1, \qquad \int u^j \rho(u)\, \mathrm{d}u = 0 \quad (1 \le j < k),$$

and put $\rho_\varepsilon(u) = \varepsilon^{-1} \rho(u/\varepsilon)$. For $k > 1$ this kernel need not be nonnegative. Integration by parts in (27.10) gives, for $0 < \varepsilon \le 1$,

$$\left| \int \rho_\varepsilon(u) (C_{F,G}(t+u) - M_{F,G})\, \mathrm{d}u \right| \le C_{h,\delta,k} A(F,G) \varepsilon^{-1} e^{-\delta t}.$$

On the other hand,

$$C_{F,G}^{(k)}(t) = \int F\, U_t V^k G\, \mathrm{d}\mu, \qquad |C_{F,G}^{(k)}(t)| \le \|F\|_2 \|V^k G\|_2.$$

Taylor's formula and the vanishing moments therefore bound the difference between the smoothed correlation and its value at t by $C_k \varepsilon^k \|F\|_2 \|V^k G\|_2$. Choosing

$$\varepsilon = e^{-\delta t/(k+1)}$$

proves (27.9). □

Only derivatives along the flow enter this last step. In particular, it does not differentiate the spatially stretched observable $G \circ g_t$ in an expanding transverse direction. This removes the aliasing difficulty from the *de-averaging step*; the difficulty of proving the weighted count itself remains.

A convenient sufficient counting hypothesis is the almost-everywhere bisector estimate

$$\begin{aligned} K_T^{F,G}(x,y) &= \frac{e^{hT} - e^h}{h} f_F(x) f_G(y) + \mathcal{E}_T^{F,G}(x,y), \\ f_F(x) &= \int_{S(x)} F\, \mathrm{d}s_x, \qquad f_G(y) = \int_{S(y)} G\, \mathrm{d}s_y, \\ |\mathcal{E}_T^{F,G}(x,y)| &\le C(x,y)\, \mathcal{N}(F) \mathcal{N}(G) e^{(h-\delta)T}, \qquad C \in L^1(D \times D, m \times m). \end{aligned} \tag{27.11}$$

Integrating gives (27.8) with $A(F,G) = \|C\|_1 \mathcal{N}(F)\mathcal{N}(G)$. A direct estimate for the averaged kernel is sufficient and is weaker than this pointwise hypothesis. Uniformity over the observables, the two basepoints, and all large real values of T is part of the assertion; one unweighted total count does not supply it.

A flow eigenfunction checks that the weighted hypothesis detects an actual spectral obstruction. If an admissible nonzero test G satisfies $U_t G = e^{i\omega t} G$, $\mu(G) = 0$, and $\|G\|_2 = 1$, take $F = \overline{G}$. Then exact unfolding gives

$$A_{F,G}(T) = \frac{e^{(h+i\omega)T} - e^{h+i\omega}}{h + i\omega},$$

whose size is comparable to e^{hT} for large T. It violates every positive-saving estimate in (27.8). This test is available only if the asserted counting estimate covers the relevant observables; an unweighted count does not test this eigenmode.

27.8 A derivative endpoint weight preserves the full exponent

For a counting estimate uniform over endpoint observables, there is a stronger converse than temporal smoothing. One can differentiate the *endpoint test* inside the discrete count and use an exact identity, without differentiating an asymptotic formula.

Theorem 27.4 (A counting converse with no exponent loss)**.** *Use the notation of Proposition 27.2. Let $F \in L^2(\mu)$ and $G \in \mathrm{Dom}(V)$, and suppose that*

$$H_G = (h + V)G$$

is an admissible terminal test for the weighted counting estimate. Assume, for some $0 < \delta \le h$, that

$$\left| A_{F,H_G}(T) - \mu(F)\mu(H_G)\frac{e^{hT} - e^{h}}{h} \right| \le A(F, H_G) e^{(h-\delta)T}, \qquad T \ge 1. \tag{27.12}$$

Then the centered correlation $\mathrm{Corr}_{F,G}(T) = C_{F,G}(T) - \mu(F)\mu(G)$ satisfies

$$\begin{aligned} |\operatorname{Corr}_{F,G}(T)| \le{}& A(F, (h+V)G) e^{-\delta T} \\ &+ e^{-h(T-1)} |\operatorname{Corr}_{F,G}(1)|. \end{aligned} \tag{27.13}$$

In particular, the counting saving δ gives the same mixing exponent δ.

Proof. Invariance of μ gives $\mu(VG) = 0$, because the integral is a continuous functional on $L^2(\mu)$ and $\partial_t U_t G = U_t V G$. Hence $\mu(H_G) = h\mu(G)$. The exact unfolding and the fundamental theorem of calculus give

$$\begin{aligned} A_{F,(h+V)G}(T) &= \int_1^T e^{ht} (h C_{F,G}(t) + C'_{F,G}(t)) \, \mathrm{d}t \\ &= e^{hT} C_{F,G}(T) - e^{h} C_{F,G}(1). \end{aligned} \tag{27.14}$$

Subtracting the main term in (27.12) therefore gives exactly

$$E_{F,(h+V)G}(T) = e^{hT} \operatorname{Corr}_{F,G}(T) - e^{h} \operatorname{Corr}_{F,G}(1). \tag{27.15}$$

Divide by e^{hT} and use the counting remainder. □

For example, if the count is uniform in a norm $\mathcal{N}$ with

$$A(F,H) \le C\mathcal{N}(F)\mathcal{N}(H),$$

the norm cost is explicit:

$$\begin{aligned} |\operatorname{Corr}_{F,G}(T)| \le C\mathcal{N}(F)\big(h\mathcal{N}(G)+\mathcal{N}(VG)\big)e^{-\delta T} \\ + e^{-h(T-1)}\|F-\mu(F)\|_2\|G-\mu(G)\|_2. \end{aligned} \tag{27.16}$$

For a suitable smooth norm on compactly supported observables, this is one additional derivative along the flow in the terminal test. On a noncompact space, any growth of the coefficients of V in the chosen norm must be retained in that estimate. The terminal weight $(h+V)G$ can have either sign, so the counting hypothesis must accommodate signed weights with a quantitative norm bound.

The distinction from Theorem 27.3 is precise. If only the primitive estimate for the particular pair (F,G) is available, temporal smoothing gives $\delta k/(k+1)$ with k flow derivatives. If the same weighted counting theorem also applies to $(F,(h+V)G)$, the exact identity above preserves δ. Thus exponent loss is not intrinsic to the converse for a sufficiently rich family of endpoint weights.

In operator notation, the weighted time integral is governed by $h+V$; inserting $(h+V)G$ removes this integral by an exact boundary evaluation. No spectral gap is needed for this algebraic step. The unresolved input is the norm-controlled remainder for the corresponding discrete weighted count.

There is a corresponding improvement of the conditional cusp calculation. Suppose bounded observables can be cut off at height H with correlation and mean errors $O(e^{-aH})$. Suppose the counting bound for the cutoff pair

$$(F_H,(h+V)G_H)$$

has constant $O(e^{bH})$, including the derivative of the cutoff. Then (27.13) gives, after controlling the boundary term,

$$|\operatorname{Corr}_{F,G}(t)| \ll e^{bH-\delta t}+e^{-aH}+e^{-ht}.$$

For $a>0$, $b\ge 0$, and $0<\delta\le h$, choose $H=\delta t/(a+b)$ to obtain

$$\eta=\frac{\delta a}{a+b}. \tag{27.17}$$

As before, unbounded observables require a norm-controlled tail estimate in place of a bound on discarded measure alone.

27.9 Comparison with the usual bisector count

The exact weight $e^{ht}/J_x(q,t)$ cannot simply be omitted. To compare normalizations, let μ_0 be the local-product normalization of [4, Section 2.2] and put $V_0=\mu_0(Y)$. Our probability normalization is $\mu=\mu_0/V_0$. The conditional probability measures s_x are unchanged, whereas $m=m_0/V_0$ and $J_x=J_{0,x}/V_0$. Use the canonical Hubbard–Masur factors $\lambda^{\pm}$ belonging to μ_0. For endpoints in a fixed compact subset of the principal stratum, and for segments spending at least half their time in a suitable larger compact set, [4, Proposition 2.5] then gives

$$J_x(q,t)=V_0^{-1}e^{ht}\lambda^-(q)\lambda^+(g_tq)+O(e^{(h-\alpha)t}). \tag{27.18}$$

If the endpoint factors are bounded away from zero on the sets under consideration, this yields

$$\frac{e^{ht}}{J_x(q,t)}=\frac{V_0}{\lambda^-(q)\lambda^+(g_tq)}+O(e^{-\alpha t}). \tag{27.19}$$

Thus the natural approximation is an ordinary bisector count with the two endpoint observables divided by the corresponding Hubbard–Masur factors, with the factor V_0. Usual bisector counts often use the direction pointing back from the terminal point; that convention replaces the terminal differential here by its negative and requires the corresponding substitution in G. There are three distinct estimates to check: the weighted bisector remainder, the accumulated error in replacing the exact Jacobian, and the contribution of segments failing the recurrence or compactness conditions. Formula (27.18) is not a global uniform asymptotic.

This optional comparison also has a dependency worth making explicit: the proof of [4, Proposition 2.5 and Section 3.2] uses Forni's contraction estimate for the Hodge norm of the Kontsevich–Zorich cocycle, quoted there as Theorem 3.3. The exact Jacobian unfolding and the elementary de-averaging theorem above do not use that estimate. Invoking the Hubbard–Masur asymptotic would, however, retain this part of the existing cocycle machinery.

Effective bisector counts are already part of [5, Section 10]; those results are obtained from exponential mixing. They locate the strength of the required counting statement, but do not give an independent proof of mixing when inserted into the converse above. An independent derivation from coefficient distributions or a different counting mechanism would have to establish the required weights and uniformity.

For completeness, the noncompact loss for the single-test smoothing argument of Theorem 27.3 can also be computed. For bounded observables, a discarded mass $O(e^{-aH})$ gives a correlation cutoff error of the same order by invariance. For general L^2 observables one must instead assume a corresponding norm-controlled tail estimate. Suppose cutoffs at height H have this error, the counting constant for cutoff observables grows as e^{bH}, and their kth flow-derivative bounds are uniform. That smoothing proof gives an error of the form

$$\varepsilon^k + \varepsilon^{-1} e^{bH-\delta t} + e^{-aH}.$$

Balancing the three terms yields

$$\eta_k = \frac{\delta}{1 + b/a + 1/k}. \tag{27.20}$$

This last formula is conditional on the stated cutoff estimates. If the product of the relevant kth derivative norm and source L^2 norm grows instead as e^{dH}, the same calculation gives

$$\eta_k = \frac{\delta}{1 + b/a + (1 + d/a)/k}.$$

For fixed growth rates with $d = 0$, increasing k approaches $\delta a/(a + b)$. These formulae do not establish any of the required cusp or cutoff estimates.

27.10 A completely explicit arithmetic two-endpoint calculation

The modular case exhibits the additional cancellation that a two-endpoint count must detect. If

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}), \qquad c > 0,$$

then

$$\gamma = \begin{pmatrix} 1 & a/c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1/c \\ c & 0 \end{pmatrix} \begin{pmatrix} 1 & d/c \\ 0 & 1 \end{pmatrix}. \tag{27.21}$$

Modulo left and right integral unipotents, the two coordinates are $a/c, d/c \pmod 1$, with the coupled congruence $ad \equiv 1 \pmod c$. They are initial and terminal horocyclic coordinates; the

corresponding cusp images are $\gamma\infty = a/c$ and $\gamma^{-1}\infty = -d/c$. They should not be identified, without an additional unfolding argument, with arbitrary phase-space flow boxes.

Let $e(x) = e^{2\pi i x}$ and, for periodic test functions f, g, set

$$D_c(f,g) = \frac{1}{\varphi(c)} \sum_{\substack{a \bmod c \\ (a,c)=1}} f(a/c)g(\overline{a}/c), \qquad a\overline{a} \equiv 1 \pmod{c}.$$

Writing $f(x) = \sum_m \widehat{f}(m)e(mx)$ and similarly for g gives the exact identity

$$D_c(f,g) - \widehat{f}(0)\widehat{g}(0) = \frac{1}{\varphi(c)} \sum_{(m,n)\neq(0,0)} \widehat{f}(m)\widehat{g}(n)S(m,n;c), \tag{27.22}$$

where

$$S(m,n;c) = \sideset{}{^*}\sum_{a \bmod c} e\big((ma + n\overline{a})/c\big).$$

The classical bound [6, Section 1.1 and equation (3.2)]

$$|S(m,n;c)| \le \tau(c)c^{1/2}(m,n,c)^{1/2}$$

therefore supplies an actual product main term and a nontrivial remainder. For example, defining

$$\mathcal{A}(f) = \sum_{m\in\mathbb{Z}} (1+|m|)^{1/2}|\widehat{f}(m)|,$$

one obtains

$$|D_c(f,g) - (\int f)(\int g)| \le \frac{\tau(c)c^{1/2}}{\varphi(c)}\mathcal{A}(f)\mathcal{A}(g) \ll_\epsilon c^{-1/2+\epsilon}\|f\|_{C^2}\|g\|_{C^2}. \tag{27.23}$$

Here $(m,n) \neq (0,0)$ implies $(m,n,c)^{1/2} \le (1+|m|)^{1/2}(1+|n|)^{1/2}$, and two integrations by parts bound $\mathcal{A}(f)$ by a constant times $\|f\|_{C^2}$. For trigonometric polynomials supported on $|m|, |n| \le M$, the sharper explicit estimate is

$$|D_c(f,g) - (\int f)(\int g)| \le \frac{\tau(c)\sqrt{cM}}{\varphi(c)}\|\widehat{f}\|_{\ell^1}\|\widehat{g}\|_{\ell^1}, \qquad M \ge 1.$$

For a prime $c = p$ and $M < p$, the numerator $\tau(c)\sqrt{cM}$ can be replaced by $2\sqrt{p}$; a frequency with exactly one zero coordinate gives the Ramanujan sum -1.

Uniformity cannot be omitted: for $f_c(x) = g_c(x) = e(cx)$, both continuous means vanish but $D_c(f_c, g_c) = 1$. The growth of the test norms in (27.23) correctly prevents a false conclusion. This is the arithmetic counterpart of the expanding-frequency problem. The result is a rigorous benchmark for endpoint-residue decorrelation. A proof of mixing of the modular geodesic flow still requires the appropriate unfolding, radial/time weights, and cusp estimates; no higher-genus analogue of the displayed Kloosterman bound has been established here.

For comparison, fix two primitive oriented torus curve types. In the corresponding matrix model, with positive oriented intersection c, the remaining residue is $a \pmod{c}$, $(a,c) = 1$, so there are exactly $\varphi(c)$ possibilities. Thus the unweighted count already has the power-saving asymptotic

$$\sum_{c\le L} \varphi(c) = \frac{3}{\pi^2}L^2 + O(L\log L).$$

Indeed, substitute $\varphi(c) = c\sum_{d|c}\mu(d)/d$ and sum first over $c = dk$; the leading term is $\frac{1}{2}L^2\sum_{d\le L}\mu(d)/d^2$ and the remainder is $O(L\log L)$. These constants belong to this oriented matrix model; no unmentioned quotient by signs or orbifold automorphism weight is being included. The exact formula $\varphi(c)$ by itself says nothing about the values of $S(m,n;c)$ for $(m,n) \neq (0,0)$.

The remaining independent estimate

We have obtained an exact encoding of the filling-pair data, an exact weighted orbit-counting identity, and a quantitative implication from a remainder for that count to mixing. What has not been established is a norm-controlled estimate of the form (27.8) from the coefficient distributions. The tensor product and the first-corner extraction do not by themselves estimate a sum over returning mapping classes. In particular, an exact or quantitative comparison between the resulting square-tiled sums and the bisector kernel has not yet been supplied.

The next analytic target can now be stated without referring vaguely to a good scalar counting error: construct that comparison and control the nonconstant two-endpoint weights in (27.6), with the exact inverse Jacobian or with a proved replacement and a controlled discarded contribution. The modular calculation shows what success looks like. Its scalar count determines the zero mode, while the Kloosterman estimate controls the coupled nonzero modes. No analogous estimate for the principal higher-genus kernel is proved here. Section 28 does close a different, explicitly defined return calculation on the exceptional marked-torus model of the nonsquare genus-two $(2, 2)$ stratum.

An independent proof of the displayed remainder would yield a new route to mixing for the specified Masur–Veech measure, with the full rate in (27.13) when derivative endpoint weights are controlled, and the cutoff rate (27.17). The single-test estimate (27.9) remains an alternative under its weaker hypothesis. It would not by itself classify invariant measures or orbit closures, nor would it establish spectral properties of a separate cocycle. Those require additional information beyond correlations of the base flow.

28 Genus two: compact polyhedra and a marked-torus model

The hyperelliptic quotient makes genus two a concrete testing ground for the counting converse. There are two useful reductions. The hyperbolic quotient gives a compact polyhedral model for every conformal base. On the nonsquare stratum with two double zeros, a second, flat quotient identifies the entire flow space, up to finite markings, with a homogeneous space of marked flat tori. The latter identification is specific to that stratum.

28.1 Six branch points and the compact polyhedron

Every genus-two Riemann surface X is hyperelliptic. Its quotient map $\pi : X \to \mathbb{P}^1$ has six branch values: the Riemann–Hurwitz formula reads $2 = 2(-2) + 6$. With compatible markings the quotient induces a holomorphic Teichmüller isometry

$$\mathcal{T}_{2,0} \cong \mathcal{T}_{0,6};$$

see, for example, [99, Corollary 1.2(3)]. Both sides have complex dimension three. The four-pointed sphere is the corresponding genus-one example.

The hyperelliptic involution is an isometry of the curvature--1 metric on X. The quotient metric is a hyperbolic sphere with *six cone angles* π, and its area is 2π. Alexandrov's hyperbolic realization theorem therefore gives a unique compact convex polyhedron in $\mathbb{H}^3$ whose intrinsic boundary metric is this quotient metric, allowing the usual degenerate case of a doubled convex polygon [100, Theorem 1.1]. There are six vertices, all with intrinsic total angle π. Conversely, take the double cover of such a cone sphere branched at its six vertices. Each cone angle becomes 2π, so the lifted metric is smooth, complete, and of curvature -1. Riemann–Hurwitz identifies its genus as two. Thus this compact-polyhedral realization is an exact model for the genus-two hyperbolic metric, with the appropriate finite labeling and marking conventions.

This metric must be distinguished from the complete hyperbolic metric on the sphere with the six branch values deleted. The latter has six cusps and area 8π; it is the metric to which the ideal-polyhedron realization applies. Passing between them uses uniformization within the same punctured conformal class. It does not identify their lengths, vertex data, or volume forms. In particular, fixed intrinsic cone angles at compact vertices are not automatically linear constraints on the dihedral angles.

28.2 An explicit formula for every quadratic differential

Choose a coordinate with all six branch values finite and write

$$X: \quad y^2 = B_6(z) := \prod_{i=1}^{6}(z - a_i). \tag{28.1}$$

The forms $\mathrm{d}z/y$ and $z\,\mathrm{d}z/y$ form a basis of $H^0(X, K_X)$. Their three symmetric products are linearly independent and, by Riemann–Roch, span $H^0(X, K_X^2)$. Hence every holomorphic quadratic differential is

$$q_X = P_2(z)\frac{\mathrm{d}z^2}{y^2} = \pi^* q_0, \qquad q_0 = \frac{P_2(z)}{B_6(z)}\,\mathrm{d}z^2, \qquad \deg P_2 \le 2. \tag{28.2}$$

In particular, every such differential is invariant under the hyperelliptic involution. At a branch point the order transforms by

$$\mathrm{ord}_p(\pi^* q_0) = 2\,\mathrm{ord}_{\pi(p)}(q_0) + 2. \tag{28.3}$$

For a generic polynomial with two distinct roots away from the branch values, this gives

$$\mathcal{Q}_2(1^4) \longleftrightarrow \mathcal{Q}_0(1^2, -1^6),$$

with finite choices of labels understood. This is the genus-two case of the hyperelliptic covering construction in [101, Section 1.1.3]. The descended flat metric has six angles π at its poles and two angles 3π at its zeros. It is different from the hyperbolic cone metric of the preceding subsection.

Pullback multiplies flat area by two. Thus the correspondence between unit-area differentials on the six-pointed sphere and on X is

$$q_0 \longmapsto \tfrac{1}{2}\pi^* q_0.$$

It commutes with $g_t = \mathrm{diag}(e^t, e^{-t})$, with no change of the time parameter. Period-lattice covolumes, finite labelings, and orbifold stabilizers must still be tracked when comparing unnormalized Masur–Veech volumes; the area factor alone is not a formula for those volumes.

28.3 The nonsquare stratum with two double zeros

Write $\mathcal{Q}_2^{\mathrm{ns}}(2,2)$ for quadratic differentials with two double zeros which are not global squares of Abelian differentials. The superscript is essential: the squares of differentials in $\mathcal{H}_2(1,1)$ have the same quadratic zero orders but form a different locus.

Formula (28.2) gives an elementary description. A double root of P_2 away from the branch values gives a global square. To obtain two double zeros without a square, the two roots must

instead be two distinct branch values, say b_1, b_2. After relabeling the remaining four branch values,

$$\begin{aligned} X: \quad & y^2 = B_4(z)(z-b_1)(z-b_2), \qquad B_4(z) = \prod_{i=1}^{4}(z-a_i), \\ & q_X = c(z-b_1)(z-b_2)\frac{\mathrm{d}z^2}{y^2} = \pi^*\left(\frac{c\,\mathrm{d}z^2}{B_4(z)}\right), \qquad c \in \mathbb{C}^*. \end{aligned} \tag{28.4}$$

The six values $a_1, \ldots, a_4, b_1, b_2$ are distinct. Downstairs $q_0 = c\,\mathrm{d}z^2/B_4(z)$ has four simple poles and is regular at the two additionally marked points b_1, b_2:

$$\mathcal{Q}_2^{\mathrm{ns}}(2,2) \longleftrightarrow \mathcal{Q}_0(-1^4, 0, 0). \tag{28.5}$$

The regular marks cannot be discarded. They are branch values of $X \to \mathbb{P}^1$, and their positions supply four real parameters.

For each fixed X, choosing the unordered pair of branch values b_1, b_2 gives $\binom{6}{2} = 15$ complex lines of nonsquare $(2,2)$ differentials in the three-dimensional cotangent space. Area normalization leaves a phase on each line. Consequently this stratum meets every conformal base, but its unit-area locus has real dimension seven, compared with eleven for the full principal unit cotangent locus. Its Masur–Veech measure is an intrinsic stratum measure, not a conditional probability obtained by restricting the principal measure to a set of measure zero.

28.4 The pillowcase cover and two marked torus points

The orientation double cover of the four-pole differential q_0 is the elliptic curve

$$E: \quad w^2 = B_4(z), \qquad \omega = \sqrt{c}\,\frac{\mathrm{d}z}{w}, \qquad \rho^* q_0 = \omega^2. \tag{28.6}$$

Choose one branch point of $E \to \mathbb{P}^1$ as the origin. The involution is then $v \mapsto -v$, and its four fixed points are $E[2]$. The two marked regular points lift to unordered pairs

$$\rho^{-1}(b_1) = \{p, -p\}, \qquad \rho^{-1}(b_2) = \{q, -q\}.$$

Their admissibility conditions are exactly

$$p, q \notin E[2], \qquad p \neq q, \qquad p \neq -q. \tag{28.7}$$

Conversely, a flat torus (E, ω) with an origin and such two points gives the sphere $E/\{\pm 1\}$ with its four poles and two regular marks. The unique connected double cover of this sphere branched at all six marked values is X, and pullback of q_0 gives (28.4). The cover is connected because its monodromy is the nontrivial homomorphism to $\mathbb{Z}/2\mathbb{Z}$ on every small loop around a branch value; six such loops satisfy the sphere relation. It is unique because those loops generate the fundamental group of the punctured sphere.

The resulting q_X is nonsquare. Indeed, its being square would force the orientation cover $E \to \mathbb{P}^1$ to factor through $X \to \mathbb{P}^1$. Both covers have degree two, but their branch sets differ: the former is branched at the four a_i, the latter also at b_1, b_2. Equivalently, the square root of B_4 is not in the function field $\mathbb{C}(z, \sqrt{B_4(z)(z-b_1)(z-b_2)})$.

Both covering maps to the sphere have degree two, so

$$\operatorname{Area}(E, \omega) = 2\operatorname{Area}(q_0) = \operatorname{Area}(X, q_X). \tag{28.8}$$

Thus the unit-area marked torus corresponds directly to the unit-area genus-two differential, with no further area scaling. The entire construction commutes with the $SL_2(\mathbb{R})$ action on flat charts. In particular its time parameter is exactly the Teichmüller time used above.

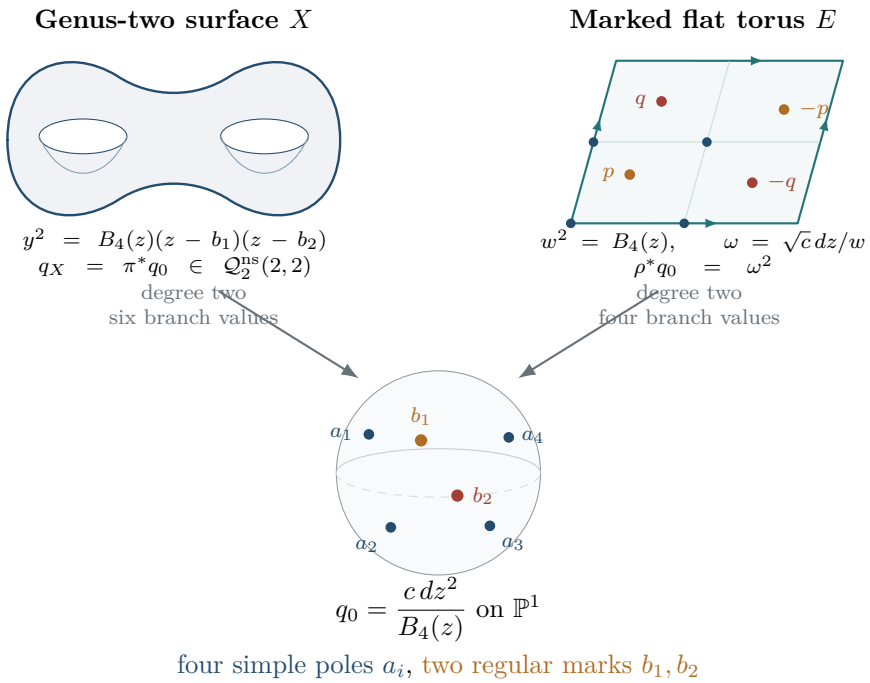


Figure 8: The two double covers used on the nonsquare genus-two stratum $\mathcal{Q}_2^{\mathrm{ns}}(2,2)$. The cover $X \to \mathbb{P}^1$ branches at all six marked values; the orientation cover $E \to \mathbb{P}^1$ branches only at the four poles. The two regular marks lift to the pairs $\{p, -p\}$ and $\{q, -q\}$. Blue points in the torus diagram represent the four classes in $E[2]$, with boundary identifications understood. This is a diagram of flat differentials, distinct from the hyperbolic six-cone-point quotient metric.

28.5 A homogeneous finite cover

Choose a full symplectic level-two structure on the torus and choose an ordering and one lift from each of the two pairs $\{\pm p\}, \{\pm q\}$. Together with the origin, level two labels the four two-torsion points and hence the four poles. Set

$$\Gamma_2 = \ker(SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/2\mathbb{Z})), \qquad G_2 = SL_2(\mathbb{R}) \ltimes (\mathbb{R}^2 \oplus \mathbb{R}^2),$$

where $SL_2(\mathbb{R})$ acts in the standard way on both vector factors. Our group law and right-quotient convention are

$$(A, u_1, u_2)(B, v_1, v_2) = (AB, u_1 + Av_1, u_2 + Av_2), \tag{28.9}$$

$$Y_2 = G_2/(\Gamma_2 \ltimes (\mathbb{Z}^2 \oplus \mathbb{Z}^2)). \tag{28.10}$$

A representative (A, u_1, u_2) describes the area-one torus $\mathbb{R}^2/A\mathbb{Z}^2$ with marked points $[u_1], [u_2]$. Right multiplication by (γ, n_1, n_2) changes the representative to

$$(A\gamma, u_1 + An_1, u_2 + An_2),$$

which is exactly change of the lattice basis preserving level two and change of the two point lifts. The presence of $-I$ in Γ_2 causes no singularity in this homogeneous quotient: right translation by a discrete subgroup on a Lie group is free. Here A and $-A$ give the same physical lattice and the same level-two data. A finer congruence level gives an additional finite cover. The spectral argument below uses level two specifically; its conclusion about the base spectrum is not automatically inherited by further covers.

Delete the locus where one of the points is two-torsion or where $[u_1] = \pm[u_2]$, and call the resulting open set Y_2°. In each torus fiber the deleted set is a finite union of real codimension-two

subtori. It has Haar measure zero and its complement is connected. Since the base $SL_2(\mathbb{R})/\Gamma_2$ is connected, so is Y_2°. Forgetting the finite marking choices gives a finite equivariant local parameterization of $\mathcal{Q}_2^{\mathrm{ns},1}(2,2)$. The finite choices include the distinguished origin over a pole, the level structure, the ordering of the two regular marks, and their signs. Their changes preserve Haar measure and commute with the $SL_2(\mathbb{R})$ action; the stratum is their finite quotient. Over the coarse-moduli locus without additional automorphisms this is an ordinary finite covering. If one retains the ubiquitous genus-two hyperelliptic involution in the orbifold stabilizers, its fixed order-two contribution must also be included in unnormalized volume conventions. No infinite marking choice is being discarded in this finite-cover statement.

The flow is left multiplication:

$$g_t[A, u_1, u_2] = [a_t A, a_t u_1, a_t u_2], \qquad a_t = \mathrm{diag}(e^t, e^{-t}). \tag{28.11}$$

The admissible set Y_2° is invariant under this action. Although the coordinates $\xi_i = A^{-1}u_i \pmod{\mathbb{Z}^2}$ are unchanged in a fixed lattice chart, a change $A \mapsto A\gamma$ changes them by $\xi_i \mapsto \gamma^{-1}\xi_i$. Therefore they are not globally fixed torus coordinates on the quotient. This monodromy is essential in the later Fourier calculation.

28.6 Haar measure and the period lattice

The group G_2 is unimodular, and its Haar measure is

$$\mathrm{d}A\ \mathrm{d}u_1\ \mathrm{d}u_2,$$

where $\mathrm{d}A$ is Haar measure on $SL_2(\mathbb{R})$. In a local lattice chart, $u_i = A\xi_i$ and $\det A = 1$, so its quotient measure is the product of modular Haar measure and Lebesgue probability on $(\mathbb{R}^2/\mathbb{Z}^2)^2$. The modular quotient has finite Haar volume and each torus fiber has volume one, hence Y_2 has finite volume.

Before area normalization, local coordinates are the two complex lattice periods and the two complex positions of the marked points. These are four complex coordinates, agreeing with

$$\dim_{\mathbb{C}} \mathcal{Q}_2^{\mathrm{ns}}(2,2) = 4.$$

The covering construction transports these coordinates to quadratic-differential period coordinates by a fixed linear map on each marked chart. This can be seen without a global classification theorem. The orientation cover of q_X is the fiber product

$$\widehat{X}: \quad w^2 = B_4(z), \qquad t^2 = (z-b_1)(z-b_2), \qquad y = wt, \qquad \widehat{\omega} = \sqrt{c}\,\frac{\mathrm{d}z}{w}.$$

Its map to E is a double cover branched at $\{p, -p, q, -q\}$. Thus its relative periods are integral linear combinations of torus periods and marked-point positions, after paths and lifts have been fixed. Transfer on absolute cycles and paths between p and $-p$, or between q and $-q$, recover the torus periods and twice the point positions, up to periods. Consequently the inverse change of coordinates is rational linear. The integral period lattices are commensurable; factors of two can occur, so the map need not be unimodular. This is the period-coordinate construction underlying the stratum measure [101, Sections 1.1.1–1.1.2].

It follows that the transported Masur–Veech measure is a constant multiple of Lebesgue measure in these cone coordinates. More explicitly, write the torus area as $s > 0$, put $A = \sqrt{s}\,B$ with $\det B = 1$, and use the point coordinates $\xi_i = A^{-1}u_i$. Up to a fixed positive constant the cone measure is

$$s^3\ \mathrm{d}s\ \mathrm{d}B\ \mathrm{d}\xi_1\ \mathrm{d}\xi_2.$$

Restricting by area as in (28.8) therefore makes the unit-area measure a constant multiple of Haar measure on Y_2. After normalizing to probability and accounting for the finite covering, the measures agree. No numerical identity between unnormalized volumes is asserted without a choice of lattice and labeling conventions. The measure-zero exclusions in $Y_2 \setminus Y_2^\circ$ do not affect this probability identification.

For observables compactly supported inside Y_2°, extension by zero gives smooth observables on the homogeneous space. Thus estimates proved there pass to the stratum after lifting the observables and, if necessary, averaging over the finite markings. For observables which grow near collisions, extension and norm control require additional estimates; a null collision set alone is not a bound on such observables.

28.7 Fourier decomposition and temperedness

The affine homogeneous model admits a direct calculation which does not use a general Teichmüller mixing theorem. Let

$$G = \mathrm{SL}_2(\mathbb{R}), \qquad \mathsf{A} = G \ltimes \mathrm{Mat}_{2\times 2}(\mathbb{R}), \qquad \Lambda = \Gamma \ltimes \mathrm{Mat}_{2\times 2}(\mathbb{Z}),$$

where $\Gamma = \Gamma(2)$ initially. The group law is

$$(g, v)(g', v') = (gg', v + gv').$$

Set $Y_2 = \mathsf{A}/\Lambda$, with Haar probability measure μ, and let $a_t = \mathrm{diag}(e^t, e^{-t})$ act by left translation. A point is a unit-area lattice together with two marked vectors modulo that lattice. If one counts the origin as a marked point, this is a torus with three ordered marked points. Coincidence loci have measure zero and may be removed. Any application to a particular quadratic stratum uses its separately established geometric identification with this model or a finite quotient; an arbitrary finite cover cannot be discarded without checking its base spectrum.

Fiber Fourier coefficients and their stabilizers

Write $v = gu$, with $u \in \mathrm{Mat}_{2\times 2}(\mathbb{R})/\mathrm{Mat}_{2\times 2}(\mathbb{Z})$. Then

$$(g, u) \sim (g\gamma, \gamma^{-1}u), \qquad a_t(g, u) = (a_t g, u).$$

For $F \in L^2(Y_2)$ expand, in the L^2 sense,

$$F(g, u) = \sum_{M \in \mathrm{Mat}_{2\times 2}(\mathbb{Z})} F_M(g) e^{2\pi i \,\mathrm{tr}(M^T u)}. \tag{28.12}$$

The equivariance relation is

$$F_M(g\gamma) = F_{\gamma^{-T}M}(g). \tag{28.13}$$

Consequently each Γ-orbit of frequency matrices contributes a G-invariant summand isometric to

$$L^2(G/\Gamma_M), \qquad \Gamma_M = \{\gamma \in \Gamma : \gamma^T M = M\}. \tag{28.14}$$

Indeed, Parseval followed by unfolding the orbit sum gives

$$\int_{G/\Gamma} \sum_{M' \in \Gamma \cdot M} |F_{M'}(g)|^2 \, \mathrm{d}g = \int_{G/\Gamma_M} |F_M(g)|^2 \, \mathrm{d}g.$$

The stabilizers are elementary. If $M = 0$, the summand is $L^2(G/\Gamma)$. If M has rank two, $\Gamma_M = \{1\}$. If M has rank one, write $M = vw^T$, with v a primitive integral column vector and

$w \neq 0$ an integral column vector. Then Γ_M fixes v under the transpose action. It is a subgroup conjugate to a discrete unipotent group, hence is amenable. For example, if $v = (1,0)^T$ and $\Gamma = \Gamma(2)$,

$$\Gamma_M = \left\{ \begin{pmatrix} 1 & 0 \\ 2n & 1 \end{pmatrix} : n \in \mathbb{Z} \right\}.$$

No transitivity assertion about primitive vectors under $\Gamma(2)$ is required: one simply sums over all its frequency orbits. The same rank distinction appears in the fiber Fourier analysis of [102, Section 4, especially Lemma 4.3].

The trivial representation of an amenable group is weakly contained in its regular representation. Induction preserves weak containment and takes the regular representation of a closed subgroup to that of G. Thus every nonzero-frequency summand in (28.14) is tempered. See [103, Lemma 2.3 and Remark 2.4]. The whole orthogonal complement of the fiber average is tempered, including its countably many frequency orbits. This gives uniform control of their sum; there is no frequency-dependent implied constant left to sum. Explicitly, for every $f \in C_c(G)$, all summands satisfy $\|\pi_{[M]}(f)\| \leq \|\lambda_G(f)\|$; their Hilbert direct sum satisfies the same inequality.

This simple stabilizer argument is enough for the present rate estimate. It does not claim a complete irreducible spectral decomposition. The corresponding finer decomposition for one relative marked vector is developed in [106]; in particular, that paper's genuine part contains no complementary series.

The base spectrum at level two

The quotient $\Gamma(2)\backslash\mathbb{H}$ is the thrice-punctured sphere. The finite-area theorem of [105, Theorem 2] states that the discrete eigenvalue λ_{2g-2+n} exceeds $1/4$, counting $\lambda_0 = 0$ and including residual as well as cuspidal eigenvalues. For $(g,n) = (0,3)$, there is therefore no positive discrete eigenvalue at or below $1/4$. The continuous spectrum starts at $1/4$. Since $-I \in \Gamma(2)$, the base representation factors through $\mathrm{PSL}_2(\mathbb{R})$. All non-tempered nontrivial representations of $\mathrm{PSL}_2(\mathbb{R})$ in this setting are spherical complementary-series representations and would produce a Laplacian eigenvalue in $(0, 1/4)$. It follows that $L^2_0(G/\Gamma(2))$ is tempered. The same conclusion holds for $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, by pullback from that quotient to the level-two cover.

Combining the base and nonzero-frequency parts proves that

$$L^2_0(Y_2) \quad \text{is a tempered } G\text{-representation} \tag{28.15}$$

for either of these two choices of Γ. The statement does not automatically hold for a further finite-index subgroup.

28.8 A global correlation bound with an explicit constant

Let $K = \mathrm{SO}(2)$, with angular generator Θ, normalized so that the nth character has derivative in. Use the norm

$$\mathcal{S}_1(F)^2 = \|F\|_2^2 + \|\Theta F\|_2^2. \tag{28.16}$$

The angle is the ordinary matrix rotation angle: on the orientation cover, ω changes to $e^{i\theta}\omega$, so the genus-two quadratic differential changes to $e^{2i\theta}q_X$. Thus this normalization is twice differentiation in the phase of q_X. Write $F^\circ = F - \mu(F)$ and $G^\circ = G - \mu(G)$. The Cowling–Haagerup–Howe bound for a tempered representation is

$$|\langle \pi(g)v, w\rangle| \leq \Xi(g)\|v\|\|w\|\sqrt{\dim Kv \, \dim Kw}$$

for K-finite vectors; see [103, Proposition 2.6] and [104]. For single K-types the two dimension factors equal one. Decomposing into K-types and using Cauchy–Schwarz gives

$$\sum_{n\in\mathbb{Z}} \|F_n\|_2 \le \left(\sum_{n\in\mathbb{Z}} (1+n^2)^{-1}\right)^{1/2} \mathcal{S}_1(F) = \sqrt{\pi\coth\pi}\,\mathcal{S}_1(F).$$

Here the subscript denotes K-type, distinct from the matrix frequency in (28.12).

In our time normalization the spherical function has the elementary integral

$$\Xi(a_t) = \frac{2}{\pi} e^{-t} \int_0^\infty \frac{\mathrm{d}u}{\sqrt{(1+u^2)(1+e^{-4t}u^2)}} \le \frac{4}{\pi}(1+t)e^{-t}, \qquad t \ge 0. \tag{28.17}$$

For the inequality, split the integral at $u = 1$ and $u = e^{2t}$. The three contributions are at most 1, $2t$, and 1.

Theorem 28.1 (An explicit marked-torus correlation bound)**.** *For Haar probability measure on Y_2 and all F, G in the domain of the rotation generator, the following inequality holds for $t \ge 0$:*

$$\left| \int_{Y_2} \overline{F}\,(G \circ a_t)\ \mathrm{d}\mu - \overline{\mu(F)}\mu(G) \right| \le 4\coth\pi\,(1+t)e^{-t}\mathcal{S}_1(F^\circ)\mathcal{S}_1(G^\circ). \tag{28.18}$$

The same estimate holds on $\mathcal{Q}_2^{\mathrm{ns},1}(2,2)$ with its probability-normalized Masur–Veech measure and the induced rotation norm.

Proof. The preceding K-type estimate and (28.17) apply to the tempered representation (28.15). Remove the constant components to obtain the displayed covariance. The geometric statement follows by pulling observables back to the finite model of Section 28.5; probability normalization preserves the norms and the correlation. □

Centering the functions does not increase $\mathcal{S}_1$. The norm needs only one rotation derivative; neither derivatives in the marked-point coordinates nor compact support are required. This is global on the finite-volume affine quotient. No compact cutoff or unquantified cusp-tail estimate enters this inequality. The polynomial factor allows every pure exponential rate $e^{-(1-\varepsilon)t}$, but not a uniform constant times e^{-t} from this argument. More explicitly, for $0 < \varepsilon \le 1$ one may replace $4\coth\pi(1+t)e^{-t}$ by $4\coth\pi\,\varepsilon^{-1}e^{-(1-\varepsilon)t}$. For the convention $a_t = \mathrm{diag}(e^{t/2}, e^{-t/2})$, replace t by $t/2$ throughout.

28.9 An exact finite lattice sum with two endpoint weights

The following unfolding supplies a concrete counting statistic. For $\varphi, \psi \in C_c^\infty(\mathsf{A})$ let

$$P\varphi(h\Lambda) = \sum_{\lambda\in\Lambda} \varphi(h\lambda), \qquad F = P\varphi, \quad G = P\psi.$$

Normalize Haar measure $\mathrm{d}h$ so that A/Λ has volume one. For fixed t define the overlap weight

$$W_t^{\varphi,\psi}(\lambda) = \int_{\mathsf{A}} \overline{\varphi(h)}\,\psi(a_t h\lambda)\ \mathrm{d}h. \tag{28.19}$$

Only finitely many λ contribute, since any such lattice element lies in the compact set $(\operatorname{supp}\varphi)^{-1}a_{-t}\operatorname{supp}\psi$. In coordinates $h = (g, v)$ and $\lambda = (\gamma, m)$, the weight is explicitly

$$W_t^{\varphi,\psi}(\gamma, m) = \int_G \int_{\mathrm{Mat}_{2\times 2}(\mathbb{R})} \overline{\varphi(g,v)}\,\psi\big(a_t g\gamma, a_t(v + gm)\big)\ \mathrm{d}v\ \mathrm{d}g. \tag{28.20}$$

Unfolding the periodization of the first observable gives exactly

$$\sum_{\lambda\in\Lambda} W_t^{\varphi,\psi}(\lambda) = \int_{Y_2} \overline{F}\,(G\circ a_t)\,\mathrm{d}\mu. \tag{28.21}$$

The weights are nonnegative when φ and ψ are nonnegative. For arbitrary observables they are signed or complex. Thus (28.21) is a finite weighted return sum, not an assertion about an unweighted sharp orbit count.

The unstable Jacobian exponent is 4: conjugation by a_t expands the upper-unipotent direction in $\mathfrak{sl}_2$ by e^{2t} and the two expanding translation directions by e^t each. Accordingly define the normalized return statistic and its cumulative version, for $T \geq 1$, by

$$\mathcal{N}_{\varphi,\psi}(t) = e^{4t}\sum_{\lambda\in\Lambda} W_t^{\varphi,\psi}(\lambda), \qquad \mathcal{A}_{\varphi,\psi}(T) = \int_1^T \mathcal{N}_{\varphi,\psi}(t)\,\mathrm{d}t.$$

This normalization uses the affine model's unstable Jacobian. It does not use the polar Jacobian of full Teichmüller space. Since $\mu(P\varphi) = \int_{\mathsf{A}} \varphi\,\mathrm{d}h$, (28.18) gives the explicit remainder

$$\left|\mathcal{A}_{\varphi,\psi}(T) - \overline{\mu(F)}\mu(G)\frac{e^{4T}-e^4}{4}\right| \leq \frac{4}{3}\coth\pi\,(1+T)e^{3T}\mathcal{S}_1(F^\circ)\mathcal{S}_1(G^\circ). \tag{28.22}$$

Equivalently, with $R = e^T$, this is an $O(R^3\log R)$ remainder against an R^4 main term, with a displayed norm bound. Here $R = e^T$ is an exponential-time parameter; identifying it with hyperbolic curve length, intersection number, or square-tiled area would require a separate comparison. The estimate follows simply from $\int_1^T (1+t)e^{3t}\,\mathrm{d}t \leq \frac{1}{3}(1+T)e^{3T}$.

These norms can also be bounded using the original compactly supported weights. If $K_0 \subset \mathsf{A}$ contains their supports and

$$M_{K_0} = \#(\Lambda \cap K_0^{-1}K_0),$$

then Cauchy–Schwarz in the locally finite periodization gives

$$\mathcal{S}_1(P\varphi) \leq \sqrt{M_{K_0}}\big(\|\varphi\|^2_{L^2(\mathsf{A})} + \|\Theta\varphi\|^2_{L^2(\mathsf{A})}\big)^{1/2},$$

and similarly for ψ. All quantities in this bound are finite for the chosen compact supports.

Define the flow generator by $V\psi(h) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\psi(a_s h)$. Periodization commutes with this generator: $V(P\psi) = P(V\psi)$. Therefore inserting the terminal weight $(4+V)\psi$ gives the exact identity

$$\begin{aligned}\mathcal{A}_{\varphi,(4+V)\psi}(T) = {} & e^{4T}\int_{Y_2}\overline{F}\,(G\circ a_T)\,\mathrm{d}\mu \\ & - e^4\int_{Y_2}\overline{F}\,(G\circ a_1)\,\mathrm{d}\mu.\end{aligned} \tag{28.23}$$

There is no convolution operator to commute past V, and no unjustified differentiation of a counting asymptotic. The quantitative norm required for this terminal test is $\mathcal{S}_1((4+V)G) \leq 4\mathcal{S}_1(G) + \mathcal{S}_1(VG)$. It includes $\|\Theta VG\|_2$; the generators Θ and V do not commute, and this mixed derivative is not omitted.

28.10 What the marked-torus calculation settles

The compact polyhedron encodes the genus-two hyperbolic base, whereas the marked flat torus encodes the nonsquare $(2,2)$ differential and its flow. They use different intrinsic metrics and

are linked by the same six branch values. The affine model gives an exact Fourier decomposition, an explicit global correlation bound, and the finite weighted return sum with remainder (28.22). These are conclusions of the calculation, not hypotheses imported from general Teichmüller mixing. The inputs are classical: amenability and induction, the Cowling–Haagerup–Howe bound, the elementary SL_2 spherical function, and the absence of positive small spectrum on the thrice-punctured sphere. Related marked-torus harmonic analysis is already developed in [102, 106]; no novelty claim for that machinery is intended.

The derivative identity (28.23) also makes the counting converse exact in this model. After subtracting the product of means, a bound for the terminal test $(4 + V)G$ recovers the correlation at T with the same exponential saving. In the proof here the direction is Fourier/representation theory to correlations and hence to counts. The identity identifies the reverse implication as well, without making the proof circular.

The restriction to the chosen stratum remains substantial. Although it meets every genus-two conformal base, its unit-area dimension is seven; the principal unit stratum has dimension eleven. The affine expanding Jacobian is e^{4t}, whereas the full genus-two Teichmüller polar kernel in Section 27 has exponent 6. The two kernels cannot be substituted for one another. Section 29 constructs selected primitive-arc observables from the coefficient data with controlled geometric norms. A comparison with the full orbit kernel, and uniform stability of the nonlinear coefficient extraction, remain open. In particular, the $O(R^3 \log R)$ estimate above is not an error term for the unweighted count of connected simple closed geodesics. A comparison with that count, or an extension to the principal stratum $\mathcal{Q}_2^1(1^4)$, is a further problem.

29 From simple-curve coefficients to controlled observables

The preceding section supplies an explicit analytic norm for its mixing estimate. We now connect that norm to a selected family of simple curves and their coefficient distributions. The genus-one norm-ball obstruction first fixes the attainable scale of sharp remainders. Primitive pole-to-pole arcs then give geometric coefficient observables with uniform rotation and terminal-derivative bounds. Exact affine Gaussian calculations provide independent tests of the analytic estimates. Throughout, a fixed-surface sharp count, a moduli-averaged covariance, and the conditioning of coefficient extraction remain separate problems.

29.1 The norm-ball obstruction and the correct error scale

The genus-one length norm gives a necessary benchmark for any passage from smooth return estimates to sharp simple-curve counting. Let X be a fixed complete hyperbolic once-punctured torus, and let B_X be the unit ball of the McShane–Rivin length norm on $H_1(X;\mathbb{R}) \simeq \mathbb{R}^2$. Put

$$A_X(L) = \#(LB_X \cap \mathbb{Z}^2), \qquad P_X(L) = \#(LB_X \cap \mathbb{Z}^2_{\mathrm{prim}}).$$

The first count includes the origin and nonprimitive classes. The second counts oriented primitive classes; the count of unoriented simple closed geodesics is $N_X(L) = \frac{1}{2}P_X(L)$. Elementary planar lattice counting and Möbius inversion give

$$A_X(L) = \mathrm{Area}(B_X)L^2 + O_X(L), \tag{29.1}$$

$$P_X(L) = \frac{\mathrm{Area}(B_X)}{\zeta(2)}L^2 + O_X(L \log L). \tag{29.2}$$

Indeed, $P_X(L) = \sum_d \mu(d)(A_X(L/d) - 1)$, with $d \ll_X L$; the boundary errors sum as $L\sum_{d \ll_X L} d^{-1}$. These are the upper estimates behind [16, Theorem 4.1], with the orientation factor made explicit here.

The next theorem in that paper concerns a dense set of hyperbolic structures. Its mechanism is an infinitely flat boundary point whose supporting normal is rational with respect to the homology lattice. The subsequent Conjecture 4.3 proposes an $O(L^{1/2+\varepsilon})$ error for the modular torus; thus the obstruction is not asserted there for every individual surface. The phrase "at least $O(L)$" in Theorem 4.2 should be separated from the literal modern assertion $\limsup |E_X(L)|/L > 0$. The elementary argument below proves the precise consequence needed here: no fixed improvement of the exponent below one is possible under the stated flat-point hypothesis. We do not use it to claim a literal $\Omega(L)$ bound.

Proposition 29.1 (A primitive-row obstruction)**.** *Let $B \subset \mathbb{R}^2$ be a strictly convex compact body with the origin in its interior. Suppose that $p \in \partial B$ has irrational radial direction, a rational supporting normal, and infinite graph-flatness. Then, for*

$$E_B(L) = \#(LB \cap \mathbb{Z}^2_{\mathrm{prim}}) - \frac{\mathrm{Area}(B)}{\zeta(2)} L^2,$$

and every $\varepsilon > 0$, there are $c_\varepsilon > 0$ and arbitrarily large L such that

$$|E_B(L)| \geq c_\varepsilon L^{1-\varepsilon}. \tag{29.3}$$

In particular $E_B(L) = O(L^{1-\eta})$ fails for every fixed $\eta > 0$.

Proof. An integral unimodular change of coordinates makes the supporting normal vertical. Write $p = (\alpha, h)$, with $h > 0$ and α/h irrational. Near p, the upper boundary has the form

$$y = h - f(x - \alpha), \qquad f(0) = 0,$$

where $f \geq 0$ is convex. Infinite graph-flatness means that for every integer $N > 2$ there are $C_N, \rho_N > 0$ such that

$$0 \leq f(s) \leq C_N |s|^N \qquad (|s| \leq \rho_N).$$

Fix N, let n run through primes, and set

$$L_n = \frac{n}{h}, \qquad \delta_n = L_n^{-1}, \qquad L_n^+ = L_n + \delta_n.$$

The line $y = n$ supports $L_n B$ at the single point $(L_n\alpha, n)$. Its first coordinate is irrational, so this row contains no lattice point of $L_n B$.

In $L_n^+ B$, the same row lies a normalized depth $h\delta_n / L_n^+$ below the supporting line. Therefore it contains the horizontal interval centered at $L_n^+\alpha$ with half-width

$$r_n = \min\left\{ \left(\frac{h}{2C_N}\right)^{1/N} (L_n^+)^{1-1/N} \delta_n^{1/N},\ \frac{n}{4} \right\}.$$

For sufficiently large n the local graph description applies, the first quantity realizes the minimum, and

$$r_n \asymp_{B,N} L_n^{1-2/N}.$$

The interval has length at most $n/2$. Among its integer points (k, n), at most one has $n \mid k$; every other point is primitive because n is prime. Thus

$$P_B(L_n^+) - P_B(L_n) \geq 2r_n - O(1).$$

The change in the quadratic main term is bounded, since

$$(L_n^+)^2 - L_n^2 = 2 + L_n^{-2}.$$

It follows that

$$\max\{|E_B(L_n)|, |E_B(L_n^+)|\} \geq c_{B,N} L_n^{1-2/N}$$

for all sufficiently large primes n. Choose $N > 2/\varepsilon$ to obtain the assertion. To rule out a proposed $O(L^{1-\eta})$ estimate, apply the assertion with $0 < \varepsilon < \eta$. □

The proof works directly with primitive classes, so no cancellation in a Möbius sum needs to be controlled. The short interval $[L_n, L_n + L_n^{-1}]$ acquires many new primitive points along a nearly flat supporting row, while its predicted main-term increment stays bounded. Infinite flatness alone does not force the row length to be a positive fraction of L_n; that is why the proof gives an exponent-one barrier, not a literal linear lower bound. The distinction between a fixed-surface estimate, an estimate uniform over moduli, and a smoothed estimate must be retained.

29.2 The Fourier and spectral meaning of the obstruction

The flatness obstruction has a direct oscillatory-integral interpretation. For illustration, suppose a localized smooth boundary arc has graph

$$y = h - f(x), \qquad f(x) = ax^k + O(x^{k+1}), \quad a > 0,$$

where $k \geq 2$ is even. Apply smooth cutoffs supported near this arc and equal to one at the supporting point. Integrating the localized indicator in the normal coordinate first gives, up to rapidly decreasing terms,

$$-\frac{e^{-2\pi irh}}{2\pi ir} \int \chi(x) e^{2\pi i r f(x)} \, \mathrm{d}x.$$

The substitution $x = r^{-1/k}u$ in the stationary-phase calculation gives

$$\widehat{\chi_0 \mathbf{1}_B}(r\nu) = c_{a,k}\, e^{-2\pi irh} r^{-1-1/k} + o(r^{-1-1/k}), \qquad c_{a,k} \neq 0. \tag{29.4}$$

where ν is the chosen normal direction and χ_0 is the localized cutoff. The familiar nondegenerate exponent is $r^{-3/2}$ when $k = 2$; larger contact order weakens it toward r^{-1}. Localization matters: the contributions of different supporting points can interfere in the Fourier transform of the whole ball. No global lower bound follows from discarding those phases.

Under dilation,

$$\widehat{\mathbf{1}_{LB}}(m) = L^2 \widehat{\mathbf{1}_B}(Lm).$$

Thus a finite-type flat point with a lattice-rational normal contributes on the scale $L^{1-1/k}$ to the corresponding localized Poisson terms. For an infinitely flat point there is no fixed finite type to insert in this formula. Proposition 29.1 supplies a direct counting argument in that case, without assuming smooth boundary regularity at all nearby points or invoking a formal stationary-phase expansion there.

There is also an exact elementary spectral model. On an auxiliary flat torus $\mathbb{R}^2/\mathbb{Z}^2$, define the positive self-adjoint Fourier multiplier

$$\mathsf{P}_X e^{2\pi i\langle m,x\rangle} = \|m\|_X e^{2\pi i\langle m,x\rangle}.$$

Norm equivalence gives a discrete spectrum tending to infinity. Its counting function is $A_X(L)$, and restriction to the closed span of primitive Fourier modes has counting function $P_X(L)$. Restricting further to the even part identifies m and $-m$ and gives $N_X(L)$. This multiplier is defined spectrally; the nonsmooth length norm is not being asserted to be a classical smooth pseudodifferential symbol. At a differentiable point p of its unit ball, a supporting covector ν satisfies

$$\nabla \|\cdot\|_X(p) = \frac{\nu}{\langle \nu, p\rangle}.$$

Thus a lattice-rational normal is exactly a rational direction for the translation velocity of the associated Hamiltonian. This explains the periodic-phase interpretation, without assuming that a global smooth trace formula applies to the length norm.

The norm-ball discrepancy is therefore a spectral-projector remainder for this particular length multiplier. Its flat normal directions produce degenerate phases in a regularized Poisson or trace calculation. This is distinct from complementary-series spectrum for the $SL_2(\mathbb{R})$ action on the marked-torus moduli space. In particular, it does not contradict the temperedness statement or the smooth correlation estimate in Section 28. The rotation-Sobolev issue depends on what is held fixed. A sharp indicator of a nonradial region in fixed physical vector coordinates generally has rotational jumps, so smoothing it introduces growing derivative norms. By contrast, hyperbolic length is unchanged when the flat charts are rotated while the underlying conformal surface stays fixed. A sharp cutoff in that invariant length can therefore lie in the rotation generator's domain, and a radial flat cutoff can be rotation invariant as well. Sharpness alone is not an obstruction to the stated mixing bound. For those invariant observables, the remaining boundary and regularity issues enter when taking flow derivatives or passing from an averaged correlation to a fixed-surface counting remainder.

29.3 A coefficient observable from primitive pole-to-pole arcs

The marked-torus construction supplies an explicit family of simple curves for which the coefficient data become observables with controlled norms. This family carries a flat-length cutoff as well as a hyperbolic-length weight. The two length parameters remain distinct.

Work on the finite labeled model Y_2 of Section 28.5. Write $E = \mathbb{R}^2/g\mathbb{Z}^2$, with its origin one of the four ramification points of the pillowcase map. Fix a nonzero parity vector $\delta \in (\mathbb{Z}/2\mathbb{Z})^2$. For a nonzero primitive integer vector $m \equiv \delta \pmod 2$, project the straight segment

$$[0, gm/2] \subset \mathbb{R}^2$$

first to E and then to $E/\{\pm 1\}$. It joins the pole at the origin to the pole corresponding to $\delta/2$.

Lemma 29.2 (Primitive directions give simple curves)**.** *Outside a Haar-null invariant subset of Y_2, the preceding segment projects to an embedded arc avoiding the two regular marked points. Its lift to the genus-two double cover is a nonseparating simple closed regular flat geodesic α_m, with*

$$\ell_{q_X}(\alpha_m) = |gm|. \tag{29.5}$$

Opposite vectors give the same curve; distinct primitive directions modulo sign give distinct free homotopy classes.

Proof. If two parameters $0 < s, t < 1$ on the segment identify on the pillowcase, then $(s-t)m \in 2\mathbb{Z}^2$ or $(s+t)m \in 2\mathbb{Z}^2$. Primitivity forces the corresponding scalar to be an even integer. The first case gives $s = t$, and the second is impossible. Likewise the interior contains no two-torsion point. The segment hits one of the additional regular marks only if one of its lifts p, q lies on the closed rational circle $\mathbb{R}gm/g\mathbb{Z}^2$. For each m this is a null condition in the point fiber. The union over the countable set of primitive m is still null and is invariant under the entire $SL_2(\mathbb{R})$ action.

The preimage of an embedded arc joining two branch values of a double cover is a circle. Removing it leaves a connected cover, since the four other branch values still have nontrivial monodromy; thus the circle is nonseparating. At its endpoints the sphere flat angle π doubles to 2π, so the two lifted segments join to a regular flat geodesic. Its length is twice $|gm|/2$. Different directions give intersecting regular geodesics with distinct tangents at the Weierstrass point over the origin. Freely homotopic regular geodesics in a nonpositively curved flat metric lie in a flat cylinder and cannot intersect transversely. This proves the last assertion. □

Write $\ell_m(q_X) = \ell_X(\alpha_m)$ for the hyperbolic length of the same topological curve. Proposition 23.3 identifies it as the position of the first atom of $\sigma^X_{\alpha_m}$. In particular, if a local cutoff χ_m isolates that atom, then for a bounded length weight H,

$$H(\ell_m) = \frac{\langle \sigma^X_{\alpha_m}, H(s)\chi_m(s)\rangle}{\langle \sigma^X_{\alpha_m}, \chi_m(s)\rangle}. \tag{29.6}$$

For arbitrary bounded Borel H, the same statement is interpreted using the already recovered first-atom location. No assertion that first-atom extraction is a bounded linear functional is being made.

Choose an even $f \in C^2_c(\mathbb{R}^2)$ supported in the unit disk, a bounded Borel function $h : (0,\infty) \to \mathbb{C}$, and parameters $R \geq 1$, $L > 0$. Define

$$\mathcal{F}^{R,L}_{\delta;f,h}(q_X) = \frac{1}{2R^2} \sum_{\substack{m\in\mathbb{Z}^2 \text{ primitive} \\ m\equiv\delta \ (2)}} f(gm/R)\, h(\ell_m(q_X)/L). \tag{29.7}$$

The factor $1/2$ identifies opposite directions. The function is well defined on Y_2: changes of lattice basis preserve the parity class and relabel the sum. Averaging over the finite labeling choices gives observables on the unlabelled stratum, with the same norm bounds. Such an average is an average of selected families; it need not count their union without multiplicities.

29.4 Uniform rotation norms, including a sharp hyperbolic cutoff

Put

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad D_B f(x) = (Bx)\cdot \nabla f(x).$$

The notation D_B here is a linear vector field on $\mathbb{R}^2$.

Lemma 29.3 (Uniform primitive-vector bound)**.** *For every unimodular lattice $g\mathbb{Z}^2$ and every $R > 0$,*

$$\#\{m \in \mathbb{Z}^2 :\ m \text{ primitive},\ |gm| \leq R\} \leq 2 + 2\pi R^2. \tag{29.8}$$

Proof. Two nonparallel lattice vectors u, v have $|\det(u,v)| \geq 1$. If both have length at most R, their unoriented angular separation therefore has sine at least R^{-2}. For $R < 1$ there is at most one primitive direction. For $R \geq 1$, placing the directions on the circle of length π shows that their number is at most $\pi/\arcsin(R^{-2}) \leq \pi R^2$. Each direction contributes only the two opposite primitive vectors. The displayed bound covers both cases without rounding conventions. □

Rotation of the flat charts fixes the underlying conformal surface and the intrinsic pole-to-pole arcs. Consequently

$$\Theta \ell_m = 0,$$

and differentiation of (29.7) under rotation only replaces f by $D_J f$. Set

$$C_* = 1 + \pi, \qquad N_{f,h} = C_* \|h\|_\infty (\|f\|^2_\infty + \|D_J f\|^2_\infty)^{1/2}.$$

The primitive-vector bound proves, uniformly in $R \geq 1$ and $L > 0$,

$$\mathcal{S}_1(\mathcal{F}^{R,L}_{\delta;f,h}) \leq N_{f,h}. \tag{29.9}$$

This is global, including the modular cusp. Although the curve labels can jump as a regular mark crosses a selected arc, the exceptional walls are SL_2-invariant. Along almost every

group orbit the labels are fixed under the natural marking, and the preceding derivatives are legitimate. Uniform bounds on the difference quotients give the corresponding strong L^2 generator statements.

In particular, Theorem 28.1 now applies to two observables of the form (29.7), with

$$|\operatorname{Corr}_t(\mathcal{F},\mathcal{G})| \le 4\coth\pi\,(1+t)e^{-t}\,N_{f,h}N_{\widetilde{f},\widetilde{h}}, \tag{29.10}$$

uniformly in their two flat cutoffs and two hyperbolic-length scales. The hyperbolic weight may be the sharp cutoff $h(s)=\mathbf{1}_{s\le 1}$: its discontinuity costs no rotation derivative. If the flat weight is radial as well, the whole observable is K-invariant. In that case a sharp disk cutoff is also allowed directly, and the single-K-type matrix-coefficient bound gives the sharper estimate

$$|\operatorname{Corr}_t(F,G)| \le \Xi(a_t)\|F^\circ\|_2\|G^\circ\|_2 \le \frac{4}{\pi}(1+t)e^{-t}\|F^\circ\|_2\|G^\circ\|_2. \tag{29.11}$$

No rotational smoothing is needed for these radial tests.

29.5 Control of the terminal derivative

For the derivative counting converse, require $h\in C^1$ and

$$H_0=\|h\|_\infty<\infty, \qquad H_1=\sup_{s>0} s|h'(s)|<\infty.$$

Wolpert's quasiconformal length inequality gives

$$|V\ell_m|\le 2\ell_m, \qquad |\Theta V\ell_m|\le 4\ell_m. \tag{29.12}$$

For completeness, the first inequality follows by lifting a K-quasiconformal map to the cyclic annular covers associated with the curve. Their conformal moduli are π/ℓ_X and π/ℓ_Y; modulus distortion gives $K^{-1}\le \ell_Y/\ell_X\le K$. The map associated with a_t has dilatation $e^{2|t|}$. For the second inequality use $\Theta\ell_m=0$ and $[\Theta,V]=\pm 2W$, where W is the generator of a rotated diagonal subgroup. The same first inequality applies to W. See also [108, Equation (1.1)] for the quasiconformal inequality and its attribution to Wolpert.

Let $F=\mathcal{F}^{R,L}_{\delta;f,h}$, and abbreviate

$$f_0=\|f\|_\infty, \quad f_K=\|D_Jf\|_\infty, \quad f_A=\|D_Af\|_\infty, \quad f_{KA}=\|D_JD_Af\|_\infty.$$

Direct differentiation and (29.12) give

$$\begin{aligned}\|VF\|_\infty &\le C_*(f_AH_0+2f_0H_1),\\ \|\Theta VF\|_\infty &\le C_*(f_{KA}H_0+2f_KH_1+4f_0H_1).\end{aligned} \tag{29.13}$$

These constants are again independent of $R\ge 1$ and $L>0$. No second derivative of h occurs, because rotation fixes ℓ_m. Thus $\mathcal{S}_1(VF)$, and hence $\mathcal{S}_1((4+V)F)$, have explicit uniform bounds. In particular the terminal identity extends to these observables:

$$\int_1^T e^{4t}\operatorname{Corr}_t(F,(4+V)G)\,\mathrm{d}t = e^{4T}\operatorname{Corr}_T(F,G)-e^4\operatorname{Corr}_1(F,G). \tag{29.14}$$

Here G has a smooth length weight as above; invariance of probability measure gives $\mu(VG)=0$. This proves applicability of the integrated counting converse to these endpoints. It does not identify an individual curve sum with the finite affine return sum of the preceding section. Sharp cutoffs require a separate smoothing or boundary argument when the terminal derivative is used.

29.6 What the bridge retains

For $h \equiv 1$, the observable reduces to a primitive parity Siegel transform on the modular base and has no nonzero marked-point Fourier mode. For a disk cutoff and fixed g, elementary lattice counting and Möbius inversion give

$$\frac{1}{2}\#\{m \text{ primitive} : m \equiv \delta\ (2),\ |gm| \le R\} = \frac{\pi}{6\zeta(2)}R^2 + O_g(R\log R) = \frac{1}{\pi}R^2 + O_g(R\log R).$$

Indeed, only odd common divisors occur in a nonzero parity class, and $\sum_{d\ modd}\mu(d)/d^2 = 4/(3\zeta(2))$. Adding the hyperbolic weight $h(\ell_m/L)$ can retain dependence on the two regular marks and uses the simple-curve first-atom data. For $h \equiv 1$ the exact mean is

$$\mu(\mathcal{F}^{R,L}_{\delta;f,1}) = \frac{1}{6\zeta(2)}\int_{\mathbb{R}^2} f(x)\ \mathrm{d}x,$$

by the primitive Siegel formula and the equal densities of the three nonzero parity classes.

There is a useful fixed-surface sparsity bound. On a fixed compact (X, q_X) put

$$C(X, q_X) = \sup_X \frac{|q_X|^{1/2}}{\mathrm{d}s_{\mathrm{hyp}}} < \infty.$$

The selected flat geodesic minimizes flat length in its free homotopy class, so comparison with the hyperbolic geodesic gives

$$|gm| = \ell_{q_X}(\alpha_m) \le C(X, q_X)\ell_X(\alpha_m).$$

Consequently the whole selected family with hyperbolic length at most L lies in the flat window $R = C(X, q_X)L$ and has size $O_{X,q_X}(L^2)$. It is a sparse family relative to the full genus-two L^6 simple-curve count. The bound is for a fixed surface and differential; no uniform comparison constant over moduli is asserted.

Equations (29.6)–(29.13) are a norm-controlled bridge for this selected family. They do not identify the full connected-simple-curve count with an affine return sum. Nor do they prove uniform stability of extracting first atoms from perturbed spectral or coefficient data: the shrinking-gap issue of Section 24.4 remains. The output observables have controlled geometric norms, while the quantitative conditioning of the nonlinear coefficient extraction is a different question.

29.7 Affine displacement sums: an exact isometric model

The homogeneous calculation can be tested on explicit geometric sums. These are the relative-period Siegel–Veech transforms discussed in [106, Section 6.1]. On the marked-torus model write a point as (g, v_1, v_2), with the physical marked points understood modulo $g\mathbb{Z}^2$. For $f \in \mathscr{S}(\mathbb{R}^2)$ define

$$S_{f,j}(g, v_1, v_2) = \sum_{m\in\mathbb{Z}^2} f(gm + v_j), \qquad F_{f,j} = S_{f,j} - \int_{\mathbb{R}^2} f(x)\ \mathrm{d}x.$$

The summands test the displacement vectors of straight segments from the torus origin to its jth marked point. These are flat geometric observables; a hyperbolic simple-curve interpretation requires a separate choice of a lift and of curve weights.

Use the Fourier transform $\widehat{f}(\xi) = \int f(x)e^{-2\pi i x\cdot\xi}\,\mathrm{d}x$. With $v_j = gu_j$, Poisson summation gives

$$F_{f,j}(g, u_1, u_2) = \sum_{n\in\mathbb{Z}^2\setminus\{0\}} \widehat{f}(g^{-T}n)e^{2\pi i n\cdot u_j}.$$

Fiber Parseval followed by the usual Siegel mean formula for all nonzero lattice vectors gives the exact identity

$$\|F_{f,j}\|^2_{L^2(Y_2)} = \|f\|^2_{L^2(\mathbb{R}^2)}. \tag{29.15}$$

This is also the one-relative-mark case of [106, Proposition 6.11]. The level-two probability cover changes neither side: the integrand descends to the full modular quotient. By polarization and equivariance,

$$\boxed{\int_{Y_2} \overline{F_{f,j}(x)} F_{h,j}(a_t x)\, \mathrm{d}\mu(x) = \int_{\mathbb{R}^2} \overline{f(x)} h(a_t x)\, \mathrm{d}x.} \tag{29.16}$$

Indeed the Fourier-side integral is $\int \overline{\widehat{f}(\xi)}\widehat{h}(a_{-t}\xi)\,\mathrm{d}\xi$. These identities extend by density to L^2 tests, using the isometric definition of the centered transform when the original sum is unavailable. For indicator tests the variance identity is also [107, Lemma 4]; the mean is Lemma 3 there. For different relative marks, the analogous centered cross-correlation vanishes by integration in the two independent torus fibers.

For the Gaussian $f(x) = e^{-\pi|x|^2}$, the formula becomes

$$\int_{Y_2} \overline{F_{f,j}} F_{f,j} \circ a_t\, \mathrm{d}\mu = \frac{1}{\sqrt{(1+e^{2t})(1+e^{-2t})}} = \frac{1}{2\cosh t}. \tag{29.17}$$

Thus an e^{-t} correlation rate is attained by a natural smooth observable. More generally, substitution $x_1 = e^{-t}u$ in (29.16) gives, for Schwartz tests,

$$e^t \int \overline{f}(x) h(a_t x)\, \mathrm{d}x \longrightarrow \left(\int_{\mathbb{R}} \overline{f(0,y)}\, \mathrm{d}y\right)\left(\int_{\mathbb{R}} h(u,0)\, \mathrm{d}u\right).$$

Schwartz domination justifies the limit.

29.8 Two marks: the rank-one divergence and its removal

The raw product $F_{f_1,1}F_{f_2,2}$ need not be in L^2, even for smooth rapidly decreasing tests. Its conditional squared norm is

$$A_1(g)A_2(g), \qquad A_j(g) = \sum_{n\neq 0} |\widehat{f}_j(g^{-T}n)|^2.$$

For Gaussians $f_1 = f_2 = e^{-\pi|x|^2}$ the divergence can be seen exactly. Write the collinear frequency pairs as (an, bn), where n is primitive modulo sign and $a, b \in \mathbb{Z}\setminus\{0\}$. Primitive Siegel unfolding gives the rank-one contribution

$$\frac{1}{2\zeta(2)} \sum_{a,b\neq 0} \int_{\mathbb{R}^2} e^{-2\pi(a^2+b^2)|x|^2}\, \mathrm{d}x = \frac{1}{4\zeta(2)} \sum_{a,b\neq 0} \frac{1}{a^2+b^2} = \infty.$$

The divergence is logarithmic. More generally, if both $\int f_j$ are nonzero, Fourier continuity gives $A_j(g) \gg \varepsilon^{-1}$ when the dual lattice has a primitive vector of length $\varepsilon \ll 1$; the cusp measure is comparable to $\varepsilon\,\mathrm{d}\varepsilon$. Centering has removed zero frequencies but leaves precisely this collinear-frequency obstruction.

There is an explicit L^2 replacement. Retain only the full-rank frequency matrices:

$$H^{(2)}_{f_1,f_2}(g,u_1,u_2) = \sum_{\substack{n_1,n_2\in\mathbb{Z}^2 \\ \det(n_1,n_2)\neq 0}} \widehat{f}_1(g^{-T}n_1)\widehat{f}_2(g^{-T}n_2) e^{2\pi i(n_1\cdot u_1 + n_2\cdot u_2)}. \tag{29.18}$$

The series is absolutely convergent on each fixed torus fiber. The rank restriction is invariant under the modular transition and under the SL_2 action. It is a fiber Fourier projection, not a positivity-preserving deletion of physical segments. It is defined fiberwise before global integrability has been proved; no orthogonal projection on the global L^2 space is being applied to the raw product.

Here is a norm computation which proves global integrability. Put $a_j = |\widehat{f_j}|^2$ and let $\sigma_1(k) = \sum_{d|k} d$. Hermite normal form gives $\sigma_1(|k|)$ full-modular orbits of integer matrices of determinant $k \neq 0$, each with trivial stabilizer. Thus Parseval and unfolding give

$$\|H^{(2)}_{f_1,f_2}\|_2^2 = \sum_{k\neq 0} \sigma_1(|k|) I_k(a_1, a_2), \tag{29.19}$$

where, with probability-covolume Haar normalization,

$$I_k(a_1,a_2) = \frac{1}{|k|\zeta(2)} \int_{\mathbb{R}^2\setminus\{0\}} \int_{\mathbb{R}} a_1(x) a_2\left(\frac{kJx}{|x|^2} + sx\right) \mathrm{d}s\ \mathrm{d}x, \qquad J(x_1,x_2) = (-x_2, x_1).$$

For completeness the Haar-coordinate formula behind this expression is

$$g = \left[x, \frac{Jx}{|x|^2} + sx\right], \qquad \mathrm{d}g = \frac{1}{\zeta(2)}\ \mathrm{d}x\ \mathrm{d}s.$$

The normalization follows by unfolding the primitive first-column Siegel transform: the stabilizer shifts s by integers, and its mean is $\zeta(2)^{-1} \int f$. At level two each full-rank modular orbit splits into six orbits while Haar measure is divided by six, leaving (29.19) unchanged.

An explicit Schwartz bound is convenient. Define

$$A(f) = \sup_{\xi\in\mathbb{R}^2} (1+|\xi|^2)^{3/2} |\widehat{f}(\xi)|.$$

For $k > 0$ the orbit integral is bounded by

$$I_{\pm k} \le A(f_1)^2 A(f_2)^2 \int_G (1 + k\,\mathrm{tr}(g^T g))^{-3}\ \mathrm{d}g = \frac{\pi^2 A(f_1)^2 A(f_2)^2}{2\zeta(2) k (1+2k)^2}.$$

One way to verify the last integral is the Gaussian identity

$$\int_G e^{-\alpha k\,\mathrm{tr}(g^T g)}\ \mathrm{d}g = \frac{\pi^2}{\alpha k \zeta(2)} e^{-2\alpha k}, \qquad \alpha, k > 0,$$

followed by the Laplace representation of $(1+u)^{-3}$. Since $\sum_{k\ge 1} \sigma_1(k) k^{-3} = \zeta(3)\zeta(2)$,

$$\boxed{\|H^{(2)}_{f_1,f_2}\|_2 \le \frac{\pi}{2}\sqrt{\zeta(3)}\, A(f_1) A(f_2).} \tag{29.20}$$

Let $Df(x) = (Jx)\cdot \nabla f(x)$ and set $B(f) = A(f) + A(Df)$. Rotation commutes with the rank projection, so

$$\Theta H^{(2)}_{f_1,f_2} = H^{(2)}_{Df_1,f_2} + H^{(2)}_{f_1,Df_2}.$$

Therefore

$$\mathcal{S}_1(H^{(2)}_{f_1,f_2}) \le \frac{\pi}{2}\sqrt{\zeta(3)}\, B(f_1) B(f_2). \tag{29.21}$$

The preceding mixing theorem now applies, with constants controlled explicitly by the four test functions. In particular,

$$\left|\left\langle H^{(2)}_{f_1,f_2}, H^{(2)}_{h_1,h_2} \circ a_t \right\rangle\right| \le \pi^2 \zeta(3) \coth \pi\, (1+t) e^{-t} B(f_1) B(f_2) B(h_1) B(h_2).$$

The mean is zero because every retained fiber frequency is nonzero.

29.9 An exact Gaussian two-mark correlation

Take $f_1 = f_2 = f = e^{-\pi|x|^2}$ and write $H^{(2)} = H^{(2)}_{f,f}$. The determinant-orbit calculation can be done at every time:

$$\boxed{\left\langle H^{(2)}, H^{(2)} \circ a_t \right\rangle = \frac{\pi}{\zeta(2)\cosh t} \sum_{k=1}^{\infty} \frac{\sigma_1(k)}{k} e^{-4\pi k \cosh t}.} \tag{29.22}$$

To see this, the product of the Fourier coefficients is

$$\exp\bigl[-\pi \operatorname{tr}(Z^T C_t Z)\bigr], \qquad C_t = I + a_{-2t}, \qquad \det Z = k.$$

Since $\sqrt{\det C_t} = 2\cosh t$, an SL_2 change of variables reduces this to the radial Gaussian integral with $\alpha = 2\pi\cosh t$. Summing the positive and negative determinant orbits gives (29.22). At $t = 0$ this gives the exact squared norm. The identity

$$\sum_{k\geq 1} \frac{\sigma_1(k)}{k} q^k = -\sum_{n\geq 1} \log(1-q^n), \qquad 0 < q < 1,$$

also expresses the answer by the ordinary partition Euler product. The resulting correlation decays faster than exponentially in t. This strong decay is a feature of these particular Gaussian tests; it is consistent with the e^{-t} one-mark example and the uniform tempered bound.

29.10 Sharp boundaries and the rotation norm

The exact isometry also displays a regularity cost of sharp cutoffs. For a bounded convex body $B \subset \mathbb{R}^2$,

$$\|F_{\mathbf{1}_B,j} \circ k_\theta - F_{\mathbf{1}_B,j}\|_2^2 = \operatorname{area}(B \triangle k_\theta B).$$

The first variation of area under rotation, interpreted using the almost-everywhere outer normal of a convex boundary, gives

$$\operatorname{area}(B \triangle k_\theta B) = |\theta| \int_{\partial B} |Jx \cdot n(x)| \, \mathrm{d}s + o(|\theta|).$$

Unless B is rotationally invariant, this rules out a finite L^2 rotation derivative. Equivalently, the sharp transform has angular Sobolev regularity of order $s < 1/2$, but not $s \geq 1/2$. The direct summation of tempered K-type estimates requires an order greater than $1/2$; the earlier theorem uses order one.

This is a generic boundary-regularity phenomenon. It is not by itself the more specific arithmetic obstruction from the unusually flat McShane–Rivin norm ball. It explains why a smooth correlation bound cannot be applied to a sharp norm-ball count while leaving the observable norm unchanged. Here the body B is held fixed in physical Euclidean coordinates. A cutoff depending only on the hyperbolic length $\ell_X(\alpha)$ is unchanged when the quadratic differential is rotated over the same conformal surface X, and does not incur this particular rotation loss. These two kinds of cutoffs must be distinguished. The fixed-surface primitive-lattice count, its Möbius inversion, and its unsmoothing error remain distinct from the affine mean-square identities proved here.

Part VI
Sharp counting and arithmetic

The local shape of a norm ball governs geometric discrepancy. Primitive counting adds a separate arithmetic cancellation problem, and exact diagnostics keep the two inputs visible.

30 Diagnosing the sharp remainder without fitting its main term

The alternatives for the sharp primitive remainder are not exhausted by L and $L\log L$. Strict convexity already gives the qualitative improvement $o(L\log L)$, proved below, but supplies no effective intermediate rate. A normalized Möbius formula identifies a sufficient criterion for $O(L)$. An exact dyadic inversion removes the leading area constant from the numerical test, and an inverse-flatness formula measures the contribution of a particular supporting row. Finally we implement the dyadic diagnostic using hyperbolic trace recursion on two tori. The proofs and the exploratory numerical observations are stated separately.

30.1 The normalized all-point discrepancy and the logarithmic loss

Rescale the norm so that its shortest nonzero integer vector has length one. Write $a = \mathrm{Area}(B)$,

$$A_0(u) = \#(uB \cap (\mathbb{Z}^2 \setminus \{0\})), \qquad D_B(u) = A_0(u) - au^2,$$

and let $P_B(L)$ count primitive vectors. For $M = \lfloor L \rfloor$, Möbius inversion gives the exact identity

$$\frac{P_B(L) - aL^2/\zeta(2)}{L} = \sum_{d\le M} \frac{\mu(d)}{d}\,\frac{D_B(L/d)}{L/d} - aL \sum_{d>M} \frac{\mu(d)}{d^2}. \tag{30.1}$$

The final term is bounded uniformly for $L \ge 2$. Thus the linear remainder problem is precisely a boundedness problem for the displayed Möbius-weighted harmonic sum. The prime number theorem alone does not bound it: the geometric weight $D_B(L/d)/(L/d)$ varies with both L and d.

For a convenient sufficient condition, put

$$\omega_B(u) = \sup_{v\ge u} \frac{|D_B(v)|}{v}, \qquad u \ge 1.$$

The elementary boundary estimate makes ω_B bounded, and it is nonincreasing. Absolute values in (30.1) give

$$|P_B(L) - aL^2/\zeta(2)| \le C_B L \left(1 + \int_1^L \omega_B(u)\,\frac{du}{u}\right). \tag{30.2}$$

Indeed, for $d \le M$,

$$\int_{L/(d+1)}^{L/d} \omega_B(u)\frac{du}{u} \ge \omega_B(L/d)\log(1+1/d) \ge \frac{\omega_B(L/d)}{2d},$$

with the harmless interval below 1 treated separately. Consequently the Dini condition

$$\int_1^\infty \omega_B(u)\,\frac{du}{u} < \infty$$

implies an $O_B(L)$ primitive remainder without any Möbius cancellation. With the prime number theorem, this improves to $o_B(L)$; see (31.4) below. In particular any all-point power saving $D_B(u) = O_B(u^{1-\eta})$, $\eta > 0$, suffices. Infinite graph-flatness does not logically exclude this route: for example a bound $u\exp(-\sqrt{\log u})$ has no fixed power saving but has an integrable normalized modulus.

Proposition 30.1 (A qualitative strict-convexity improvement)**.** *For every fixed strictly convex compact body $B \subset \mathbb{R}^2$ containing the origin in its interior,*

$$D_B(L) = o_B(L), \qquad P_B(L) - \frac{a}{\zeta(2)}L^2 = o_B(L\log L).$$

These statements supply no quantitative rate of little-oh convergence.

Proof. Fix a unit vector ν and let $w_\nu(s)$ be the length of the chord $B \cap \{x : x \cdot \nu = s\}$. Its support is an interval $[a_\nu, b_\nu]$. Convexity makes w_ν concave, and strict convexity makes both extreme supporting chords singletons, so $w_\nu(a_\nu) = w_\nu(b_\nu) = 0$. A finite continuous concave function on a compact interval is absolutely continuous: its monotone one-sided derivative is integrable, and the fundamental theorem on interior subintervals extends to the endpoints. The zero extension therefore belongs to $W^{1,1}(\mathbb{R})$. Moreover

$$\|w_\nu'\|_1 = 2\max_s w_\nu(s) \le 2\,\mathrm{diam}(B).$$

Fubini and integration by parts give

$$\widehat{\mathbf{1}_B}(r\nu) = \widehat{w}_\nu(r), \qquad r\widehat{\mathbf{1}_B}(r\nu) = \frac{1}{2\pi i}\widehat{w_\nu'}(r) \longrightarrow 0$$

by the Riemann–Lebesgue lemma. The same formula gives the uniform bound $|\widehat{\mathbf{1}_B}(\xi)| \le C_B/|\xi|$.

Fix a nonnegative smooth function ρ_ϵ of integral one, supported in the Euclidean disk of radius ϵ. Choose $r_0 > 0$ with $r_0\mathbb{D} \subset B$, and put $\delta = \epsilon/r_0$. Convexity gives the pointwise sandwich

$$\mathbf{1}_{(L-\delta)B} * \rho_\epsilon \le \mathbf{1}_{LB} \le \mathbf{1}_{(L+\delta)B} * \rho_\epsilon.$$

For this fixed ϵ, Poisson summation shows that either smoothed lattice count equals its area plus $o_{B,\epsilon}(L)$: after division by L, each nonzero-frequency term tends to zero, and the summable majorant is a constant times $|m|^{-1}|\widehat{\rho}_\epsilon(m)|$. The two areas differ from aL^2 by $O_B(\epsilon L) + O_B(\epsilon^2)$. Thus

$$\limsup_{L\to\infty} |D_B(L)|/L \le C_B\epsilon.$$

Letting $\epsilon \downarrow 0$ proves the first claim. It implies $\omega_B(u) \to 0$, and logarithmic Cesàro averaging in (30.2) proves the second. □

The qualitative improvement excludes a fixed positive multiple of $L\log L$ along a subsequence, but still allows $L\log\log L$, $L(\log L)^{0.9}$, and many other intermediate scales. Quantitative control of ω_B, or signed cancellation between its Möbius blocks, is needed to reach $O_B(L)$.

30.2 A rational-polygon comparison

The intermediate scales cannot be dismissed merely because the $O(L\log L)$ proof takes absolute values. For a rational polygon $\mathcal{A}$ that is star-shaped about the origin, Bárány, Martin, Naslund, and Robins prove

$$E_{\mathcal{A}}(L) = \Omega_\pm(L\sqrt{\log\log L}),$$

while their upper bound for every rational polygon is

$$E_{\mathcal{A}}(L) = O_{\mathcal{A}}(L(\log L)^{2/3}(\log\log L)^{4/3}).$$

These are Theorems 1.4 and 1.2, respectively, in *Primitive points in rational polygons* [109]. The polygons have genuine straight sides. These results do not transfer to a strictly convex McShane–Rivin norm ball; they show why an unspecified arithmetic cancellation is not a proof of a linear error. Even the elementary square illustrates the relevant arithmetic: at integer $N \geq 1$,

$$\#([-N,N]^2 \cap \mathbb{Z}^2_{\mathrm{prim}}) = 8\sum_{n=1}^{N}\varphi(n).$$

There is also a numerical warning. Freezing a rational polygonal approximation to a strictly convex ball can introduce the polygon's own totient fluctuations. A fixed area error δa contributes $\delta a L^2$; resolving a linear remainder requires area accuracy $o(L^{-1})$, together with control of points near the two boundaries. The trace enumeration below avoids replacing the norm ball by a polygon.

30.3 An exact test independent of the leading coefficient

Suppose $N(L)/L^2 \to c$, put $E(L) = N(L) - cL^2$, and define

$$\Delta_2 N(L) = N(2L) - 4N(L) = E(2L) - 4E(L).$$

Iterating the latter equality and using $E(L) = o(L^2)$ gives

$$E(L) = -\sum_{j=0}^{\infty} 4^{-j-1}\Delta_2 N(2^j L). \tag{30.3}$$

The series is understood as the limit of its telescoping partial sums; it is absolutely convergent under any of the bounds below. Consequently

$$E(L) = O(L) \quad\Longleftrightarrow\quad \Delta_2 N(L) = O(L),$$

and likewise

$$E(L) = O(L\log L) \quad\Longleftrightarrow\quad \Delta_2 N(L) = O(L\log L).$$

For example, if $|\Delta_2 N(u)| \leq Cu$ for every $u \geq L$, then

$$|E(L)| \leq \frac{CL}{4}\sum_{j\geq 0} 2^{-j} = \frac{CL}{2}.$$

These are asymptotic equivalences, not finite-computation certificates. They eliminate contamination by an inaccurately estimated value of c.

If an estimated coefficient $\widetilde{c} = c + \delta c$ is subtracted instead, then

$$\frac{N(L) - \widetilde{c}L^2}{L} = \frac{E(L)}{L} - \delta c L.$$

Thus a small coefficient error creates a growing normalized drift. The dyadic statistic removes this source of error exactly. It does not remove the possibility of rare large excursions beyond the computed range. For measurements over a block define

$$M(T) = \sup_{T\leq L\leq 2T} \frac{|\Delta_2 N(L)|}{L},$$
$$Q(T) = \left(\frac{1}{T}\int_T^{2T} \left|\frac{\Delta_2 N(L)}{L}\right|^2 \mathrm{d}L\right)^{1/2}. \tag{30.4}$$

The first detects the largest computed excursions, while the second measures their contribution over the whole block. Between events at $L = \ell_\gamma$ and $L = \ell_\gamma/2$, $\Delta_2 N(L)$ is constant. Hence its normalized absolute value decreases between events, and the block maximum is determined by event updates and the block endpoints. Its squared integral over such an interval $[a, b]$ is exactly $D^2(1/a - 1/b)$. A coarse integer mesh can miss narrow excursions.

30.4 A quantitative row-crossing formula

The exponent-one obstruction does not specify the slowly varying factor in the discrepancy. The local graph itself gives a more precise statement. Let $B \subset \mathbb{R}^2$ be strictly convex, compact, and contain the origin in its interior. After an integral unimodular change of coordinates, suppose its unique top point is $p = (\alpha, h)$, with $h > 0$ and $\alpha/h \notin \mathbb{Q}$. Write the upper boundary near p as

$$y = h - f(x - \alpha), \qquad f(0) = 0.$$

For sufficiently small $u \geq 0$, define the one-sided inverse widths

$$\rho_+(u) = \sup\{s \geq 0 : f(s) \leq u\}, \qquad \rho_-(u) = \sup\{s \geq 0 : f(-s) \leq u\},$$
$$W(u) = \rho_+(u) + \rho_-(u),$$

where the suprema are restricted to the local graph interval. Strict convexity implies $W(u) \to 0$ as $u \downarrow 0$. The functions $\rho_\pm$, and hence W, are concave on a sufficiently small interval: they are inverses of increasing convex functions vanishing at zero. No symmetry between the two sides of the graph is assumed.

Let

$$P_B(L) = \#(LB \cap \mathbb{Z}^2_{\mathrm{prim}}), \qquad c_B = \frac{\mathrm{Area}(B)}{\zeta(2)}, \qquad E_B(L) = P_B(L) - c_B L^2.$$

Proposition 30.2 (Inverse flatness controls a primitive-row burst)**.** *Let n tend to infinity through primes, put $L = n/h$, and let $\delta = \delta(L) > 0$ satisfy $\delta = o(L)$. Set $T = L + \delta$. For all sufficiently large n, the intersection of TB with the row $y = n$ is the interval with endpoints*

$$a_T = T\alpha - T\rho_-\left(\frac{h\delta}{T}\right), \qquad b_T = T\alpha + T\rho_+\left(\frac{h\delta}{T}\right).$$

Its width is

$$w_n(\delta) = TW\left(\frac{h\delta}{T}\right). \tag{30.5}$$

The number of primitive lattice points on this row is $w_n(\delta) + O(1)$, with an absolute rounding constant, and

$$\max\{|E_B(L)|, |E_B(L+\delta)|\} \geq \frac{1}{2}\left[w_n(\delta) - c_B(2L\delta + \delta^2) - 2\right]. \tag{30.6}$$

In particular, with $\delta = L^{-1}$,

$$\max\{|E_B(L)|, |E_B(L+L^{-1})|\} \geq \frac{L}{2}W(h/L^2) - O_B(1). \tag{30.7}$$

Proof. At scale $L = n/h$, the row touches LB only at $(L\alpha, n)$. The first coordinate is irrational, so there is no lattice point there. At scale T, the normalized row has height n/T and thus depth

$$h - \frac{n}{T} = \frac{h\delta}{T}$$

below the supporting line. Inverting the two sides of the graph gives the displayed endpoints. Since $\delta/T \to 0$ and $W(u) \to 0$,

$$\frac{w_n(\delta)}{n} = \frac{T}{hL} W\left(\frac{h\delta}{T}\right) \longrightarrow 0.$$

Eventually $w_n(\delta) < n$. The number of integers in a closed interval of width w differs from w by at most one. Such an interval contains at most one multiple of the prime n. Every other integer k gives a primitive pair (k, n). Therefore the row contains at least $w - 2$ and at most $w + 1$ primitive lattice points.

Since $LB \subset TB$, all these points are new, while points elsewhere cannot disappear. Thus

$$E_B(T) - E_B(L) \geq w_n(\delta) - 2 - c_B(T^2 - L^2),$$

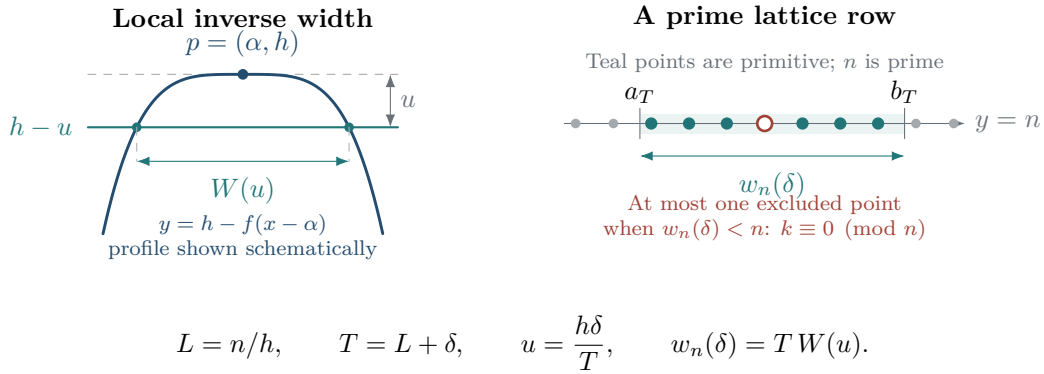
which implies (30.6).

For $\delta = L^{-1}$, set $u_0 = h/L^2$. Concavity and $W(0) = 0$ give

$$W\left(\frac{h}{L^2+1}\right) \geq \frac{L^2}{L^2+1} W(u_0).$$

Multiplying by $T = (L^2+1)/L$ yields $w_n(L^{-1}) \geq LW(h/L^2)$. The change in the quadratic main term is $c_B(2 + L^{-2})$, proving (30.7). □

The theorem remains useful without a model formula for f. It measures the width of the graph at depth L^{-2} and converts that width into a lower bound for the discrepancy. Infinite flatness implies, for every N, a lower bound $W(u) \gg_N u^{1/N}$ and recovers Proposition 29.1. A specified modulus yields a specified slowly varying factor.



$$L = n/h, \qquad T = L + \delta, \qquad u = \frac{h\delta}{T}, \qquad w_n(\delta) = T\,W(u).$$

Figure 9: How a supporting cap becomes a primitive-row contribution. At the contact scale $L = n/h$, the hypothesis $\alpha/h \notin \mathbb{Q}$ leaves the row empty. Increasing the scale to $T = L + \delta$ cuts an interval of width $TW(h\delta/T)$ from that row. When n is prime and the width is smaller than n, all but at most one of its integer points are primitive. The local profile and row spacing are schematic; the displayed width relation is exact.

30.5 Model flatness profiles and what they do not determine

For illustration, suppose the two-sided local graph satisfies

$$f(s) \leq C \exp\left(-\frac{a}{|s|^\beta}\right), \qquad a, \beta > 0,$$

for sufficiently small nonzero s. Then

$$W(u) \geq 2\left(\frac{a}{\log(C/u)}\right)^{1/\beta},$$

and (30.7) gives, along arbitrarily large prime-row contact scales,

$$\max\{|E_B(L)|, |E_B(L+L^{-1})|\} \gg_{B,a,\beta} \frac{L}{(\log L)^{1/\beta}}. \tag{30.8}$$

If the graph has matching upper and lower estimates of this form, the row width itself has this order. This does not prove that the full discrepancy has exactly this order; other rows and other supporting directions can contribute.

Other infinitely flat local profiles give different slowly varying factors. For example, the symmetric model $f(s) = \exp[-\exp(|s|^{-\beta})]$ gives a row width of order $L/(\log\log L)^{1/\beta}$. The model

$$f(s) = \exp[-(\log(1/|s|))^{\gamma}], \qquad \gamma > 1,$$

gives a width of order

$$L\exp[-(2\log L)^{1/\gamma}].$$

These are convex for sufficiently small $|s|$ and illustrate local possibilities. They are not asserted here to be the precise profiles of a hyperbolic torus norm ball.

For a strictly convex body, $W(u) \to 0$. Consequently the single-row argument with a microscopic scale increment yields a quantity $o(L)$, although it can exceed every fixed power $L^{1-\varepsilon}$. It neither proves a literal $\Omega(L)$ lower bound nor distinguishes an $O(L)$ global remainder from an $O(L\log L)$ remainder. Its role is to quantify one geometric obstruction, rather than to identify the entire sharp remainder.

30.6 An exact row diagnostic and its precision requirements

For an arbitrary positive integer n and a row interval $[a,b]$, the primitive count is exactly

$$\#\{k \in \mathbb{Z} \cap [a,b] : \gcd(k,n) = 1\} = \sum_{d|n} \mu(d)\left(\left\lfloor \frac{b}{d}\right\rfloor - \left\lceil \frac{a}{d}\right\rceil + 1\right). \tag{30.9}$$

In particular, it equals

$$(b-a)\frac{\varphi(n)}{n} + O(2^{\omega(n)}).$$

The formula includes endpoints exactly. For prime n and $b-a<n$, the simpler absolute-error comparison with $b-a$ used above applies.

One can therefore record the local statistic

$$J_n = P_{B,n}(n/h + h/n), \qquad P_{B,n}(T) = \#\{k \in \mathbb{Z} : \gcd(k,n) = 1,\ (k,n) \in TB\},$$

at prime n. Since $L = n/h$ and $h/n = L^{-1}$, the proposition gives

$$J_n = (L + L^{-1})W\left(\frac{h}{L^2+1}\right) + O(1).$$

This statistic needs no estimate of $\mathrm{Area}(B)$; it tests the local flatness mechanism directly. It does require a verified lattice-rational supporting normal at an irrational radial point. Finding such a direction is a separate geometric step and is not automatic for an arbitrary fixed surface.

The height h must also be known accurately. An error Δh moves the contact scale n/h by approximately $-L\Delta h/h$. Resolving the window of width L^{-1} thus requires $|\Delta h| \ll hL^{-2}$. Using a larger window relaxes this requirement but changes the depth sampled by (30.5) and increases the background main-term increment. The local statistic is consequently complementary to a global main-term-free discrepancy diagnostic; it does not replace one.

30.7 A trace-based numerical experiment

The count here is $N_X(L)$ for unoriented essential simple closed geodesics on a once-punctured torus. A positive trace triple satisfies

$$x^2 + y^2 + z^2 = xyz, \qquad x, y, z > 2,$$

and a Farey edge replaces z by $xy - z$. The identification of Farey regions with simple curves and this edge relation are described in [110, Sections 1–2]. For the experiment take the reduced triples

$$(x, y, z) = (3, 3, 3), \qquad (x, y, z) = \left(\frac{5}{2}, \frac{10 - \sqrt{11}}{2}, 4\right). \tag{30.10}$$

The first is the modular torus; the second is an asymmetric example. Both satisfy $2 < x \le y \le z \le xy/2$.

The three root curves and the three outward binary sectors enumerate every unoriented primitive slope once. After introducing an outward trace $c \ge \max(a, b) > 2$, the two new traces are $ac - b$ and $bc - a$, both strictly larger than c. Therefore pruning a sector when its new curve exceeds the cutoff does not omit shorter descendants. Equal traces are retained with their curve multiplicities; they are not deduplicated. The hyperbolic length associated with trace x is $2\,\mathrm{arcosh}(x/2)$.

To avoid overflow the computation uses logarithmic traces. If $u_a = \log a$, $u_b = \log b$, and $u_c = \log c$, then

$$\log(ab - c) = u_a + u_b + \log\bigl(1 - \exp(u_c - u_a - u_b)\bigr).$$

On the chosen outward edges the subtracted ratio is at most $1/2$, so this computation avoids catastrophic cancellation. The algorithm uses extended-precision arithmetic for the recursion and stores the resulting lengths in double precision for sorting. This is floating-point enumeration, not an interval certificate of every event location.

We enumerated through hyperbolic length 8192, obtaining 18 191 508 curves for the modular torus and 18 486 116 for the asymmetric torus. This permits evaluating the dyadic block statistics through $T = 2048$, because the largest count used in a block is $N_X(4T)$. Selected output is as follows; all entries are rounded.

	Modular torus		Asymmetric torus	
T	$M(T)$	$Q(T)$	$M(T)$	$Q(T)$
32	2.1686	0.6015	0.8134	0.2734
128	1.1130	0.3501	0.5285	0.1494
512	0.7415	0.1697	0.3756	0.0950
2048	0.4342	0.0950	0.2492	0.0612

The reported maxima use all computed length events, including the half-length events, rather than an integer mesh. As precision checks, an independent direct-trace computation at 100 decimal digits reproduces all lengths through 64 within 3.6×10^{-15}, and thirty selected long recursion paths reaching lengths near 8192 agree within 7.5×10^{-13}. The latter include long repeated-twist paths. Every dyadic cutoff used for the integer counts is separated from the computed spectrum by at least 8.3×10^{-5}.

A separate sensitivity calculation treats nearby events as having an unknown internal order and allows errors up to 10^{-6} in each event position. Over all blocks the resulting lower and upper bounds for $M(T)$ differ by less than 7.1×10^{-7}. This is a conditional sensitivity check, not a proof that every computed length has that error bound. The source, full block table,

precision report, and sensitivity tables accompany this monograph as a reproducible diagnostic archive.

The displayed finite-range normalized excursions decrease overall; they do not establish an asymptotic linear or sublinear bound. The modular torus was already singled out as a candidate for exceptional behavior in [16, Conjecture 4.3]. Neither example is asserted to have a verified lattice-rational supporting normal at an infinitely flat irrational point. Sparse resonances can occur beyond the tested range. The numerical conclusion is only that this main-term-free diagnostic can be computed and is stable under the stated precision checks.

30.8 A recent qualification of irrational flatness

Doan–Li–Nguyen's recent preprint [111], version 3 dated 9 October 2026, gives a Diophantine refinement of the irrational flatness statement. Its Theorem 4.7 states that infinite graph-flatness holds precisely when the exponential rational-approximation exponent vanishes; superexponentially approximable irrational directions can have graph-flatness order only one. In slope coordinates the relevant quantity is $\limsup_j (\log q_{j+1})/q_j$, where q_j are continued-fraction denominators. The local convexity mechanism is that a nearby rational corner can remain visible at a much smaller spatial scale than its exponentially small angle.

We have checked the local Fricke and convexity mechanism relevant to this distinction, but have not audited the entire preprint. None of its global counting assertions is used in the proofs above. In particular, Propositions 29.1 and 30.2 retain their explicit flat-point hypotheses, and Proposition 30.1 uses only strict convexity. The preprint also continues to list the McShane–Rivin $O_X(L \log L)$ bound as the cumulative counting benchmark; its multiplicity estimate concerns exact length levels, a different question.

31 Arithmetic cancellation and the geometric remainder

31.1 A transfer estimate with both inputs explicit

Let $B \subset \mathbb{R}^2$ be a fixed convex body containing the origin in its interior, and normalize its gauge so that the shortest nonzero integer vector has gauge one. Put

$$A_B(u) = \#(uB \cap (\mathbb{Z}^2 \setminus \{0\})), \qquad \Delta_B(u) = A_B(u) - a_B u^2, \qquad a_B = \mathrm{Area}(B),$$

and write

$$P_B(L) = \#(LB \cap \mathbb{Z}^2_{\mathrm{prim}}), \qquad E_B(L) = P_B(L) - \frac{a_B}{\zeta(2)} L^2.$$

All primitive vectors are counted with their signs; an unoriented count for a centrally symmetric body is half this count. The notation Δ_B here is the all-point discrepancy D_B of Section 30.1.

There are two inputs in the primitive problem: the geometric discrepancy Δ_B, and cancellation in the Möbius function. A quantitative bound for the second input does not by itself specify the first. The following elementary estimate keeps their roles separate.

Proposition 31.1 (A large-divisor transfer estimate)**.** *Suppose*

$$M(x) := \sum_{n \le x} \mu(n), \qquad |M(x)| \le Cx\eta(x),$$

where η is nonnegative and nonincreasing on $[1, \infty)$. Then, for every integer $1 \le D \le L$,

$$|E_B(L)| \ll_{B,C} \sum_{d \le D} |\Delta_B(L/d)| + \frac{L^2}{D}\eta(D). \tag{31.1}$$

In particular, $M(x) \ll x^\theta$, with $0 \le \theta < 1$, gives

$$|E_B(L)| \ll_{B,\theta} \sum_{d \le D} |\Delta_B(L/d)| + L^2 D^{\theta-2}. \tag{31.2}$$

Proof. Möbius inversion gives

$$P_B(L) = \sum_{d \le L} \mu(d) A_B(L/d).$$

Set $U = L/D$. Reversing the order of summation in its large-divisor part gives the exact identity

$$\sum_{d > D} \mu(d) A_B(L/d) = \sum_{\substack{m \in \mathbb{Z}^2 \setminus \{0\} \\ \|m\|_B \le U}} \left(M\Big(\frac{L}{\|m\|_B}\Big) - M(D) \right).$$

The summands with $\lfloor L/\|m\|_B \rfloor = D$ are zero. Since $A_B(u) \ll_B u^2$ for $u \ge 1$, partial summation gives

$$\sum_{\substack{m \ne 0 \\ \|m\|_B \le U}} \|m\|_B^{-1} \ll_B U.$$

Monotonicity of η therefore bounds the large-divisor part by

$$\eta(D) \left(L \sum_{0 < \|m\|_B \le U} \|m\|_B^{-1} + D A_B(U) \right) \ll_B \frac{L^2}{D} \eta(D).$$

The area term has the same tail bound, because partial summation yields

$$\left| \sum_{d > D} \frac{\mu(d)}{d^2} \right| \ll \frac{\eta(D)}{D}.$$

Subtract the area main term in the small-divisor sum and bound its remaining terms absolutely. This proves the first estimate; take $\eta(x) = x^{\theta-1}$ for the second. □

31.2 An unconditional sublinear consequence of the Dini condition

Define the nonincreasing geometric modulus

$$\omega_B(u) = \sup_{v \ge u} \frac{|\Delta_B(v)|}{v}.$$

For $L/(D+1) \ge 1$, direct comparison on the intervals $[L/(d+1), L/d]$ gives

$$\sum_{d \le D} |\Delta_B(L/d)| \le 2L \int_{L/(D+1)}^{L} \omega_B(u) \frac{du}{u}. \tag{31.3}$$

Consequently,

$$\int_1^\infty \omega_B(u) \frac{du}{u} < \infty \quad \Longrightarrow \quad E_B(L) = o_B(L). \tag{31.4}$$

Here the arithmetic input is only the prime number theorem, in the equivalent form $M(x) = o(x)$, not RH. Indeed, use the nonincreasing envelope $\eta(x) = \sup_{y \ge x} |M(y)|/y$, fix $U \ge 2$, and choose $D = \lfloor L/U \rfloor$. After division by L, the large-divisor bound tends to zero for this fixed U, whereas the small-divisor bound is at most $2 \int_{U/2}^\infty \omega_B(u)\, du/u$ for all sufficiently large L. Let $L \to \infty$ and then $U \to \infty$. This strengthens the earlier absolute-value Dini argument, which gave only $O_B(L)$.

31.3 Fixed zero-free half-planes and the recent quasi-RH claims

For the Riemann zeta function and $1/2 \le \theta_0 < 1$, the standard zero-free-half-plane formulation is equivalent to

$$\zeta(s) \neq 0 \quad (\Re s > \theta_0) \quad \Longleftrightarrow \quad M(x) = O_\varepsilon(x^{\theta_0+\varepsilon}) \quad \text{for every } \varepsilon > 0. \tag{31.5}$$

For the forward direction, polynomial growth of ζ and branches of $\log\zeta$ on zero-free disks at sufficiently large height, followed by Borel–Carathéodory and a three-circles estimate, give $1/\zeta(\sigma+it) \ll_{\delta,\varepsilon} (1+|t|)^\varepsilon$ for $\sigma \ge \theta_0+\delta$. The pole at one is harmless for the reciprocal. A truncated Perron integral for $1/\zeta$, shifted to this line, proves the Mertens estimate after shrinking the small parameters. Conversely, partial summation makes $s\int_1^\infty M(x)x^{-s-1}\,dx$ holomorphic on $\Re s > \theta_0$ and equal to $1/\zeta(s)$ on $\Re s > 1$. This equivalence uses the analytic properties of ζ; zero-freeness of an arbitrary Dirichlet series would not alone suffice.

The OpenAI manuscript dated 30 September 2026 claims the half-plane $\Re s > 7/8$ for all finite-order Hecke L-functions over $\mathbb{Q}(\sqrt{-3})$, and hence for Dirichlet L-functions, including ζ [114, Theorem 1.1]. The manuscript dated 5 October 2026 gives a separate proof claiming $\Re s > 11/12$ and describes it as a simpler intermediate result toward the September theorem [115, Theorem 1.1 and the following paragraph]. We have checked those statements, not independently verified either proof or its formalization. All applications here are conditional on the explicit arithmetic hypothesis in Proposition 31.1. The corresponding Mertens exponents would be $7/8+\varepsilon$, $11/12+\varepsilon$, and, under full RH, $1/2+\varepsilon$. The generalized assertions about other L-functions are unnecessary for this primitive planar count.

31.4 Logarithmic geometric savings

The next comparison is conditional on an additional estimate for the all-point discrepancy:

$$|\Delta_B(u)| \ll_B \frac{u}{(\log u)^\beta}, \qquad u \ge 3, \qquad \beta > 0. \tag{31.6}$$

No such estimate is asserted here for a McShane–Rivin norm ball.

First suppose $M(x) \ll x^\theta$ for some fixed $\theta < 1$. Choose any fixed d with $1/(2-\theta) < d < 1$, and take $D = \lfloor L^d \rfloor$. The large-divisor contribution in (31.2) is $O(L^{2-d(2-\theta)})$, with exponent strictly smaller than one. The small-divisor contribution is

$$\ll L\int_{L^{1-d/2}}^{L} \frac{du}{u(\log u)^\beta} \ll_{\beta,d} L(\log L)^{1-\beta}.$$

Thus

$$M(x) \ll x^\theta \ (\theta<1), \quad \Delta_B(u) \ll u(\log u)^{-\beta} \quad \Longrightarrow \quad E_B(L) \ll L(\log L)^{1-\beta}. \tag{31.7}$$

In particular, the geometric threshold $\beta = 1$ gives $O_B(L)$. Any fixed power saving for M suffices for this conclusion: full RH does not improve the logarithmic exponent in this elementary argument.

For comparison, the classical unconditional Vinogradov–Korobov estimate, recorded for example in Zhai [112, eq. (1.5)], has the form

$$M(x) \ll x\exp\Big(-c(\log x)^{3/5}(\log\log x)^{-1/5}\Big).$$

Write $F(L) = (\log L)^{3/5}(\log\log L)^{-1/5}$ and take $D = \lfloor L\exp(-aF(L))\rfloor$ with a sufficiently small fixed $a > 0$. Since $F(D) \sim F(L)$, the large-divisor part is $O(L\exp(-c'F(L)))$. The

small-divisor part is bounded by

$$L\int_{\exp(aF(L))/2}^{L}\frac{du}{u(\log u)^{\beta}}.$$

The resulting elementary comparison is

geometric exponent β	unconditional VK input	$M(x)\ll x^{\theta},\ \theta<1$
$0<\beta<1$	$O\Big(L(\log L)^{1-\beta}\Big)$	$O\Big(L(\log L)^{1-\beta}\Big)$
$\beta=1$	$O(L\log\log L)$	$O(L)$
$\beta>1$	$O\Big(LF(L)^{1-\beta}\Big)$	$O\Big(L(\log L)^{1-\beta}\Big)$.

These are consequences of one elementary splitting argument, not assertions of optimality. The constants can depend on B, β, the arithmetic exponent, and the chosen cutoff exponent.

31.5 Power-saving geometric discrepancies and established benchmarks

If instead $|\Delta_B(u)|\ll u^{\alpha}$ with $0\le\alpha<1$, then

$$\sum_{d\le D}|\Delta_B(L/d)|\ll L^{\alpha}D^{1-\alpha}.$$

Balancing this with $L^2D^{\theta-2}$ gives

$$D\asymp L^{(2-\alpha)/(3-\alpha-\theta)},\qquad E_B(L)\ll L^{\kappa(\alpha,\theta)},\qquad \kappa(\alpha,\theta)=\frac{2-\alpha\theta}{3-\alpha-\theta}. \tag{31.8}$$

For example, $\alpha=2/3$ and a Mertens exponent arbitrarily close to $1/2$ give $O_\varepsilon(L^{10/11+\varepsilon})$. Exponents arbitrarily close to $7/8$ and $11/12$ give, respectively, $O_\varepsilon(L^{34/35+\varepsilon})$ and $O_\varepsilon(L^{50/51+\varepsilon})$ by this method. These examples concern a hypothetical or separately established geometric estimate with $\alpha=2/3$; they do not assign that exponent to a McShane–Rivin ball. Stronger results for smooth curved domains use additional structure and should not be confused with this general transfer bound.

With the same power-saving geometric input, the unconditional VK estimate already gives

$$E_B(L)\ll L\exp(-c_BF(L))$$

for some $c_B>0$, by choosing $D=L\exp(-aF(L))$ and balancing the two exponential factors. Thus sublinear primitive errors in suitable geometries do not require RH.

This unconditional sublinear conclusion is a classical one for sufficiently smooth, positively curved domains: Zhai records it in equation (1.4) of [112]. His parameter is the area scale $x=L^2$, so his $O(x^{1/2}\exp(-cF(x)))$ becomes $O(L\exp(-c'F(L)))$ here. The hypotheses include finite nonzero curvature throughout and a C^4 Gauss map. Those hypotheses do not hold for a McShane–Rivin ball.

Under RH, Baker [113, Theorem 1] proves $O_\varepsilon(L^{10/13+\varepsilon})$ for a specified class of piecewise C^4 domains whose arcs have nonzero curvature, with additional reciprocal-curve regularity and endpoint-tangent conditions. This is an established comparison result, not a claim about the best current exponent or about the flat norm balls considered here. In the opposite direction, the unconditional polygon obstruction $\Omega_\pm(L\sqrt{\log\log L})$ recalled in Section 30.2 persists regardless of RH. It belongs to rational polygons and is not a lower bound for strictly convex bodies. These examples show why the shape hypothesis must accompany any arithmetic hypothesis.

31.6 The remaining geometric obstruction

Strict convexity supplies $\Delta_B(u) = o(u)$ and hence $\omega_B(u) \to 0$, but no specified rate for that convergence. Under any fixed arithmetic exponent $\theta < 1$, the preceding power cutoff yields a bound of the form

$$|E_B(L)| \ll L \log L\, \omega_B(cL^{1-d}) + L^{2-d(2-\theta)}.$$

It is again $o(L \log L)$, without a universal logarithmic saving specified by this argument. This is a limitation of the stated inputs and this transfer estimate, not an impossibility theorem for more special convex bodies.

One cannot simply multiply a bound for M by the supremum of the weights $\Delta_B(L/d)$. Abel summation also sees their variation. On a dyadic interval $d \asymp D$, the counting part of these weights can have total variation of order $(L/D)^2$, despite a much smaller pointwise discrepancy. The large-divisor estimate above pays for that variation explicitly. Further cancellation would require additional information about the geometric weights or their correlation with $\mu(d)$.

Equivalently, for $\Re s > 2$ the gauge zeta functions satisfy

$$Z_B(s) = \sum_{m \neq 0} \|m\|_B^{-s}, \qquad Z_B^{\mathrm{prim}}(s) = \frac{Z_B(s)}{\zeta(s)}.$$

The estimate $\Delta_B(u) = o(u)$ provides the representation

$$Z_B(s) = \frac{a_B s}{s-2} + s \int_1^\infty \Delta_B(u) u^{-s-1}\, du, \qquad \Re s > 1.$$

A zero-free result for the denominator does not on its own continue the geometric numerator through the line $\Re s = 1$, nor provide the vertical growth bounds needed for a sharp counting inversion. This is the analytic reason to keep the arithmetic and geometric inputs separate.

For a body satisfying the infinite-flatness and rational-normal hypotheses of Proposition 29.1, the prime-row construction gives $\Omega(L^{1-\varepsilon})$ along an unbounded sequence for every $\varepsilon > 0$. This excludes a fixed sublinear power bound for that body even if RH holds. It does not exclude $O(L)$, or even $o(L)$: the quantitative lower bound is $LW(h/L^2)$ up to constants, and $W(h/L^2) \to 0$ under strict convexity. Thus the useful next geometric target is a quantitative modulus, such as $\Delta_B(u) = O(u/\log u)$, rather than a power saving incompatible with the flat-point obstruction. Establishing that target for any specified McShane–Rivin ball remains a separate problem.

The all-point target has a numerical diagnostic independent of the area constant, just as in Section 30.3. Define

$$\mathcal{D}_A(L) = A_B(2L) - 4A_B(L).$$

The dyadic inverse identity proves the equivalence

$$\Delta_B(L) = O(L/\log L) \quad \Longleftrightarrow \quad \mathcal{D}_A(L) = O(L/\log L). \tag{31.9}$$

Indeed, the reverse implication follows from $\Delta_B(L) = -\sum_{j \geq 0} 4^{-j-1} \mathcal{D}_A(2^j L)$ and $\log(2^j L) \geq \log L$. For a symmetric torus norm ball, let λ_γ be the simple-geodesic length in the same normalization as the gauge of B. The length list already used in Section 30.7, rescaled when imposing the unit-shortest-vector convention, gives the exact identity

$$A_B(L) = 2 \sum_{\substack{\gamma \text{ unoriented simple} \\ \lambda_\gamma \leq L}} \left\lfloor \frac{L}{\lambda_\gamma} \right\rfloor.$$

Thus the proposed geometric threshold can be investigated without estimating the norm-ball area or introducing a new enumeration of curve types. No finite numerical range proves the uniform bound in (31.9).

Part VII
Open problems and conjectural directions

The comparisons suggest concrete further estimates, integrations, and gluing problems. Their hypotheses identify what the next results would need to establish.

32 Questions suggested by the calculations

The ideal volume proposal already has an established measure-theoretic realization in (2.5), and the tree methods described in Section 4 also compute bordered genus-zero volumes. The following questions isolate further work suggested by the discussion; they are proposed tasks, not claims that no related results exist.

Question 32.1 (Direct hyperideal integration)**.** Can the fixed-L domains and the density (4.5) be integrated directly, with all cell and symmetry factors, to recover $2\pi^2 + \frac{1}{2}\sum_i L_i^2$? Which change of variables makes the cancellation or polynomiality visible?

Question 32.2 (A usable measure on general polyhedral cells)**.** Can (4.7) be expressed in dihedral-angle variables in a way that makes higher-dimensional cell integrals computationally useful? In the ideal case the measure is particularly simple. What replaces that simplicity for fixed hyperideal boundary lengths or compact cone openings?

Question 32.3 (Gluing and topology)**.** Can a polyhedral description be combined with Fenchel–Nielsen gluing to give a practical calculation for arbitrary (g, n)? A proposed formula must specify both the cutting multiplicity and the full mapping-class quotient; integrating over all pants lengths and twist periods alone overcounts moduli space.

Question 32.4 (Curvature corrections in period coordinates)**.** Can the correction class Ξ_μ in (8.8) be represented by an explicit first curvature correction to the WP form in the actual finite-area conformal coordinates? Equation (12.19) answers the local four-point calculation in a blown-up character chart, whose conversion to conformal coordinates is still required. How does the local density behave in a joint limit toward a balanced collision wall, where the final-chamber expansion is not uniform in the weights?

Question 32.5 (Beyond the regular quantum degeneration)**.** Can Proposition 15.1 be extended uniformly over compact sets of nondegenerate Euclidean tetrahedral shapes, with a geometric description of the required scalar normalization? Can a geometric state integral in that regime produce the Thurston–Deligne–Mostow measure on the boundary-shape space? The model (11.8) specifies a classical comparison, not an identification with an existing quantum invariant. The exact Liouville identity (12.6) and the torsion limit (12.14) identify two precise boundary constructions; the auxiliary-tetrahedron obstruction and the different regimes in (12.24) explain what remains to connect them to that joint asymptotic. The finite one-screening fusion calculation in Section 13 computes the finite period and residue-sector comparison, with a logarithmic limit at balanced weights. Its direct infinite-spectral derivation uses the continuation and tail-control input (13.3); Section 14 fixes the quantum-group normalization, proves the rank-two obstruction, and computes the classical rank-three fold in the regular compact model. Section 15 corrects the integration-coordinate dictionary, removes the external parameter poles, and proves the uniform regular-family Airy formula with its first correction. General shape dependence, global gluing, and identification of the resulting measure remain to be established.

Question 32.6 (Uniform quantum Dehn filling)**.** For a knot complement admitting the positive geometric optimization of Section 16, can an exact quantum surgery representation be controlled uniformly in a complex neighborhood of complete peripheral holonomy, with a nonzero geometric contribution and bounds on all competing terms? Proposition 16.1 already fixes every classical long-slope coefficient, and (16.26) specifies the torsion check. The further issue is the analytic input of Proposition 16.2, especially in a joint limit where the slope varies with the quantum level.

Question 32.7 (Certified realization and quantitative angle bounds)**.** For the planar problem of Section 9, can positive feasibility, initial centering, inexact solves, and geometric error certification be combined into a complete complexity theorem with explicit input and precision dependence? For a three-manifold angle polytope whose maximum is interior, which additional geometric hypotheses bound its smallest angle in terms of the combinatorial input and a quantitative feasible margin? The barrier iteration bound is available; the planar Laplacian sign structure does not extend automatically to the full three-dimensional Hessian.

Question 32.8 (A continuum comparison of second variations)**.** Can a family of refined polyhedral functionals, with specified geometric convergence and normalization, have a second-variation limit related to a WP potential? The exact cotangent identity, Schur complement, and Gaussian determinant in Section 9 provide discrete objects for such a comparison. Control of their limits and an explicit identification of the target metric are still required.

Question 32.9 (Continuation of a coefficient Laplace transform)**.** Does $D_C(s)$ from Section 23 admit a meromorphic continuation beyond the proved half-plane $\Re s > 1$? A proof must control coefficient extraction in the continued character-valued expression, since it is not bounded for the uniform norm on unitary characters.

Question 32.10 (Control of the globally filtered remainder)**.** Proposition 25.1 gives a bounded global cancellation of every one-crossing primitive and all its iterates. Can the signed residual $\mathcal{R}_X$ in (25.6) be estimated with constants controlled by the geometry of X? The exact obstruction (25.7) rules out a perfect fixed linear selector; length-dependent cancellation or the nonlinear diagonal extraction in (25.10) retains information that such a selector discards.

Question 32.11 (Effective selection of connected supports)**.** Can the signed coefficient relations produce the smoothed estimate (24.22) with specified η, m and geometric constants? The elementary lattice remainder, the stable finite support projection (24.19), and the constructive isolation estimate are available. The global one-crossing filter and the length-valued Euler product now provide concrete aggregate objects. An estimate for the higher-intersection residual or for the nonlinear first-atom extraction, strong enough to yield a counting saving, remains to be proved.

Question 32.12 (From weighted curve counts to mixing)**.** Can the filling-pair coefficient construction of Section 27 produce a norm-controlled estimate for the weighted orbit kernel (27.6) without assuming mixing? Proposition 27.2 gives the exact unfolding, and Theorem 27.4 preserves the counting exponent when derivative endpoint weights are controlled. The independent two-endpoint estimate, its Jacobian weights, and the noncompact tail remain to be established for that principal-stratum kernel; a scalar area or length count is insufficient. Section 28 settles the separate affine return estimate on the nonsquare $(2, 2)$ stratum, and Section 29 gives selected primitive-arc coefficient observables with uniform geometric norms. Can their class be enlarged to control the full two-endpoint kernel, with quantitative stability of the nonlinear coefficient extraction, and can the calculation be extended beyond the marked-torus locus?

Question 32.13 (The sharp norm-ball remainder)**.** Can the qualitative $o_X(L \log L)$ remainder of Proposition 30.1 be made effective through a decay modulus satisfying (30.2), or through

cancellation in (30.1)? The inverse-flatness formula (30.7) quantifies a local lower obstruction, while (30.3) gives a test independent of the leading coefficient. Determining whether the primitive remainder is $O_X(L)$, or has an unbounded slowly varying factor, remains open here. Section 31 makes the two inputs explicit: $\Delta_B(u) = O(u/\log u)$ together with any fixed Mertens power saving gives $O_X(L)$, whereas an integrable geometric modulus and the prime number theorem already give $o_X(L)$. The arithmetic and flatness hypotheses must be specified for each surface; finite-range numerical observations do not decide them.

Question 32.14 (Beyond closure operations)**.** Which positive constant series have a geometric partition or a useful functional equation behind them? The figure-eight kernel (18.7) and the twist kernel (20.9) provide two concrete models. Their different index sets suggest separating global curve identities from identities for a distinguished twist orbit.

Question 32.15 (A boundary phase-space formulation)**.** Can the endpoint-and-direction matching relations in Section 21 be organized into an integration scheme for general filling curves? Orthogeodesic lengths give the canonical local geometry, while the endpoint phase variables retain the missing gluing information. A useful result would control both the sum over itineraries and the integration over those variables.

These questions concern concrete integrals or pointwise identities. The distinctions established above—fixed parameters, geometric chambers, correct curve types, and specified measures—are part of their statements, rather than technicalities to be added after a formal calculation.

A A trace-coordinate check of the WP factor

A separate two-dimensional calculation is useful for checking the factor 2 in the local densities. Let $A, B, C, D \in \mathrm{SL}_2(\mathbb{R})$ satisfy $ABCD = I$, and set

$$a = \operatorname{tr} A,\ b = \operatorname{tr} B,\ c = \operatorname{tr} C,\ d = \operatorname{tr} D, \qquad x = \operatorname{tr}(AB),\ y = \operatorname{tr}(AC),\ z = \operatorname{tr}(BC).$$

With the four peripheral traces fixed, the character surface satisfies the Fricke relation

$$\begin{aligned}\Phi(x,y,z) = {} & x^2 + y^2 + z^2 + xyz - (ab+cd)x - (ac+bd)y \\ & - (ad+bc)z + a^2 + b^2 + c^2 + d^2 + abcd - 4 = 0.\end{aligned} \tag{A.1}$$

This is the standard four-holed-sphere trace surface; see [58].

On a chart where $x = 2\cosh(\ell/2) > 2$, put $s = \sinh(\ell/2)$ and $X = AB$. Twisting across the AB cuff fixes A, B and conjugates C, D by

$$\exp\left(\frac{\tau}{2} H_X\right), \qquad H_X = \frac{X - (x/2)I}{s}.$$

The eigenvalues of H_X are $1, -1$, so this is a displacement of length τ along the cuff axis. Differentiating $y = \operatorname{tr}(AC)$ gives

$$\frac{\partial y}{\partial \tau} = \frac{1}{2s} \operatorname{tr}(A[X, C]) = -\frac{\Phi_z}{2s}.$$

For the last equality, Cayley–Hamilton gives $\operatorname{tr}(A^2BC) = ad - z$, while the two-matrix trace identity gives $\operatorname{tr}(ACAB) = xy - bc + z$. Their difference is $ad + bc - xy - 2z = -\Phi_z$.

Since $\mathrm{d}x = s\ \mathrm{d}\ell$ and $\omega_{\mathrm{WP}} = \mathrm{d}\ell \wedge \mathrm{d}\tau$,

$$\boxed{\mathrm{d}V_{\mathrm{WP}} = \frac{2|\ \mathrm{d}x \wedge \mathrm{d}y|}{|\Phi_z|}.} \tag{A.2}$$

At a zero of Φ_z use a different pair of trace coordinates. The trace formula is local on the marked character surface; it does not describe a global mapping-class fundamental domain.

B Conventions at a glance

Quantity	Convention in this monograph
Surface/ambient hyperbolic curvature	-1; the area-one rescaling has curvature $-A$.
Complex hyperbolic curvature	Holomorphic sectional curvature -1.
Flat weights	Deficit $2\pi\mu_i$; opening $2\pi(1-\mu_i)$; $\sum \mu_i = 2$.
Three volume scales	A is boundary area, W is enclosed volume, and V_{WP} is moduli volume.
Flat WP limit	$A^{-1}\omega_{\mathrm{WP}} \to \omega_{\mathrm{CH}}/2$ at fixed weights, locally over moduli.
WP form	$\sum \mathrm{d}\ell_i \wedge \mathrm{d}\tau_i$; a full twist has period ℓ_i.
Boundary components	Labeled, unparametrized, with lengths fixed.
Volumes	Orbifold/stack volumes, including ineffective stabilizers.
Punctured-torus volume	$V_{1,1}(0) = \pi^2/12$; effective volume $\pi^2/6$.
Four-punctured-sphere volume	$V_{0,4}(0) = 2\pi^2$.
Cone parameter	Opening angle θ, represented by $L = i\theta$; the deficit is $2\pi - \theta$.
Ideal tetrahedral z	Cross-ratio of ambient ideal vertices, not intrinsic conformal punctures.
Exterior angle θ_e	An edge length in the dual metric map; an ideal vertex has angle sum 2π.
Lambda length	$e^{\delta/2}$ for signed horocycle distance δ.
Closed-curve sums	Unoriented free-homotopy classes unless stated otherwise.
Coefficient basis	$T_C = \prod_{\alpha \subset C} \operatorname{tr} \rho(\alpha)$; parallel copies are multiplicities.
Twisted Selberg sum	Unoriented primitive δ, iterate k: weight $\ell_\delta / \sinh(k\ell_\delta/2)$.
Coefficient heat transform	$H_C(t) = e^{-t/4}(4\pi t)^{-1/2} \int e^{-L^2/(4t)}\, \mathrm{d}\sigma_C(L)$.
Coefficient Laplace parameter	Kernel e^{-sL}; the product parameter is $v = s + 1/2$.
Total multicurve count	All nonzero integral laminations, including disconnected and parallel supports.
Primitive lamination	Gcd of distinct-support weights is one; connectedness is an additional condition.
Support variable z	Exponent counts distinct geodesic supports, not parallel copies.
One-crossing curves	Not proper iterates; both hyperelliptic and non-hyperelliptic torus types are included in (18.8).
Basmajian sum	Coefficient 2 for oriented orthogeodesics, 4 for unoriented arcs counted once.
Cusp orthospectrum	Simple cusp-to-cusp arcs, with both ends truncated at a specified horocycle.
Peripheral logarithms	Logs of $\mathrm{PSL}_2(\mathbb{C})$ multipliers; smooth filling has $pu + qv = 2\pi i$.
Real filling source	Adding jLx fixes conjugate source $j = \ell_{\mathrm{core}}/2$; the cone angle is Lx.
Holomorphic filling action	$F' = v$ and $\Delta V = \frac{1}{2}\Im \mathcal{S}^*$, with $\mathcal{S}$ as in (16.10).

Source conventions. Penner's logarithmic triangle form, Mondello's ordered Poisson sum, and the effective versus stack convention for $\mathcal{M}_{1,1}$ can introduce different factors if imported without conversion. The formulas here are calibrated by (1.1), the direct tetrahedral integral, and the trace calculation in Appendix A. In particular, the proportionality constants in the Delaunay-form conventions of [50] should not be substituted unchanged for those of (3.1); the metric-map volume normalization in (2.5) is the one specified by [48].

References

[1] R. A. Wentworth, *Energy of harmonic maps and Gardiner's formula*, in *In the Tradition of Ahlfors–Bers, IV*, Contemp. Math. **432** (2007), 221–229. DOI: 10.1090/conm/432/08311; Author's version.

[2] F. Arana-Herrera, *Counting square-tiled surfaces with prescribed real and imaginary foliations and connections to Mirzakhani's asymptotics for simple closed hyperbolic geodesics*, J. Mod. Dyn. **16** (2020), 81–107. DOI: 10.3934/jmd.2020004; arXiv:1902.05626.

[3] F. Arana-Herrera, *Effective count of square-tiled surfaces with prescribed real and imaginary foliations in connected components of strata*, Ergodic Theory Dynam. Systems **45** (2025), 649–662. DOI: 10.1017/etds.2024.54; arXiv:2207.04091.

[4] J. Athreya, A. Bufetov, A. Eskin and M. Mirzakhani, *Lattice point asymptotics and volume growth on Teichmüller space*, Duke Math. J. **161** (2012), 1055–1111. arXiv:math/0610715.

[5] F. Arana-Herrera, *Effective mapping class group dynamics I: Counting lattice points in Teichmüller space*, Duke Math. J. **172** (2023), 1437–1529. arXiv:2010.03123.

[6] P. Humphries, *Distributing points on the torus via modular inverses*, Quart. J. Math. **73** (2022), 1–16. DOI: 10.1093/qmath/haab015; arXiv:2003.09955.

[7] M. Pollicott and R. Sharp, *Zeta functions in higher Teichmüller theory*, Math. Z. **306** (2024), article 37. DOI: 10.1007/s00209-024-03437-4.

[8] D. Borthwick, C. Judge and P. A. Perry, *Selberg's zeta function and the spectral geometry of geometrically finite hyperbolic surfaces*, Comment. Math. Helv. **80** (2005), 483–515. DOI: 10.4171/CMH/23.

[9] S. Edwards and H. Oh, *Spectral gap and exponential mixing on geometrically finite hyperbolic manifolds*, Duke Math. J. **170** (2021), 3417–3458. DOI: 10.1215/00127094-2021-0051.

[10] A. Gorodnik and A. Nevo, *On discrepancy, intrinsic Diophantine approximation, and spectral gaps*, J. Théor. Nombres Bordeaux **36** (2024), 127–174. DOI: 10.5802/jtnb.1275; arXiv:2203.02802.

[11] R. Sharp, *Convergence of zeta functions for amenable group extensions of shifts*, in *Dynamics: Topology and Numbers*, Contemp. Math. **744** (2020), 245–257. Author's version.

[12] M. Babillot, *On the mixing property for hyperbolic systems*, Israel J. Math. **129** (2002), 61–76. Full text.

[13] D. Dolgopyat, *On decay of correlations in Anosov flows*, Ann. of Math. (2) **147** (1998), 357–390. DOI: 10.2307/121012.

[14] A. Eskin, M. Mirzakhani and A. Mohammadi, *Effective counting of simple closed geodesics on hyperbolic surfaces*, J. Eur. Math. Soc. **24** (2022), 3059–3108. DOI: 10.4171/JEMS/1144.

[15] F. Luo and R. Stong, *Lengths of simple loops on surfaces with hyperbolic metrics*, Geom. Topol. **6** (2002), 495–521. DOI: 10.2140/gt.2002.6.495.

[16] G. McShane and I. Rivin, *Simple curves on hyperbolic tori*, C. R. Acad. Sci. Paris Sér. I Math. **320** (1995), 1523–1528. arXiv:math/0005220.

[17] I. Rivin, *A simpler proof of Mirzakhani's simple curve asymptotics*, arXiv:math/0512066 (2005).

[18] I. Agol, J. Hass and W. Thurston, *The computational complexity of knot genus and spanning area*, arXiv:math/0205057, 2004 revision, Theorem 12 and Corollary 13.

[19] Y. Colin de Verdière, *Spectre du laplacien et longueurs des géodésiques périodiques. II*, Compositio Math. **27** (1973), 159–184. Numdam.

[20] L. Charles and J. Marché, *Multicurves and regular functions on the representation variety of a surface in* SU(2), Comment. Math. Helv. **87** (2012), 409–431. doi:10.4171/CMH/258.

[21] F. Naud, *Random covers of compact surfaces and smooth linear spectral statistics*, Ann.

Henri Poincaré **27** (2026), 347–373. arXiv:2209.07941.

[22] D. P. Thurston, *A positive basis for surface skein algebras*, Proc. Natl. Acad. Sci. USA **111** (2014), no. 27, 9725–9732. arXiv:1310.1959.

[23] M. Basok and D. Chelkak, *Tau-functions à la Dubédat and probabilities of cylindrical events for double-dimers and CLE(4)*, J. Eur. Math. Soc. **23** (2021), 2787–2832. doi:10.4171/JEMS/1072.

[24] S. Wolpert, *An elementary formula for the Fenchel–Nielsen twist*, Comment. Math. Helv. **56** (1981), 132–135. doi:10.1007/BF02566203.

[25] T. Aougab, M. Lahn, M. Loving, and N. Miller, *Unmarked simple length spectral rigidity for covers*, arXiv:2210.16706, revised 2023.

[26] Y. U. Taylor and C. T. Woodward, *6j symbols for $U_q(\mathfrak{sl}_2)$ and non-Euclidean tetrahedra*, preprint (2003), revised 2005. arXiv:math/0305113.

[27] V. Turaev and A. Virelizier, *On two approaches to 3-dimensional TQFTs*, 2010. arXiv:1006.3501.

[28] R. M. Kashaev, *The hyperbolic volume of knots from the quantum dilogarithm*, Lett. Math. Phys. **39** (1997), 269–275. arXiv:q-alg/9601025.

[29] H. Murakami, *An introduction to the volume conjecture*, in *Interactions between hyperbolic geometry, quantum topology and number theory*, Contemp. Math. **541** (2011), 1–40. arXiv:1002.0126.

[30] L. D. Faddeev and R. M. Kashaev, *Quantum dilogarithm*, Mod. Phys. Lett. A **9** (1994), 427–434. arXiv:hep-th/9310070.

[31] T. Dimofte and S. Garoufalidis, *The quantum content of the gluing equations*, Geom. Topol. **17** (2013), 1253–1315. arXiv:1202.6268.

[32] E. Witten, *Quantum field theory and the Jones polynomial*, Comm. Math. Phys. **121** (1989), 351–399. doi:10.1007/BF01217730.

[33] E. Witten, *Analytic continuation of Chern–Simons theory*, 2010. arXiv:1001.2933.

[34] S. Gukov, *Three-dimensional quantum gravity, Chern–Simons theory, and the A-polynomial*, 2003. arXiv:hep-th/0306165.

[35] T. Dimofte, *Quantum Riemann surfaces in Chern–Simons theory*, 2011. arXiv:1102.4847.

[36] W. D. Neumann and D. Zagier, *Volumes of hyperbolic three-manifolds*, Topology **24** (1985), no. 3, 307–332. doi:10.1016/0040-9383(85)90004-7.

[37] D. Futer and F. Guéritaud, *From angled triangulations to hyperbolic structures*, in *Interactions between hyperbolic geometry, quantum topology and number theory*, Contemp. Math. **541** (2011), 159–182. arXiv:1004.0440.

[38] Q. Chen and T. Yang, *Volume conjectures for the Reshetikhin–Turaev and the Turaev–Viro invariants*, Quantum Topol. **9** (2018), no. 3, 419–460. doi:10.4171/QT/111.

[39] D. Gang, M. Romo, and M. Yamazaki, *All-order volume conjecture for closed 3-manifolds from complex Chern–Simons theory*, Comm. Math. Phys. **359** (2018), 915–936. arXiv:1704.00918.

[40] K. H. Wong and T. Yang, *On the volume conjecture for hyperbolic Dehn-filled 3-manifolds along the figure-eight knot*, Quantum Topol. **16** (2025), no. 3, 419–491. doi:10.4171/QT/240.

[41] T. Dimofte and R. van der Veen, *A spectral perspective on Neumann–Zagier*, 2014. arXiv:1403.5215.

[42] D. V. Mathews and J. S. Purcell, *A symplectic basis for 3-manifold triangulations*, 2022. arXiv:2208.06969.

[43] D. Alessandrini, L. Liu, A. Papadopoulos, and W. Su, *The behaviour of Fenchel–Nielsen distance under a change of pants decomposition*, preprint (2011). arXiv:1105.0202.

[44] L. Anagnostou, S. Mullane, and P. Norbury, *Weil–Petersson volumes, stability conditions and wall-crossing*, preprint (2023). arXiv:2310.13281.

[45] L. Anagnostou and P. Norbury, *Volumes of moduli spaces of hyperbolic surfaces with*

cone points, preprint (2022), version 4 (2025). arXiv:2212.13701.

[46] X. Bao and F. Bonahon, *Hyperideal polyhedra in hyperbolic 3-space*, Bull. Soc. Math. France **130** (2002), no. 3, 457–491. doi:10.24033/bsmf.2426.

[47] G. Belletti and T. Yang, *Asymptotics of quantum* $6j$*-symbols and generalized hyperbolic tetrahedra*, J. Topol. **18** (2025), no. 3, e70033. doi:10.1112/topo.70033.

[48] T. Budd, *Irreducible metric maps and Weil–Petersson volumes*, Comm. Math. Phys. **394** (2022), 887–917. doi:10.1007/s00220-022-04418-6.

[49] T. Budd, T. Meeusen, and B. Zonneveld, *A tree bijection for the moduli space of genus-0 hyperbolic surfaces with boundaries*, preprint (2025). arXiv:2512.09722.

[50] S. Charbonnier, F. David, and B. Eynard, *Local properties of the random Delaunay triangulation model and topological 2D gravity*, preprint (2017), version 3 (2018). arXiv:1701.02580.

[51] Q. Chen and J. Murakami, *Asymptotics of quantum* $6j$ *symbols*, preprint (2017), revised 2021. arXiv:1706.04887.

[52] F. Costantino, *6j-symbols, hyperbolic structures and the Volume Conjecture*, Geom. Topol. **11** (2007), 1831–1854. arXiv:math/0611399.

[53] P. Deligne and G. D. Mostow, *Monodromy of hypergeometric functions and non-lattice integral monodromy*, Publ. Math. IHES **63** (1986), 5–89. Numdam original article.

[54] NIST Digital Library of Mathematical Functions, *Airy functions, hypergeometric functions, elliptic integrals, and coalescing saddles*, Sections 9.5, 15.10, 19.12, and 36.12. Airy contour integrals; Hypergeometric connection formulas; elliptic logarithmic expansions; uniform approximations of integrals.

[55] N. Do and P. Norbury, *Weil–Petersson volumes and cone surfaces*, Geom. Dedicata **141** (2009), 93–107. arXiv:math/0603406; doi:10.1007/s10711-008-9345-y.

[56] L. Eberhardt and G. J. Turiaci, *2D dilaton gravity and the Weil–Petersson volumes with conical defects*, preprint (2023). arXiv:2304.14948.

[57] F. Fillastre, *Existence and uniqueness theorem for convex polyhedral metrics on compact surfaces*, survey (2010). arXiv:1011.3123.

[58] W. M. Goldman, *Mapping class group dynamics on surface group representations*, in *Problems on mapping class groups and related topics*, Proc. Sympos. Pure Math. **74**, Amer. Math. Soc. (2006), 189–214. arXiv:math/0509114.

[59] C. Guillarmou, A. Kupiainen, R. Rhodes, and V. Vargas, *Conformal bootstrap in Liouville theory*, preprint (2020), version 2. arXiv:2005.11530.

[60] M. Heusener and J. Porti, *Holomorphic volume forms on representation varieties of surfaces with boundary*, Ann. H. Lebesgue **3** (2020), 341–380. arXiv:1804.07504.

[61] C. D. Hodgson, *Degeneration and regeneration of geometric structures on 3-manifolds*, Ph.D. thesis, Princeton University (1986), vii+169 pp. Bibliographic metadata is confirmed by the author's publication list; no numbered theorem from the thesis is used here.

[62] H. Hu and S. P. Tan, *New identities for small hyperbolic surfaces*, Bull. London Math. Soc. **46** (2014), no. 5, 1021–1031. arXiv:1402.1573; doi:10.1112/blms/bdu056.

[63] S. P. Kerckhoff, *The Nielsen realization problem*, Ann. of Math. (2) **117** (1983), 235–265. doi:10.2307/2007076.

[64] S. Kojima, *Complex hyperbolic cone structures on the configuration spaces*, Rend. Istit. Mat. Univ. Trieste **32**, suppl. 1 (2001), 149–163. arXiv:math/9907147; published article.

[65] V. Koziarz and D.-M. Nguyen, *Complex hyperbolic volume and intersection of boundary divisors in moduli spaces of pointed genus zero curves*, Ann. Sci. Ec. Norm. Super. (4) **51** (2018), 1549–1597. arXiv:1601.05252.

[66] A. Kupiainen, R. Rhodes, and V. Vargas, *Integrability of Liouville theory: proof of the DOZZ formula*, Ann. of Math. **191** (2020), 81–166. arXiv:1707.08785.

[67] D. Lei and Y. Zhang, *A note on Rivin's interpretation of McShane's identity as identities*

for closed geodesics with one self-intersection, J. Knot Theory Ramifications **32** (2023), no. 6, 2350049, 10 pp. doi:10.1142/S0218216523500499. The abstract and publisher preview were consulted for the correction; the identities used here are independently derived in Section 18.

[68] F. Luo and S. P. Tan, *A dilogarithm identity on moduli spaces of curves*, preprint (2011). arXiv:1102.2133.

[69] R. Mazzeo and H. Weiss, *Teichmüller theory for conic surfaces*, preprint (2015). arXiv:1509.07608.

[70] C. T. McMullen, *The Gauss–Bonnet theorem for cone manifolds and volumes of moduli spaces*, author's manuscript, revised 24 April 2015. Author's PDF.

[71] G. McShane, *Simple geodesics and a series constant over Teichmüller space*, Invent. Math. **132** (1998), 607–632. doi:10.1007/s002220050235.

[72] G. McShane, *The double life of pants*, draft (October 2004). Author's PDF. See the cusp-ray and gap descriptions in Sections 3, 5, and 9.

[73] M. Mirzakhani, *Simple geodesics and Weil–Petersson volumes of moduli spaces of bordered Riemann surfaces*, Invent. Math. **167** (2007), 179–222. doi:10.1007/s00222-006-0013-2.

[74] M. Mirzakhani, *Weil–Petersson volumes and intersection theory on the moduli space of curves*, J. Amer. Math. Soc. **20** (2007), no. 1, 1–23. doi:10.1090/S0894-0347-06-00526-1.

[75] M. Mirzakhani, *Growth of the number of simple closed geodesics on hyperbolic surfaces*, Ann. of Math. (2) **168** (2008), 97–125. Journal article.

[76] M. Mirzakhani, *Counting mapping class group orbits on hyperbolic surfaces*, preprint (2016). arXiv:1601.03342.

[77] G. Mondello, *Triangulated Riemann surfaces with boundary and the Weil–Petersson Poisson structure*, J. Differential Geom. **81** (2009), no. 2, 391–436. arXiv:math/0610698; doi:10.4310/jdg/1231856265.

[78] G. Mondello, *Poisson structures on the Teichmüller space of hyperbolic surfaces with conical points*, preprint (2008), version 2 (2009). arXiv:0812.1602.

[79] G. D. Mostow, *Generalized Picard lattices arising from half-integral conditions*, Publ. Math. IHES **63** (1986), 91–106. Numdam original article.

[80] H. Murakami and J. Murakami, *The colored Jones polynomials and the simplicial volume of a knot*, Acta Math. **186** (2001), 85–104. arXiv:math/9905075.

[81] J. Murakami and M. Yano, *On the volume of a hyperbolic and spherical tetrahedron*, Comm. Anal. Geom. **13** (2005), no. 2, 379–400. Author's manuscript.

[82] H. Parlier, *Geodesic and orthogeodesic identities on hyperbolic surfaces*, preprint (2020). arXiv:2004.09078.

[83] R. C. Penner, *Weil–Petersson volumes*, J. Differential Geom. **35** (1992), no. 3, 559–608. doi:10.4310/jdg/1214448257.

[84] J. Porti and H. Weiss, *Deforming Euclidean cone 3-manifolds*, Geom. Topol. **11** (2007), 1507–1538. arXiv:math/0510432.

[85] I. Rivin, *Intrinsic geometry of convex ideal polyhedra in hyperbolic 3-space*, in *Analysis, algebra, and computers in mathematical research (Luleå, 1992)*, Lecture Notes in Pure and Applied Mathematics **156**, Dekker, New York (1994), 275–291. arXiv:math/0005234.

[86] I. Rivin, *A characterization of ideal polyhedra in hyperbolic 3-space*, Ann. of Math. (2) **143** (1996), 51–70. doi:10.2307/2118652.

[87] I. Rivin, *Volumes of degenerating polyhedra—on a conjecture of J. W. Milnor*, Geom. Dedicata **131** (2008), 73–85. arXiv:math/0512065; doi:10.1007/s10711-007-9217-x.

[88] I. Rivin, *Geodesics with one self-intersection, and other stories*, Adv. Math. **231** (2012), 2391–2412. arXiv:0901.2543.

[89] J. Roberts, *Classical 6j-symbols and the tetrahedron*, Geom. Topol. **3** (1999), 21–66. arXiv:math-ph/9812013.

[90] M. Saadi and B. Zwiebach, *Closed string field theory from polyhedra*, Ann. Physics **192**

(1989), 213–227. doi:10.1016/0003-4916(89)90126-7.

[91] J.-M. Schlenker, *Métriques sur les polyèdres hyperboliques convexes*, J. Differential Geom. **48** (1998), no. 2, 323–405. doi:10.4310/jdg/1214460799.

[92] B. Springborn, *The hyperbolic geometry of Markov's theorem on Diophantine approximation and quadratic forms*, L'Enseignement Math. (2) **63** (2017), 333–373. doi:10.4171/LEM/63-3/4-5.

[93] J. Teschner, *Liouville theory revisited*, Classical Quantum Gravity **18** (2001), R153–R222. arXiv:hep-th/0104158.

[94] J. Teschner and G. S. Vartanov, *6j symbols for the modular double, quantum hyperbolic geometry, and supersymmetric gauge theories*, preprint (2012). arXiv:1202.4698.

[95] W. P. Thurston, *Shapes of polyhedra and triangulations of the sphere*, Geom. Topol. Monogr. **1** (1998), 511–549. arXiv:math/9801088.

[96] X. Wang, *The complex hyperbolic form as a Weil–Petersson form*, preprint (2022), revised 2024. arXiv:2209.10842; doi:10.1142/S0219199724500068.

[97] B. Zonneveld, *A tree bijection for cusp-less planar hyperbolic surfaces*, preprint (2025). arXiv:2512.10673.

[98] B. Zwiebach, *Consistency of closed string polyhedra from minimal area*, Phys. Lett. B **241** (1990), 343–349. doi:10.1016/0370-2693(90)91654-T.

[99] F. Benirschke and C. A. Serván, *Isometric embeddings of Teichmüller spaces are covering constructions*, Advances in Mathematics **452** (2024), 109817. Journal article. For the classical genus-two exception see Corollary 1.2(3).

[100] F. Fillastre and I. Izmestiev, *Hyperbolic cusps with convex polyhedral boundary*, Geometry & Topology **13** (2009), 457–492. doi:10.2140/gt.2009.13.457. Alexandrov's sphere realization theorem is recalled as Theorem 1.1.

[101] E. Lanneau, *Connected components of the strata of the moduli spaces of quadratic differentials*, Annales scientifiques de l'École Normale Supérieure (4) **41** (2008), 1–56. Publisher text. See Sections 1.1.1–1.1.3 for period coordinates and the hyperelliptic covering constructions.

[102] A. Strömbergsson and P. Vishe, *An effective equidistribution result for $SL(2,\mathbb{R}) \ltimes (\mathbb{R}^2)^{\oplus k}$ and application to inhomogeneous quadratic forms*, Journal of the London Mathematical Society (2) **102** (2020), 143–204. doi:10.1112/jlms.12316; arXiv:1811.10340. The fiber Fourier orbit types used here are described in Section 4.

[103] Y. Benoist and T. Kobayashi, *Tempered homogeneous spaces II*, in D. Fisher, D. Kleinbock, and G. Soifer (eds.), *Dynamics, Geometry, Number Theory: The Impact of Margulis on Modern Mathematics*, University of Chicago Press, 2022, pp. 213–245. arXiv:1706.10131. See Lemma 2.3, Remark 2.4, and Proposition 2.6 for induction, amenability, and the matrix-coefficient estimate.

[104] M. Cowling, U. Haagerup, and R. Howe, *Almost L^2 matrix coefficients*, Journal für die reine und angewandte Mathematik **387** (1988), 97–110. doi:10.1515/crll.1988.387.97.

[105] J.-P. Otal and E. Rosas, *Pour toute surface hyperbolique de genre g, $\lambda_{2g-2} > 1/4$*, Duke Mathematical Journal **150** (2009), 101–115. doi:10.1215/00127094-2009-048; arXiv:0905.0506. Theorem 2 is the finite-area statement, including residual and cuspidal eigenvalues.

[106] J. S. Athreya, J. Lagacé, M. Möller, and M. Raum, *Spectral decomposition and Siegel–Veech transforms for strata: the case of marked tori*, Journal of Spectral Theory **15** (2025), 895–959. doi:10.4171/JST/563; arXiv:2404.06597. Its one-relative-mark model provides a finer spectral comparison; the present affine model has two relative marked vectors.

[107] J. S. Athreya, *Random affine lattices*, in *Geometry, Groups and Dynamics*, Contemporary Mathematics **639** (2015), 169–174. doi:10.1090/conm/639/12793; arXiv:1409.6355. Lemmas 3 and 4 give the affine mean and variance formulas.

[108] H. Shiga, *On the hyperbolic length and quasiconformal mappings*, Complex Variables, Theory and Application **50** (2005), 123–130. doi:10.1080/02781070412331328206. Equation (1.1) recalls Wolpert's quasiconformal length inequality.

[109] I. Bárány, G. Martin, E. Naslund, and S. Robins, *Primitive points in rational polygons*, Canadian Mathematical Bulletin **63** (2020), 850–870. doi:10.4153/S0008439520000090; arXiv:1509.02201.

[110] B. H. Bowditch, *Markoff triples and quasifuchsian groups*, Proceedings of the London Mathematical Society (3) **77** (1998), 697–736. doi:10.1112/S0024611598000604.

[111] N. M. Doan, X. Li, and V. Nguyen, *Stable-norm boundary geometry on once-punctured tori*, preprint, version 3, 9 October 2026. arXiv:2605.14574v3. The local flatness refinement is discussed in Section 4.2 of that preprint.

[112] W. Zhai, *On primitive lattice points in planar domains*, Acta Arithmetica **109** (2003), 1–26. doi:10.4064/aa109-1-1. The geometric hypotheses and unconditional Mertens estimate are on pp. 1–2; the paper uses the area parameter $x = L^2$.

[113] R. C. Baker, *Primitive lattice points in planar domains*, Acta Arithmetica **142** (2010), 267–302. doi:10.4064/aa142-3-4. Theorem 1 is conditional on RH and the geometric hypotheses (H1)–(H4).

[114] OpenAI, *The Quasi-Riemann Hypothesis: A Zero-Free Half-Plane* $\Re s > 7/8$, manuscript dated 30 September 2026. Official manuscript. The claim in Theorem 1.1 is used only conditionally in this monograph; its proof and formalization have not been independently verified here.

[115] OpenAI, *The Quasi-Riemann Hypothesis*, manuscript dated 5 October 2026. Official manuscript. Theorem 1.1 claims a zero-free half-plane $\Re s > 11/12$ and is described there as an intermediate result toward the September claim.

[116] *Duffin's resistor networks, the volume of ideal polyhedra, and nearly-linear-time Laplacian solvers*, companion working study, 18 September 2026; supplied snapshot of the `igorrivin/ipoly` repository, manuscript and numerical-verification scripts.

[117] R. J. Duffin, *Distributed and lumped networks*, J. Math. Mech. **8** (1959), 793–826.

[118] I. Rivin, *Euclidean structures on simplicial surfaces and hyperbolic volume*, Ann. of Math. **139** (1994), 553–580. https://doi.org/10.2307/2118572.

[119] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004, especially Chapter 11. https://stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf.

[120] D. A. Spielman and S.-H. Teng, *Nearly-linear time algorithms for preconditioning and solving symmetric, diagonally dominant linear systems*, SIAM J. Matrix Anal. Appl. **35** (2014), 835–885. https://arxiv.org/abs/cs/0607105.

[121] S. Rippa, *Minimal roughness property of the Delaunay triangulation*, Comput. Aided Geom. Design **7** (1990), 489–497. https://doi.org/10.1016/0167-8396(90)90011-F.

[122] V. Disarlo and H. Parlier, *The geometry of flip graphs and mapping class groups*, Trans. Amer. Math. Soc. **372** (2019), 3809–3844. https://arxiv.org/abs/1411.4285.

[123] J. F. Brock, *The Weil–Petersson metric and volumes of* 3*-dimensional hyperbolic convex cores*, J. Amer. Math. Soc. **16** (2003), 495–535.

[124] J. F. Brock, *Weil–Petersson translation distance and volumes of mapping tori*, Comm. Anal. Geom. **11** (2003), 987–999. https://arxiv.org/abs/math/0109050.